*Journal of Statistical Physics, Vol. 78, Nos. 1/2, 1995*

# Solvable Models in Statistical Mechanics, from Onsager Onward

**R. J. Baxter**[1]

*Received October 7, 1993*

There is now a whole field in mathematical physics concerned with solvable models in statistical mechanics, field theory, and related areas. We indicate the influence that Onsager's solution of the planar Ising model has had, and continues to have, on this field.

**KEY WORDS:** Statistical mechanics; lattice models; Ising model; solvable models; integrable systems; Yang–Baxter relations.

There is now a large class of models in statistical mechanics—mostly two-dimensional, but with some three-dimensional members—that have been solved by the Bethe ansatz or functional relation method. The first such model was the Ising model, solved by Onsager in 1944.[1] This is a model on a square lattice of $L$ columns and $M$ rows. The partition function can be written as

$$Z = \mathrm{Trace}(V_1 V_2)^M \tag{1}$$

where $V_1$, $V_2$ are $2^L$ by $2^L$ matrices, known as "transfer matrices":

$$\begin{aligned} V_1 &= (2 \sinh 2H)^{L/2} \exp(-H^* A_0) \\ V_2 &= \exp(H' A_1) \end{aligned} \tag{2}$$

Here $H$, $H'$ are real parameters, $H^* = \frac{1}{2} \ln \coth H$, and $A_0$, $A_1$ are the matrices

$$A_0 = -\sum_{j=1}^{L} \sigma_j^x, \qquad A_1 = \sum_{j=1}^{L} \sigma_j^z \sigma_{j+1}^z \tag{3}$$

$\sigma_j^x$, $\sigma_j^y$, $\sigma_j^z$ are the usual Pauli spin matrices at site $j$.

[1] Department of Theoretical Physics I.A.S. and School of Mathematical Sciences, Australian National University, Canberra, A.C.T. 0200, Australia RJB105@phys 0.anu.edu.au.

Onsager [Ref. 1, Eqs. (60)–(61)] showed that one could recursively generate two sets of matrices $A_k$, $G_k$ satisfying the commutation relations

$$[A_k, A_m] = 4G_{k-m}$$
$$[G_m, A_k] = 2A_{k+m} - 2A_{k-m} \tag{4}$$
$$[G_m, G_k] = 0$$

They also satisfy the periodicity condition

$$A_{m+L} = -CA_m \tag{5}$$

$C = \sigma_1^x \sigma_2^x \cdots \sigma_L^x$ is the operator that reverses all spins.

Taken together, these Lie algebra relations (4) and (5) enabled Onsager to calculate all the eigenvalues of $V_1 V_2$. They have a simple direct product property and he was thus able to calculate $Z$ and hence the free energy. In the same paper he went on to calculate the interfacial tension, correlation length, and specific heat. He then turned his attention to the even harder problem of calculating the spontaneous magnetization. He announced the result (but not the derivation) at a conference in Florence in 1949.[2,3] Yang independently studied the problem in 1951. The calculation was long and intricate, but the pieces suddenly fitted together, giving an amazingly simple result (ref. 4; ref. 5, p. 12).

Other methods for calculating the free energy were found later: Kaufman[6] also solved the problem algebraically, but using spinor operators. Kasteleyn[7,8] and Fisher[9] and Temperley[10] used the combinatorial Pfaffian method to solve the related dimer problem. Schultz *et al.*[11] showed how the problem could be solved algebraically using fermion operators.

The next development was not in two-dimensional lattice models, but in one-dimensional quantum problems. One was the $N$-body problem with delta-function interactions:

$$\mathscr{H} = -\sum_{i=1}^{N} \partial^2/\partial x_i^2 + 2c \sum_{i<j} \delta(x_i - x_j) \tag{6}$$

The other was the anisotropic Heisenberg spin chain of $L$ sites:

$$\mathscr{H} = -\tfrac{1}{2} \sum_{i=1}^{L} \{\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z\} \tag{7}$$

where $\sigma_i^x$, $\sigma_i^y$, $\sigma_i^z$ are the Pauli spin matrices at site $i$. In both cases one wants to calculate the eigenvalues of $\mathscr{H}$, in particular the lowest eigenvalue, which is the ground-state energy $E$.

