An algorithm to study tunnelling in a wide class of one-dimensional multiwell potentials. I

1989 J. Phys. A: Math. Gen. 22 1027

(http://iopscience.iop.org/0305-4470/22/8/018)

J. Phys. A: Math. Gen. **22** (1989) 1027-1052. Printed in the UK

An algorithm to study tunnelling in a wide class of one-dimensional multiwell potentials. I

Filippo Cesi

Dipartimento di Fisica, Universitá di Roma 'La Sapienza', Piazzale Aldo Moro 2, 00185 Roma, Italy, and Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Roma, Italy

Received 16 May 1988, in final form 24 October 1988

Abstract. We prove a theorem which gives an algorithmic solution to the problem of finding the logarithmic derivative of the ground-state wavefunction of one-dimensional systems. By means of this quantity, as is well known, one can determine the lowest part of the spectrum of the Hamiltonian by probabilistic methods. We show that, in some natural classes of potentials, the complexity of our algorithm is less than N^3, where N is the number of the absolute minima of the potential. Our approach allows a systematic treatment of cases of much greater complexity than those analysed so far in the literature and it can be useful in the study of physical systems like, for example, long molecular chains or superlattice structures.

1. Introduction

In this paper we study some features of the one-dimensional Schrödinger operator in the semiclassical limit ($\hbar \to 0$). In particular we are interested in the properties of localisation of the ground-state wavefunction and in the splitting of the lowest eigenvalues of the Hamiltonian.

Consider the Schrödinger equation in one dimension $H\psi = E\psi$, where

$$H = -\frac{\hbar^2}{2}\frac{d^2}{dx^2} + V(x).$$

If the potential $V(x)$ has only one absolute minimum in the point $x = a$ then we expect that, for $\hbar$ small enough:

(i) the ground-state wavefunction ψ_0 is concentrated near that minimum;

(ii) the ground-state energy E_0 is about $\hbar\omega/2$ where $\omega = \sqrt{V''(a)}$;

(iii) the splitting of the lowest eigenvalues of H is of order $\hbar$.

If, on the contrary, V is a multiwell potential with several equal minima (figure 1), then ψ_0 is localised around a certain subset of these minima. Moreover, due to tunnelling, the lowest part of the spectrum of the Hamiltonian is characterised by a certain number of energy levels whose distance from E_0 is exponentially small in $\hbar$, i.e.

$$E_i - E_0 \sim \exp(-d_i/\hbar) \qquad \text{when } \hbar \to 0. \tag{1.1}$$

Problem. We want to determine the behaviour of the ground-state wavefunction and to compute the values of the constants d_i, which give the leading term of the splitting in the semiclassical limit.

0305-4470/89/081027+26$02.50

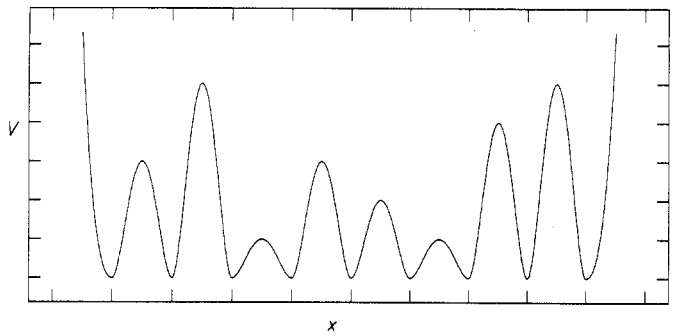

Figure 1. A multiwell potential.

The prototypical situation is the symmetric double well. In this case one finds (see for example [1, 2]) that the single-well ground-state energy level splits into two, E_0 and E_1, whose difference is given by

$$E_0 - E_1 \sim \exp\left(-\frac{1}{\hbar}\int_{a_1}^{a_2}\sqrt{2V(x)}\,dx\right)$$

where a_1 and a_2 are the two minima of the potential.

The analysis of more complicated cases required the development of new techniques. In 1981 Jona-Lasinio *et al* proposed a new approach [3] based on the theory of diffusion processes. By this method they discovered that tunnelling is very sensitive to small localised perturbations of the potential V, in some interesting situations like the symmetric double well and the periodic potential over a finite interval. The idea of the instability of quantum tunnelling proved to be a key point in understanding the physical mechanism of Anderson localisation in disordered electronic systems [4-6]. In the approach of [3] the problem is divided into two steps. The first step consists in the study of the logarithmic derivative of the ground-state wavefunction. Once one knows this quantity, one can determine (this is the second step) the splitting of the energy levels by probabilistic methods, using some powerful results of Ventzel and Freidlin [7, 8]. However the solution of the first step relied on a clever use of two integral equations, and became very tricky when the complexity of the potential increased. In a subsequent paper [9] a general strategy was developed for solving this problem for most of those potentials (which will be called *binary potentials*) that can be constructed as sequences of two types of barriers (figure 2).

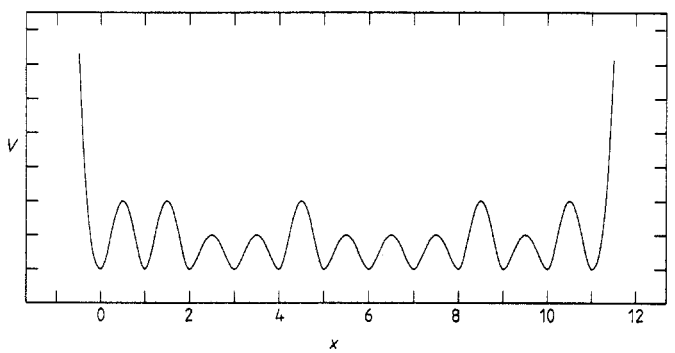

Figure 2. A binary potential.

The aim of this work is to prove a theorem which enables us to construct algorithmically the solution of the first step. The advantages of this algorithm are its simplicity of use and its high speed. To give an idea, after some practice one can solve the problem posed above for a potential with, say, 30 absolute minima, in about five minutes without any electronic aid. Much more complicated cases can be easily handled by means of a computer, since the complexity of the algorithm is polynomial (more precisely it is no worse than N^3).

Unfortunately this approach is not of general applicability, since there are potentials which cannot be solved by our method. These potentials, however, are *few*, at least in some classes where it is easy to *count* them. In a related paper [10] we give, for these suitable classes, a numerical estimate of the percentage of potentials that can be solved. For this purpose we have written a computer program that implements the algorithm. We have determined, for example, that more than 99% of all binary potentials can be treated by this method.

The analysis of multiwell potentials in the semiclassical limit has been extended by Helffer and Sjöstrand to arbitrary dimension in [11, 12] and rederived by Graffi *et al* [13] and by Simon [14] using functional analytic methods.

The organisation of the paper is as follows. A general outline of the approach is contained in § 2; in § 3 we define a mapping, Ext, which is the fundamental object of our construction; in § 4 the theorem which determines the logarithmic derivative of the ground-state wavefunction is stated; in § 5 we show that the complexity of the algorithm is at most N^3; § 6 contains some examples that show how to apply the algorithm in practice; § 7 is dedicated to the proof of the main theorem; the proofs of two basic lemmas are collected in an appendix.

2. Outline of the approach

Consider the Hamiltonian

$$H = -\frac{\hbar^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x).$$

We assume that

(1) $V \in C^\infty(\mathbb{R})$,
(2) V has several points of absolute minimum $x_1, \ldots, x_N$,
(3) $V(x) \geq 0$ and $V(x) = 0$ if and only if $x = x_i$ for some i,
(4) $V(x)$ goes to $+\infty$ faster than x^2 when $|x| \to \infty$.

Let ψ_0 and E_0 be the ground-state eigenfunction and eigenvalue of H and let

$$b_\hbar(x) \equiv \frac{\hbar}{2}\frac{\mathrm{d}}{\mathrm{d}x}\log(\psi_0(x))^2.$$

One can show that the operator

$$\frac{1}{\hbar}(H - E_0) \qquad \text{on } L^2(\mathbb{R}, \mathrm{d}x)$$

is unitarily equivalent to the operator $-L_\hbar$, where

$$L_\hbar \equiv \frac{\hbar}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b_\hbar(x)\frac{\mathrm{d}}{\mathrm{d}x} \qquad \text{on } L^2(\mathbb{R}, \psi_0^2\,\mathrm{d}x).$$

In turn, adapting a theorem by Ventzel [7], the study of the lower part of the spectrum of $-L_\hbar$ can be reduced, when $\hbar \to 0$, to the study of the spectrum of a finite matrix. The problem is that we do not know ψ_0, so we do not know $b_\hbar$, an essential ingredient of $L_\hbar$. In conclusion, within this approach the problems posed in the introduction can be solved if one *only* knows the function

$$b_0(x) = \lim_{\hbar \to 0} b_\hbar(x).$$

What we want to do in this paper is the describe an algorithm to construct b_0. The procedure to calculate the constants d_i that appear in the splitting of the eigenvalues (1.1) is straightforward when one knows b_0 and is described in [9, 10].

2.1. General properties of b_0

In the following we will make extensive use of some general properties of the functions $b_\hbar$ and b_0 that have been derived in [3, 9]. Their starting point is the observation that from the Schrödinger equation for ψ_0 one obtains

$$\hbar \frac{\mathrm{d}}{\mathrm{d}x} b_\hbar(x) + (b_\hbar(x))^2 = 2(V(x) - E_0(\hbar)). \tag{2.1}$$

From this equation, equipped with appropriate conditions at infinity which ensure that $\psi_0 \in L^2(\mathbb{R}, \mathrm{d}x)$, they prove the following.

