

A Phase Cell Cluster Expansion for Euclidean Field Theories*

GUY A. BATTLE III[†] AND PAUL FEDERBUSH

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109

Received January 6, 1982

We adapt the cluster expansion first used to treat infrared problems for lattice models (a mass zero cluster expansion) to the usual field theory situation. The field is expanded in terms of special block spin functions and the cluster expansion given in terms of the expansion coefficients (phase cell variables); the cluster expansion expresses correlation functions in terms of contributions from finite coupled subsets of these variables. Most of the present work is carried through in d space time dimensions (for ϕ_4^4 the details of the cluster expansion are pursued and convergence is proven). Thus most of the results in the present work will apply to a treatment of ϕ_3^4 to which we hope to return in a succeeding paper. Of particular interest in this paper is a substitute for the stability of the vacuum bound appropriate to this cluster expansion (for $d=2$ and $d=3$), and a new method for performing estimates with tree graphs. The phase cell cluster expansions have the renormalization group incorporated intimately into their structure. We hope they will be useful ultimately in treating four dimensional field theories.

INTRODUCTION

Let $\phi(x)$ be a field arising in classical statistical mechanics or in Euclidean quantum field theory. For suitable expansion functions $u_k(x)$ we write

$$\phi(x) = \sum_k \alpha_k u_k(x). \quad (0.1)$$

We call the α_k "phase cell variables". A phase cell cluster expansion expresses correlation functions in terms of contributions from (finite) coupled subsets of the α_k . A step in the expansion process tries to decouple interactions between the variables already coupled and additional variables. This cluster expansion is quite different from the usual expansion that decouples interactions between regions of space time [4]. (Each such region would involve an infinite number of phase cell variables.) In [2], a phase cell cluster expansion was applied to study infrared behavior of some lattice systems. Further development in this direction should include a similar (but

* This work was supported in part by the National Science Foundation under Grant No. PHY 79-05688.

[†] On leave from Mathematics Department, Texas A & M University, College Station, Texas 77843.

much more difficult) treatment of the lattice dipole gas and massless $(\bar{\nabla}\phi)^4$ models. We have postponed this extension to treat the field theory models ϕ_2^4 and ϕ_3^4 as being possibly a more important and interesting application, and certainly an easier one.

For our particular expansion functions, presented in Section 2, we find two useful positivity properties. (The second positivity property actually depends only on the orthonormality of the ψ_i introduced in that section.) The first shown in Section 6 is a relation

$$\sum_k c(k) |\alpha_k|^{4-\epsilon} \leq \int \phi^4 + \int (\bar{\nabla}\phi)^2 \tag{0.2}$$

(holding for any $\epsilon > 0$ and suitable $c(k) > 0$). This positivity holds in any dimension d . A second positivity property is proven in Section 7, and replaces the “stability of the vacuum” bounds necessary in usual studies of $P(\phi)$ models:

$$\int : \phi^4 : \geq -cN. \tag{0.3}$$

Here ϕ and the normal ordering are constructed keeping only a finite number of modes, N , of the field. This result we prove for $d < 4$. This of course is analogous to dominating

$$\int e^{-J_V : \phi^4 :} d\mu(\phi) \tag{0.4}$$

by e^{cV} , V the volume. It is interesting as a stability bound in the number of modes, not volume, and as being true for $d = 3$. It is the only stability bound we need to study ϕ_3^4 .

In [2] a technique of “spatial sequence interpolation” was introduced to help control combinatoric estimates. Unfortunately our positivity estimate (0.2) does not survive this interpolation process; but it does, trivially, survive the usual interpolation of the total interaction. It turns out that the estimates of [2] could have been performed using the usual interpolation procedure. This hinges on a new tree graph estimate given in Section 8; to those familiar with the language [4, p. 40]:

For η an (ordered) tree graph with n vertices one has the well known inequality

$$\sum_n \int d\sigma f(\eta, \sigma) \leq e^n. \tag{0.5}$$

Let P be a permutation on the n vertices leaving the first vertex fixed. We say P is compatible with a given η , $P : \eta$, if $P^{-1}\eta P$ is also an (ordered) tree graph. Our new observation asserts, for each fixed η ,

$$\sum_{P:\eta} \int d\sigma f(P^{-1}\eta P, \sigma) = 1. \tag{0.6}$$

Professor R. W. Robinson has shown us how one may deduce from the new relation (0.6) an identity—to replace the inequality—for (0.5)

$$\sum_n \int d\sigma f(\eta, \sigma) = \frac{n^{n-1}}{n!}. \quad (0.7)$$

While this relation does not improve cluster expansion estimates, it is surprising such a simple formula has remained hidden so long!

Through the scaling properties of the $u_k(x)$, the renormalization group is intimately embedded in our treatment of field theories. Already the properties of the variables living on different scales are used in many ways in [2], and even in the treatment of ϕ_2^4 herein. The importance of looking at scales of momentum and spatial localization was first apparent in Glimm and Jaffe's original treatment of ϕ_3^4 [3]. We hope ultimately that renormalization group properties of the procedure will help handle renormalization problems with four dimensional field theories. There are glimmerings of ideas in this direction.

In structure the present work is a natural development, combining the machinery of the Glimm–Jaffe–Spencer cluster expansion [4] with renormalization group ideas. The renormalization group of Wilson and Kadanoff has been incorporated into constructive field theory through a sequence of important papers, of which we select Glimm and Jaffe's treatment of ϕ_3^4 [3], the work of Gallavotti *et al.* [7], and that of Kupiainen and Gawędzki [8]. We have been inspired by these works. It would be profitable to relate, if possible, the analysis of the Kosterlitz–Thouless transition by Fröhlich and Spencer [9] to our efforts.

1. THE SYSTEM

For a d -dimensional field theory we wish to study quantities such as

$$\int d\phi P(\phi) e^{-I(\phi)}, \quad (1.1)$$

$$I(\phi) = \frac{1}{2} \int (\bar{\nabla}\phi)^2 + \frac{M^2}{2} \int \phi^2 + V(\phi) = I_0 + V. \quad (1.2)$$

We approach this by selecting $\{\psi_k\}$ to be an orthonormal set of functions living in a volume V , a union of unit cubes, a subset of the unit lattice. These are defined in Section 2. We then set

$$\phi(x) = \sum_k \alpha_k \left(\frac{1}{\sqrt{-\Delta + M^2}} \psi_k \right) (x), \quad (1.3)$$

$$u_k(x) = \left(\frac{1}{\sqrt{-\Delta + M^2}} \psi_k \right) (x) \quad (1.4)$$

and we will require

$$M \sim 1. \quad (1.5)$$

Notice that

$$I_0 = \frac{1}{2} \sum \alpha_k^2. \quad (1.6)$$

We may restrict the sum in (1.3) so that modes below some scale (see Section 2) do not contribute. This provides an ultraviolet cutoff. Ultimately one will want to take the limit $V \rightarrow \infty$ and remove the ultraviolet cutoff.

We now define for $p(\alpha)$ a polynomial in the α_k 's:

$$[p(\alpha)] = \prod_k \left(\int d\alpha_k \right) e^{-(1/2) \sum \alpha_k^2 + \nu(\alpha)}, \quad (1.7)$$

$$Z = [1], \quad (1.8)$$

$$\langle p(\alpha) \rangle = [p(\alpha)]/Z. \quad (1.9)$$

The cluster expansion will be an expression of the form

$$\langle p(\alpha) \rangle = \sum_{A \supset D} K_A(p) Z^{A^c}/Z. \quad (1.10)$$

Here each A is a subset of the α_k containing D , the α 's in $p(\alpha)$ (the distinguished variables). Z^{A^c} is the same as Z but where only variables not in A are kept. (Variables in A are neither summed over nor integrated over.) We restrict ourselves in this paper to find estimates of convergence for (1.10) (for $d = 2$) that are independent of the cutoffs, but we do not discuss removing cutoffs.

2. THE EXPANSION FUNCTIONS

Our construction of the orthonormal set of functions $\{\psi_k(x)\}_k$ is closely patterned after the construction in [2], to which we refer the reader. There are compatible lattices of cubes in R^d , of edge size 2^r , $r = 0, -1, -2, \dots$. k is associated to a triple, $k \leftrightarrow (r_k, \gamma_k, t_k)$, where $\psi_k(x)$ lives on a cube of edge size $L_k = 2^{r_k}$, and γ_k labels which such cube it lives on. (We understand ψ_k to "live" on the smallest cube wherein it is supported.) The t_k specifies which function of the finite set living in the cube is denoted. We have fixed throughout a nonnegative integer s and a number M (each determining the other), where there are M linearly independent polynomials of degree s in R^d . The cardinality of the set of $\{t_k\}$ for given γ and r is

$$\begin{aligned} \#t &= 2^d M - M & r < 0 \\ &= 2^d M & r = 0. \end{aligned} \quad (2.1)$$

r determines the "scale" of the function: lower r value means "smaller scale."

The ψ_k may be constructed satisfying, in addition to the above properties, the following additional conditions

(I) $\{\psi_k\}_k$ form a basis.

$$(II) \quad \int \psi_k x^\alpha = 0 \quad \text{if } |\alpha| \leq s \quad \text{and} \quad r_k < 0 \quad (2.2)$$

(here $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_d^{\alpha_d}$).

(III) $\{\psi_k\}$ with $r_k < 0$ are all translates and scalings of the finite set of ψ_k living in any $r < 0$ cube. $\{\psi_k\}$ with $r_k = 0$ are all translates of the finite set of ψ_k living in any $r = 0$ cube.

$$(IV) \quad \|\psi_k\|_\infty \leq c \frac{1}{L_k^{d/2}}. \quad (2.3)$$

(V) Let $0 \leq \bar{s} \leq s + 1$ and $r < 0$; then

$$\psi_k = \sum_{|\alpha|=\bar{s}} D^\alpha G^{(\alpha)}, \quad (2.4)$$

where $G^{(\alpha)}$ lives on the same cube and

$$\|G^{(\alpha)}\|_\infty \leq c \frac{1}{L_k^{d/2}} L_k^{\bar{s}}. \quad (2.5)$$

For all these properties see [2, Sect. 3, 4].

3. BASIC NUMERICAL ESTIMATES

In this section we collect some estimates that will be useful in our applications.

Estimate 3.1. If $x \neq y$

$$\left| \left(A_\alpha \frac{1}{\sqrt{-\Delta + M^2}} \right)_{xy} \right| \leq \frac{c}{|x - y|^{|\alpha| + d - 1}}. \quad (3.1)$$

Estimate 3.2. If $\hat{d} \geq L_k$

$$\left| \frac{1}{\sqrt{-\Delta + M^2}} \psi_k \right| \leq c \left(\frac{L_k}{\hat{d}} \right)^{d+s} \frac{1}{L_k^{d/2-1}}. \quad (3.2)$$

Here \hat{d} is the distance to the cube on which ψ_k lives.

Estimate 3.3. If $\hat{d} \leq L_k$

$$\left| \frac{1}{\sqrt{-\Delta + M^2}} \psi_k \right| \leq c \frac{1}{L_k^{d/2-1}}. \quad (3.3)$$

We now wish to study

$$I = \int |u_1(x) u_2(x) u_3(x) u_4(x)|, \tag{3.4}$$

where we are abbreviating $u_{k_i} = u_i$. We introduce some notation. We let d_{ij} be the distance between the centers of the cubes on which ψ_i and ψ_j live. We also set

$$\begin{aligned} h_r(x) &= 1 & x \leq 1 \\ &= \frac{1}{x^r} & x \geq 1. \end{aligned} \tag{3.5}$$

Estimate 3.4. Assume

$$L_1, L_2, L_3 \geq L_4;$$

then if s is large enough

$$I \leq cL_4^d \prod_i \left(\frac{1}{L_i^{d/2-1}} \right) \prod_{j \neq 4} h_{s/4} \left(\frac{d_{j4}}{L_j} \right). \tag{3.6}$$

Estimate 3.5. Let $D = \sqrt{-\Delta + M^2}$ and $j < 1$; then

$$|D^j \psi_k| \leq cL_k^{-d/2-j}, \quad \hat{d} \leq L_k, \tag{3.7}$$

$$|D^j \psi_k| \leq c \left(\frac{L_k}{\hat{d}} \right)^{d+s+1+j} L_k^{-d/2-j}, \quad \hat{d} \geq L_k. \tag{3.8}$$

4. OPERATIONS

In constructing the cluster expansions, only two basic operations are employed: interpolation and integration by parts.

Interpolation.

$$e^{v(1)} = e^{v(0)} + \int_0^1 ds \frac{d}{ds} e^{v(s)}. \tag{4.1}$$

Integration by parts.

$$\int d\alpha e^{-(1/2)\alpha^2} \alpha P(\alpha) = \int d\alpha e^{-(1/2)\alpha^2} P'(\alpha). \tag{4.2}$$

Developing the cluster expansion in [2] required only (4.1); for ϕ_2^4 and ϕ_3^4 , (4.2) is used to exhibit renormalization cancellations, including Wick ordering. We conjecture only (4.1) will be used to treat four dimensional field theories whose renormalization terms will be cancelled "implicitly."

5. INTERPOLATION OF THE INTERACTION

In this section we discuss interpolating the interaction for ϕ_2^4 . (The interpolation for ϕ_3^4 , in particular for the mass counterterm, will be treated quite analogously.) We write (abbreviating k_i as i)

$$\lambda \int \phi^4 = \sum_{1,2,3,4} w(1, 2, 3, 4) \alpha_1 \alpha_2 \alpha_3 \alpha_4, \tag{5.1}$$

where w is symmetric in all four variables. We let V be the normal ordered expression corresponding to (5.1).

$$\begin{aligned} V &= \sum_{1,2,3,4} v(1, 2, 3, 4) \\ &= \sum_{1,2,3,4} w(1, 2, 3, 4) [\alpha_1 \alpha_2 \alpha_3 \alpha_4 - 6\delta_{34} \alpha_1 \alpha_2 + 3\delta_{12} \delta_{34}]. \end{aligned} \tag{5.2}$$

We describe the interpolations iteratively as in [2]. At the onset we have

$$V = V(s, 0)$$

where this term has no actual s dependence. Previous to the n th step we have (for $v(1)$ in (4.1))

$$V(s, n - 1),$$

where here $s = (s_1, s_2, \dots, s_{n-1})$. At the n th step we will have

$$\hat{V}(s, n - 1)$$

(again with $s = (s_1, \dots, s_{n-1})$) for the $v(0)$ term in (4.1) and for $v(s)$ in (4.1)

$$V(s, n) = s_n V(s, n - 1) + (1 - s_n) \hat{V}(s, n - 1); \tag{5.3}$$

we need specify the construction of $\hat{V}(s, n - 1)$. We consider $V(s, n - 1)$; it is a linear combination of terms $W(S)$:

$$W(S) = \sum_{1,2,3,4 \in S} v(1, 2, 3, 4). \tag{5.4}$$

At the onset of step n , we are given two sets of variables, I^n and O^n , with zero intersection and union the total set of variables. These sets will be specified in Section 9. $\hat{V}(s, n - 1)$ is derived from $V(s, n - 1)$ by replacing $W(S)$ by

$$W(S \cap O^n) + W(S \cap I^n) \tag{5.5}$$

for each such term $W(S)$.