The first problem (for the simple bosonic case) was solved by Lieb and Liniger in 1963,[12] the second by Yang and Yang in 1966.[13]

Both problems can be interpreted as particles moving on a line, interacting when they come together. Both can be solved by the Bethe ansatz approach. In this method one first notes that as long as the particles are separate, they do not interact: this suggests using a product form for the eigenvector:

$$\psi(x_1,..., x_N) = \phi_1(x_1)\,\phi_2(x_2)\cdots\phi_N(x_N)$$

provided $x_1 < x_2 < \cdots < x_N$. In fact, since the system is translation-invariant, the single-particle functions are plane waves, giving

$$\psi(x_1,..., x_N) = \exp(k_1 x_1 + k_2 x_2 + \cdots + k_N x_N) \tag{8}$$

where $k_1,..., k_N$ are $N$ wave numbers that are to be determined. The corresponding eigenvalue or energy $E$ of $\mathscr{H}$ is a symmetric function of $k_1,..., k_N$.

The next step is to take account of the interactions that occur when two particles come together, say when $x_1 = x_2$. Then the wave function involves $k_1,..., k_N$ only via the momenta $k_1 + k_2$ and $k_3,..., k_N$. These and the total energy $E$ are unaltered by interchanging $k_1$ with $k_2$ (and there are no other choices of $k_1,..., k_N$ that do this). This suggest trying a superposition of plane wave solutions (8):

$$\psi(x_1,..., x_N) = \sum_P A(k_1,..., k_N)\exp(k_1 x_1 + \cdots + k_N x_N)$$

where the sum is over all $N!$ permutations $P$ of $k_1,..., k_N$.

Then the boundary conditions that arise when two adjacent particles come together are satisfied, provided $A(k_1,..., k_N)$ satisfies relations of the form

$$S(k_i, k_{i+1})\,A(k_1,..., k_i, k_{i+1},..., k_N) + S(k_{i+1}, k_i)\,A(k_1,..., k_{i+1}, k_i,..., k_N) = 0 \tag{9}$$

Here $S(k, l)$ is some function of $k$ and $l$ that is determined by the equation and its symmetries; (9) is to hold for $i = 1,..., N-1$ and all permutations of $k_1,..., k_N$.

There are $(N-1)\,N!/2$ equations (9) and it is not obvious that they can all be satisfied. If $S$ and $A$ are scalar functions (as in the bosonic $N$-body problem and the ansisotropic Heisenberg chain), then this problem is easily solved by noting that the equations have the explicit solution

$$A(k_1,..., k_N) = \varepsilon_P \sum_{i>j} S(k_i, k_j)$$

where $k_1,..., k_N$ is some permutation $P$ of the original $N$ wave numbers, and $\varepsilon_P$ is the sign ($\pm$) of $P$.

The corresponding fermionic $N$-body proble requires a more sophisticated Bethe ansatz. It was first studied by McGuire[14, 17] and Gaudin.[18] The problem is simple if all the spins point upwards, or if only one points down. Flicker and Lieb[19] showed that the problem could be solved when two of the spins were down, and predicted that their method could be extended to the general case. Yang[20, 21] considered the $N$-body problem with arbitrary symmetry, and noted that each $S(k, l)$ becomes a matrix $S_i(k, l)$, dependent on the position $i$ on which it acts; defining

$$Y_i(k, l) = -S_i(l, k)^{-1} S_i(k, l)$$

we find that Eqs. (9) become

$$A(k_1,..., k_{i+1}, k_i,..., k_N) = Y_i(k_i, k_{i+1}) A(k_1,..., k_i, k_{i+1},..., k_N) \tag{10}$$

For $N=3$ we deduce that

$$\begin{aligned} A(m, l, k) &= Y_1(l, m) A(l, m, k) \\ &= Y_1(l, m) Y_2(k, m) A(l, k, m) \\ &= Y_1(l, m) Y_2(k, m) Y_1(k, l) A(k, l, m) \end{aligned}$$

and that

$$\begin{aligned} A(m, l, k) &= Y_2(k, l) A(m, k, l) \\ &= Y_2(k, l) Y_1(k, m) A(k, m, l) \\ &= Y_2(k, l) Y_1(k, m) Y_2(l, m) A(k, l, m) \end{aligned}$$

These two equations are consistent if, for all $k, l, m$,

$$Y_1(l, m) Y_2(k, m) Y_1(k, l) = Y_2(k, l) Y_1(k, m) Y_2(l, m) \tag{11}$$

and indeed if this is satisfied (with suffixes 1, 2 generalized to $i, i+1$), then the relations (10) are mutually consistent for all $N$ [pp. 631–633 of ref. 14, Eq. (8) of ref. 20].