(i) Let I be a finite closed interval; then $\forall \eta > 0\ \exists \hbar_0 > 0$ such that if $\hbar < \hbar_0$ then

$$\max_{x \in I} |b_\hbar(x)| \leqslant \max_{x \in I} \sqrt{2V(x)} + \eta$$

(the functions $b_\hbar$ are uniformly bounded on every finite closed interval for small $\hbar$).

(ii) $N-1$ points $y_i \in [x_i, x_{i+1}]$, $i = 1, \ldots, N-1$, exist such that if $x \neq y_i$ for each i then $b_0(x) \equiv \lim_{\hbar \to 0} b_\hbar(x)$ exists and

$$\begin{aligned}
b_0(x) &= +\sqrt{2V(x)} && \forall x \in (-\infty, x_1) \\
b_0(x) &= -\sqrt{2V(x)} && \forall x \in [x_N, +\infty) \\
b_0(x) &= -\sqrt{2V(x)} && \forall x \in [x_i, y_i) \\
b_0(x) &= +\sqrt{2V(x)} && \forall x \in (y_i, x_{i+1}].
\end{aligned}$$

(iii) Let I be a finite closed interval and let $\alpha \in (0, 1)$. Then $\exists C > 0$ such that if $x \in I$ and $|x - y_i| > \hbar^\alpha$, $\forall i$, then

$$|b_\hbar(x) - b_0(x)| < C\hbar^\alpha$$

(i.e. the convergence of $b_\hbar$ to b_0 is uniform on every finite closed interval that does not contain any of the jump points y_i).

From propositions (i) and (iii) it follows that if I is *any* finite closed interval then

$$\lim_{\hbar \to 0} \int_I b_\hbar(x)\, \mathrm{d}x = \int_I b_0(x)\, \mathrm{d}x. \tag{2.2}$$

We see from proposition (ii) that b_0 can exhibit several (at most $N-1$) points of discontinuity where it jumps from $-\sqrt{2V(x)}$ to $+\sqrt{2V(x)}$, and that the knowledge of the locations of such points is equivalent to the knowledge of the function b_0. We remark that, since

$$y_i = x_{i+1} = y_{i+1}$$

is allowed, the number of *different* y_i can be less than $N-1$. Moreover when $y_i = x_i$ or $y_i = x_{i+1}$ there is no discontinuity in y_i, since $V(x_i) = V(x_{i+1}) = 0$. Thus the true jump points are those y_i such that $y_i \in (x_i, x_{i+1})$.

The complete determination of b_0 has been carried out in [3, 9] in many interesting cases with the help of the following equations.

Let $t \in \mathbb{R}$ and $f: \mathbb{R} \to \mathbb{R}$; we denote by $\tilde{f}$ the reflection of f about the point t and by $\bar{f}$ the translation of f

$$\tilde{f}(x) \equiv f(2t-x) \qquad \text{and} \qquad \bar{f}(x) \equiv f(x+t).$$

From (2.1) one obtains, $\forall x, y \in \mathbb{R}$,

$$(b_\hbar + \tilde{b}_\hbar)(x) = (b_\hbar + \tilde{b}_\hbar)(y) \exp\left(\frac{1}{\hbar}\int_x^y (b_\hbar - \tilde{b}_\hbar)(u)\,du\right) - \frac{2}{\hbar}\int_x^y (V-\tilde{V})(w) \exp\left(\frac{1}{\hbar}\int_x^w (b_\hbar - \tilde{b}_\hbar)(u)\,du\right) dw \tag{2.3}$$

and

$$(b_\hbar - \bar{b}_\hbar)(x) = (b_\hbar - \bar{b}_\hbar)(y) \exp\left(\frac{1}{\hbar}\int_x^y (b_\hbar + \bar{b}_\hbar)(u)\,du\right) - \frac{2}{\hbar}\int_x^y (V-\bar{V})(w) \exp\left(\frac{1}{\hbar}\int_x^w (b_\hbar + \bar{b}_\hbar)(u)\,du\right) dw. \tag{2.4}$$

The idea is that if you choose the values of x, y and t cleverly, you can get information on b_0 from the above equations without solving them explicitly, but simply using (2.3) and (2.4) as consistency conditions.

2.2. Outline of the algorithm

Before stating the precise formulation of the algorithm that enables us to determine b_0, we give some rough indications about it.

As we remarked above, the problem is to decide whether $b_0(x) = +\sqrt{2V(x)}$ or $b_0(x) = -\sqrt{2V(x)}$ for each x that does not coincide with any of the jump points y_i. For this purpose we will define recursively a sequence, $\{\Lambda_j\}$, of nested subsets of $\mathbb{R}$

$$\Lambda_0 \subset \Lambda_1 \subset \Lambda_2 \subset \ldots \subset \Lambda_k \subset \ldots \subset \mathbb{R}.$$

As regards the geometrical structure of $\{\Lambda_j\}$, each Λ_k is the disjoint union of a finite number of closed intervals, i.e.

$$\Lambda_k = [a_1, b_1] \cup \ldots \cup [a_n, b_n]$$

and, in particular, Λ_0 is the set of the points of absolute minimum of V (figure 3)

$$\Lambda_0 = \{x_1\} \cup \ldots \cup \{x_N\}.$$

Then we will prove a theorem by induction; at the step k of the induction this theorem tells us which is the sign of b_0 on the set Λ_k. When the algorithm is successfully implemented we arrive at $\Lambda_k = \mathbb{R}$ after a finite number of steps (the number of steps is always less than $2N$).

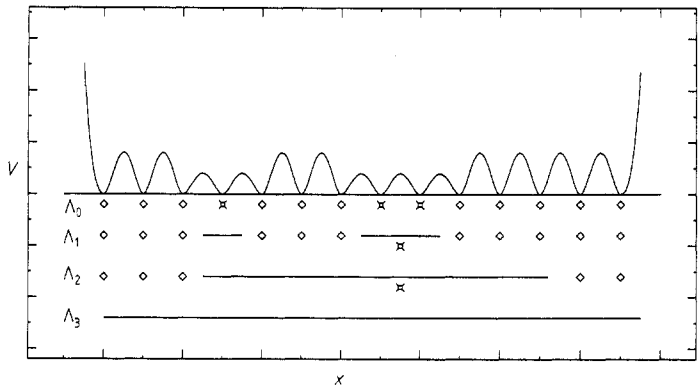

Figure 3. An example of Λ_k sets.

For the recursive definition of the set Λ_{k+1} starting from Λ_k, we proceed roughly as follows. First we look at the closed intervals which make up Λ_k and choose some of these intervals, which are said to be *maximal.* They are chosen in such a way that the potential V on a maximal interval is, in a certain sense, lower than the potential on a non-maximal interval, so that the maximal intervals are the best candidates to represent the localisation set of the ground-state wavefunction (in figure 3 maximal intervals are marked with a ⌶). Then we *extend* in some way all the maximal intervals of Λ_k, while non-maximal intervals are left unchanged. This new set that we have obtained is just Λ_{k+1}.

3. The mapping Ext

In this section we construct a mapping, which we call Ext, which allows us to define recursively a sequence $\{\Lambda_j\}$ of nested subsets of $\mathbb{R}$

$$\Lambda_0 \subset \Lambda_1 \subset \Lambda_2 \subset \ldots \subset \Lambda_k \subset \ldots \subset \mathbb{R}$$

setting $\Lambda_{k+1} \equiv \mathrm{Ext}(\Lambda_k)$.
This is by far the most important step in our algorithm because once we know the sequence $\{\Lambda_j\}$, we can get b_0 by means of an extremely simple rule.

Ext acts on those subsets of $\mathbb{R}$ which can be obtained as a union of a finite number of closed intervals, i.e. sets like

$$\Lambda = I_1 \cup \ldots \cup I_n \tag{3.1}$$

where

$$I_i = [a_i, b_i]. \tag{3.2}$$

We denote this family of subsets of $\mathbb{R}$ by Ω, so that

$$\mathrm{Ext}: \Omega \to \Omega.$$

There are of course infinitely many ways of representing an element of Ω as a union of closed intervals, but in only one representation can the intervals be taken disjoint. So, if

$$I_i \cap I_j = \varnothing \qquad \text{when } i \neq j \tag{3.3}$$

then (3.1) is said to be the *standard representation* of Λ and the intervals I_i are called the *components* of Λ. In the following we always assume that (3.3) holds, unless there is an explicit statement to the contrary.

3.1. The maximal components

The aim of this subsection is to give three definitions which are necessary to understand how Ext acts. The three terms we define are equivalence, preference and maximal components. We assume that a potential V satisfying the hypothesis of § 2 is given.

Definition of equivalence. Given two closed intervals $[a, b]$ and $[c, d]$ we say that they are *equivalent* (and write $[a, b] \approx [c, d]$) if they have the same length and if the potential V on $[c, d]$ is the translation or the reflection of the potential V on $[a, b]$, i.e. if

$$b - a = d - c$$

and either

$$V(x) = V_T(x) \equiv V(c - a + x) \qquad \forall x \in [a, b]$$

or

$$V(x) = V_R(x) \equiv V(a + d - x) \qquad \forall x \in [a, b]$$

where V_T is the potential obtained translating V form $[c, d]$ to $[a, b]$, while V_R is the reflected potential.

We say also that $[a, b] \geqslant [c, d]$ if

$$b - a = d - c$$

and either

$$V(x) \geqslant V_T(x) \equiv V(c - a + x) \qquad \forall x \in [a, b]$$

or

$$V(x) \geqslant V_R(x) \equiv V(a + d - x) \qquad \forall x \in [a, b].$$

Finally $[a, b] > [c, d]$ means $[a, b] \geqslant [c, d]$ and $[a, b] \not\approx [c, d]$.