One should observe that the interpolation procedure ensures that the following properties hold for the fourth order terms in $V(s, n)$:

(I.P.1) These terms are manifestly positive. To see this we note that the fourth order terms in $W(S)$ are exactly

$$\int \left(\sum_{i \in S} \alpha_i u_i(x) \right)^4. \quad (5.6)$$

(I.P.2) Fourth order terms in $V(s, n)$ that contain at least one α from O^n and at least one α from I^n contain a factor s_n . (An analogous statement holds for the second order terms in $V(s, n)$ when one considers contributions to the normal ordering constant as well as α 's.)

(I.P.3) The fourth order terms in $V(s, n)$ are a *convex* linear combination of expressions of the form

$$\sum_j \int \left(\sum_{i \in S_j} \alpha_i u_i \right)^4 \quad (5.7)$$

with $\cup_j S_j$ the total set of $\{\alpha\}$.

6. α -POSITIVITY

This section is devoted to a proof of the following theorem.

THEOREM 6.1 (α -Positivity). *For any $\varepsilon > 0$ (and any d) there are $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_1 < \varepsilon$, $\varepsilon_2 < \varepsilon$, and $c > 0$, such that*

$$\int \phi^4 + \int (\bar{\nabla} \phi)^2 \geq c \sum_k |\alpha_k|^{4-\varepsilon_1} L_k^{4-d+\varepsilon_2}. \quad (6.1)$$

We define

$$D = \sqrt{-\Delta + M^2} \quad (6.2)$$

and as a first step in the proof of (6.1) present an interpolation inequality

PROPOSITION 6.2. *For $0 < j < 1$ one has*

$$|D^j u|_r \leq c |u|_4^{1-j} |Du|_2^j, \quad (6.3)$$

where

$$\frac{1}{4} + \frac{j}{4} = \frac{1}{r}. \quad (6.4)$$

This is a standard interpolated Sobolev inequality. It may be proved using the three lines theorem [6] and Theorem 2 from [1].

From the numerical inequality, for $x, y \geq 0$

$$xy \leq \frac{x^\alpha}{\alpha} + \frac{y^\beta}{\beta},$$

where $1/\alpha + 1/\beta = 1$, $\alpha, \beta > 1$. One deduces from (6.3)

PROPOSITION 6.3.

$$\int (D^j \phi)^r \leq c \left(\int \phi^4 + \int (\bar{\nabla} \phi)^2 \right) \tag{6.5}$$

with r and j as in Proposition 6.2.

The remainder of the proof is an application of the Riesz–Thorin convexity theorem [6] to a mapping we now define. We let T be the mapping from functions $D^j \phi$ on R^d to sequences labelled by $\{k\}$ of Section 2.

$$T: D^j \phi \rightarrow \left\{ \frac{1}{L_k^{j+d/2-1}} \alpha_k \right\}_k = \{w_k \alpha_k\}_k, \tag{6.6}$$

where of course

$$\alpha_k = \int (D^j \phi)(D^{1-j} \psi_k). \tag{6.7}$$

For the measure on R^d we take Lebesgue measure; and for the measure on the sequence space, a mass $m_k = L_k^{d+\epsilon_3}$ at point k ($\epsilon_3 > 0$).

PROPOSITION 6.4. T is bounded as a mapping from L_∞ to l_∞ .

This follows easily from the dimensional inequality

$$\int |D^{1-j} \psi_k| \leq c L^{j+d/2-1}. \tag{6.8}$$

A slightly less trivial result supplies the other bounded mapping we need to apply the convexity theorem.

PROPOSITION 6.5. T is bounded as a mapping from L_1 to l_1 .

This will follow immediately provided we know

$$\sum_k m_k w_k |D^{1-j} \psi_k| \in L_\infty. \tag{6.9}$$

This follows from (3.7) and (3.8) with straightforward estimation. We deduce Theorem 6.1 from Proposition 6.2, Proposition 6.3, and the convexity theorem yielding $T: L_r \rightarrow l_r$ bounded. Picking j close enough to 4 and ε_3 close enough to zero, one can make ε_1 and ε_2 arbitrarily small.

Observation 6.6. Proposition 6.1 remains true with $\int \phi^4$ replaced by the fourth order terms in the interpolated interaction.

This follows from Proposition 6.1 and I.P.3 of Section 5.

7. α -STABILITY

We begin this section with a proposition and its proof, from which α -stability will quickly follow.

PROPOSITION 7.1. *For $d < 4$ there is a constant c , such that for any N , if $\{\psi_i(x)\}$ are N orthonormal functions on R^d , and $u_i(x)$ is defined as*

$$u_i(x) \equiv \left(\frac{1}{\sqrt{-\Delta + 1}} \psi_i \right) (x) \tag{7.1}$$

then one has

$$\sum_i \sum_j \int u_i^2(x) u_j^2(x) d^d x \leq cN. \tag{7.2}$$

We do not know if the result is true for $d = 4$; it is not true for $d > 4$. For simplicity we carry through the proof in an arbitrarily large box Ω with periodic boundary conditions. In $L_2(\Omega) \times L_2(\Omega) = \mathcal{H}$ we let

$$\phi(x, y) = \sum_{i=1}^N \psi_i(x) \psi_i(y). \tag{7.3}$$

We let K be an integral operator on this space, \mathcal{H} , with kernel

$$k(x_1, x_2, x_3, x_4) = \int dx \prod_{i=1}^4 \left(\frac{1}{\sqrt{-\Delta + 1}} (x, x_i) \right). \tag{7.4}$$

With this notation the left side of (7.2) may be written as

$$\langle \phi, K\phi \rangle. \tag{7.5}$$

We note that $|\phi| = \sqrt{N}$, so that Proposition 7.1 will follow if we know K is a bounded operator. The operator K is symmetric, and preserves momentum; we need only show that its restriction to constant momentum subspaces is bounded uniformly. For $d < 4$ these restrictions are represented by Hilbert–Schmidt kernels with norm

easily estimated (in Fourier transform space) as bounded uniformly with respect to the channel and the volume.

THEOREM 7.2 (α -Stability). *Let $d < 4$, and c_1 be fixed; then there is a c , such that for any N , if ϕ is constructed with the N modes $\{\psi_i\}_{1 \leq i \leq N}$ then*

$$\int d^d x (\phi^4(x) - c_1 \phi^2(x) c(x, x)) \geq -cN, \tag{7.6}$$

where

$$c(x, x) = \sum_{i=1}^N u_i(x) u_i(x) \tag{7.7}$$

in the notation of (7.1).

This follows from Proposition 7.1 upon completing the square in (7.6). The relation between ϕ^4 , or V of (5.2), and (7.6), (7.7) is clear.

8. CONNECTIVITY STRUCTURE AND TREE GRAPHS

In this section we present a procedure for studying the cluster expansion of a simplified system. It may be immediately adapted to deal with more complicated systems as that of the present paper. For estimation purposes it replaces the spatial sequence interpolation of [2]. As an alternative to the usual tree graph estimation procedure of cluster expansions, it is more powerful and flexible. This section may be read independently of the rest of the paper.

For our simplified system involving a set A of variables $\{\alpha_a\}$ we define

$$\begin{aligned} I &= I_0 + V \\ &= \frac{1}{2} \sum \alpha_a^2 + \frac{1}{2} \sum_{a \neq b} v_{a,b} \alpha_a \alpha_b \end{aligned} \tag{8.1}$$

and specify a cluster expansion for a polynomial in the single variable α_0 :

$$\langle p(\alpha_0) \rangle = \sum_{\substack{A \\ \alpha_0 \in A \\ A \subset \Lambda}} K_A(p) Z^{A^c} / Z. \tag{8.2}$$

The cluster expansion has been developed by attempting to decouple variables one at a time

$$K_A(p) = \sum_G \prod_{a \in A} \left(\int_{-\infty}^{\infty} d\alpha_a \right) \prod_{i=1}^{N-1} \left(\int_0^1 ds_i \right) p(\alpha_0) f_G(s) B_G(v) e^{-I_0 - V_G(s)}. \tag{8.3}$$

Here A is a subset of N elements containing α_0 . G is an “ordered connectivity graph,” an object we now define.

DEFINITION 8.1. An *ordered connectivity graph* is a mapping G of the integers $2, 3, \dots, N$ into ordered pairs of elements of A , where with $G(i) = (g_1(i), g_2(i))$, the mapping must satisfy

$$(1) \quad g_2(i) \neq g_2(j) \quad \text{if } i \neq j \quad (8.4)$$

and

$$(2) \quad g_1(i) \in \left(\bigcup_{j < i} \{g_2(j)\} \right) \cup \{\alpha_0\}. \quad (8.5)$$

We notice that G states what interaction terms are differentiated down from the exponential during interpolation, and in what order.

Continuing to define terms in (8.3),

$$B_G(v) = \prod_{i=2}^n (-v_{g_1(i), g_2(i)} \alpha_{g_1(i)} \alpha_{g_2(i)}). \quad (8.6)$$

We set $g_2(1) = \alpha_0$ and define a mapping of integers $2, 3, \dots, N$ into $1, 2, \dots, N-1, \eta$, by

$$j = \eta(i) \quad \text{if } g_1(i) = g_2(j), \quad j < i. \quad (8.7)$$

$\eta = \eta_G$ depends on the ordered connectivity graph G and is itself an “(ordered) tree graph.”

DEFINITION 8.2. An *(ordered) tree graph* is a mapping η from $2, 3, \dots, N$ to $1, 2, \dots, N-1$ that satisfies

$$\eta(i) < i. \quad (8.7a)$$

In terms of η_G we have expressions for $f_G(s)$ and $V_G(s)$ from (8.3),

$$f_G(s) = \prod_{i=2}^N (s_{\eta_G(i)} \cdots s_{i-3} \cdot s_{i-2}), \quad (8.8)$$

where we understand the term in parentheses to be 1 if $\eta(i) = i-1$. And

$$V_G(s) = \sum_{1 \leq i < j \leq N} (s_i \cdots s_{j-2} \cdot s_{j-1}) v_{g_2(i), g_2(j)} \alpha_{g_2(i)} \alpha_{g_2(j)}. \quad (8.9)$$

Notice that $f_G = f_{G'}$ if $\eta_G = \eta_{G'}$, so we will write $f(\eta, s)$. The usual method of using tree graphs to estimate (8.3) is based on the inequality [4, p. 40]

$$\sum_{\eta} \prod_i \left(\int_0^1 ds_i \right) f(\eta, s) \leq e^N. \tag{8.10}$$

We take another route. We define a type of graph, an “unordered connectivity graph,” T , that specifies which interactions have been differentiated down from the exponential (to form $B_G(v)$) but not the order in which they were differentiated down.

DEFINITION 8.3. An *unordered connectivity graph* is just the range of some ordered connectivity graph.

Thus, an unordered connectivity graph is a subset of $A \times A$ such that no two ordered pairs have the same second element, α_0 can appear only as a first element, and every other member of A appearing as a first element must also appear as a second. If one regards the ordered pairs of points as line segments connecting them, one obtains a connected graph containing α_0 .

We note that each G specifies an (ordered) tree graph $\eta = \eta_G = P_{r_1}(G)$ and an unordered connectivity graph $T = T_G = P_{r_2}(G)$. The following simple identity holds, application of which replaces the use of (8.10).

$$\sum_{\substack{G \\ P_{r_2}(G) = T}} \prod_i \left(\int_0^1 ds_i \right) f_G(s) = 1. \tag{8.11}$$

We use (8.11) to estimate (8.3) as follows:

$$|K_A(p)| \leq \sum_T \sup_{P_{r_2}(G) = T} \sup_{s_i} \prod_a \left(\int_{-\infty}^{\infty} d\alpha_a \right) e^{-I_0 - V_G(s)} |p(\alpha_0) B_T(v)| \tag{8.12}$$

(noting that $B_G(v)$ depends only on $P_{r_2}(G) = T$). The power of (8.12) lies in its ability to be generalized to more complicated systems and interpolation procedures, such as in [2], and in Section 10. (The reader familiar with [2] may recognize close analogies between Representations 2 and 3, and numerical estimates therein, and ordered connectivity graphs and estimate (8.12) here.)

We mention again here that R. W. Robinson has shown us that (8.11) may be used to prove

$$\sum_{\eta} \prod_i \left(\int_0^1 ds_i \right) f(\eta, s) = \frac{n^{n-1}}{n!}, \tag{8.13}$$

replacing (8.10) by an equality!

In Appendix A we provide proofs of (8.11) and (8.13).

9. THE CLUSTER EXPANSION FOR ϕ_2^4

In this section we define our phase cell cluster expansion for ϕ_2^4 , and state the main convergence theorem. In succeeding sections we will prove the convergence theorem. Our construction of the cluster expansion is by an inductive process, a very straightforward use of the operations in Section 4. In succeeding sections we will have alternate descriptions of the cluster expansion more useful for combinatoric estimates, and, some may feel, more elegant and abstract. (We never use the symmetry of the interaction; ϕ^3 terms could easily be accommodated.)