Things then began to move very rapidly, for it was quickly realized that the Bethe ansatz method can be used to solve two-dimensional six-vertex models in statistical mechanics. In these, one places arrows on the edge of the square lattice, subject to the "ice" rule that there are two arrows into each vertex and two arrows out. This gives six possible configurations of arrows at a site, which one can order in the standard way.[22, 23] To these one assigns weights $\omega_1,\ldots,\omega_6$. If the model is reflection-symmetric, then the weights are equal in pairs and we can define

$$a=\omega_1=\omega_2, \qquad b=\omega_3=\omega_4, \qquad c=\omega_5=\omega_6$$

One can regard the down arrows in a row as "particles." Because of the ice rule, their number is conserved and one can try a Bethe ansatz for the eigenvectors of the transfer matrix. It works. Lieb solved three special but archetypal cases.[22-25] The general solution was published by Sutherland,[26] who was Yang's first graduate student at Stony Brook. Yang and Yang developed a finite-temperature field theory for the thermodynamics of the anisotropic Heisenberg chain in 1969.[27]

These were exciting times: McGuire was now at Florida Atlantic, Yang and Sutherland at Stony Brook, Lieb and Wu at Northeastern in Boston, though Lieb was about to move to MIT. All the developments were occurring in the eastern United States, in particular in the northeast.

I was fortunate enough to join Lieb in 1968 at MIT, and I well remember the thrill of walking into one of Joel Lebowitz's Yeshiva University meetings in New York to see gathered there so many of the people whose papers I had read.

With Lieb, I looked at a number of potential Bethe ansatz problems, in particular inhomogeneous six-vertex models. Some worked, some did not. Finally I sat down and considered a completely inhomogeneous model in which $\omega_1,\ldots,\omega_6$ were all allowed to vary arbitrarily from site to site. It was straightforward enough to deduce the conditions under which an appropriately modified Bethe ansatz would still work—I wrote this up in the MIT applied math journal.[28] Then in 1970, after two great years in Boston, my wife Elizabeth and I left for a five-month holiday in England before taking the P & O liner Arcadia back to Australia.

Toward the end of those five months, I picked up my MIT article and realized something that was in fact present in Sutherland's 1967 paper: the eigenvectors of the transfer matrix depended on $a, b, c$ only via the parameter

$$\Delta=(a^2+b^2-c^2)/(2ab)$$

(In fact they were identical with the eigenvectors of the anisotropic Heisenberg chain: an observation that had been made and used by Lieb.) This meant that transfer matrices with different $a, b, c$, but the same $\Delta$, commuted. If one fixed $\Delta$, then there was still one nontrivial degree of freedom that could be varied from row to row without affecting the common eigenvectors: for row $i$, call this variable $p_i$. There are corresponding variables $q_j$ associated with the columns $j$: these do affect the eigenvectors, but in a very simple way, and the Bethe ansatz still goes through.

An explicit parametrization that manifests these properties is to set, for the site in row $i$ and column $j$,

$$a : b : c = \sinh(q_j - p_i) : \sinh(\lambda - q_j + p_i) : \sinh \lambda \tag{12}$$

where

$$\Delta = -\cosh \lambda \tag{13}$$

Then the eigenvectors are independent of $p_i$ and the eigenvalues are Laurent polynomials in $\exp(p_i)$. Note that the weights $a$, $b$, $c$ depend on the two rapidities $p_i$, $q_j$ only via their difference $q_j - p_i$.

Could this commutation property be established directly? If so, could it be extended to the more general eight-vertex model, in which the number of down arrows is not conserved and the Bethe ansatz is not immediately applicable?

Sutherland[29] considered this possibility in early 1970. He explicitly showed that any eight-vertex model transfer matrix commutes with an $XYZ$ Hailtonian

$$\mathscr{H} = -\tfrac{1}{2} \sum_{i=1}^{L} \{J_x \sigma_i^x \sigma_{i+1}^x + J_y \sigma_i^y \sigma_{i+1}^y + J_z \sigma_i^z \sigma_{i+1}^z\}$$

which is obviously a generalization of the anisotropic Heisenberg chain (7).