Definition of preference. Given any closed interval $I = [a, b]$, we define

$$I^{(\varepsilon)} \equiv [a - \varepsilon, b + \varepsilon].$$

Let I and J be two closed intervals. I is said to be *preferred* to J if one of the following two propositions (p_1 or p_2) is true:

p_1: $I \approx J$ and $\exists \varepsilon_0 > 0$ such that $\forall \varepsilon < \varepsilon_0$ $I^{(\varepsilon)} < J^{(\varepsilon)}$;

p_2: there exists a closed interval K strictly contained in I such that $K \approx J$ and $K^{(\varepsilon)} \leqslant J^{(\varepsilon)}$ for some $\varepsilon > 0$.

If p_1 is true we also say that I is p_1-preferred to J; in the second case we say that I is p_2-preferred to J.

Example. Look at figure 2 and consider the intervals $[2, 3]$ and $[9, 10]$. Clearly we have

$$[2, 3] \approx [9, 10]$$

$$[2 - \varepsilon, 3 + \varepsilon] < [9 - \varepsilon, 10 + \varepsilon] \qquad \forall \varepsilon < 1.$$

So $[2, 3]$ is p_1-preferred to $[9, 10]$. Consider now $[2, 4]$ and $[5, 8]$. $[5, 8]$ turns out to be p_2-preferred to $[2, 4]$. In fact if we choose $K = [5, 7] \subset [5, 8]$ then it is easy to check that

$$[5, 7] \approx [2, 4]$$

$$[5-\varepsilon, 7+\varepsilon] \leqslant [2-\varepsilon, 4+\varepsilon] \qquad \forall \varepsilon < 1.$$

Definition of maximal components. Let $\Lambda \in \Omega$ be given by (3.1) and (3.2). By means of preference we can (partially) order the components I_i of Λ. The maximal components of Λ are those intervals which are maximal with respect to preference. In order words I_i is said to be maximal in Λ if there is no other component I_j such that I_j is preferred to I_i.

The maximal intervals are especially relevant since they are the candidates to represent the sets where the ground-state wavefunction is concentrated in the semi-classical limit.

3.2. The mapping Ext

Given a closed interval $[a, b]$ and a real number $\alpha \geqslant 0$, we define the *α-extension* of $[a, b]$ as

$$\mathrm{Ext}_\alpha([a, b]) \equiv [c, d]$$

where $[a, b] \subset [c, d]$ and c, d are determined by the conditions

$$\int_c^a \sqrt{2V(x)}\, \mathrm{d}x = \int_b^d \sqrt{2V(x)}\, \mathrm{d}x = \alpha. \tag{3.4}$$

The quantity $|\int_c^a \sqrt{2V(x)}\, \mathrm{d}x|$ is called the *Agmon distance* between a and c and will be denoted by $\rho(a, c)$.

If we have not just a single interval, but a set like Λ given by (3.1) then we define

$$\mathrm{Ext}_\alpha(\Lambda) \equiv \hat{I}_1 \cup \ldots \cup \hat{I}_n$$

where

$$\hat{I}_i = \begin{cases} \mathrm{Ext}_\alpha(I_i) & \text{if } I_i \text{ is maximal} \\ I_i & \text{if } I_i \text{ is non-maximal.} \end{cases}$$

Remark. The intervals $\hat{I}_1, \ldots, \hat{I}_n$ can of course intersect, depending on the value of α, so $\hat{I}_1 \cup \ldots \cup \hat{I}_n$ is not necessarily the standard representation of $\mathrm{Ext}_\alpha(I_1 \cup \ldots \cup I_n)$.

The last point required in order to define the mapping Ext (without index) is how to choose the value of α in Ext_α. The rule is the following. Let Λ, as usual, be given by (3.1) and suppose that the intervals I_i are ordered in such a way that $I_1, \ldots, I_m$ are maximal, while $I_{m+1}, \ldots, I_n$ are non-maximal. In this way we can write

$$\mathrm{Ext}_\alpha(\Lambda) = [a_1(\alpha), b_1(\alpha)] \cup \ldots \cup [a_m(\alpha), b_m(\alpha)] \cup [a_{m+1}, b_{m+1}] \cup \ldots \cup [a_n, b_n]$$

where $a_i(\alpha)$ and $b_i(\alpha)$ are clearly given by

$$\int_{a_i(\alpha)}^{a_i} \sqrt{2V(x)}\, \mathrm{d}x = \int_{b_i}^{b_i(\alpha)} \sqrt{2V(x)}\, \mathrm{d}x = \alpha.$$

Now we associate a real number, which we call $\alpha(\Lambda)$, with the set Λ. $\alpha(\Lambda)$ is defined as the largest non-negative real number α with the following conditions.

(i) The intervals that make up $\mathrm{Ext}_\alpha(\Lambda)$, i.e.

$$[a_1(\alpha), b_1(\alpha)], \ldots, [a_m(\alpha), b_m(\alpha)], [a_{m+1}, b_{m+1}], \ldots, [a_n, b_n]$$

do not overlap. More precisely, they are allowed to be disjoint or adjacent;

(ii) the α-extended maximal intervals

$$[a_1(\alpha), b_1(\alpha)], \ldots, [a_m(\alpha), b_m(\alpha)]$$

are all equivalent.

If there is no such value of α then we set $\alpha(\Lambda) \equiv 0$.

Finally, the mapping Ext is defined by

$$\mathrm{Ext}: \Lambda \to \mathrm{Ext}_{\alpha(\Lambda)}(\Lambda).$$

Thus we have

$$\mathrm{Ext}(\Lambda) = [a_1(\beta), b_1(\beta)] \cup \ldots \cup [a_m(\beta), b_m(\beta)] \cup [a_{m+1}, b_{m+1}] \cup \ldots \cup [a_n, b_n] \tag{3.5}$$

where $\beta = \alpha(\Lambda)$.

Remark. Owing to condition (i), some of the intervals which make up $\mathrm{Ext}(\Lambda)$ can be adjacent. In the standard representation of $\mathrm{Ext}(\Lambda)$ each group of adjacent intervals which appears in (3.5) must be replaced by one interval that covers the whole group.

Now let

$$\mathrm{Ext}(\Lambda) = J_1 \cup \ldots \cup J_n$$

be the standard representation of $\mathrm{Ext}(\Lambda)$ in terms of disjoint closed intervals J_i. Then the general structure of each J_i is clearly

$$J_i = \mathrm{Ext}_\beta(K_1) \cup \ldots \cup \mathrm{Ext}_\beta(K_p) \cup L_1 \cup \ldots \cup L_q \tag{3.6}$$

where K_i are some maximal components of Λ and L_i are some non-maximal components. The $p+q$ intervals K_i and L_i are of course all adjacent. Moreover the p intervals $\mathrm{Ext}_\beta(K_i)$ are all equivalent.

It will be useful in the following to also define the left and the right extensions of Λ given by

$$\mathrm{Lext}(\Lambda) \equiv (a_1(\beta), a_1) \cup \ldots \cup (a_m(\beta), a_m)$$

$$\mathrm{Rext}(\Lambda) \equiv (b_1, b_1(\beta)) \cup \ldots \cup (b_m, b_m(\beta))$$

with $\beta \equiv \alpha(\Lambda)$. Clearly we have

$$\mathrm{Ext}(\Lambda) = \Lambda \cup \mathrm{cl}\,\mathrm{Lext}(\Lambda) \cup \mathrm{cl}\,\mathrm{Rext}(\Lambda)$$

where cl stands for the topological closure.

4. The algorithm for constructing b_0

We assume that a potential V satisfying the hypothesis of § 2 is given and that $x_1, \ldots, x_N$ are the points of absolute minimum of V. We set

$$\Lambda_0 = \{x_1\} \cup \ldots \cup \{x_N\} \qquad L_0 \equiv R_0 \equiv \varnothing$$

and then define recursively

$$\Lambda_k \equiv \mathrm{Ext}(\Lambda_{k-1}) \qquad L_k \equiv \mathrm{Lext}(\Lambda_{k-1}) \qquad R_k \equiv \mathrm{Rext}(\Lambda_{k-1}).$$

For simplicity of notation we also set

$$\alpha_k \equiv \alpha(\Lambda_k).$$

As we observed at the end of previous section, we have

$$\Lambda_k = \Lambda_{k-1} \cup \mathrm{cl}\, L_k \cup \mathrm{cl}\, R_k$$

which implies

$$\Lambda_k = \Lambda_0 \cup \left(\bigcup_{j \leqslant k} \mathrm{cl}\, L_j \right) \cup \left(\bigcup_{j \leqslant k} \mathrm{cl}\, R_j \right).$$

Thus it is clear that if one knows b_0 on the sets L_j and R_j $\forall j \leqslant k$ then one also knows b_0 on Λ_k (with the exception of a finite number of isolated points, but this is not relevant since we are always interested in quantities like $\int_a^b b_0(x)\,\mathrm{d}x$). The key theorem is then the following.

Theorem 4.1. For each natural number k we have

$$b_0(x) = +\sqrt{2V(x)} \qquad \text{if } x \in L_k$$

$$b_0(x) = -\sqrt{2V(x)} \qquad \text{if } x \in R_k.$$

The algorithm has a successful implementation if $\Lambda_k = \mathbb{R}$ for some k. We will show in a related paper [10] that this situation is very likely, at least in some interesting classes of potentials.

It is easy to convince oneself that the algorithm must stop (that is $\Lambda_k = \Lambda_{k-1}$) after a number of steps which is not greater than $2N$. In fact, from conditions (i) and (ii) that define the quantity $\alpha(\Lambda)$ in § 3.2, it follows that, at each step, either (at least) two components of Λ_k become adjacent or (at least) one maximal component becomes inequivalent to the other ones. In the first case the total number of components of Λ_k decreases, while in the second case the number of maximal components must decrease. But this can happen at most $2N$ times since we start with Λ_0 which has N components and at most N maximals.