We proceed to describe the expansion of an expectation and present some notations. With use of the operations, interpolation and integration by parts, terms are differentiated down from the exponent. Referring to (5.1) and (5.2), these terms will be one of the forms

$$\begin{aligned}
 a(1, 2, 3, 4) &= w(1, 2, 3, 4) \alpha_1 \alpha_2 \alpha_3 \alpha_4, \\
 b(1, 2, 3) &= w(1, 2, 3, 4) \alpha_1 \alpha_2 \delta_{34}, \\
 \dot{a}(1, 2, 3, 4) &= w(1, 2, 3, 4) \alpha_2 \alpha_3 \alpha_4, \\
 \dot{b}(1, 2, 3) &= w(1, 2, 3, 4) \alpha_2 \delta_{34}, \\
 c(1, 2, 3, 4) &= w(1, 2, 3, 4) \delta_{12} \delta_{34},
 \end{aligned}
 \tag{9.1}$$

each possibly multiplied by a numerical factor and a polynomial in the s 's. Any of the six types of terms we will call a *unit*. Each use of an operation from Section 4 will be called a *move*, and the "moves" associated to an "interpolation" will be called *steps*. A unit is said to be *introduced* at the "move" in which it is differentiated down. Each of the variables 1, 2, 3, 4 occur in the "units" in (9.1). Those cases in which the corresponding a is also present are *real* "occurrences" (such as 1 in $a(1, 2, 3, 4)$). Those cases in which the corresponding a is not present are *ghost* "occurrences" (such as 3 in $b(1, 2, 3)$ or 1 in $\dot{a}(1, 2, 3, 4)$). A variable may have real and ghost occurrences in the same unit.

Our construction of $K_A(p)$ is as a sum of many terms. A single term will correspond to a unique sequence of operations and units. Thus, say, in an interpolation, a sum of many units is differentiated down from the exponent. Each such unit will appear in a different term; the inductive construction of the cluster expansion branches into a different term for each choice of a unit. (In Section 10 we will introduce notation for describing the sequence of units and moves.) At the onset of any move the *interior* variables are the union of the distinguished variables with all variables occurring in units previously introduced. The *exterior* variables are the complementary set.

When we later specify which variables are integrated by parts, we will allow a variable to be integrated by parts only if it appears for the first time as an interior variable in the last unit introduced. We now divide all possible outcomes of a move into classes.

Class 1. At an (interpolation) step the term $e^{v(0)}$ is selected in (4.1).

Class 2. At an (interpolation) step a unit is introduced.

Class 3. At a move in which integration by parts takes place a unit is introduced.

Class 4. At a move in which integration by parts takes place one differentiates a variable that has been differentiated down (rather than an appearance of the variable in the exponent).

The result of a Class 1 move is a *completed* term; the result of all other moves are *remainder* terms. $K_A(p)$ (more properly $K_A(p) Z^{A^c}$) is the sum of all completed terms whose interior variables at the last move are the set A . The result of a Class 4 move is to remove two identical α 's from the last unit introduced; such terms will be involved in the normal ordering renormalization. A specific term in the cluster expansion will involve a specification at each move of the class, of the choice of unit in the Class 2 and 3 situations, and the variable integrated by parts in the Class 3 and Class 4 situations. Such an allowable sequence of moves is called a *path*. $K_A(p)$ will be written as a sum of terms each labelled by a path.

From Section 5 we know how to define the interpolation of interactions, and thus how to perform a move corresponding to interpolation. To complete our specification we need only state when to integrate by parts, and which variables to integrate by parts.

Rules for integration by parts. A variable α is integrated by parts at a move if

- (1) It had a real occurrence in the unit, u , introduced at the previous move, and
- (2) it was an exterior variable at the onset of the previous move, and
- (3) it is smaller scale than any variables occurring in u that were interior variables at the onset of the previous move, and
- (4) no other variables occurring in u are smaller scale than α , and
- (5) no other variable occurring in u has been integrated by parts giving rise to a Class 3 move.

If more than one α meets the requirements of the rule, only one is chosen to be integrated by parts at the given move.

MAIN THEOREM. *There is an s_0 , such that given any $s \geq s_0$ and any $c > 0$, there is a $\lambda_0(s, c) > 0$, such that*

$$\sum_A |K_A(p)| e^{c|A|} \leq c_1(p) \quad (9.2)$$

provided $0 \leq \lambda \leq \lambda_0$.

Here $|A|$ is the number of elements in A , and s is the parameter used in the construction of the ψ_k . λ is defined in (5.1).

A bound on Z^{A^c}/Z , needed for convergence of the cluster expansion, (1.10), may be obtained by the standard procedure, p. 335 of [4].

10. INTEGRATING BY PARTS AND RENORMALIZATION

Having given a complete verbal description of the expansion procedure, we now need to automate it with notation, formulae, and convenient terminology.

Let \mathcal{K} denote the set of triples that label the cells introduced in Section 2, where an arbitrary cutoff in both scale and volume is understood to be in effect. Thus \mathcal{K} will be a finite set, but of course we must obtain estimates that are independent of the cutoff. For convenience we introduce an arbitrary but fixed linear ordering on \mathcal{K} that is lexicographic with respect to scale; i.e., if k precedes k' , then $L_{k'} \leq L_k$. We will often call the elements of \mathcal{K} *vertices* because they label the cell variables α_k upon which our expansion is based and will therefore appear as vertices in our graphs.

As we indicated in Section 5, the ϕ_2^4 interaction with the above cutoff is given by

$$\lambda \int : \phi(x)^4 : dx = \sum_{1,2,3,4} w(1, 2, 3, 4) : \alpha_1 \alpha_2 \alpha_3 \alpha_4 :, \tag{10.1}$$

where we still abbreviate the vertex variable k_i with i . Since each interaction term is symmetric with respect to the vertices, we collect identical terms by disregarding order but counting multiplicities. Given a quadruple $(1, 2, 3, 4)$ we let $\tau_{1,2,3,4}$ be the weight function on \mathcal{K} such that $\tau_{1,2,3,4}(k)$ is merely the number of times that k appears in $(1, 2, 3, 4)$. Clearly, the terms labeled by $(1, 2, 3, 4)$ and $(1', 2', 3', 4')$ are identical if and only if $\tau_{1,2,3,4} = \tau_{1',2',3',4'}$. For an arbitrary nonnegative-integer-valued function τ on \mathcal{K} such that

$$\sum_{k \in \mathcal{K}} \tau(k) = 4 \tag{10.2}$$

we let $M(\tau)$ denote the number of quadruples $(1, 2, 3, 4)$ for which $\tau_{1,2,3,4} = \tau$. Thus, $M(\tau) = 4!$ if $\tau(k) \leq 1$, $M(\tau) = 2 \cdot 3!$ if τ assumes the value 2 once and only once, etc. We refer to nonnegative-integer-valued functions τ on \mathcal{K} satisfying (10.2) as *elementary graphs*, and we set

$$g(\tau) \equiv M(\tau) w(1, 2, 3, 4), \quad \tau_{1,2,3,4} = \tau, \tag{10.3}$$

which is well-defined by the symmetry of w . Hence, we may rewrite (10.1) as

$$\lambda \int : \phi(x)^4 : dx = \sum g(\tau) : \alpha^\tau :, \tag{10.4}$$

where τ ranges over elementary graphs and we have adopted the familiar multi-index notation

$$\alpha^\tau \equiv \prod_{k \in \mathcal{K}} \alpha_k^{\tau(k)}. \tag{10.5}$$

We may check using integration by parts (4.2) and the definition of normal ordering (5.2) the following generalized integration by parts formulas.

$$[P(\alpha) : \alpha_1 \alpha_2 \alpha_3 \alpha_4 :] = \sum_{\tau} g(\tau) \left[: \alpha_2 \alpha_3 \alpha_4 : P(\alpha) \frac{\partial}{\partial \alpha_1} : \alpha^\tau : \right], \tag{10.6}$$

$$[P(\alpha) : \alpha_1 \alpha_2 \alpha_3 :] = \sum_{\tau} g(\tau) \left[: \alpha_2 \alpha_3 : P(\alpha) \frac{\partial}{\partial \alpha_1} : \alpha^\tau : \right]. \tag{10.7}$$

The notation suppresses the fact that $g(\tau)$ may actually involve an interpolated interaction; and we have assumed $P(\alpha)$ does not depend on α_1 , the only situation in which we employ these formulas. One should note that the derivation of (10.6), (10.7) may involve a Class 4 move as well as a Class 3 move, if there are repetitions in the set of normal ordered α 's on the left side of the equation. Keeping units in combinations involving normal ordered expressions, from now on we will need only (10.6), (10.7) (and not (4.2)); we will not use the terms Class 4 moves or ghost variables henceforth. One is merely combining certain terms as defined in Section 9, a very natural procedure that does not do violence to any of our concepts, in particular, the definition of interior and exterior variables, Class 1, 2, or 3 moves.

If τ is an elementary graph, then the support of τ is its set of vertices. For a given τ , we denote its last vertex (with respect to the order on \mathcal{K}) by k_τ , and set $L_\tau = L_{k_\tau}$, $\alpha_\tau = \alpha_{k_\tau}$. We also set δ_τ to be the delta function on \mathcal{K} located at k_τ . We will use (10.6), (10.7) in the form

$$[P(\alpha) : \alpha^\tau :] = \sum_{\tau'} g(\tau') \left[: \alpha^{\tau - \delta_\tau} : P(\alpha) \frac{\partial}{\partial \alpha_\tau} : \alpha^{\tau'} : \right] \tag{10.8}$$

or

$$\left[P(\alpha) \frac{\partial}{\partial \alpha_{\bar{k}}} : \alpha^\tau : \right] = \sum_{\tau'} g(\tau') \left[\frac{\partial}{\partial \alpha_{\bar{k}}} : \alpha^{\tau - \delta_\tau} : P(\alpha) \frac{\partial}{\partial \alpha_\tau} : \alpha^{\tau'} : \right], \tag{10.9}$$

where $L_\tau < L_{\bar{k}}$.

If the remainder term of a Class 2 move is naturally written as the left side of (10.8) or the remainder term of a Class 3 move is written as the left side of (10.9) we may restate the rules for integration by parts of Section 9 as follows:

Rules for integration by parts (new form). Employ integration by parts, (10.8) or (10.9), on a given remainder term if

- (1) α_τ was an exterior variable previous to the last move.
 (2) α_τ is strictly smaller scale than any variables in α^τ that were interior variables at the onset of the previous move.

11. REPRESENTATION ONE GRAPHS

Let $p(\alpha)$ be an arbitrary polynomial in the cell variables. The problem is to control the expectation $\langle p(\alpha) \rangle$ with our cluster expansion, and our first step is to write the expansion in terms of graphs that mirror the expansion procedure. Such graphs will be called *representation one graphs*.

Combining the terminology introduced in Section 9, with that of Section 10, we note that the distinguished vertices are the vertices whose variables appear in $p(\alpha)$, the interior vertices at a given point in the expansion process are the distinguished vertices and the vertices of those elementary graphs that have been differentiated down, and the exterior vertices are the vertices that are in the complementary set. Since \mathcal{X} is a finite set, the expansion procedure must terminate because the interior set grows with at least half the moves that choose a remainder term and so a completed (decoupled) term must be chosen sooner or later when one builds a path of terms. We wish to describe the end result of our process, which is a complicated expression for $\langle p(\alpha) \rangle$ involving only completed terms. (Our formula (1.10) is only schematic and we have to be more explicit.) The elementary graphs are the key to our description, and we define types of composite graphs in analogy to what is done in Section 8. Our definitions are more complicated, of course, but the idea is the same: the completed terms are labeled by histories that obey the rules of the expansion process.

DEFINITION 11.1. An *m*th-degree chain ($m > 1$) is a sequence $C \equiv (C^1, \dots, C^m)$ of elementary graphs such that

$$k_{Ci} \in \text{supp } C^{i+1}, \quad 1 \leq i < m, \quad (11.1)$$

$$L_{Ci} < L_{C^{i-1}}, \quad 1 < i < m. \quad (11.2)$$

Chains are designed to describe sequences of Class 3 moves. We adopt the abuse of notation $\text{supp } C \equiv \bigcup_{i=1}^m \text{supp } C^i$ for the next definition. Note first, however, that (11.2) implies

$$k_{Ci} \notin \bigcup_{j < i} \text{supp } C^j, \quad 1 \leq i < m, \quad (11.3)$$

and that (11.1) implies $L_{C^m} \leq L_{C^{m-1}}$. If $L_{C^m} = L_{C^{m-1}}$, we say that C is *type 1*; otherwise, C is *type 2*.

DEFINITION 11.2. Let \mathcal{D} be the set of distinguished vertices. A *representation one*

graph rooted on \mathcal{D} is a finite sequence $G \equiv (G_1, \dots, G_N)$ of chains and elementary graphs such that the following conditions are satisfied:

- (a) If G_l is an elementary graph, then $\text{supp } G_l$ is not contained in $\mathcal{D} \cup \bigcup_{j < l} \text{supp } G_j$ and the intersection contains a vertex whose scale is equal to L_{G_l} .
- (b) If G_l is a m th-degree chain, then $\text{supp } G_l^1$ meets $\mathcal{D} \cup \bigcup_{j < l} \text{supp } G_j$ and for $1 \leq i < m$ the intersection

$$\text{supp } G_l^i \cap \left(\mathcal{D} \cup \bigcup_{j < l} \text{supp } G_j \right)$$

contains only vertices whose scales are strictly greater than $L_{G_l^i}$. In particular, (11.3) is supplemented with

$$k_{G_l^i} \notin \mathcal{D} \cup \bigcup_{j < l} \text{supp } G_j, \quad 1 \leq i < m. \quad (11.4)$$

- (c) If G_l is an m th-degree chain of type 2, then the intersection

$$\text{supp } G_l^m \cap \left(\mathcal{D} \cup \bigcup_{j < l} \text{supp } G_j \right)$$

contains a vertex whose scale is equal to $L_{G_l^m}$.

Remark. Condition (a) represents a decision: if a Class 2 move yields an elementary graph violating condition (a) and therefore requiring the next move to be Class 3, then we count that elementary graph as the first one in the chain representing the subsequent Class 3 moves. We make this choice because we want to associate the interpolation parameter of the Class 2 move with the whole chain. Conditions (b) and (c) are just the constraints our expansion procedure has imposed on the history represented by G . Note that condition (c) is concerned with the possibility that a Class 3 move can bring down an elementary graph whose scale is less than that of its predecessor but which contains a smallest scale vertex that was interior before the chain was triggered, i.e., an “old variable.” Such an event terminates the chain and requires the next move to be Class 1 or 2.

For $1 \leq l \leq N$ we define ε_G^l as follows:

- (a) If G_l is an elementary graph we define

$$\begin{aligned} \varepsilon_G^l &= 1, & \text{supp } G_l \text{ meets } \mathcal{D}, \\ &= \min\{j \mid \text{supp } G_l \text{ meets } \text{supp } G_{j-1}\}, & \text{otherwise.} \end{aligned} \quad (11.5)$$

- (b) If G_l is an m th degree chain then ε_G^l is a mapping from $\{1, \dots, m\}$ to $\{1, \dots, l\}$, where $\varepsilon_G^l(i)$ is defined as in (11.5) but with G_l replaced by G_l^i .