Sitting in Frinton-on-Sea in England, unaware of Sutherland's work, I looked at the rather more general problem of the commutation of two eight-vertex transfer matrices. I soon realized this was possible, provided only that a certain local relation is satisfied. This involves the four-by-four matrices that add vertices to the lattice [Eq. (B6) of ref. 30]. If the lattice is drawn diagonally and the vertex in column $n$ has weights given by (12), then one can write the corresponding matrix as $Y_n(p_i, q_j)$ (regarding $\lambda$ as a given constant). It acts on the arrows on columns $n$ and $n+1$. Then the relation that ensures commutativity is.

$$Y_n(q, r)\, Y_{n+1}(p, r)\, Y_n(p, q) = Y_{n+1}(p, q)\, Y_n(p, r)\, Y_{n+1}(q, r) \tag{14}$$

for all $p$, $q$, $r$ [Eq. (10.4.31) of ref. 31]. Sutherland had used a limiting case of this.

Obviously this is the same equation (apart from notation) as (11). It is true that here the $Y_n(p, q)$ are vertex operators of the lattice, while in (11) they are operators internal to the Bethe ansatz. There are intimate connections between these, as is brought out in the QISM approach of Faddeev.[32,33]

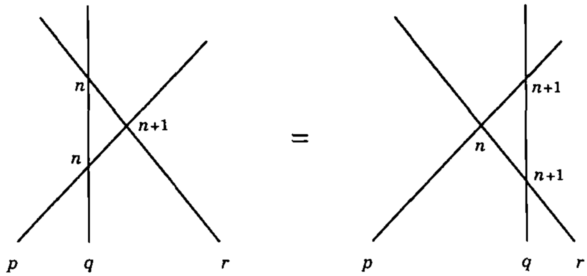


Fig. 1. Graphical representation of the Yang–Baxter relation (14).

The parameters $p, q, r$ are known (by analogy with field theory) as "rapidities." They are associated with the lines of the lattice. A vertex is the intersection of two lines, and its weight involves the corresponding two rapidities. Equation (14) can be interpreted graphically (Fig. 1) as meaning that the effect of three lines intersecting is independent of the order in wich they cross one another (Fig. 7 of ref. 14, Fig. 10.1 of ref. 34, Fig. 9.3 of ref. 31).

I believe it was Faddeev's group who coined the name "Yang–Baxter" relation for Eqs. (11) and (14). However, it should be noted that this relation takes different forms according to whether one is considering a lattice model with spins on sites interacting along edges, a vertex model [ref. 31, Eq. (9.6.8)], or a spin model with interactions round a face [ref. 31, Eq. (13.3.6)]. For the first type of model (in particular, for the Ising model), the relation becomes the "star-triangle" relation (refs. 35, 36; ref. 31, §6.4). This was used as long ago as 1945 by Wannier,[37] and subsequently by Houtappel,[38] to locate the critical point of the triangular and honeycomb lattice Ising models. Indeed, it is clear from Onsager's article of 1971 that he was aware of the relation before he wrote his famous 1944 paper, and realized its implications for the commutation of transfer matrices.

This relation places one well on the road to solving a model, and I was able to solve the eight-vertex model. A large number of models have since been solved by this technique. Some notable examples are the three-spin model,[39–41] the hard-hexagon model,[42, 43] the Fateev–Zamolodchikov model,[44, 45] the Kashiwara–Miwa model,[46–48] the ABF model,[49, 50] and the $A_n^{(1)}$ face models.[51] (A more comprehensive list is given in ref. 52.) Usually one can go on to show that the transfer matrices satisfy a functional matrix relation [e.g., ref. 53, Eq. (4)] and that this relation (together with the commutation properties) determines their eigenvalues. From these one can in principle calculate the free energy, the correlation length or mass gap, and the interfacial tension.

The Yang-Baxter relation has fascinating implications for correlations within a model.[54, 55] Indeed, for the Ising model (and the chiral Potts model discussed in the next paragraph) it is possible to calculate the free energy solely from the relation.[56, 57] [The spontaneous magnetization can be calculated reasonably simply using corner transfer matrices; ref. 31, Eq. (13.7.21)]. A whole field seems thereby to be arising, connecting solvable models, integrable systems, knot theory, Lie algebras, and quantum groups.