5. Complexity of the algorithm

It is well known to information theorists, as well as to most people who have bought CPU time from some computational centre, that one of the main features of any algorithm is its complexity, i.e. the number of elementary operations (we also call this number *time* or *cost*) which are necessary to solve a problem as a function of the *dimension* of the problem. Thus we dedicate this section to estimating the time T_N that our algorithm takes to solve a potential, as a function of the number N, of the absolute minima of the potential.

Of course, a true algorithmic solution (i.e. in a finite number of steps) is possible only for those potentials which can be specified by giving a finite amount of information, so we consider in the following the (quite natural) class of potentials which are sequences of a certain number of barriers of fixed shape. More precisely consider k types of potential barriers V_i with support in $[0, 1]$ which satisfy

$$\begin{aligned} &0 < V_1(x) < V_2(x) < \ldots < V_k(x) \qquad \forall x \in (0, 1) \\ &V_i(0) = V_i(1) = 0 \qquad \forall i = 1, \ldots, k. \end{aligned} \tag{5.1}$$

Now to each sequence of N integers $\{h_i\}_{i=1}^N$ with $1 \leq h_i \leq k$, we associate a potential V on $[0, N]$ given by

$$V(x) = V_{h_i}(x - [x]) \qquad \text{if } x \in [i-1, i]$$

where $[x]$ denotes the integral part of x. This is just the potential that one obtains if one puts the barriers $V_{h_1}, V_{h_2}, \ldots, V_{h_N}$ in sequence. V is defined outside $[0, N]$ in such a way that it increases faster than x^2 when $|x| \to \infty$ (the particular shape of V in this region is not relevant). The family of all the potentials that can be obtained in this way is denoted by P_k^N. It clearly contains k^N elements. If V belongs to P_k^N for some N and k, we call it a *string potential.*

Now we want to show that the time T_N, required by the algorithm described in the previous section to solve a string potential, obeys

$$T_N \leq CN^3 \tag{5.2}$$

for some constant C.

Before reading the following argument it could be useful to glance at table 1, where some experimental values of T_N are collected.

Table 1. CPU time required to solve a string potential for some values of N.

N	CPU time (s) VAX 8650
100	0.018
300	0.120
1000	1.0
3000	8.0

Since a detailed proof of (5.2) is quite involved, we will make use of some simplifying assumptions and intuitive statements. Moreover we think that, by means of arguments more refined than ours (which are actually very rough), bound (5.2) can be improved. This is also suggested by the data of table 1. Our primary interest was, however, to show that the algorithm has a polynomial complexity.

As we remarked at the end of the previous section, the number of recursive steps cannot exceed $2N$, so

$$T_N \leq 2Nt$$

where t is the maximum time required to construct Λ_{p+1} starting from Λ_p. We decompose t as

$$t = t_1 + t_2 + t_3$$

where t_1 is the time needed to find the maximal components of Λ_p, t_2 is the time taken to compute α_p and t_3 is the time necessary to construct the α_p-extension of all maximals.

As usual we let

$$\Lambda_p = I_1 \cup \ldots \cup I_n.$$

In order to simplify the discussion we assume that each component I_i consists of an integral number of barriers (this is not true in general, but eliminates some boundary N-independent complications). In this way, with each component I_i of length l_i, we can associate a l_i-ple

$$(h_1^{(i)}, \ldots, h_{l_i}^{(i)})$$

which gives the heights of the barriers that make up I_i. We also define

$$m = \text{number of maximal components in } \Lambda_p.$$

Let us begin by estimating t_1, t_2 and t_3.

t_3 is the easiest to compute, since it is simply given by the product of the number of maximals with something which is independent of N. So we have

$$t_3 \sim A_3 m \leq A_3 N.$$

As regards t_1, we know that, in order to find the maximals, we have to compare each pair (I_i, I_j) of components of Λ_p and check if I_i is preferred to I_j (of course, in most cases we do not need to compare *each* pair). If we set

$$p_{ij} = \text{cost of } (I_i, I_j) \text{ preference test}$$

we obtain

$$t_1 = \sum_{i,j=1}^{n} p_{ij} = \sum_{i,j=1}^{n} (p_{ij}^{(1)} + p_{ij}^{(2)})$$

where we have separated the cost of checking p_1-preference from the cost of checking p_2-preference. The cost of p_1-preference is (apart from boundary N-independent complications) equal to the cost of *equivalence*. In turn I_i is equivalent to I_j if and only if

$$l_i = l_j$$

and either

$$h_a^{(i)} = h_a^{(j)} \qquad \forall a = 1, \ldots, l_i$$

or

$$h_a^{(i)} = h_{l_i - a + 1}^{(j)} \qquad \forall a = 1, \ldots, l_i.$$

In the last of these conditions we have supposed that each barrier $V_i(x)$ given by (5.1) is symmetric about its middle point, otherwise reflection equivalence would not be possible. From the above conditions it follows that the cost of the equivalence test is given by (twice) the number of barriers that make up each interval if they have the same length, while it is an N-independent constant if the two intervals have different lengths, so

$$p_{ij}^{(1)} \sim l_i \delta(l_i, l_j)$$

where $\delta(\cdot, \cdot)$ is the Kronecker delta.

p_2-preference is a little more complicated. Assume in fact $l_i > l_j$. We have to look for a subinterval K of I_i, of length l_j, such that K is equivalent to I_j. Again we observe that the second condition

$$K^{(\varepsilon)} \leqslant I_j^{(\varepsilon)} \qquad \text{for some } \varepsilon > 0$$

is a boundary condition with a fixed cost, so we do not worry about it. The number of all possible subintervals is, *a priori*, $l_i - l_j + 1$, since the absolute minima of K must be put in correspondence with those of I_j. Thus, since the cost of equivalence is l_j, we get

$$p_{ij}^{(2)} \sim (l_i - l_j + 1) l_j$$

and

$$\begin{aligned} t_1 &= \sum_{i,j=1}^{n} (p_{ij}^{(1)} + p_{ij}^{(2)}) \\ &\sim \sum_{i,j=1}^{n} (l_i \delta(l_i, l_j) + |l_i - l_j + 1| \min\{l_i, l_j\}) \\ &\leqslant \sum_{i,j=1}^{n} (l_i + l_i l_j) \leqslant 2 \sum_{i,j=1}^{n} l_i l_j = 2\left(\sum_{i=1}^{n} l_i \right)^2 \leqslant 2N^2 \end{aligned}$$

so that

$$t_1 \leqslant A_1 N^2.$$

It remains to estimate the time necessary to find $\alpha(\Lambda_p)$. To compute $\alpha(\Lambda_p)$ we proceed roughly as follows. Consider the maximal components of Λ_p and:

(1) add one barrier on the left and one barrier on the right of each of them (again we are simplifying, since the α extension must be taken symmetric in the Agmon distance, but this does not introduce any relevant complication);

(2) check the conditions (i) and (ii) which define the quantity $\alpha(\Lambda)$;

(3) if they are both satisfied then go back to (1), else stop.

Now let L be the (common) length of the α_p extension of the maximal components and g be number of iterations of the above procedure (1) → (2) → (3), i.e. is the number of barriers added to the left (or to the right) of each maximal. Also let $s_{(i)}$ and $s_{(ii)}$ be the costs of checking conditions (i) and (ii) of § 3.2, respectively. Clearly we have

$$t_2 \sim g(m + s_{(i)} + s_{(ii)}).$$

In order to test condition (i), we can check, for example, if the first maximal is equivalent to each of the remaining $m-1$ maximals, so we need $m-1$ equivalence tests on intervals which have a length not greater than L. This implies

$$s_{(i)} \sim mL \leqslant N.$$

Condition (ii) requires $2m$ operations to see if each maximal has intersected any of its two neighbouring intervals, so

$$s_{(ii)} \sim m \leqslant N.$$

Besides, we observe that the total number of added barriers cannot exceed N, so we can write

$$2gm \leqslant N$$

and obtain

$$t_2 \sim g(m + mL + m) \leq g(3N) \leq 3N^2/2m \leq A_2 N^2.$$

Now we have all the ingredients required to estimate T_N

$$T_N \leq 2N(t_1 + t_2 + t_3) \leq 2N(A_1 N^2 + A_2 N^2 + A_3 N) \leq CN^3.$$

6. Some examples

Here we give some examples to see how to apply the algorithm in practice, with the help of some figures. We always consider a potential V which is a sequence of barriers of two types (binary potentials) or three types (ternary potentials).

The closed intervals that make up Λ_k are represented in the figures by means of horizontal lines, except the isolated points which are never drawn, since they have no effect after the first step.

A ⌶ posed under an interval of Λ_k denotes that such an interval is maximal.

As soon as we find Λ_k we write the sign of b_0 following the rule to theorem 4.1: *plus* on L_k and *minus* on R_k.

Example 1 (*figure* 4). Here we have a binary potential. The maximal components of Λ_0 are the three minima which lie between two adjacent low barriers. They are marked with a ⌶. Now we construct Λ_1 by means of the function Ext: we have to find three closed intervals centred (in the Agmon metric) on each maximal component of Λ_0, as large as possible, but they must be equivalent and they must be disjoint or adjacent. Now we put a + sign on the left extension of each of these intervals and a − sign on the right extension.