The ε_G^l are designed to keep track of the interpolation parameters that are brought down from the exponent by the moves. For example, if G_l is a chain, then for $i > 1$ the Class 3 move which differentiates down the interaction term associated with the elementary graph G_l^i also brings down the product $s_l s_{l-1} \cdots s_{\varepsilon_G^l(i)}$, where s_j denotes the interpolation parameter associated with the j th interpolation step. To collect this information we define the functions $b_G^l(s)$ as follows:

- (a) If G_l is an elementary graph, then

$$b_G^l(s) = s_{l-1} s_{l-2} \cdots s_{\varepsilon_G^l}, \quad (11.6)$$

where, as is standard, empty products are taken to be 1.

- (b) If G_l is an m th-degree chain, then

$$b_G^l(s) = (s_{l-1} s_{l-2} \cdots s_{\varepsilon_G^l(1)}) \prod_{i=2}^m (s_l s_{l-1} \cdots s_{\varepsilon_G^l(i)}). \quad (11.7)$$

To indicate the powers of α introduced in the various moves we define $B_G^l(\alpha)$ as follows:

- (a) If G_l is an elementary graph, then $B_G^l(\alpha) = : \alpha^{G_l} :$.
 (b) If G_l is an m th-degree chain, then

$$B_G^l(\alpha) = : \alpha^{\tau_1 - \delta_1} : \left(\prod_{i=2}^{m-1} \frac{\partial}{\partial \alpha_{i-1}} : \alpha^{\tau_i - \delta_i} : \right) \frac{\partial}{\partial \alpha_{m-1}} : \alpha^{\tau_m} :, \quad (11.8)$$

where for notational expediency we have set $\tau_i = G_l^i$, $k_i = k_{\tau_i}$, $\alpha_i = \alpha_{k_i}$, and $\delta_i = \delta_{k_i}$.

We are now ready to write the expectation in terms of representation one graphs. Let \tilde{G} be the sequence of elementary graphs induced by G if we ignore the chain structure, and let A_G be the set of interior vertices for the completed term labeled by G .

THEOREM 11.3.

$$\begin{aligned} \langle p(\alpha) \rangle &= \sum_G \prod_i (-g(\tilde{G}^i)) \frac{Z^{\mathcal{X} \setminus A_G}}{Z} \prod_l \left(\int_0^1 ds_l \right) \prod_{k \in A_G} \left(\int_{-\infty}^{\infty} d\alpha_k \right) \\ &\times \exp \left(-\frac{1}{2} \sum_{k \in A_G} \alpha_k^2 \right) \prod_l B_G^l(\alpha) \prod_l b_G^l(s) p(\alpha) e^{-V_G(s, \alpha)}, \quad (11.9) \end{aligned}$$

where $V_G(s, \alpha)$ is the final (interpolated) form of the interaction (depending only on the interior variables) for the history associated with G .

Remark. The proof is “trivial” in the sense that it involves only collecting the procedures, formulae, definitions, and notation.

12. HOLDER'S INEQUALITY

For convenience we introduce the notation $H_G(p)$ for the integral expression in (11.9) so that Theorem 11.3 may be written as

$$\langle p(\alpha) \rangle = \sum_{G: A_G \in \mathcal{X}} \prod_i (-g(\tilde{G}^i)) \frac{Z^{\mathcal{X}-A_G}}{Z^{\mathcal{X}}} H_G(p). \tag{12.1}$$

We will later prove the following generalized form of the *Main Theorem* (Section 9).

THEOREM 12.1. *There is an s_0 , and $\varepsilon > 0$, and a $\lambda_0(s)$ such that for any $s > s_0$ and any $0 \leq \lambda \leq \lambda_0(s)$ one has*

$$\sum_G \prod_i |g(\tilde{G}^i)|^{1-\varepsilon} |H_G(p)| \leq c(p), \tag{12.2}$$

where the sum is over all representation one graphs (rooted on \mathcal{D}) in the infinite set of all vertices.

We find it convenient to introduce two new expectations. We let $\langle \cdot \rangle_0$ be the free expectation ($\lambda = 0$) and $[\cdot]_{G,s}$ the expectation with respect to the measure

$$e^{-(1/2) \Sigma \alpha^2 - V_G(\alpha,s)} \prod_{k \in A_G} d\alpha_k. \tag{12.3}$$

We find the basic estimate

$$|H_G(p)| \leq \langle p^2(\alpha) \rangle_0^{1/2} e^{c|A_G|} \prod_i \int ds_i \prod_l b_G^l \left[\prod_l (B_G^l)^2 \right]_{G,s}^{1/2}, \tag{12.4}$$

where we have here used the Schwartz inequality and α -stability (which automatically holds also for the interpolated interaction).

We finally introduce an expectation $[\cdot]_G$ with respect to the measure

$$e^{-(1/4) \Sigma \alpha^2 - c\lambda \Sigma L_k^{2+\varepsilon'} |\alpha_k|^{4-\varepsilon'}} \prod_{k \in A_G} d\alpha_k \tag{12.5}$$

and using α -positivity and α -stability find

$$[\cdot]_{G,s} \leq e^{c|A_G|} [\cdot]_G. \tag{12.6}$$

We also define

$$\hat{\alpha}_k = \max\{1, |\alpha_k|\} \tag{12.7}$$

and set

$$\hat{\alpha}^r = \prod_k \hat{\alpha}_k^{r(k)}. \tag{12.8}$$

The principal result of this section is the estimate:

LEMMA 12.2. *There is a λ_0 such that if $0 \leq \lambda \leq \lambda_0$ and n is the length of \tilde{G}*

$$|H_G(p)| \leq \langle p^2 \rangle_0^{1/2} c^n [\hat{\alpha}^{2\Sigma \tilde{G}^i}]_G^{1/2} \prod_l \int ds_l \prod_l b_G^l. \tag{12.9}$$

This follows from elementary estimates

$$\left| \prod_{i=1}^l \alpha_i \right| \leq c^l \prod_{i=1}^l \hat{\alpha}_i \tag{12.10}$$

and the inequality $|A_G| \leq |\mathcal{D}| + 4n$.

13. REPRESENTATION TWO GRAPHS

Our next step is to wash out some of the ordering in the objects that we are summing over. More specifically, we wish to rid ourselves of the order in which chains and elementary graphs appear but to keep the ordered structure of the chains themselves.

DEFINITION 13.1. *A representation two graph rooted on \mathcal{D} is the set of chains and elementary graphs occurring in some representation one graph rooted on \mathcal{D} .*

We denote the representation two graph associated with a representation one graph G by T_G . Now, since

$$A_G = \mathcal{D} \cup \bigcup_l \text{supp } G_l,$$

A_G depends only on T_G and so A_T is well-defined for a representation two graph T . Similarly, if n is the length of \tilde{G} , n depends only on T_G and

$$\sum_i \tilde{G}(i) = \sum_{\tau \in T'_G} \tau + \sum_{C \in T''_G} \sum_i C^i, \tag{13.1}$$

$$\prod_i |g(\tilde{G}^i)| = \left(\prod_{\tau \in T'_G} |g(\tau)| \right) \prod_{C \in T''_G} \prod_i |g(C^i)|, \tag{13.2}$$

where T' (T'') denotes the set of all elementary graphs (chains) in T .

It follows from Lemma 12.2 that

$$\begin{aligned} & \sum_G \prod_i |g(\tilde{G}^i)|^{1-\epsilon} |H_G(p)| \\ & \leq \langle p(\alpha)^2 \rangle_0^{1/2} \sum_T e^{cn_T} \prod_{k \in A_T} [\xi^{2\Omega_T(k)}]_k^{1/2} \\ & \cdot \left(\prod_{\tau \in T'} |g(\tau)|^{1-\epsilon} \right) \prod_{C \in T''} \prod_i |g(C^i)|^{1-\epsilon} \sum_{G: T_G = T} \prod \left(\int_0^1 ds_i \right) \prod_l b_G^l(s), \end{aligned} \tag{13.3}$$

where

$$\Omega_T = \sum_{\tau \in T'} \tau + \sum_{C \in T''} \sum_l C^l \tag{13.4}$$

and we have factored measure (12.5); i.e., \int_k denotes integration with respect to the measure

$$\exp \left(-\frac{\zeta^2}{4} - c\lambda L_k^{2+\epsilon'} |\zeta|^{4-\epsilon'} \right) d\zeta. \tag{13.5}$$

For a given representation one graph $G = (G_1, \dots, G_N)$ we define η_G as the mapping from $\{2, \dots, N\}$ into $\{1, \dots, N\}$ given by

$$\begin{aligned} \eta_G(l) &= \varepsilon_G^l && \text{if } G_l \text{ is elementary graph} \\ &= \text{Min}_i \{ \varepsilon_G^l(i) \} && \text{if } G_l \text{ is a chain.} \end{aligned} \tag{13.6}$$

Let

$$f_G(s) \equiv \prod_{l=2}^N (s_{l-1} \cdots s_{\eta_G(l)}). \tag{13.7}$$

Clearly,

$$\prod_l b_G^l(s) \leq f_G(s). \tag{13.8}$$

THEOREM 13.2. *If T is a representation two graph rooted on \mathcal{D} , then*

$$\sum_{G: T_G = T} \prod_l \left(\int_0^1 ds_l \right) f_G(s) \leq 1.$$

We prove this theorem in Appendix B; it is a ϕ_2^4 analogue of the tree graph identity proven in Appendix A, and the idea of the proof is similar if one treats whole chains as steps. We have an inequality instead of an identity because of restrictions like (11.4) on the order in which chains appear.

Combining (13.3) (13.8), and Theorem 13.2, we obtain

THEOREM 13.3. *For $0 \leq \lambda \leq \lambda_0$*

$$\begin{aligned} \sum_G \prod_i |g(\tilde{G}^i)|^{1-\epsilon} |H_G(p)| &\leq c(p) \sum_T e^{cn_T} \prod_{k \in A_T} [\hat{\zeta}^{2\Omega_T(k)}]^{1/2} \\ &\cdot \left(\prod_{\tau \in T'} |g(\tau)|^{1-\epsilon} \right) \prod_{C \in T''} \prod_i |g(C^i)|^{1-\epsilon}. \end{aligned} \tag{13.9}$$

Bounding the right side of this equation occupies the next four sections.

14. REPRESENTATION THREE GRAPHS

Having eliminated some of the ordering in our graphs, we now wish to introduce an ordering that is better suited for our estimation of (13.9). The resulting graphs will be called *representation three graphs*. First, for each representation two graph T , let us pick a representative G_T from the class of representation one graphs G for which $T_G = T$. We fix G_T permanently, as some of our terminology will depend on this arbitrary choice.

Now for a given element $G_{T,l}$ in the sequence G_T we define the *attachment* of $G_{T,l}$ (relative to T) as the last vertex in

$$\text{supp } G_{T,l} \cap \left(\emptyset \cup \bigcup_{j < l} \text{supp } G_{T,j} \right)$$

with respect to our scale-lexicographic order on the vertices. It immediately follows from Definition 11.2 that if $G_{T,l}$ is either an elementary graph or a type 2 chain, then the attachment of $G_{T,l}$ is a smallest-scale vertex in $\text{supp } G_{T,l}$; this is not the case for type 1 chains.

For the induced sequence of vertices that appear as attachments of elements of G_T , consider the subsequence of first appearances. For each vertex list every element of G_T attached to that vertex (i.e., for which the vertex is an attachment), where, say, the elementary graphs are listed first, the type 1 chains next, and the type 2 chains last, but otherwise the order is arbitrary. Then adjoin these sequences in the order that their attachments appear in the subsequence. Such a sequence is clear by a reordering of G_T , and we call it a *representation three graph of T* . For each vertex k , let $r_T(k)$ (resp. $r'_T(k)$, $r''_T(k)$) denote the number of elementary graphs (resp. type 1 chains, type 2 chains) that are attached to k relative to T . Clearly, the total number of representation three graphs that arise from T is

$$\prod_k [r_T(k)! r'_T(k)! R''_T(k)!]. \tag{14.1}$$

Thus, a term in the sum in (13.9) can be trivially rewritten as

$$\frac{e^{c n_T}}{\prod_k [r_T(k)! r'_T(k)! r''_T(k)!]} \sum_{J: T_J = T} \prod_i |g(\tilde{J}^i)|^{1-\epsilon} \prod_{k \in A_J} [\xi^{2\Omega_J(k)}]_k^{1/2}, \tag{14.2}$$

where we adopt much of the notation for representation three that we did for representation one: if J is a representation three graph, then \tilde{J} denotes the sequence of elementary graphs induced by J if we ignore the chain structure, T_J is just the set of all chains and elementary graphs that occur in J (i.e., the representation two graph that J arises from), and we introduce the abuse of notation $A_J \equiv A_{T_J}$ and $\Omega_J \equiv \Omega_{T_J}$.

We wish to warn against two possible sources of confusion. Firstly, most of the nomenclature and concepts of the present paper have parallels in the organization of

[2], but our transition from Representation 2 to Representation 3 is rather different from that in [2]. Secondly, we should emphasize that although in the paragraph above we opted, for notational convenience, to put a (largely arbitrary) linear ordering on the elementary graphs in $\tilde{\mathcal{J}}$, in fact the only additional information in Representation 3 over Representation 2 is a specific order to the attachments at each vertex.