The model that has most recently been solved by this route is the chiral Potts model—a generalization of the Ising model that also satisfies a star-triangle relation. We have functional relations determining the transfer matrix eigenvalues[58–60] and the free energy and interfacial tension have been calculated.[61, 62] However, unlike most previous models, the chiral Potts model does not have the "rapidity difference property" mentioned after (12). This makes it technically much more difficult.

Like the six-vertex model, the transfer matrices of the chiral Potts model commute with a one-dimensional Hamiltonian. There is a special "superintegrable" case of the model[63–65] when the Hamiltonian becomes a linear combination of two operators $A_0$ and $A_1$ that generate *precisely* the algebra (4) found by Onsager for the Ising model, except that one no longer has the periodicity condition (5). Again the eigenvalues have simple direct product structures, but now there are many sets of such structures.

There is an extremely strong conjecture (based on series expansions) for the analog of the spontaneous magnetization $M_0$ [ref. 66, Eq. (1.20)], which implies that it has a very simple form—very like that of the Ising model, but with a different exponent. We should of course like to prove this, but as yet have not been able to do so.

This is rather intriguing—we are in much the same position with the chiral Potts model as we were with the Ising model in the late 1940s: much has been done, but much remains to do, and Onsager's pioneering work continues to provide inspiration.

## REFERENCES


1. L. Onsager, *Phys. Rev.* **65**:117–149 (1944).
2. L. Onsager, *Nuovo Cimento (Suppl.)* **6**:261 (1949).
3. L. Onsager, In *Critical Phenomena in Alloys, Magnets and Superconductors*, R. E. Mills, E. Ascher, and R. I. Jaffee, eds. (McGraw-Hill, New York, 1971).
4. C. N. Yang, *Phys. Rev.* **85**:808–816 (1952).
5. C. N. Yang, *Selected Papers 1945-1980 with Commentary* (Freeman, San Francisco, 1983).
6. B. Kaufman, *Phys. Rev.* **76**:1232–1243 (1949).
7. P. W. Kasteleyn, *Physica* **27**:1209–1225 (1961).
8. P. W. Kasteleyn, *J. Math. Phys.* **4**:287–293 (1963).