If two or more intervals have come into contact they will be considered as one interval only, so Λ_1 is made up (apart from isolated points) of two intervals. We call the left interval J and the right interval I. It is easy to check that I is p_2-preferred to J (the role of K can be played by either of the two halves of I), so I is the only maximal component in Λ_1. Once we remain with one only maximal element the game is over. In fact the maximal interval will begin to extend and at each step of the iteration it will merge with at least one non-maximal element of Λ_k. In this way, after

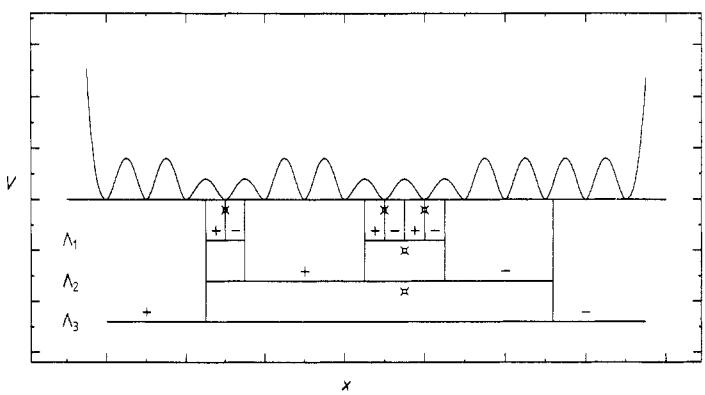

Figure 4. Example 1; application of the algorithm to a binary potential.

a finite number of steps Λ_k (in our example Λ_2) itself will be just one interval, so that $\Lambda_{k+1} = \mathbb{R}$. Thus the sign of b_0 is *plus* on the left of the maximal interval and *minus* on its right (except, of course, in those regions of $\mathbb{R}$ where the sign has been already found in a previous step).

In this example we have skipped all the steps where one interval comes into contact with an isolated point of Λ_k, since that has no consequences in the procedure.

Example 2 (*figure 5*). We again have a binary potential. Here Λ_2 contains (apart from isolated points) three intervals, but the first one (marked with a ⌶) is the only maximal interval: in fact it is p_1-preferred to the second interval and p_2-preferred to the third one. Therefore, as explained in example 1, we can skip the intermediate steps and establish the sign of b_0 everywhere. We repeat for clarity that the + and − which appear on the Λ_3 line have to be considered valid only in the regions not already covered by Λ_2 or by Λ_1.

Example 3 (*figure 6*). Finally we would like to finish with a failure! We have chosen a ternary potential, since the analogous situation with a binary potential would require at least 20 barriers. In this case we meet the following problem: Λ_1 contains (apart from points) two closed intervals that we call I and J. We see from figure 6 that, of course

$$I \approx J$$

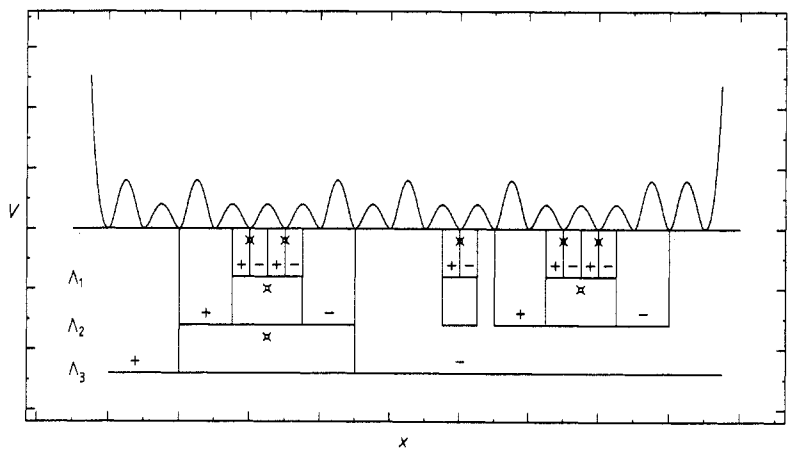

Figure 5. Example 2; application of the algorithm to another binary potential.

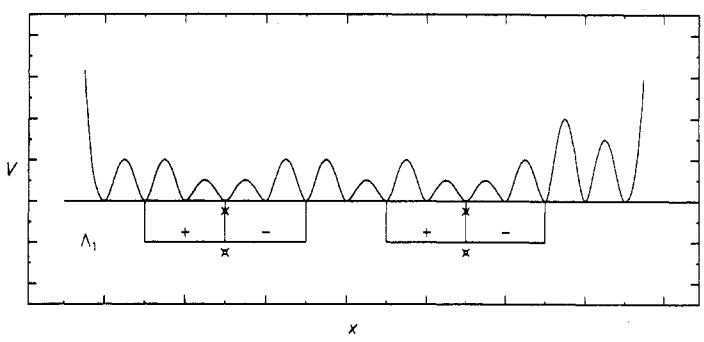

Figure 6. Example 3; failed application of the algorithm to a tenary potential.

but, for each $\varepsilon > 0$, the two statements

$$I^{(\varepsilon)} \geqslant J^{(\varepsilon)} \tag{6.1}$$

$$I^{(\varepsilon)} \leqslant J^{(\varepsilon)} \tag{6.2}$$

are both false. Moreover

$$\mathrm{Ext}_\alpha(I) \not\approx \mathrm{Ext}_\alpha(J) \qquad \forall \alpha > 0. \tag{6.3}$$

Expressions (6.1) and (6.2) tell us that neither is I preferred to J nor J preferred to I. So the two intervals are both maximal. But (6.3) implies that $\alpha_1 \equiv \alpha(\Lambda_1) = 0$, and, in consequence, $\Lambda_2 \equiv \mathrm{Ext}(\Lambda_1) = \Lambda_1$. We see that, in this case, the algorithm stops before Λ_k has covered $\mathbb{R}$.

7. Proof of the theorem

Let $I = [a, b]$ be any closed interval. In the proof of theorem 4.1 it will be essential to know the sign of b_0 on some neighbourhood of I when I is one of the components of Λ_k. Since the number of points where b_0 changes its sign is finite we can always find two points $z_1 < a$ and $z_2 > b$ such that b_0 has a definite sign on (z_1, a) and on (b, z_2). In this way we can divide the closed intervals into four classes

$$(++) \quad \text{if} \quad b_0(x) = +\sqrt{2V(x)} \qquad \forall x \in (z_1, a) \cup (b, z_2)$$

$$(--) \quad \text{if} \quad b_0(x) = -\sqrt{2V(x)} \qquad \forall x \in (z_1, a) \cup (b, z_2)$$

$$(-+) \quad \text{if} \begin{cases} b_0(x) = -\to \sqrt{2V(x)} & \forall x \in (z_1, a) \\ b_0(x) = +\sqrt{2V(x)} & \forall x \in (b, z_2) \end{cases}$$

$$(+-) \quad \text{if} \begin{cases} b_0(x) = +\sqrt{2V(x)} & \forall x \in (z_1, a) \\ b_0(x) = -\sqrt{2V(x)} & \forall x \in (b, z_2). \end{cases}$$

We want also to quantify the amount of plus sign on the left of I and the amount of minus sign on its right. So let $\alpha > 0$ and $[c, d] \equiv \mathrm{Ext}_\alpha(I)$. We define

$$p_\alpha(I) \equiv \inf\{y \in [c, a] \mid b_0(x) = +\sqrt{2V(x)}\ \forall x \in (y, a)\} \tag{7.1}$$

$$m_\alpha(I) \equiv \sup\{y \in [b, d] \mid b_0(x) = -\sqrt{2V(x)}\ \forall x \in (b, y)\} \tag{7.2}$$

and

$$P_\alpha(I) \equiv \rho(p_\alpha(I), a) \qquad M_\alpha(I) \equiv \rho(m_\alpha(I), b)$$

where ρ is the Agmon distance defined in (3.4). Clearly

$$c \leqslant p_\alpha(I) \leqslant a \qquad b \leqslant m_\alpha(I) \leqslant d$$

$$0 \leqslant P_\alpha(I) \leqslant \alpha \qquad 0 \leqslant M_\alpha(I) \leqslant \alpha.$$

Finally we set

$$Z_\alpha(I) = \min\{P_\alpha(I), M_\alpha(I)\}.$$

$Z_\alpha(I)$ is positive, of course, if and only if I is a $(+-)$ interval.

Remark. It is worth recalling here that b_0 is the (semiclassical) log derivative of the ground-state wavefunction ψ_0, so a large value of $Z_\alpha(I)$ is intuitively connected to an expectation of finding ψ_0 concentrated in I. One of the main steps in the proof of the theorem is to show that only the *maximal components* in Λ_k are allowed to have a value of $Z_\alpha(\cdot)$ greater than zero.

In order to find the sign of b_0 it will be crucial to recognise that some subsets of $\mathbb{R}$ have the following interesting property.

Definition 7.1. Let A be any subset of $\mathbb{R}$. We say that A is *transitive* if the following statements are true:

(i) if $b_0(x) > 0$ for some $x \in A$, then $b_0(u) = +\sqrt{2V(u)} \geqslant 0 \; \forall u \in A, \; u > x$;

(ii) if $b_0(x) < 0$ for some $x \in A$, then $b_0(u) = -\sqrt{2V(u)} \leqslant 0 \; \forall u \in A, \; u < x$.

In other words transitive sets can contain at most one point where the sign of b_0 changes from *minus* to *plus* and no points where the sign goes from *plus* to *minus*. Proposition (ii) of § 2 just says that the intervals $[x_i, x_{i+1}]$ are *transitive*.

Now let (a, b) and (c, d) be two disjoint open intervals (for example $b < c$) and assume that they are both transitive. It is easy to see that if $[b, c]$ is of type $(++)$, $(--)$ or $(-+)$ then the set $(a, b) \cup (c, d)$ is also transitive. So the intervals which belong to the above three classes can *connect* transitive intervals and give more complex transitive sets. For simplicity we give them a specific name.