15. PREPARATORY TO THE FINAL RECKONING

With the use of an estimate like (3.6) we see that to bound the r.h.s. of (13.9) it is sufficient to show, for small $\varepsilon, \lambda > 0$ and large s , that

$$\begin{aligned}
 & c^n \lambda^{n(1-\varepsilon)} \sum_{\mathcal{J}: n_{\mathcal{J}}=n} \frac{1}{\prod_k [r_{\mathcal{J}}(k)! r'_{\mathcal{J}}(k)! r''_{\mathcal{J}}(k)!]} \prod_{k \in A_{\mathcal{J}}} [\tilde{\zeta}^{2\Omega_{\mathcal{J}}(k)}]_k^{1/2} \\
 & \cdot \prod_i \left[L_i^{2-2\varepsilon} \cdot \prod_{\substack{k, k' \in \text{supp } \tilde{\mathcal{J}}_i \\ L_{k'} \leq L_k}} h_{\bar{s}} \left(\frac{1}{L_k} d_{k, k'} \right) \right] \tag{15.1}
 \end{aligned}$$

is finite and $O(\lambda^{\varepsilon n})$ uniformly in n , where we have set

$$r_{\mathcal{J}}^{\#}(k) = r_{T_{\mathcal{J}}}^{\#}(k), \quad \# = ', ''$$

where we have adopted the further abuses of notation $k_i \equiv k_{\mathcal{J}_i}$ and $L_i \equiv L_{k_i}$, and where we have used the fact that $n_{\mathcal{J}} = n_{T_{\mathcal{J}}}$ is actually the length of the sequence $\tilde{\mathcal{J}}$. We also have $\bar{s} = cs$. For what is to follow, we must introduce the following notation:

DEFINITION 15.1. Let $\tilde{\mathcal{J}}^{N+1}, \dots, \tilde{\mathcal{J}}^{N+m}$ be the elementary graphs in an m th-degree type 1 chain attached to a vertex v , and set

$$t = \max\{j \mid N + 1 \leq j \leq N + m, v \in \text{supp } \tilde{\mathcal{J}}^j\}.$$

Then $\tilde{\mathcal{J}}^t$ is the *attachment graph* of the chain.

In estimating (15.1) appropriately there are two divergences that must be cancelled by the small factors that are available to us. One of them is the counting divergence, which is generally the “easier” divergence to cancel in a cluster expansion, and we postpone this estimation to Section 17. The other divergence is the number divergence, i.e., the vertex-wise factorial growth of

$$\prod_{k \in A_{\mathcal{J}}} [\tilde{\zeta}^{2\Omega_{\mathcal{J}}(k)}]_k^{1/2}, \tag{15.2}$$

and this divergence is more difficult to cancel.

We cannot cancel the number divergence if we make the simple over-estimation

$$[\zeta^{2\Omega_J(k)}]_k^{1/2} \leq (\Omega_J(k)!)^{1/2} c^{\Omega_J(k)} \tag{15.3}$$

based on throwing away the $|\zeta|^{4-\varepsilon'}$ term in the exponent of (13.5). The $|\zeta|^{4-\varepsilon'}$ term arising from α -positivity represents a milder number divergence and we need this advantage. On the other hand, using this term and neglecting the ζ^2 term in the exponent would create a problem with powers of λ : each power of α_k would bring down a power of $\lambda^{-(1+\frac{1}{4})}$ ($(1+\frac{1}{4}) \equiv 1/(4-\varepsilon')$) in such an estimate. (One would also have trouble with powers of L_k .)

To handle these difficulties we recall that every move in our expansion procedure gives rise to at least one “new variable” (i.e., makes an exterior vertex interior) except possibly those Class 3 moves that terminate chains. Roughly, the idea is to apply, say, the Schwarz inequality so that these “new variables” are integrated against (13.5) by neglecting the $|\zeta|^{4-\varepsilon'}$ term in the exponent and the remaining variables are integrated by throwing away the ζ^2 term instead. Clearly, the power counting problem for λ is eliminated by this choice of estimation. This strategy does not reduce the number divergence, of course, but since a variable can be “new” only once, the integration against $e^{-(1/4)\zeta^2} d\zeta$ does not make it worse, either. Moreover, this notion of “new variable” represents the additional information about Ω_J that we need for cancelling the number divergence, as our assignment of small factors will demonstrate.

The concept of “new variables” is not easily dealt with in the language of representation three graphs—at least, not directly. Accordingly, we consider again our fixed choice of a representation one graph for a given representation two graph. Since a given representation three graph J arises from only one representation two graph (namely, T_J) we may denote the corresponding representative by G_J . Consider the induced sequence \tilde{G}_J of elementary graphs and notice that $\text{supp } \tilde{G}_J^i$ is *not* contained in $\mathcal{D} \cup \bigcup_{j < i} \text{supp } \tilde{G}_j^i$ except possibly when \tilde{G}_J^i is the last elementary graph of a chain. Therefore we define the sequence $(\sigma_1^J, \dots, \sigma_n^J)$ of weight functions as follows:

(i) If \tilde{G}_J^i is neither the last elementary graph of a chain in G_J nor the attachment graph of a type 1 chain, then

$$\begin{aligned} \sigma_1^J(k) &\equiv \tilde{G}_J^i(k), & k \notin \bigcup_{j < i} \text{supp } \tilde{G}_j^i \cup \mathcal{D}, \\ &\equiv 0, & \text{otherwise.} \end{aligned} \tag{15.4}$$

(ii) If \tilde{G}_J^i is the attachment graph of a type 1 chain $G_{J,\mu}$ in G_J , then

$$\begin{aligned} \sigma_1^J(k) &\equiv \tilde{G}_J^i(k), & k \notin \mathcal{D} \cup \bigcup_{\mu < i} \text{supp } G_{J,\mu}, \\ &\equiv 0, & \text{otherwise.} \end{aligned} \tag{15.5}$$

(iii) If \tilde{G}_J^i is the last elementary graph of a chain in G_J and—in the type 1 case—is not the attachment graph of that chain, then

$$\sigma'_i(k) \equiv (15.4) + \delta_{\tilde{G}_J^{i-1}}(k). \tag{15.6}$$

Thus, in case (i), σ'_i is just the restriction of the elementary graph \tilde{G}_J^i to its “new” vertices, in case (ii), σ'_i is the restriction of \tilde{G}_J^i to the “new” vertices of the whole type 1 chain that \tilde{G}_J^i occurs in, and in case (iii) we restrict \tilde{G}_J^i to its “new” vertices and add the delta function located at the vertex corresponding to the variable with respect to which the move integrates by parts. The reason for defining σ'_i by (15.5) in case (ii) will become apparent in the next section, but notice that a vertex can be “new” for at most one type 1 chain and that a chain has at most one attachment graph. Hence, such a contribution cannot “pile up” at a given vertex. As for case (iii), our expansion process integrates by parts with respect to a given variable no more than once, so in this case $k_{\tilde{G}_J^{i-1}}$ is “as good” as a “new” vertex for \tilde{G}_J^i in handling our problem. Of course, this vertex is indeed a “new” vertex for \tilde{G}_J^{i-1} , so the supports of the weight functions σ'_{i-1} , σ'_i are definitely not disjoint in this case, but there is obviously no “pile-up” in this case either. The purpose of (15.6) is to insure the condition

$$\int \sigma'_i \equiv \sum_k \sigma'_i(k) \geq 1 \tag{15.7}$$

for all i .

Now, since \tilde{J} is just a reordering of \tilde{G}_J , consider the permutation of $\{1, \dots, n\}$ whose action takes \tilde{G}_J to \tilde{J} and let it act on $(\sigma'_1, \dots, \sigma'_n)$. We will denote the resulting sequence by $(\hat{\sigma}'_1, \dots, \hat{\sigma}'_n)$. This sequence of weight functions bears no reasonable relation to the notion of “new variable” in the attachment scheme, but by inspection of cases it has the important property that

$$\sum_{i=1}^n \hat{\sigma}'_i(k) \leq c, \quad k \in A_J. \tag{15.8}$$

Let

$$\Omega'_J = \Omega_J - \sum_{i=1}^n \hat{\sigma}'_i. \tag{15.9}$$

It follows from (15.7) that

$$\int \Omega'_J \leq 3n, \tag{15.10}$$

and it is easy to check that $\Omega'_J \geq 0$. Since $\xi \geq 1$, it follows from (15.8) and (15.9) that

$$[\xi^{2\Omega_J(k)}]_k \leq (c\lambda^{-(1+4)} L_k^{-(1+2)})^{2\Omega'_J(k)} (\Omega'_J(k)!)^{(1+4)} \cdot c, \tag{15.11}$$

where $(1 + 1/2) \equiv (2 + \varepsilon')/(4 - \varepsilon')$ and we have applied the Schwarz inequality in the manner described above. Therefore (15.1) is dominated by

$$c^n \lambda^{\varepsilon n} \sum_{J: n_J = n} \frac{1}{\prod_k \prod_{\#} r^{\#}(k)!} \prod_{k \in A_J} L_k^{-(1+1/2)\Omega'_J(k)} (\Omega'_J(k)!)^{(1+1/4)} \cdot \prod_{i=1}^n \left[L_i^{2^-} \prod_{\substack{k, k' \in \text{supp } \tilde{J}^i \\ L_{k'} \leq L_k}} h_{\bar{s}} \left(\frac{1}{L_k} d_{k, k'} \right) \right], \tag{15.12}$$

where $2^- \equiv 2 - 2\varepsilon$.

16. CANCELLING THE NUMBER DIVERGENCE—THE NITTY-GRITTY

We consider an arbitrary representation three graph J rooted on \mathcal{D} such that $n_J = n$ and introduce the abbreviation (suppressing \bar{s})

$$h(k, k') \equiv h_{\bar{s}} \left(\frac{1}{L_k} d_{k, k'} \right). \tag{16.1}$$

Let $(1_i, 2_i, 3_i)$ (resp. $(1_i, 2_i), 1_i, \emptyset$) be the unique triple (resp. pair, single, empty set) of vertices whose weight function is $\tilde{J}^i - \hat{\sigma}_i^J$ in the case $\int \hat{\sigma}_i^J = 1$ (resp. $\int \hat{\sigma}_i^J = 2, \int \hat{\sigma}_i^J = 3, \int \hat{\sigma}_i^J = 4$) and which respects our order on vertices. The summand in (15.12) may be rewritten as

$$\frac{1}{\prod_k \prod_{\#} r^{\#}(k)!} \prod_k ((\Omega'(k)!)^{1+1/4} \prod_{i=1}^n \left[L_i^{2^-} \prod_{l=1}^{\varepsilon_i} L_{i,l}^{-1+1/2} \prod^{(i)} h(k, k') \right]), \tag{16.2}$$

where $\prod^{(i)}$ denotes $\prod_{k, k' \in \text{supp } \tilde{J}^i, L_{k'} \leq L_k}$,

$$\varepsilon_i \equiv 4 - \int \hat{\sigma}_i^J, \tag{16.3}$$

$$L_{i,l} \equiv L_{l_i}, \tag{16.4}$$

and our notational suppression of the J -dependence has been broadened.

At this point we introduce a concept that will be important to us—the concept of *pinning*. For a given elementary graph \tilde{J}^i we define the pinning of \tilde{J}^i as the vertex $k \in \text{supp } \tilde{J}^i$ such that

- (i) If \tilde{J}^i occurs in J , then k is the attachment of \tilde{J}^i .
- (ii) If \tilde{J}^i is the last elementary graph in a type 2 chain, then k is the attachment of that chain.
- (iii) If \tilde{J}^i occurs in a type 2 chain but is not the last elementary graph, then $k = k_i$.

- (iv) If \tilde{J}^i occurs in a type 1 chain and
 - (iva) If \tilde{J}^i is the attachment graph of the chain, then k is the attachment of the chain.
 - (ivb) If \tilde{J}^i occurs after the attachment graph, then $k = k_{i-1}$.
 - (ivc) If \tilde{J}^i occurs before the attachment graph, then $k = k_i$.

Let k be a vertex in the support of Ω' and consider an arbitrary \tilde{J}^i such that $\tilde{J}^i(k) \neq \hat{\sigma}_i^j(k)$. It will be true in all cases that the scale of the pinning will be less than or equal to L_k .

- Case 1. \tilde{J}^i is a case (i) graph.
- Case 2. \tilde{J}^i is a case (ii) graph.
- Case 3. \tilde{J}^i is a case (iva) graph.
- Case 4. \tilde{J}^i either case (iii) or case (ivc).
- Case 5. \tilde{J}^i is a case (ivb) graph.

We wish to classify the contributions to $\Omega'(k)$ by considering the various cases. We let $N_v(k)$ denote the contribution to $\Omega'(k)$ by those \tilde{J}^i that satisfy Case v . Since

$$\Omega'(k) = \sum_{v=1}^5 N_v(k), \tag{16.5}$$

we have

$$\Omega'(k)! \leq c^{\Omega'(k)} \prod_v N_v(k)!, \tag{16.6}$$

so (16.2) is dominated by

$$\frac{1}{\prod_k \prod_{\#} r^{\#}(k)!} c^n \prod_k \prod_v (N_v(k)!)^{1+1/4} \prod_{i=1}^n \left[L_i^{2-} \prod_{l=1}^{q_i} L_{i,l}^{-(1+1/2)} \prod^{(i)} h(k, k') \right]. \tag{16.7}$$

Now notice that in Cases 4 and 5 the pinning of \tilde{J}^i corresponds to a variable with respect to which we integrate by parts and bring down \tilde{J}^{i+1} and \tilde{J}^i , respectively, in the history of moves labeled by G_j . Since this can happen to a given variable no more than once, it follows that

LEMMA 16.2. *No vertex can be the pinning of more than one Case 4 or Case 5 graph.*

On the other hand, there is no such lemma for Case 1, 2, and 3 graphs. However, note that $r(k)$ (resp. $r'(k)$, $r''(k)$) is the number of Case 1 (resp. Case 3, Case 2) graphs that are pinned to k (i.e., for which k is the pinning), so we expect the $r^{\#}(k)!$ factors in the denominator to help cancel the number divergences in (16.7). To this

end let $r(k, k')$ (resp. $r'(k, k')$, $r''(k, k')$) denote the evaluation at k' of the sum of all Case 1 (resp. Case 3, Case 2) graphs that are pinned to k . Clearly,

$$\sum_{k'} \left(r^{\#}(k, k') - \sum_{\text{such } i} \hat{\sigma}_i^j(k') \right) \leq 3r^{\#}(k). \tag{16.8}$$