9. M. E. Fisher, *Phys. Rev.* **124**:1664–1672 (1961).
10. H. N. V. Temperley and M. E. Fisher, *Phil. Mag.* **6**:1061–1063 (1961).
11. T. D. Schultz, D. C. Mattis, and E. H. Lieb, *Rev. Mod. Phys.* **36**:856–871 (1964).
12. E. H. Lieb and W. Liniger, *Phys. Rev.* **130**:1605–1615 (1963).
13. C. N. Yang and C. P. Yang, **150**:321–327 (1966); **150**:327–339 (1966); **151**:258–264 (1966).
14. J. B. McGuire, *J. Math. Phys.* **5**:622–636 (1964).
15. J. B. McGuire, *J. Math. Phys.* **6**:432–439 (1965).
16. J. B. McGuire, *J. Math. Phys.* **7**:123–132 (1966).
17. J. B. McGuire and C. A. Hurst, *J. Math. Phys.* **13**:1595–1607 (1972).
18. M. Gaudin, *Phys. Rev.* **24A**:55–56 (1967).
19. M. Flicker and E. H. Lieb, *Phys. Rev.* **161**:179–188 (1967).
20. C. N. Yang, *Phys. Rev. Lett.* **19**:1312–1315 (1967).
21. C. N. Yang, *Phys. Rev.* **168**:1920–1923 (1968).
22. E. H. Lieb, *Phys. Rev. Lett.* **18**:1046–1048 (1967).
23. E. H. Lieb, *Phys. Rev. Lett.* **19**:108–110 (1967).
24. E. H. Lieb, *Phys. Rev. Lett.* **18**:692–694 (1967).
25. E. H. Lieb, *Phys. Rev.* **162**:162–172 (1967).
26. B. Sutherland, *Phys. Rev. Lett.* **19**:103–104 (1967).
27. C. N. Yang and C. P. Yang, *J. Math. Phys.* **10**:1115–1122 (1969).
28. R. J. Baxter, *Studies Appl. Math.* **50**:51–69 (1971).
29. B. Sutherland, *J. Math. Phys.* **11**:3183–3186 (1970).
30. R. J. Baxter, *Ann. Phys.* **70**:193–228 (1972).
31. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, 1982).
32. L. A. Takhtadzhan and L. D. Faddeev, *Russ. Math. Surv.* **34**(5):11–68 (1979).
33. L. D. Faddeev, *Sov. Sci. Rev. Math. Phys.* *C*1:107–155 (1980).
34. M. Gaudin, *La fonction d,onde de Bethe* (Masson, Paris, 1983).
35. R. J. Baxter, J. H. H. Perk, and H. Au-Yang, *Phys. Lett. A* **128**:138–142 (1988).
36. R. J. Baxter, Corner transfer matrices and parametrizations, *Int. J. Mod. Phys. B* **7** (1993).
37. G. H. Wannier, *Rev. Mod. Phys.* **17**:50–60 (1945).
38. R. M. F. Houtappel, *Physica* **16**:425–455 (1950).
39. R. J. Baxter and F. Y. Wu, *Phys. Rev. Lett.* **31**:1294–1297 (1973).
40. R. J. Baxter and F. Y. Wu, *Aust. J. Phys.* **27**:357–367 (1974).
41. R. J. Baxter, *Aust. J. Phys.* **27**:369–381 (1974).
42. R. J. Baxter, *J. Phys. A* **13**:L61–70 (1980).
43. R. J. Baxter and P. A. Pearce, *J. Phys. A* **15**:897–910 (1982).
44. V. A. Fateev and A. B. Zamolodchikov, *Phys. Lett.* **92A**:37–39 (1982).
45. R. N. Onody and V. Kurak, *J. Phys. A* **17**:L615–619 (1984).
46. M. Kashiwara and T. Miwa, *Nucl. Phys. B* **275**[FS17]:121–134 (1986).
47. K. Hasegawa and Y. Yamada, *Phys. Lett. A* **146**:387–396 (1990).
48. M. Gaudin, *J. Phys. I* (Paris) **1**:351–361 (1991).
49. G. E. Andrews, R. J. Baxter, and P. J. Forrester, *J. Stat. Phys.* **35**:193–266 (1984).
50. P. J. Forrester and R. J. Baxter, *J. Stat. Phys.* **38**:435–472 (1985).
51. M. Jimbo, A. Kuniba, T. Miwa, and M. Okado, *Commun. Math. Phys.* **119**:543–565 (1988).
52. M. Wadati, T. Deguchi, and Y. Akutsu, *Phys. Rep.* **180**:248–332 (1989).
53. R. J. Baxter, *Phys. Rev. Lett.* **26**:832–833 (1971).
54. R. J. Baxter, *Phil. Trans. R. Soc. Lond.* **289**:315–346 (1978).
55. R. J. Baxter, *Ann. Israel Phys. Soc.* **2**(1):37–47 (1978).
56. R. J. Baxter and I. G. Enting, *J. Phys. A* **11**:2463–2473 (1978).
57. R. J. Baxter, *J. Stat. Phys.* **52**:639–667 (1988).

58. G. Albertini, B. M. McCoy, and J. H. H. Perk, *Phys. Lett. A* **135**:159–166 (1989).
59. R. J. Baxter, V. V. Bazhanov, and J. H. H. Perk, *Int. J. Mod. Phys. B* **4**:103–170 (1990).
60. R. J. Baxter, *Phys. Lett. A* **146**:110–114 (1990).
61. R. J. Baxter, In *Proceedings Fourth Asia-Pacific Physics Conference*, S. H. Ahn, Il-T. Cheon, S. H. Choh, and C. Lee, eds. (World Scientific, Singapore, 1991), Vol. 1, pp. 42–57.
62. R. J. Baxter, Interfacial tension of the chiral Potts model, *J. Phys. A*, submitted.
63. H. Au-Yang, B. M. McCoy, J. H. H. Perk, and S. Tlng, In *Algebraic Analysis*, M. Kashiwara and T. Kawai, eds. (Academic Press, 1988), pp. 29–39.
64. G. von Gehlen and V. Rittenberg, *Nucl. Phys. B* **257** [FS14]:351–370 (1985).
65. L. Dolan and M. Grady, *Phys. Rev. D* **25**:1587–1604 (1982).
66. G. Albertini, B. M. McCoy, J. H. H. Perk, and S. Tang, *Nucl. Phys. B* **314**:741–763 (1989).