The matrizant and Saxon - Hutner theorem

1997 J. Phys. A: Math. Gen. 30 1689

(http://iopscience.iop.org/0305-4470/30/5/029)

J. Phys. A: Math. Gen. **30** (1997) 1689–1694. Printed in the UK PII: S0305-4470(97)74981-3

The matrizant and Saxon–Hutner theorem

Ivaïlo M Mladenov†

Institute of Biophysics, Bulgarian Academy of Sciences, Acad. G Bonchev Str., Bl. 21, Sofia-1113, Bulgaria

Received 29 May 1996, in final form 9 October 1996

Abstract. Relying on the one-to-one correspondence between real localized potentials and transfer matrices well known from inverse scattering theory, the Saxon–Hutner conjecture is reformulated initially as a group-theoretical and subsequently as a Lie-algebraic problem. A very basic Lie theory, in conjunction with time-reversal symmetry of the time-independent Schrödinger equation, leads to a fairly general condition which ensures the validity of the Saxon–Hutner theorem.

No matter which approach is followed, the main difference between classical and quantum mechanics is the fundamental role assigned to non-commuting operators in quantum physics. The algebra of non-commuting operators has been the subject of intensive studies over the years by both mathematicians and physicists in various contexts and great many profound results have been derived which are of potential value for quantum considerations. A problem which is often met in concrete situations is that of expressing the product of two matrizants (exponential operators) in a suitable equivalent form. Our purpose in the present paper is to extract useful information about the electronic spectra of binary alloys from the spectra of the alloying constituents by making use of the properties of matrizants. The basic conjecture with which we will deal here was made originally by Saxon and Hutner [1] and concerns the coupling of impurities introduced into an infinite one-dimensional lattice. It states that:

> 'Forbidden energies that are common to the pure A crystal and pure B crystal (with the same lattice constant) will always be forbidden energies in any arrangement of A and B atoms in a substitutional solid solution.'

In that paper the authors verified their conjecture only for the special case of a B atom which appears periodically in an infinite lattice of A. Luttinger [2] was the first to present the proof of such a theorem for the special case of two δ-potentials situated symmetrically in the lattice of a fixed cell constant, (no spatial, only strength disorder). Guided by the well known fact that the δ-function lattice lacks some important qualitative characteristics found in real crystals, while lattices built up of square-well potentials do not possess these characteristics, Landauer and Helland [3] checked the validity of the theorem in the latter case. Their findings forced them to conclude that the original theorem cannot be extended beyond δ-function potentials, and Kerner [4] had tried to prove this rigorously. However, he was unable to show more than that the Saxon–Hutner theorem is not satisfied automatically,

† Research partially supported by Bulgarian NSF-Project F-644.

and suggested a criterion based on polarity considerations which seems to restore the theorem in a modified form. Numerical work shows that this polarity criterion is of restricted generality, as Sah and Srivastava [5] have exhibited explicit examples of its failure. The more general question as to which properties of a mixed lattice can be inferred from a knowledge of the individual band structures of the pure lattices has received more definite answers after the appearance of papers by Matsuda [6], Hori [7], Dworin [8], Tong and Tong [9] and Mladenov [10]. The basic research interest there was the conditions under which the Saxon–Hutner theorem is valid. Several sufficient conditions have been found which guarantee its accuracy. In the meantime, Subramanian and Bhagwat [11] transferred Luttinger's result into the relativistic situation and later on a general criterion which covers both relativistic and non-relativistic domains was derived in [12]. A number of such criteria and their interrelations have been analysed by Hermann [13]. The work of Erdös and Herndon [14] provides a comprehensive review of methods, both analytical and numerical, for the non-relativistic case of the Saxon–Hutner theorem, as well as an extensive list of references.

The purpose of the present paper is to derive a novel condition which ensures the validity of the Saxon–Hutner theorem. The transfer-matrix formalism will be used, as this allows encoding of this genuine physical problem into mathematical form—initially as a group-theoretical and subsequently as a Lie-algebraic one. The necessary mathematical apparatus can be found in any book that deals with applications of group-theoretical methods in physics (see, e.g., Fässler and Stiefel [15]).

By definition, the transfer matrix relates the wavefuunction on the left- and right-hand sides of the potential barrier. The crucial point in this formalism is the observation that the real localized potentials and transfer matrices are in one-to-one correspondence. Furthermore, it turns out that they are elements of one of the isomorphic Lie groups $SU(1,1)$, $SL(2,R)$ or $Sp(2,R)$. Which group they form depends on the choice of the basis in the space of solutions of the time-independent one-dimensional Schrödinger equation. The relationships between these three-dimensional groups, the group $SO(2,1)$ which they cover twice and the Schrödinger equation on the line have been clarified by Peres [16]. He proved that $SL(2,R)$ is the most appropriate of the above mentioned groups (cf also Kerner [17]). This statement will be made more precise at the end of the paper.