But $\sum_i \hat{\sigma}_i^j(k') \leq 4$, so

$$\sum_{k'} (r^{\#}(k, k') - 4) \leq 3r^{\#}(k). \tag{16.9}$$

Hence

$$r^{\#}(k)! \geq \prod_{k'} (c^{-r^{\#}(k, k')} r^{\#}(k, k')!)^{1/3}, \tag{16.10}$$

so (16.7) is bounded by

$$c^n \prod_{k'} \left[\frac{1}{\prod_{\#} \prod_k (r^{\#}(k, k')!)^{1/3} \prod_v (N_v(k')!)^{1+1/4}} \right] \cdot \prod_{i=1}^n \left[L_i^{2-\varepsilon_i} \prod_{l=1}^{\varepsilon_i} L_i^{-1+1/2} \prod^{(i)} h(k, k') \right] \tag{16.11}$$

because

$$\sum_{\#} \sum_{k, k'} r^{\#}(k, k') \leq 4n. \tag{16.12}$$

Now the problem is to extract enough small factors from the product $\prod_{i=1}^n$ for each large factor in the product $\prod_{k'}$ to effect a cancellation. More specifically, we need to construct a function w_v on the vertices from quantities in the product $\prod_{i=1}^n$ such that $w_v(k)$ is small enough to take care of Case v . We will define these functions inductively with respect to \bar{J} (a direct definition is also possible): initially, $w_v(k) = 1$, and at the i th stage we assign a certain additional factor to the product we already have for $w_v(k)$ at the $(i - 1)$ th stage. The rules are determined by cases, but one rule we stipulate applies to all cases: if $k \notin \text{supp}(\bar{J}^i - \hat{\sigma}_i^j)$, then the $w_v(k)$ are assigned the factor 1 at the i th stage—e.g., if $\varepsilon_i = 3$, then only $w_v(1_i)$, $w_v(2_i)$, and $w_v(3_i)$ are candidates for receiving small factors at the i th stage. Moreover, if a vertex $k \in \text{supp}(\bar{J}^i - \hat{\sigma}_i^j)$ and \bar{J}^i satisfies Case v_0 , then only $w_{v_0}(k)$ is a candidate at the i th stage. The rules are as follows:

Case 1. In this case \bar{J}^i occurs in J , so L_i is the scale of the pinning (although the pinning need not be k_i). Only $w_1(k)$ is to receive a small factor for each vertex $k \in \text{supp}(J^i - \hat{\sigma}_i^j)$. We assign it

$$\left(\frac{L_i}{L_k} \right)^{1+1/2(\bar{J}^i(k) - \hat{\sigma}_i^j(k))} h_{\bar{s}} \left(\frac{1}{L_k} d_{k,i} \right), \tag{16.13}$$

where $d_{k,i}$ denotes the distance between k and the pinning of \bar{J}^i , and the value of

$\bar{s} = cs$ is to be chosen later. After extracting these factors from the corresponding factor in the product $\prod_{i=1}^n$, we are left with

$$L_i^{2-(1+1/2)\epsilon_i} \prod^{(i)} \tilde{h}(k, k') \tag{16.14}$$

(\tilde{h} is h with \bar{s} replaced by \tilde{s}) which we propose to save for the counting problem.

Case 2. This case is parallel to Case 1 and we make the same kind of assignment to w_2 with (16.14) left over for the counting.

Case 3. In this case \tilde{J}^i is the attachment graph of a type 1 chain. Only $w_3(k)$, $k \in \text{supp}(\tilde{J}^i - \hat{\sigma}_i^j)$, will receive small factors. If we let p_i denote the pinning of \tilde{J}^i , then we are assigning to $w_3(k)$ the factor

$$\left(\frac{L_{p_i}}{L_k}\right)^{1+1/2(\tilde{J}^i(k) - \hat{\sigma}_i^j(k))} h_{\tilde{s}}\left(\frac{1}{L_k} d_{k,i}\right) \tag{16.15}$$

for such k . After extracting these from the corresponding factor, we have

$$L_i^{2-} L_{p_i}^{-(1+1/2)\epsilon_i} \prod^{(i)} \tilde{h}(k, k') \tag{16.16}$$

left over for the counting problem.

Case 4. This case is parallel to Case 1 and we make assignment (16.13) to w_4 with (16.14) left over for the counting.

Case 5. This case is parallel to Case 3 and we make assignment (16.15) to w_5 with (16.16) saved for the counting.

Having defined the functions w_v , we now state the estimate that effects the cancellation of the number divergence.

THEOREM 16.3. *There is a universal constant $c > 0$ such that*

$$\frac{1}{\prod_{\neq} \prod_k (r^{\#}(k, k')!)^{1/3}} \prod_v [w_v(k')(N_v(k')!)^{1+1/4}] \leq c^{\Omega(k')}. \tag{16.17}$$

By (16.5) it is obvious that the theorem will follow from the inequalities

$$\frac{1}{\prod_k (r(k, k')!)^{1/3}} w_1(k')(N_1(k')!)^{1+1/4} \leq c^{N_1(k')}, \tag{16.18}$$

$$\frac{1}{\prod_k (r'(k, k')!)^{1/3}} w_3(k')(N_3(k')!)^{1+1/4} \leq c^{N_3(k')}, \tag{16.19}$$

$$\frac{1}{\prod_k (r''(k, k')!)^{1/3}} w_2(k')(N_2(k')!)^{1+1/4} \leq c^{N_2(k')}, \tag{16.20}$$

$$w_v(k')(N_v(k')!)^{1+1/4} \leq c^{N_v(k')}, \quad v = 4, 5. \tag{16.21}$$

Now it follows from inspection of our inductive construction of the w_ν that for arbitrary ν and vertex k' there is a sequence (v_1, \dots, v_N) of vertices (i.e., the relevant pinnings) such that $N = N_\nu(k')$, $L_{v_i} \leq L_{k'}$, and

$$w_\nu(k') = \prod_{i=1}^N \left(\frac{L_{v_i}}{L_{k'}} \right)^{1+2} h_{\tilde{s}} \left(\frac{1}{L_{k'}} d_{k', v_i} \right). \tag{16.22}$$

It also follows from inspection that for $\nu = 1$ (resp. $\nu = 3, \nu = 2$), $r(k, k')$ (resp. $r'(k, k')$, $r''(k, k')$) is just the multiplicity of k in (v_1, \dots, v_N) .

By (16.22) we see that the proofs of (16.18)–(16.20) will be identical, so we will prove only (16.18) and (16.21).

Proof of (16.18). Recall that $(1^+ / 2) = (2 + \varepsilon') / (4 - \varepsilon')$ and $(1^+ / 4) = 1 / (4 - \varepsilon')$. Hence

$$\sum_{k: L_k \leq L_{k'}} \left[\left(\frac{L_k}{L_{k'}} \right)^{1+2} h_{\tilde{s}} \left(\frac{1}{L_{k'}} d_{k', k} \right) \right]^{4-\varepsilon'} \leq c \tag{16.23}$$

for sufficiently large \tilde{s} because the power of the scale ratio is high enough for dominating the sum over lower scales with a convergent geometric series. (The sum over one scale uses up the small factor $(L_k / L_{k'})^2$.) Now consider the following variation on a factorial estimate [2]:

LEMMA 16.4. *Let $c_1, \dots, c_N > 0$ and let m_1, \dots, m_N be positive integers. Then there is a constant c such that*

$$\prod_{i=1}^N \left(\frac{c_i^{m_i}}{m_i!} \right) \leq \frac{c^{\sum_{i=1}^N m_i}}{(\sum_{i=1}^N m_i)!} \left(\sum_{i=1}^N c_i \right)^{\sum_{i=1}^N m_i}. \tag{16.24}$$

(It is easy to show $c = e$ works; we suspect $c = 1$ suffices.) Since $r(k, k')$ is the multiplicity of k in the sequence (v_1, \dots, v_N) corresponding to vertex k' and $\nu = 1$, we see by combining this lemma with (16.22) and (16.23) that

$$\frac{1}{\prod_k r(k, k')!} w_1(k')^{4-\varepsilon'} N_1(k')! \leq c^{N_1(k')} \tag{16.25}$$

because

$$\sum_k r(k, k') = N_1(k').$$

The desired inequality follows immediately. ■

Proof of (16.21). We apply (16.22) and (16.23) again. This time, however, there are no r factorials to help us. On the other hand, it follows from Lemma 16.2 that the sequence (v_1, \dots, v_N) corresponding to vertex k' and $\nu = 4, 5$ contains no multiplicities greater than 3. Thus, we merely apply Lemma 16.4 in the special case $m_i = 3$. ■

17. COUNTING—THE FINAL RECKONING—A PIECE OF CAKE

The immediate corollary of Theorem 16.3 and our inductive definition of the functions w_ν is that (16.7), and therefore (16.2), is dominated by $c^n \prod_{i=1}^n q_i^J$, where

$$q_i^J \equiv (16.14), \quad \tilde{J}^i \text{ is Case 1, 2 or 4,} \tag{17.1}$$

$$q_i^J \equiv (16.16), \quad \tilde{J}^i \text{ is Case 3 or 5.} \tag{17.2}$$

Thus, to complete the proof of Theorem 12.1 we need only show that

$$\sum_{J: n_J = n} \prod_{i=1}^n q_i^J \leq c^n, \tag{17.3}$$

where the sum is understood to be over representation three graphs rooted on the set \mathcal{D} of distinguished vertices. The proof of (17.3) will involve the use of the small factors q_i^J in cancelling the counting divergence that is implicit in the summation.

DEFINITION 17.1. A *J-suspension* is that final segment of a type 1 chain (in J) whose first elementary graph is the attachment graph of the chain.

Our first step is to rewrite the product of small factors associated with a J -suspension. The point is that to make the estimation by breaking up chains into their elementary graphs with the present correspondence between small factors and elementary graphs is a mistake. Simple power-counting shows that if \tilde{J}^i is Case 3 or 5, then q_i^J is not small enough to cancel the counting divergence unless \tilde{J}^i happens to be the last elementary graph in the J -suspension. It is the small factor associated with this last elementary graph that really effects the cancellation for the whole J -suspension, so we must propagate this smallness up the chain in order to exploit it.

As in the preceding section, we let p_i denote the pinning of \tilde{J}^i ; now suppose that $(\tilde{J}^m, \dots, \tilde{J}^{m+N})$, $N > 0$, is a J -suspension. Since $p_i = k_{i-1}$ for $m < i \leq m + N$, we may write

$$\prod_{i=m}^{m+N} q_i^J = q_m^J L_m^{1/2-\delta} \left[\prod_{i=m+1}^{m+N-1} q_i^J \left(\frac{L_i}{L_{p_i}} \right)^{1/2-\delta} \right] q_{m+N}^J L_{p_{m+N}}^{-1/2+\delta}, \tag{17.4}$$

where $\delta > 0$ will be chosen later. The significance of this telescoping product is clear: since \tilde{J}^{m+N} is the last elementary graph in a type 1 chain, we know that $L_{p_{m+N}} = L_{m+N}$, so by (17.2) we have

$$q_{m+N}^J L_{p_{m+N}}^{-1/2+\delta} \leq L_{m+N}^{\delta-} \prod \tilde{h}^{(m)}(k, k') \tag{17.5}$$

because $\varepsilon_i \leq 3$. Also, if we stipulate $\delta < 1/2 - 2\varepsilon$, then,

$$q_m^J L_m^{1/2-\delta} \leq L_m^\rho \left(\frac{L_m}{L_{p_m}} \right)^{2+\rho} \prod \tilde{h}^{(m)}(k, k'), \tag{17.6}$$

$$q_i^J \left(\frac{L_i}{L_{p_i}} \right)^{1/2-\delta} \leq L_i^\rho \left(\frac{L_i}{L_{p_i}} \right)^{2+\rho} \prod^{(i)} \tilde{h}(k, k'), \tag{17.7}$$

provided $0 < 2\rho < 1/2 - \delta - 2\varepsilon$.

We have still to consider the special case of a type 1 chain in J whose attachment graph is the last elementary graph. If \tilde{J}^m is this elementary graph, then

$$q_m^J = L_m^{2-} L_{p_m}^{-(1+1/2)\varepsilon_m} \prod^{(m)} \tilde{h}(k, k'), \tag{17.8}$$

$$q_{m-1}^J = L_{m-1}^{2--(1+1/2)\varepsilon_{m-1}} \prod^{(m-1)} \tilde{h}(k, k'). \tag{17.9}$$

But since \tilde{J}^m is the last elementary graph in a type 1 chain, we know that $L_{m-1} = L_m$, so

$$q_{m-1}^J q_m^J \leq \left[L_{m-1}^\rho \prod^{(m-1)} \tilde{h}(k, k') \right] \left[\left(\frac{L_m}{L_{p_m}} \right)^{2+\rho} \prod^{(m)} \tilde{h}(k, k') \right]. \tag{17.10}$$

Now for an arbitrary elementary graph τ and vertex $k_0 \in \text{supp } \tau$ we define $\kappa(\tau, k_0)$ as follows:

(i) If $L_{k_0} = L_\tau$, then

$$\kappa(\tau, k_0) \equiv L_\tau^\rho \prod_{\substack{k' \in \text{supp } \tau \\ L_{k'} \leq L_k}} h_{\tilde{s}} \left(\frac{1}{L_k} d_{k,k'} \right). \tag{17.11}$$

(ii) If $L_{k_0} > L_\tau$, then

$$\kappa(\tau, k_0) \equiv L_\tau^\rho \left(\frac{L_\tau}{L_{k_0}} \right)^{2+\rho} \prod_{\substack{k, k' \in \text{supp } \tau \\ L_{k'} < L_k}} h_{\tilde{s}} \left(\frac{1}{L_k} d_{k,k'} \right). \tag{17.12}$$

Combining (17.1), (17.2) with (17.4)–(17.12), we see that

$$\prod_{i=1}^n q_i^J \leq \prod_{i=1}^n \kappa(\tilde{J}^i, p_i) \tag{7.13}$$

provided that $\rho < \delta^-$ also.