Definition 7.2. Closed intervals of class $(++)$, $(--)$ and $(-+)$ are said to be *connective*.

It is clear that a closed interval $[a, b]$ is connective if and only if

$$Z_\alpha([a, b]) = 0 \qquad \forall \alpha > 0.$$

we can now state two lemmas (for their proofs see the appendix) that constitute the basic instruments for determining b_0.

Lemma 7.1. Let I and J be two closed intervals and let $\alpha > 0$ be such that

$$\mathrm{Ext}_\alpha(I) \approx \mathrm{Ext}_\alpha(J) \qquad \text{and} \qquad \int_I b_0(x)\, dx = \int_J b_0(x)\, dx = 0.$$

Then $Z_\alpha(I) = Z_\alpha(J)$.

Lemma 7.2. Let I and J be two closed intervals such that

$$I \text{ is } p_1\text{-preferred to } J \qquad \text{and} \qquad \int_I b_0(x)\, dx = \int_J b_0(x)\, dx = 0.$$

Then J is connective or, equivalently, $Z_\alpha(J) = 0 \; \forall \alpha > 0$.

The relevence of the above two lemmas is the following. The problem of finding the ground-state wavefunction in a multiple-well potential is a global one. We mean that the knowledge of the potential V on some finite interval cannot give, by itself (we stress that in our approach we never need to know E_0), any information about the behaviour of ψ_0 on that interval. For this reason, any device which allows us, under certain conditions, to subdivide the global problem into several local problems is of fundamental utility. Lemmas 7.1 and 7.2 do just that. In fact, if we can find two intervals I and J which satisfy the second condition of each lemma, then we can get information on b_0 (which is contained in the quantity $Z_\alpha(\cdot)$), simply by looking at the potential in some neighbourhood of the two intervals.

The key observation at this point is the following. If we assume theorem 4.1 to hold, then the second condition of lemmas 7.1 and 7.2 is satisfied for each component of Λ_k, since each component consists of an equal number of left and right extensions. Moreover the second condition is trivially satisfied if I and J are components of Λ_0. This strongly suggests that an inductive proof could have a good chance of success.

7.1. Proof of theorem 4.1.

In order to prove the theorem will prove each of the following statements for each k:

(A) if (x, y) is any open interval such that $(x, y) \cap \Lambda_k = \varnothing$ then (x, y) is transitive;

(B) $b_0(x) = +\sqrt{2V(x)}$ if $x \in L_k$ and $b_0(x) = -\sqrt{2V(x)}$ if $x \in R_k$ or, equivalently, $Z_{\alpha_{k-1}}(I) = \alpha_{k-1}$ if I is a maximal component in Λ_{k-1};

(C) if I is any closed interval which is equivalent to some component of Λ_k (in particular I can be itself a component of Λ_k) then $\int_I b_0(x)\,dx = 0$;

(D) if J is a non-maximal component of Λ_k then J is connective;

(E) if (x, y) is any open interval that does not intersect any maximal component of Λ_k then $(x, y) \setminus \Lambda_k$ is transitive.

Proposition B is just theorem 4.1.

The proof is by induction on k and is organised in several steps. We denote by A_k the proposition A at the kth step of the induction, and similarly B_k, C_k, etc. The following steps enable, as one can check, the proof of all the statements A-E for each k.

(1) Proof of A_0, B_0, C_0, D_0

(2) $A_k + D_k \Rightarrow E_k$

(3) $A_{k-1} \Rightarrow A_k$

(4) $C_{k-1} + E_{k-1} \Rightarrow B_k$

(5) $C_{k-1} + B_k \Rightarrow C_k$

(6) $D_{k-1} + B_k + C_k \Rightarrow D_k$.

Remark. We are aware that the following proof, though elementary, requires a good deal of patience, so we suggest that first the reader concentrate his attention just on step 4, which is the core of the proof. Then, if he is interested, he can go back to the other steps.

Step 1. A_0 follows from proposition (ii) of § 2, since Λ_0 is the set of the minima of V. B_0 and C_0 are trivial. As regards D_0, we note that if I is the component of Λ_0 which is preferred to J, since I and J are both points, then I is p_1-preferred to J. Thus we can apply lemma 7.2 and obtain D_0.

Step 2. $A_k + D_k \Rightarrow E_k$. Since Λ_k is the union of a finite number of closed intervals

$$\Lambda_k = [a_1, b_1] \cup \ldots \cup [a_n, b_n]$$

then $(x, y) \setminus \Lambda_k$ has the form

$$(x, y) \setminus \Lambda_k = (b_j, a_{j+1}) \cup \ldots \cup (b_l, a_{l+1})$$

where the intervals $[a_i, b_i]$, $i = j+1, \ldots, l$, are, by the hypothesis, non-maximal in Λ_k and, by proposition D_k, connective. So $(x, y) \setminus \Lambda_k$ is the union of transitive (by proposition A_k) open intervals separated by connective intervals. This implies, as we remarked above, that $(x, y) \setminus \Lambda_k$ is itself transitive.

Step 3. $A_{k-1} \Rightarrow A_k$. Since $\Lambda_{k-1} \subset \Lambda_k$, proposition A_k is weaker that A_{k-1}.

Step 4. $C_{k-1} + E_{k-1} \Rightarrow B_k$. This is the core of the proof. Let

$$K_i = [x_i, y_i] \qquad i = 1, \ldots, p$$

be the p maximal components of Λ_{k-1}. We assume that they are in increasing order, i.e. any point inside K_l is greater than any point inside K_m if $l > m$. Also let

$$[a_i, b_i] \equiv \mathrm{Ext}_\beta(K_i)$$

where we have set $\beta = \alpha_{k-1}$. So the left and the right extensions of Λ_{k-1} are given by

$$L_k = \bigcup_{i=1}^{p} (a_i, x_i) \qquad R_k = \bigcup_{i=1}^{p} (y_i, b_i)$$

and

$$\rho(a_i, x_i) = \rho(y_i, b_i) = \beta.$$

Assume now that $\exists w \in (a_j, x_j)$ such that $b_0(w) < 0$ and let

$$W \equiv \rho(w, x_j) < \beta.$$

We want to show that this leads to a contradiction. The strategy is the following. Choose one point w_i in each open interval (a_i, x_i) determined by the condition

$$\rho(w_i, x_i) = W$$

so that $w = w_j$. Since $b_0(w) < 0$ we have $p_\beta(K_j) \geqslant w_j$ (see equation (7.1)) and, by consequence

$$P_\beta(K_j) \leqslant W < \beta. \tag{7.3}$$

The key point is to prove that (7.3) implies $P_\beta(K_{j-1}) \leqslant W$. In fact applying this argument $j-1$ times, we obtain $P_\beta(K_1) \leqslant W$. In that case we could find a point $z \in (a_1, x_1)$ such that $b_0(z) < 0$. But this is impossible. In fact the interval $(-\infty, x_1)$ does not intersect any maximal component of Λ_{k-1} and so the set $(-\infty, x_1) \backslash \Lambda_{k-1}$ is transitive by proposition E_{k-1}. But since z is contained in $(-\infty, x_1) \backslash \Lambda_{k-1}$ and $b_0(z) < 0$, then we conclude that

$$b_0(x) = -\sqrt{2V(x)} \qquad \forall x \in (-\infty, \inf \Lambda_{k-1})$$

in contradiction with proposition (ii) of § 2.

Now we prove that $P_\beta(K_{j-1}) \leqslant W$ follows from (7.3). Consider, in fact, the interval (y_{j-1}, x_j). It does not intersect any maximal component of Λ_{k-1}, so $(y_{j-1}, x_j) \backslash \Lambda_{k-1}$ is transitive. Clearly

$$(y_{j-1}, b_{j-1}) \cup (a_j, x_j) \subset (y_{j-1}, x_j) \backslash \Lambda_{k-1}.$$

Since $w \in (a_j, x_j)$, from $b_0(w) < 0$ it follows, by transitivity, that

$$b_0(x) = -\sqrt{2V(x)} \qquad \forall x \in (y_{j-1}, b_{j-1})$$

so that

$$M_\beta(K_{j-1}) = \beta. \tag{7.4}$$

We know that all α_{k-1} extensions of the maximal components of Λ_{k-1} must be equivalent and, in particular,

$$\mathrm{Ext}_\beta(K_{j-1}) \approx \mathrm{Ext}_\beta(K_j).$$

Besides, proposition C_{k-1} tells us that

$$\int_{K_j} b_0(x)\,dx = \int_{K_{j-1}} b_0(x)\,dx = 0.$$

Thus we can apply lemma 7.1 and obtain

$$Z_\beta(K_{j-1}) = Z_\beta(K_j) \leqslant P_\beta(K_j) \leqslant W < \beta$$

which, combined with (7.4), gives

$$P_\beta(K_{j-1}) \leqslant W.$$

In this way we have shown that b_0 is non-negative on L_k. Analogously one proves that $b_0(x) = -\sqrt{2V(x)}\ \forall x \in R_k$.