LEMMA 17.2. For an arbitrary vertex k_0 (if \tilde{s} is large enough)

$$\sum_{\tau: k_0 \in \text{supp } \tau} \kappa(\tau, k_0) \leq c. \tag{17.14}$$

Proof. Clearly

$$\sum_{\tau: k_0 \in \text{supp } \tau} \kappa(\tau, k_0) \leq \sum_{\substack{1,2,3 \\ L_1 > L_2 > L_3}} \kappa(\tau_{123k_0}, k_0); \tag{17.15}$$

the case (i) terms are given by the condition $L_3 \geq L_{k_0}$ and the case (ii) terms are given by $L_3 < L_{k_0}$. A case (i) term is obviously dominated by

$$L_{k_0}^\rho h_{\bar{s}} \left(\frac{1}{L_1} d_{1,k_0} \right) h_{\bar{s}} \left(\frac{1}{L_2} d_{2,k_0} \right) h_{\bar{s}} \left(\frac{1}{L_3} d_{3,k_0} \right); \tag{17.16}$$

since the sum of, say, $h_{\bar{s}}((1/L_1) d_{1,k_0})$ over 1 for a fixed scale is bounded by a constant we know that

$$\sum_{1:L_1 \geq L_{k_0}} h_{\bar{s}} \left(\frac{1}{L_1} d_{1,k_0} \right) \tag{17.17}$$

grows linearly with respect to the scale parameter of k_0 (r value) and therefore only logarithmically with respect to L_{k_0} , so it follows that the sum over case (i) terms is dominated by a universal constant. A case (ii) term is dominated by

$$L_3^\rho \left(\frac{L_3}{L_{k_0}} \right)^{2+\rho} h_{\bar{s}} \left(\frac{1}{L_{k_0}} d_{k_0,3} \right) h_{\bar{s}} \left(\frac{1}{L_2} d_{2,3} \right) h_{\bar{s}} \left(\frac{1}{L_1} d_{1,3} \right). \tag{17.18}$$

This time we first sum over 1 and 2, holding 3 fixed: since $L_1 \geq L_2 \geq L_3$, the situation is essentially the same as before—i.e., the sum grows logarithmically with respect to L_3 , and L_3^ρ kills that growth. Finally, we sum over 3: the power of L_3/L_{k_0} is high enough to insure convergence. ■

It is becoming obvious that we will need to consider *pointed graphs*—i.e., elementary graphs paired with single vertices in their supports.

DEFINITION 17.3. An m th *degree string* is a sequence $F \equiv \{(\tau_1, v_1), \dots, (\tau_m, v_m)\}$ of pointed graphs with one of the vertices v_l such that

- (a) If $j < l$, then $v_j \in \text{supp } \tau_{j+1} \setminus \{v_{j+1}\}$.
- (b) If $j > l$, then $v_j \in \text{supp } \tau_{j-1} \setminus \{v_{j-1}\}$.

We say that F is *nailed to* v_l and write $F \succ -v_l$; note that nailings of strings are not unique in general. The motivation for introducing this notion is plain. Let J be a representation three graph and pair each elementary graph in \tilde{J} with its pinning. Then every chain in J is a string nailed to the attachment of the chain. The usefulness of the notion is contained in the following lemma:

LEMMA 17.4. For an arbitrary vertex k_0 and positive integer m ,

$$\sum_{\substack{F:F \succ -k_0 \\ \text{deg } F = m}} \prod_{i=1}^m \kappa(\tau_i, v_i) \leq c^m. \tag{17.19}$$

Proof. The idea is to apply the “sums into sups” trick that is so characteristic of

counting. Since $\text{supp } \tau_l \setminus \{v_l\}$ has at most three vertices in it, we can make the estimation

$$\begin{aligned} & \sum_{\substack{F:F \succ k_0 \\ \text{deg } F = m}} \prod_{i=1}^m \kappa(\tau_i, v_i) \\ &= \sum_{l=1}^m \sum_{\substack{F:F \succ v_l \\ \text{deg } F = m \\ v_l = k_0}} \prod_{i=1}^m \kappa(\tau_i, v_i) \\ &\leq 6 \sum_{l=1}^m \left(\sum_{\tau: k_0 \in \text{supp } \tau} \kappa(\tau, k_0) \right) \left(\sup_k \sum_{\substack{F':F' \succ v'_{l-1} \\ \text{deg } F' = l-1 \\ v'_{l-1} = k}} \prod_{i=1}^{l-1} \kappa(\tau'_i, v'_i) \right) \\ &\quad \times \left(\sup_k \sum_{\substack{F'':F'' \succ v''_1 \\ \text{deg } F'' = m-l \\ v''_1 = k}} \prod_{i=1}^{m-l} \kappa(\tau''_i, v''_i) \right). \end{aligned}$$

By Lemma 17.2 the first factor is dominated by a universal constant. Since each of the other two factors involves a sum over strings that are nailed at one end or the other, the remainder of the iteration is purely inductive. ■

We will refer to the empty set paired with an arbitrary vertex k as the *null string nailed to k* . We will include null strings from this point on when we speak of strings.

DEFINITION 17.5. Let A be a set of vertices. A *counting graph rooted on A* is a sequence $\mathcal{F} \equiv (F_1, \dots, F_N)$ of strings with specified nailings μ_1, \dots, μ_N , respectively, such that for a given l , one of the following holds:

- (a) $\mu_l = \mu_{l-1}$.
- (b) μ_l is the first vertex in $A \cup \bigcup_{j < l} \text{supp } F_j \setminus \{\mu_1, \dots, \mu_{l-1}\}$ with respect to the scale-lexicographic order on the vertices.

Remark. As in the case of chains, the support of a string is defined as the union of the supports of the elementary graphs in it.

For a given representation three graph J rooted on \mathcal{D} there is a canonical counting graph \mathcal{F}_J rooted on \mathcal{D} defined in the following way. Note that if we regard J as a sequence of strings with specified nailings (where the pointed graphs are the elementary graphs in \tilde{J} paired with their pinnings, the elementary graphs occurring in J are regarded as 1st degree strings, and the attachments of the elements of J are the specified nailings of the corresponding strings), then all of the strings for which a given vertex is the specified nailing appear consecutively. This follows from the way we defined a representation three graph in Section 14. It is easy to see (but tedious to demonstrate) that there is a unique inductive definition of \mathcal{F}_J that in general shuffles the order of these blocks of strings in J while preserving the integrity of those blocks. (We adopt the convention that if a vertex arises which is not the specified nailing of

any string in J , then we assign the null string nailed at that vertex and proceed to the next vertex. Also, we terminate \mathcal{F}_J when all of the vertices in A_J have been exhausted.) Obviously:

LEMMA 17.6. $J \mapsto \mathcal{F}_J$ is one-to-one; if $\mathcal{F}_J \equiv (F_1, \dots, F_N)$, then

$$\sum_{l=1}^N \deg F_l = n_J, \tag{17.20}$$

$$N \leq |\mathcal{D}| + 5n_J. \tag{17.21}$$

The inequality follows from the fact that the number of nonnull strings in \mathcal{F}_J is dominated by n_J and that $|A_J| \leq |\mathcal{D}| + 4n_J$.

In summary, we will have proved the estimate (17.3) if we prove

THEOREM 17.7. Let $K_F \equiv \prod_{i=1}^m \kappa(\tau_i, v_i)$ for an arbitrary string $F \equiv ((\tau_1, v_1), \dots, (\tau_m, v_m))$. Then

$$\sum_{\substack{\mathcal{F}: \mathcal{F} \text{ rooted on } \mathcal{D} \\ \sum_{l=1}^N \deg F_l = n \\ N \leq |\mathcal{D}| + 5n}} \prod_{l=1}^N K_{F_l} = O(c^n). \tag{17.22}$$

Proof. Fix N and a sequence (m_1, \dots, m_N) of nonnegative integers such that $\sum_{l=1}^N m_l = n$. Consider the sum over all $\mathcal{F} \equiv (F_1, \dots, F_N)$ rooted on \mathcal{D} such that $\deg F_l = m_l$. If μ_1, \dots, μ_N are the specified nailings of F_1, \dots, F_N , respectively, for a given \mathcal{F} , note that μ_1 is always the first vertex in \mathcal{D} . The constrained sum is dominated by

$$\sum_{M=0}^N \prod_{l=1}^M \left(\sum_{\substack{F: F \succ \mu_1 \\ \deg F = m_l}} \prod_{i=1}^{m_l} \kappa(\tau_i, v_i) \right) \\ \times \sup_{\substack{F_1, \dots, F_m \succ \mu_1 \\ \deg F_l = m_l}} \sum_{\substack{\mathcal{F}': \mathcal{F}' \text{ rooted on } \mathcal{D} \cup \bigcup_{l=1}^m \text{supp } F_l \setminus \{\mu_1\} \\ \deg F'_i = m_{l+M}}} \prod_{l=1}^{N-M} K_{F'_l},$$

where we have applied the ‘‘sums into sups’’ trick. Iterating this estimation and applying Lemma 17.4, we obtain

$$c^n \left| \left\{ (M_1, \dots, M_{N'}) \mid M_i \text{ is a nonnegative integer, } N' \leq N, \text{ and } \sum_{i=1}^{N'} M_i = N \right\} \right|.$$

One can show by an elementary combinatoric estimate that the set cardinality is $O(c^N)$. Thus the l.h.s. of (17.22) is dominated by

$$c^n \sum_{N \leq |\mathcal{Q}| + 5n} c^N \left| \left\{ (m_1, \dots, m_N) \mid m_i \text{ is a nonnegative integer and } \sum_{i=1}^N m_i = n \right\} \right|.$$

But the same combinatoric estimate controls this set cardinality in the same manner; it is $O(c^N)$. ■

APPENDIX A: THREE GRAPH IDENTITIES

In this appendix we prove (8.11), (8.13), and some additional results of some possible interest concerning tree graphs. Much of this work is joint with R. W. Robinson.

We add here to the definitions of Section 8 for ordered connectivity graphs, (ordered) tree graphs, and unordered connectivity graphs. We replace the set A of variables with distinguished variable α_0 by a set S of points (where $|S| = |A| = N$) with distinguished point p_1 . For an (ordered) tree graph η , we let $\pi(\eta)$ be the number of permutations P of the domain $\{2, \dots, N\}$ of η for which $\eta \circ P$ is also an (ordered) tree graph. With the notation $\eta_G = P_{r_1}(G)$, $T_G = P_{r_2}(G)$, and $f_G(s) = f(\eta_G; s)$ introduced in Section 8, we first state our results:

THEOREM A.1. *Let T be an unordered connectivity graph on (S, p_1) :*

$$\sum_{\substack{G \\ T_G = T}} \prod_i \left(\int_0^1 ds_i \right) f_G(s) = 1. \tag{A.1}$$

(8.11)

THEOREM A.2. *Let η be an (ordered) tree graph:*

$$\prod_i \left(\int_0^1 ds_i \right) f(\eta; s) = \frac{1}{\pi(\eta)}. \tag{A.2}$$

COROLLARY A.3. *With T an unordered connectivity graph on (S, p_1) ,*

$$\sum_{\substack{G \\ T_G = T}} \frac{1}{\pi(\eta_G)} = 1. \tag{A.3}$$

THEOREM A.4. *Let \mathcal{N} be the set of all (ordered) tree graphs η mapping $\{2, \dots, N\}$ into $\{1, \dots, N-1\}$. Then*

$$\sum_{\eta \in \mathcal{N}} \prod_i \left(\int_0^1 ds_i \right) f(\eta; s) \leq e^N. \tag{A.4}$$

THEOREM A.5.

$$\sum_{\eta \in \mathcal{A}'} \prod_i \left(\int_0^1 ds_i \right) f(\eta; s) = \frac{N^{N-1}}{N!}. \tag{A.5}$$

((8.13))

COROLLARY A.6.

$$\sum_{\eta \in \mathcal{A}'} \frac{1}{\pi(\eta)} = \frac{N^{N-1}}{N!}. \tag{A.6}$$

Theorem A.4 is well known and we do not supply a separate proof, but of course it follows from Theorem A.5. Given Theorem A.2, we have the equivalences (A.1) \Leftrightarrow (A.3) and (A.5) \Leftrightarrow (A.6). Theorem A.2 will be proven in [10].

Proof of Theorem A.1. With each element $(p, q) \in T$ we associate the complex variable $\mu_{p,q}$ and consider the function $\exp \sum_{(p,q) \in T} \mu_{p,q}$. Abstracting the notion of interior and exterior random variables to this situation, where p_1 is regarded as the “distinguished variable,” we first try to “decouple” p_1 from the other points in S and we get

$$\begin{aligned} \exp \sum_{(p,q) \in T} \mu_{p,q} &= \exp \sum_{\substack{(p,q) \in T \\ p \neq p_1}} \mu_{p,q} + \sum_{(p_1,q) \in T} \mu_{p_1,q} \int_0^1 ds_1 \\ &\times \exp \left(\sum_{\substack{(p,q') \in T \\ p \neq p_1}} \mu_{p,q'} + s_1 \sum_{(p_1,q') \in T} \mu_{p_1,q'} \right) \end{aligned}$$

from our first interpolation step. Our application of the second interpolation step to the term labeled by (p_1, q) is based on regarding p_1 and q as the “interior variables.” Thus

$$\begin{aligned} &\exp \left(\sum_{\substack{(p,q') \in T \\ p \neq p_1}} \mu_{p,q'} + s_1 \sum_{(p_1,q') \in T} \mu_{p_1,q'} \right) \\ &= \exp \left(\sum_{\substack{(p,q') \in T \\ p \neq p_1, q}} \mu_{p,q'} + s_1 \mu_{p_1,q} \right) \\ &\quad + s_1 \sum_{\substack{(p_1,q') \in T \\ q' \neq q}} \mu_{p_1,q'} \int_0^1 ds_2 \exp V(q; \mu; s_1, s_2) \\ &\quad + \sum_{(q,q') \in T} \mu_{q,q'} \int_0^1 ds_2 \exp V(q; \mu; s_1, s_2), \end{aligned}$$

where $V(q; \mu; s_1, s_2)$ is just the convex combination in s_2 of the exponent of the l.h.s. with the exponent of the first term on the r.h.s. Obviously, our interpolation

procedure is an abstract mimicry of that mentioned in Section 8 and described in greater detail in [4]. The collection of points in S that serve to index those complex variables that have been differentiated down for a given term at a given stage also serve as “interior” points for the next interpolation step, whose first term will not include those complex variables that are labeled by both “interior” and “exterior” points. Clearly, the expansion procedure terminates after $N - 1$ steps, and the final remainder terms are the only terms that depend on *all* of the $\mu_{p,q}$. Hence

$$\begin{aligned} & \frac{\partial^{N-1}}{\prod_{(p,q) \in T} \partial \mu_{p,q}} \exp \sum_{(p,q) \in T} \mu_{p,q} \\ &= \frac{\partial^{N-1}}{\prod_{(p,q) \in T} \partial \mu_{p,q}} (\text{sum of all final remainder terms}). \end{aligned} \tag{A.7}$$

Now the set of all final remainder terms is labeled by the set of all ordered connectivity graphs G for which $T_G = T$ because the division between “interior” and “exterior” points at each step builds the unordered connectivity graph T in a certain order for a given path of interpolations in the expansion. Moreover, the differentiation with respect to s_{i-1} associated with the $(i - 1)$ th interpolation brings down the product

$$\mu_{G(i)} \prod_{k=\eta_G(i)}^{i-2} s_k$$

from the exponent, where G is the ordered connectivity graph associated with the given path of interpolations. Hence, we may write the sum of all final remainder terms as

$$\prod_{(p,q) \in T} \mu_{p,q} \sum_{T_G=T} \prod_{i=1}^{N-1} \left(\int_0^1 ds_i \right) f_G(s) \exp V_G(s_1, \dots, s_{N-1}; \mu), \tag{A.8}$$

where $V_G(s_1, \dots, s_{N-1}; \mu)$ is the final form of the exponent for the path of interpolations determined by G . Since this function is linear in the $\mu_{p,q}$, we have

$$\exp V_G(s_1, \dots, s_{N-1}; \mu) = \prod_{(p,q) \in T} (1 + O(\mu_{pq})). \tag{A.9}$$

Combining (A.7)–(A.9), we see that

$$\begin{aligned} 1 &= \frac{\partial^{N-1}}{\prod_{(p,q) \in T} \partial \mu_{p,q}} \exp \sum_{(p,q) \in T} \mu_{p,q} \Big|_{\mu=0} \\ &= \sum_G \prod_{i=1}^{N-1} \left(\int_0^1 ds_i \right) f_G(s). \quad \blacksquare \end{aligned}$$

Another (probabilistic) proof of this theorem will be given in [10].