Step 5. $C_{k-1} + B_k \Rightarrow C_k$. Let I be a closed interval equivalent to some component, J, of Λ_k. We want to prove that $\int_I b_0(x)\,dx = 0$. Let, as usual, $\beta \equiv \alpha_{k-1}$. From (3.6) we know that the general structure of J is

$$J = \mathrm{Ext}_\beta(K_1) \cup \ldots \cup \mathrm{Ext}_\beta(K_p) \cup F_1 \cup \ldots \cup F_q$$

where K_i are some maximal component of Λ_{k-1}, while F_i are some non-maximal components. Since $I \approx J$ we can decompose I as

$$I = \mathrm{Ext}_\beta(M_1) \cup \ldots \cup \mathrm{Ext}_\beta(M_p) \cup N_1 \cup \ldots \cup N_q \tag{7.5}$$

in such a way that $M_i \approx K_i$, $\mathrm{Ext}_\beta(M_i) \approx \mathrm{Ext}_\beta(K_i)$ and $N_i \approx F_i$. From proposition C_{k-1} we obtain

$$\int_{M_i} b_0(x)\,dx = 0 \qquad \int_{N_i} b_0(x)\,dx = 0. \tag{7.6}$$

Now we apply lemma 7.1 to each couple of intervals K_i and M_i and get

$$Z_\beta(M_i) = Z_\beta(K_i) = \beta$$

where the last equality comes from proposition B_k (remember that K_i are maximal). But this implies that

$$\int_{\mathrm{Ext}_\beta(M_i)} b_0(x)\,dx = \int_{M_i} b_0(x)\,dx = 0. \tag{7.7}$$

Equations (7.5)-(7.7) give finally

$$\int_I b_0(x)\,dx = 0.$$

Step 6. $D_{k-1} + B_k + C_k \Rightarrow D_k$. Let J be a non-maximal component of Λ_k. We have to prove that J is connective. Since J is non-maximal, a component I of Λ_k must exist such that I is preferred to J. Proposition C_k tells us that

$$\int_I b_0(x)\,dx = 0 \qquad \text{and} \qquad \int_J b_0(x)\,dx = 0.$$

If I is p_1-preferred to J we apply lemma 7.2 and obtain our thesis directly. Consider therefore the p_2 case. By definition, a closed interval K strictly contained in I exists such that

(a) $K \approx J$

(b) $\exists \varepsilon > 0$ such that $K^{(\varepsilon)} \leqslant J^{(\varepsilon)}$.

Proposition (b) can be written as either

(b_1) $\exists \varepsilon > 0$ such that $\forall \delta < \varepsilon$, $K^{(\delta)} < J^{(\delta)}$

or

(b_2) $\exists \varepsilon > 0$ such that $K^{(\varepsilon)} \approx J^{(\varepsilon)}$.

Since K is equivalent to a component of Λ_k, we know from proposition C_k that

$$\int_K b_0(x)\,dx = 0.$$

If (b_1) is true then we have only to observe that, in this case, K is p_1-preferred to J, so we can use lemma 7.2, and conclude that J is connective. The (b_2) case is much more intriguing and a detailed proof of statement D in this case would be very cumbersome and not particularly instructive, so we prefer to give only a sketch of how to carry it out.

The idea is to show that

$$Z_\alpha(K) = 0 \qquad \forall \alpha > 0 \tag{7.8}$$

and then to apply lemma 7.1 to J and K (this is possible, since if α is sufficiently small then $\mathrm{Ext}_\alpha(J) \approx \mathrm{Ext}_\alpha(K)$). In order to get (7.8), it is necessary to represent I and J as unions of adjacent intervals

$$I = \mathrm{Ext}_{\beta_1}(M_1) \cup \ldots \cup \mathrm{Ext}_{\beta_p}(M_p)$$

$$J = \mathrm{Ext}_{\gamma_i}(N_1) \cup \ldots \cup \mathrm{Ext}_{\gamma_\vartheta}(N_q)$$

where β_i, $\gamma_i \in \{\alpha_1, \ldots, \alpha_{k-1}\}$ and M_i and N_i are closed intervals which are maximal in some Λ_j with $j < k$. Such a representation always exists and is unique (this follows from (3.6) and the induction principle). Now is possible to prove that from $K \subset I$ and $K^{(\varepsilon)} \approx J^{(\varepsilon)}$ it follows that K must be of the form

$$K = \mathrm{Ext}_{\beta_r}(M_r) \cup \ldots \cup \mathrm{Ext}_{\beta_s}(M_s)$$

where $r \geqslant 1$ and $s \leqslant p$, but if $r = 1$ then $s < p$.

If we assume $s < p$ (the case $s = p$, $r > 1$ is analogous) then we see that K is adjacent on its right-hand side to $\mathrm{Ext}_{\beta_{s+1}}(M_{s+1})$. By proposition B_k this means that K has a certain amount of plus sign on its right, i.e. K is of class $(-+)$ or $(++)$, so it is connective and (7.8) follows.

Acknowledgments

I would like to thank G Jona-Lasinio, F Martinelli and E Scoppola for having suggested the problem and for helpful suggestions and discussions.

Appendix

We give below the proofs of the two lemmas we used in § 7.

Lemma 7.1. Let I and J be two closed intervals and let $\alpha > 0$ be such that

$$\mathrm{Ext}_\alpha(I) \approx \mathrm{Ext}_\alpha(J) \qquad \text{and} \qquad \int_I b_0(x)\,\mathrm{d}x = \int_J b_0(x)\,\mathrm{d}x = 0.$$

Then $Z_\alpha(I) = Z_\alpha(J)$.

Proof. We suppose that $\mathrm{Ext}_\alpha(I)$ and $\mathrm{Ext}_\alpha(J)$ are equivalent by translation (the reflection case is analogous) and set (see figure 7)

$$I = [i_3, i_4] \qquad \mathrm{Ext}_\alpha(I) = [i_1, i_6] \qquad i_2 = p_\alpha(I)$$

$$i_5 = m_\alpha(I) \qquad P_I = P_\alpha(I) = \rho(i_2, i_3) \qquad M_I = M_\alpha(I) = \rho(i_4, i_5)$$

and analogously for J

$$J = [j_3, j_4] \qquad \mathrm{Ext}_\alpha(J) = [j_1, j_6] \qquad j_2 = p_\alpha(J)$$

$$j_5 = m_\alpha(J) \qquad P_J = P_\alpha(J) = \rho(j_2, j_3) \qquad M_J = M_\alpha(J) = \rho(j_4, j_5).$$

From the first hypothesis of the lemma we know that

$$V(x) = V(x+t) \qquad \forall x \in [i_1, i_6]$$

where $t \equiv j_1 - i_1 = j_6 - i_6 = j_3 - i_3 = j_4 - i_4$. We have to prove that

$$\min\{P_I, M_I\} = \min\{P_J, M_J\}.$$

Without loss of generality we assume

$$P_J \leqslant M_J$$

and show that if

$$\min\{P_I, M_I\} - P_J = a \neq 0 \tag{A1}$$

a contradiction follows. Again we can assume $a > 0$ (the opposite case can be treated analogously).

Consider equation (2.4) with $\bar{b}_\hbar(x) \equiv b_\hbar(x+t)$ and $x, y \in [i_1, i_6]$. In this case we have

$$V(w) - \bar{V}(w) = V(w) - V(w+t) = 0 \qquad \forall w \in (x, y)$$

by hypothesis and so the second term in (2.4) can be dropped, yielding

$$(b_\hbar - \bar{b}_\hbar)(x) = (b_\hbar - \bar{b}_\hbar)(y) \exp\left(\frac{1}{\hbar}\int_x^y (b_\hbar + \bar{b}_\hbar)(u)\,\mathrm{d}u\right) \tag{A2}$$

for each $x, y \in [i_1, i_6]$. We know also that, when $\hbar \to 0$,

$$(b_\hbar - \bar{b}_\hbar) \to (b_0 - \bar{b}_0)(x) \qquad (b_\hbar - \bar{b}_\hbar)(y) \to (b_0 - \bar{b}_0)(y) \tag{A3}$$

and

$$\int_x^y (b_\hbar + \bar{b}_\hbar)(u)\,\mathrm{d}u \to \int_x^y (b_0 + \bar{b}_0)(u)\,\mathrm{d}u$$

(see equation (2.2)). The idea is to show that, unless $Z_\alpha(I) = Z_\alpha(J)$, it is possible to find two values x_0 and y_0 such that

$$c \equiv (b_0 - \bar{b}_0)(x_0) = +2\sqrt{2V(x_0)} > 0 \tag{A4}$$

and

$$d \equiv \int_{x_0}^{y_0} (b_0 + \bar{b}_0)(u)\,\mathrm{d}u < 0. \tag{A5}$$

This would give an immediate contradiction since, if we assume (A5), then (A2) would imply

$$c \equiv (b_0 - \bar{b}_0)(x_0) = (b_0 - \bar{b}_0)(y_0) \lim_{\hbar \to 0} [\exp(d/\hbar)] = 0.$$

Remark. If x_0 (the same remark applies to y_0) coincides with one of the jump points of b_0, the limit (A3) in general does not exist. But since the number of jump points is finite (less than N), we can always find $\tilde{x}_0$ such that it does not coincide with any jump point and it is sufficiently close to x_0 in order that (A4) holds for $\tilde{x}_0$ too.

Since $P_J < \min\{P_I, M_I\}$, we have $P_J < P_I$ so that $i_2 + t < j_2$ (see figure 7). Moreover we know that b_0 is negative on some interval on the left of j_2, by definition of $p_\alpha(J)$, so it is possible to find $z \in (i_2 + t, j_2)$ such that

$$b_0(u) = -\sqrt{2V(u)} \leqslant 0 \qquad \forall u \in [z, j_2).$$

In particular we choose z such that it does not coincide with any of the absolute minima of V, so $b_0(z)$ is strictly negative. If we set $x_0 = z - t$ then clearly $x_0 \in (i_2, i_3)$ so we have $b_0(x_0) > 0$, and

$$c \equiv b_0(x_0) - \bar{b}_0(x_0) = b_0(x_0) - b_0(z) = +2\sqrt{2V(x_0)} > 0. \tag{A6}$$

Since

$$(b_\hbar - \bar{b}_\hbar)(y) = (b_\hbar - \bar{b}_\hbar)(x_0) \exp\left(-\frac{1}{\hbar}\int_{x_0}^{y} (b_\hbar + \bar{b}_\hbar)(u)\,\mathrm{d}u\right)$$

(A6) implies that $(b_\hbar - \bar{b}_\hbar)(y)$ must be positive when $\hbar \to 0$, and so

$$b_0(y) - \bar{b}_0(y) \geqslant 0 \qquad \forall y \in [i_1, i_6].$$

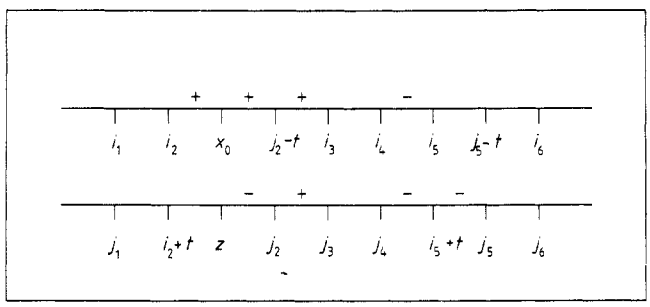

Figure 7. Sign of b_0.

This means that $b_0(y+t)$ is negative whenever $b_0(y)$ is negative and, in consequence,

$$M_I \leqslant M_J$$

or equivalently $j_5 \geqslant i_5 + t$.

It is now straightforward to show that we get $d<0$ if we choose $y_0 \equiv i_5$. In fact, divide the interval $[x_0, y_0]$ as

$$[x_0, y_0] = [x_0, j_2 - t] \cup [j_2 - t, i_3] \cup [i_3, i_4] \cup [i_4, i_5]$$

and express the integral d over $[x_0, y_0]$ as a sum

$$d = d_1 + d_2 + d_3 + d_4$$

over the corresponding subintervals. Thus we obtain (see figure 7, where we have reported the sign of b_0 in the relevant intervals)

$$d_1 = \int_{x_0}^{j_2-t} (b_0 + \bar{b}_0)(u)\,\mathrm{d}u = \int_{x_0}^{j_2-t} \sqrt{2V(u)}\,\mathrm{d}u - \int_z^{j_2} \sqrt{2V(u)}\,\mathrm{d}u = 0$$

$$d_2 = \int_{j_2-t}^{i_3} (b_0 + \bar{b}_0)(u)\,\mathrm{d}u = \rho(j_2 - t, i_3) + \rho(j_2, j_3) = 2P_J$$

$$d_4 = \int_{i_4}^{i_5} (b_0 + \bar{b}_0)(u)\,\mathrm{d}u = -\rho(i_4, i_5) - \rho(j_4, i_5 + t) = -2M_I$$

$$d_3 = \int_{i_3}^{i_4} (b_0 + \bar{b}_0)(u)\,\mathrm{d}u = \int_I b_0(u)\,\mathrm{d}u + \int_J b_0(u)\,\mathrm{d}u = 0.$$

Finally, using (A1), we have

$$d = d_1 + d_2 + d_3 + d_4 = 2(P_J - M_I) < 2(P_J - \min\{P_I, M_I\}) = -2a < 0.$$

So the proof is complete. If we had chosen $a<0$ at the beginning, then the proof would have been similar, with I and J exchanged.

Lemma 7.2. Let I and J be two closed intervals such that

$$I \text{ is } \mathrm{p}_1\text{-preferred to } J \qquad \text{and} \qquad \int_I b_0(x)\,\mathrm{d}x = \int_J b_0(x)\,\mathrm{d}x = 0.$$

Then J is connective or, equivalently, $Z_\alpha(J) = 0$, $\forall \alpha > 0$.

Proof. (As in the proof of lemma 7.1, we suppose that I and J are equivalent by translation.) Let $I = [i_1, i_2]$ and $J = [j_1, j_2]$ and $t \equiv j_1 - i_1 = j_2 - i_2$. By hypothesis we have

$$I \approx J \qquad \text{and} \qquad \exists \varepsilon_0 > 0 \text{ such that } \forall \varepsilon < \varepsilon_0,\ I^{(\varepsilon)} < J^{(\varepsilon)}.$$

Assume now the thesis to be false. This means that J is a $(+-)$ interval, so it is possible to find $\eta > 0$ such that

$$b_0(x) = +\sqrt{2V(x)} \qquad \forall x \in (j_1 - \eta, j_1) \tag{A7}$$

$$b_0(x) = -\sqrt{2V(x)} \qquad \forall x \in (j_2, j_2 + \eta). \tag{A8}$$

We can suppose $\eta < \varepsilon_0$, so that

$$I^{(\delta)} < J^{(\delta)} \qquad \forall \delta \leqslant \eta. \tag{A9}$$

This condition, together with $I \approx J$, implies that (at least) one of the two following propositions is true:

(c_1) $\exists$ a sequence $\{y_n\}$ with $y_n \in (i_1 - \eta, i_1)$ such that $y_n \to i_1$ and $V(y_n) < V(y_n + t)$ for each n;

(c_2) $\exists$ a sequence $\{y_n\}$ with $y_i \in (i_2, i_2 + \eta)$ such that $y_n \to i_2$ and $V(y_n) < V(y_n + t)$ for each n.

Assume that (c_2) is satisfied (the other case is analogous). In this case (c_2) and (A8) yield, for each n,

$$(b_0 - \bar{b}_0)(y_n) \equiv b_0(y_n) - b_0(y_n + t) \geqslant \sqrt{2V(y_n + t)} - \sqrt{2V(y_n)} > 0. \tag{A10}$$

Now let x be any point inside $(i_1 - \eta, i_1)$. From equation (2.4), remembering that $V(w) \leqslant V(w + t)$, $\forall w \in I^{(\eta)}$, we obtain

$$(b_\hbar - \bar{b}_\hbar)(x) \geqslant (b_\hbar - \bar{b}_\hbar)(y_n) \exp\left(\frac{1}{\hbar} \int_x^{y_n} (b_\hbar + \bar{b}_\hbar)(u)\, \mathrm{d}u\right) \tag{A11}$$

which, together with (A10), tells us that $(b_\hbar - \bar{b}_\hbar)(x)$ is positive when $\hbar \to 0$, so that

$$(b_0 - \bar{b}_0)(x) = b_0(x) - b_0(x + t) \geqslant 0 \qquad \forall x \in (i_1 - \eta, i_1). \tag{A12}$$

In turn, by (A12) and (A7),

$$b_0(x) = +\sqrt{2V(x)} \qquad \forall x \in (i_1 - \eta, i_1). \tag{A13}$$

Consider now some fixed point, x_0, inside $(i_1 - \eta, i_1)$. We know that the LHS of (A11) is bounded (proposition (i) of § 2), so by (A10), we must require that

$$d^{(n)} \equiv \int_{x_0}^{y_n} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u \leqslant 0 \qquad \forall n \tag{A14}$$

in order to avoid an explosion of the RHS of (A11) when $\hbar \to 0$. We are going to show that the condition (A14) is necessarily violated.

We decompose the integral in (A14) as

$$\begin{aligned} d^{(n)} &= \int_{x_0}^{y_n} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u \\ &= \int_{x_0}^{i_1} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u + \int_{i_1}^{i_2} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u + \int_{i_2}^{y_n} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u \\ &= d_1 + d_2 + d_3^{(n)}. \end{aligned}$$

Now we note that d_2 vanishes owing to the second hypothesis of the lemma, while $d_3^{(n)}$ goes to zero since $y_n \to i_2$. As regards d_1, we use (A7) and (A13) and obtain

$$d_1 = \int_{x_0}^{i_1} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u = \int_{x_0}^{i_1} (\sqrt{2V(u)} + \sqrt{2V(u + t)})\, \mathrm{d}u$$

which is a finite positive number. So

$$\int_{x_0}^{y_n} (b_0 + \bar{b}_0)(u)\, \mathrm{d}u > 0$$

for n sufficiently large, in contradiction with (A14).

References

[1] Harrel E M 1978 *Commun. Math. Phys.* **60** 73
[2] Harrel E M 1980 *Commun. Math. Phys.* **75** 239
[3] Jona-Lasinio G, Martinelli F and Scoppola E 1981 *Commun. Math. Phys.* **80** 223
[4] Frölich J and Spencer T 1983 *Commun. Math. Phys.* **88** 151
[5] Jona-Lasinio G, Martinelli F and Scoppola E 1985 *Ann. Inst. H Poincaré* **42** 73; 1984 *J. Phys. A: Math. Gen.* **17** L635
[6] Fröhlich J, Martinelli F, Scoppola E and Spencer T 1985 *Commun. Math. Phys.* **101** 21
[7] Ventzel A D 1973 *Theor. Prob. Appl.* **18** 1
[8] Ventzel A D and Freidlin M I 1984 *Random Perturbations of Dynamical Systems* (Berlin: Springer)
[9] Jona-Lasinio G, Martinelli F and Scoppola E 1986 *Mathematics and Physics, Lectures on Recent Results* vol II ed L Streit (Singapore: World Scientific) pp 227-60
[10] Cesi F 1988 An algorithm to study tunnelling in a wide class of one-dimensional multiwell potentials. II *Preprint* 622, Università di Roma (submitted to *Ann. Phys., N.Y.*)
[11] Helffer B and Sjöstrand J 1984 *Commun. PDE* **9** 337
[12] Helffer B and Sjöstrand J 1985 *Ann. Inst. H Poincaré* **42** 127
[13] Graffi S, Grecchi V, and Jona-Lasinio G 1984 *J. Phys. A: Math. Gen.* **17** 2935
[14] Simon B 1985 *J. Funct. Anal.* **63** 123