Before giving the proof (due to Robinson) of Theorem A.5, we need to consider the effect of permutations of underlying sets on ordered and unordered connectivity graphs. Let h be a permutation of S such that $h(p_1) = p_1$ and let G be an ordered connectivity graph on (S, p_1) . Clearly, $(h \times h) \circ G$ is an ordered connectivity graph on (S, p_1) , $P_{r_2}((h \times h) \circ G)$ is the image of $P_{r_2}(G) = T_G$ under $h \times h$, and $P_{r_1}((h \times h) \circ G) = P_{r_1}(G) = \eta_G$. Indeed, if G and G' are any two ordered connectivity graphs on (S, p_1) , then

$$\eta_G = \eta_{G'} \Leftrightarrow G' = (h \times h) \circ G$$

for some permutation h of (S, p_1) . (A.10)

DEFINITION A.7. An *automorphism* of an unordered connectivity graph T is a permutation h of S such that $h(p_1) = p_1$ and $h \times h$ maps T onto itself.

Thus, with the proviso that $T_G = T_{G'} = T$, (A.10) becomes

$$\eta_G = \eta_{G'} \Leftrightarrow G' = (g \times g) \circ G$$

for some automorphism g of T .

Finally, if T and T' are unordered connectivity graphs on (S, p_1) , we say that $T \sim T'$ if and only if T' is the image of T under $h \times h$ for some permutation h of (S, p_1) . Let $n(T)$ denote the number of unordered connectivity graphs in the resulting equivalence class of T . If $s(T)$ is the number of automorphisms of T (including the identity), then

$$s(T) n(T) = (N - 1)! \tag{A.11}$$

because $(N - 1)!$ is the number of permutations of (S, p_1) (the order of the symmetric group S_{N-1}) and because there is a one-to-one correspondence between the unordered connectivity graphs equivalent to T and the left cosets of the automorphism group of T in the symmetric group S_{N-1} .

Proof of Theorem A.5. Let T be an unordered connectivity graph on (S, p_1) . By Theorem A.1,

$$\sum_{T_G=T} \prod_i \left(\int_0^1 ds_i \right) f(\eta_G; s) = 1,$$

and by the above remarks we know that (in the set we are summing over) $\eta_G = \eta_{G'}$ if and only if $(g \times g) \circ G = G'$ for some automorphism g of T . Hence

$$\sum_{\eta \in \mathcal{N}(T)} \prod_i \left(\int_0^1 ds_i \right) f(\eta; s) = \frac{1}{s(T)},$$

where

$$\mathcal{N}(T) = \{\eta_G \mid T_G = T\}.$$

By (A.11),

$$\sum_{\eta \in \mathcal{N}(T)} \prod_i \left(\int_0^1 ds_i \right) f(\eta; s) = \frac{n(T)}{(N-1)!}.$$

Now pick a representative from each equivalence class of unordered connectivity graphs and index the representatives: T_1, \dots, T_m . But notice that the sets $\mathcal{N}(T_1), \dots, \mathcal{N}(T_m)$ of (ordered) tree graphs are also mutually disjoint and that $\mathcal{N} = \bigcup_{k=1}^m \mathcal{N}(T_k)$, so

$$\sum_{\eta \in \mathcal{N}} \prod_i \left(\int_0^1 ds_i \right) f(\eta; s) = \frac{1}{(N-1)!} \sum_{k=1}^m n(T_k).$$

Since $\sum_{k=1}^m n(T_k)$ is the total number of unordered connectivity graphs on (S, p_1) and since (in the language of graph theory) an unordered connectivity graph on (S, p_1) is a rooted tree on (S, p_1) , it follows from Cayley's Theorem [5] that $\sum_{k=1}^m n(T_k) = N^{n-2}$. ■

We conclude this appendix by showing the relation between (0.6) and Theorem A.1. For a given unordered connectivity graph T , any two ordered connectivity graphs G and G' for which $T_G = T_{G'} = T$ are related by some permutation P :

$$G' = G \circ P.$$

This follows from the fact that G and G' are bijections between T and $\{2, \dots, N\}$. Hence in (8.11) we may fix G and sum over all permutations P of $\{2, \dots, N\}$ for which $G \circ P$ is an ordered connectivity graph. But with $G'_{(i)} \equiv (g'_1(i), g'_2(0))$ it follows from the definition of $\eta_{G'} = Pr_1(G')$ that

$$\begin{aligned} \eta_{G'}(i) &= g'_2{}^{-1}(g'_1(i)) = (g_2 \circ P)^{-1}(g_1(P(i))) \\ &= P^{-1}(g_2^{-1}(g_1(P(i)))) = P^{-1}(\eta_G(P(i))), \end{aligned}$$

so we are effectively fixing η and summing over all P for which $P^{-1} \circ \eta \circ P$ is an (ordered) tree graph.

APPENDIX B: THE PENULTIMATE TREE GRAPH IDENTITY

In this appendix we prove a generalization of Theorem A.1 from which Theorem 13.2 immediately follows. Let S be an arbitrary but fixed set of points and fix $A \subset S$.

DEFINITION B.1. An n th order generalized ordered connectivity graph rooted on A is a finite sequence $G \equiv (A_1, \dots, A_N)$ of distinct subsets of S such that

$$A_l \cap \left(A \cup \bigcup_{j=1}^{l-1} A_j \right) \neq \emptyset, \quad 1 \leq l \leq N. \tag{B.1}$$

Remark. The sequence of supports of chains and elementary graphs induced by a representation one graph rooted on a set \mathcal{D} of distinguished vertices is a generalized o.c.g. rooted on \mathcal{D} .

DEFINITION B.2. An N th order generalized unordered connectivity graph rooted on A is the range of some generalized o.c.g. rooted on A .

Let T_G denote the generalized u.c.g. arising from the generalized o.c.g. G . For an N th order generalized o.c.g. G rooted on A , we consider the mapping η_G from $\{1, \dots, N\}$ into $\{1, \dots, N\}$ defined by

$$\begin{aligned} \eta_G(l) &\equiv 1, & A_l \cap A \neq \emptyset, \\ &\equiv \min\{j \mid A_l \cap A_{j-1} \neq \emptyset\}, & \text{otherwise.} \end{aligned} \tag{B.2}$$

It is easy to check that this definition agrees with our definition of η_G in the case of a representation one graph, so we define f_G exactly as we did in Section 13:

$$f_G(s_1, \dots, s_{N-1}) \equiv \prod_{l=2}^N (s_{l-1} s_{l-2} \cdots s_{\eta_G(l)}). \tag{B.3}$$

THEOREM B.3. Let T be an N th-order generalized u.c.g. rooted on A . Then

$$\sum_{\substack{G \text{ rooted on } A \\ T_G = T}} \prod_{l=1}^{N-1} \left(\int_0^1 ds_l \right) f_G(s_1, \dots, s_{N-1}) = 1. \tag{B.4}$$

Proof. The argument differs from the proof of Theorem A.1 in an important respect—the “decoupling” process we use is not as closely related to the notion of “interior.” To each set Q in T we assign a complex variable μ_Q and consider the function $\exp \sum_{Q \in T} \mu_Q$. Our first step in the expansion procedure is to interpolate between $\sum_Q \mu_Q$ and $\exp \sum_{Q: Q \cap A = \emptyset} \mu_Q$,

$$\begin{aligned} \exp \sum_Q \mu_Q &= \exp \left(\sum_{Q: Q \cap A = \emptyset} \mu_Q + \sum_{Q: Q \cap A \neq \emptyset} \mu_Q \int_0^1 ds_1 \right) \\ &\quad \times \exp \left(\sum_{Q: Q \cap A = \emptyset} \mu_Q + s_1 \sum_{Q: Q \cap A \neq \emptyset} \mu_Q \right), \end{aligned} \tag{B.5}$$

and then to pick either the completed term or one of the remainder terms. Note in the “interior–exterior” sense that the completed term is not a “decoupled” term, but a

term involving only those complex variables associated with sets of “exterior” points. If we choose a remainder term and denote this choice of Q by Q_1 then the second step is to interpolate between this remainder term and the expression one obtains from the remainder term if all μ_Q terms for which Q meets $A \cup Q_1$ and $Q \neq Q_1$ are dropped from the exponent and then to pick either the latter term or one of the remainder terms resulting from this second interpolation. Suppose we have chosen remainder terms $i - 1$ times and have thus introduced the sets Q_1, \dots, Q_{i-1} in T : our i th step is to interpolate between the remainder term and the expression obtained from it if one drops all μ_Q terms for which $Q \cap (A \cup \bigcup_{j=1}^{i-1} Q_j) \neq \emptyset$ and $Q \neq Q_1, \dots, Q_{i-1}$, and then to pick either the completed term or a remainder term. Now, since T has N elements and Q is always taken from T when a remainder term is chosen and must be distinct from the Q 's that have already appeared, the expansion must terminate after N steps. Thus, every term is completed and represents a path—in this case a sequence (Q_1, \dots, Q_i) of elements of T . The point is that the term depends only on the complex variables $\mu_{Q_1}, \dots, \mu_{Q_i}$ together with those μ_Q for which

$$Q \cap \left(A \cup \bigcup_{j=1}^i Q_j \right) = \emptyset;$$

if there are such Q , then the connectivity property of T implies that the term cannot depend on all of the complex variables; on the other hand, if there are no such Q , then the term depends only on $\mu_{Q_1}, \dots, \mu_{Q_i}$. In any case, the term cannot depend on all of the complex variables if $i < N$. Hence

$$\begin{aligned} \exp \sum_{Q \in T} \mu_Q &= \frac{\partial^N}{\prod_{Q \in T} \partial \mu_Q} \exp \sum_{Q \in T} \mu_Q \\ &= \frac{\partial^N}{\prod_{Q \in T} \partial \mu_Q} (\text{sum of all terms whose paths have length } N). \end{aligned} \tag{B.6}$$

But it is obvious from our expansion procedure that the paths of length N that we have built are exactly the generalized o.c.g. G rooted on A for which $T_G = T$. Moreover, the differentiation with respect to s_l associated with the l th interpolation step that builds $G \equiv (A_1, \dots, A_N)$ brings down the product

$$\mu_{A_1} \prod_{j=n_G(l)}^{l-1} s_j \tag{B.7}$$

from the exponent for the choice A_l of remainder term. Thus, we may write the sum of terms whose paths have length N as

$$\prod_{Q \in T} \mu_Q \sum_{\substack{G \text{ rooted on } A \\ T_G = T}} \prod_{l=1}^{N-1} \left(\int_0^1 ds_l \right) f_G(s) \exp V_G(s_1, \dots, s_N; \mu), \tag{B.8}$$

where $V_G(s; \mu)$ denotes the final form of the exponent for the sequence of interpolations determined by G . Since this function is linear in the μ_Q , we have

$$\exp V_G(s_1, \dots, s_N; \mu) = \prod_{Q \in T} (1 + O(\mu_Q)). \quad (\text{B.9})$$

Hence, by (B.6)

$$1 = \exp \sum_{Q \in T} \mu_Q \Big|_{\mu=0} = \sum_{\substack{G \text{ rooted on } A \\ T_G = T}} \prod_{l=1}^{N-1} \left(\int_0^1 ds_l \right) f_G(s). \quad \blacksquare$$

Remark. This proof is obviously parallel to the proof of Theorem A.1 in many respects. The important generalization is that in our “decoupling” procedure we regard sets in T that are purely “interior” as “interacting” if they have not yet appeared. Such a situation does not arise in the case of a tree (o.c.g.), so the usual notion of “decoupling” coincides with our generalized notion in this case.

Note added in proof. Elliott Lieb has given a proof of α -stability for ϕ_4^4 .

Unfortunately Eq. (3.7), and thus the proof of α -positivity, is incorrect. Without changes of a fundamental nature, the procedure can be corrected by expanding ϕ in terms of the ψ_k instead of the u_k . This will be developed in a future paper.

REFERENCES

1. A. P. CALDERON, *Proc. Symp. Pure Math.* **5** (1961).
2. P. FEDERBUSH, *Commun. Math. Phys.* **81** (1981), 327–340, 341–360.
3. J. GLIMM AND A. JAFFE, *Fortschr. Phys.* **21** (1973).
4. J. GLIMM AND A. JAFFE, “Quantum Physics: A Functional Integral Point of View,” Springer-Verlag, New York/Berlin, 1981.
5. F. HARARY AND E. M. PALMER, “Graphical Enumeration,” Academic Press, New York, 1973.
6. M. REED AND B. SIMON, “Fourier Analysis, Self-Adjointness,” Academic Press, New York, 1975.
7. G. BENFATTO, M. CASSANDRO, G. GALLAVOTTI, F. NICOLO, E. OLIVIERI, E. PRESUTTI, AND E. SCACCIATELLI, *Commun. Math. Phys.* **59** (1978), 143–166.
8. K. GAWEDZKI, AND A. KUPIAINEN, *Commun. Math. Phys.* **77** (1980), 31–64.
9. J. FRÖHLICH AND T. SPENCER, On the statistical mechanics of classical coulomb and dipole gases, preprint. J. FRÖHLICH AND T. SPENCER, The Kosterlitz–Thouless transition in two-dimensional abelian spin systems and the coulomb gas, preprint.
10. G. BATTLE, P. FEDERBUSH, R. W. ROBINSON, in preparation.