Now we turn to the mathematical formulation of the Saxon–Hutner theorem. The real transfer matrices in question are of the form:

$$M(E) = \begin{pmatrix} a(E) & b(E) \\ c(E) & d(E) \end{pmatrix} \qquad \det(M(E)) = a(E)\,d(E) - b(E)\,c(E) = 1. \tag{1}$$

As examples, we write down the transfer matrices corresponding to the δ-potential potential scatterer $V(x) \equiv \eta\,\delta(x)$ (below a denotes the lattice constant):

$$M = \begin{pmatrix} \cos(ka) & \sin(ka)/k \\ -k\sin(ka) + \eta\cos(ka) & \cos(ka) + \eta\sin(ka)/k \end{pmatrix} \qquad k^2 = E \tag{2}$$

and the potential step $V \equiv$ constant

$$M = \begin{pmatrix} \cosh(\gamma a) & \sinh(\gamma a)/\gamma \\ \gamma\sinh(\gamma a) & \cosh(\gamma a) \end{pmatrix} \qquad \gamma = (V - E)^{1/2}. \tag{3}$$

Let us now consider an arbitrary binary linear lattice composed of two types of atoms A and B, each having $r_i, s_i \in \mathbb{Z}^+$ in the ith period,

$$A^{r_1} B^{s_1} \dots A^{r_k} B^{s_k}. \tag{4}$$

The group nature of the individual transfer matrices M_A and M_B representing A and B atoms makes possible to define the total transfer matrix $M_{A^{r_1}B^{s_1}\ldots A^{r_k}B^{s_k}}$ of an arbitrary linear chain (4) as the product

$$M_B^{s_k} M_A^{r_k} \cdots M_B^{s_1} M_A^{r_1}. \tag{5}$$

If $M(E)$ is the transfer matrix for the unit cell in the periodic lattice the forbidden energies for electrons propagating there are given by the condition

$$|\operatorname{tr}(M(E))| > 2. \tag{6}$$

Using the transfer-matrix approach, the Saxon–Hutner theorem may be formulated as follows:

Is it true that for any arrangements $A^{r_1}B^{s_1}\ldots A^{r_k}B^{s_k}$ of A and B atoms we have

$$|\operatorname{tr}(M_B^{s_k} M_A^{r_k} \cdots M_B^{s_1} M_A^{r_1})| > 2 \tag{7}$$

provided that $|\operatorname{tr}(M_A)|,\ |\operatorname{tr}(M_B)| > 2$?

The mathematical difficulties in proving such a statement stem from the fact that each transfer matrix is described generally by four real parameters (a, b, c, d) under a constraint $(ad - bc = 1)$ instead of three independent ones. This parametrization redundancy is the main obstacle to proving the Saxon–Hutner theorem and similar results. Here we will use the so-called exponential form (matrizant) of the group $SL(2, R)$ introduced above. Its Lie algebra $\mathfrak{sl}(2, R)$ comprises 2×2 matrices with zero trace. Despite of the well known fact that the exponential map is not surjective for this group, any of its elements can be written in one of the following two forms:

$$g = \pm \operatorname{Exp}(\tilde{P}) \qquad \tilde{P} = \begin{pmatrix} w & u \\ v & -w \end{pmatrix} \in \mathfrak{sl}(2, R). \tag{8}$$

From now on we will not make any distinction between the atoms and the transfer matrices representing them, i.e. A, B will be used to denote M_A and M_B. Also, as we are interested in the form which the transfer matrix (say A for definiteness), takes for forbidden energies we will consider just the latter case. This further specifies its form:

$$A = \pm \operatorname{Exp}(\lambda_A \tilde{A}) \qquad \tilde{A} \in \mathfrak{sl}(2, R) \qquad \tilde{A}^2 = I \qquad \det \tilde{A} = -1 \qquad \lambda_A \in \mathbb{R}^+ \tag{9}$$

where I denotes 2×2 identity matrix. An alternative to this algebraic description is to consider $\tilde{A}$ as a point on the single-sheeted hyperboloid H^2 in R^3.

Taking into account all the above, we may write

$$\begin{aligned} A &= \pm \Big(I \operatorname{ch}(\lambda_A) + \tilde{A} \operatorname{sh}(\lambda_A) \Big) \\ \lambda_A &= \operatorname{arch} |\operatorname{tr}(A/2)| \\ \tilde{A} &= \pm (A \mp I \operatorname{ch}(\lambda_A)) \,/\, \operatorname{sh}(\lambda_A). \end{aligned} \tag{10}$$

The nice feature of this representation is that up to a sign we have

$$A^n \sim (\operatorname{Exp}(\lambda_A \tilde{A}))^n = \operatorname{Exp}(n\lambda_A \tilde{A}) = I \operatorname{ch}(n\lambda_A) + \tilde{A} \operatorname{sh}(n\lambda_A)\,. \tag{11}$$

This sign is actually of no importance here, as we have to compute

$$\begin{aligned} |\operatorname{tr}(M)| &= \left| \operatorname{tr} \left[\operatorname{Exp}(s_k\lambda_B \tilde{B})\ \operatorname{Exp}(r_k\lambda_A \tilde{A}) \cdots \operatorname{Exp}(s_1\lambda_B \tilde{B})\ \operatorname{Exp}(r_1\lambda_A \tilde{A}) \right] \right| \\ &= \left| \operatorname{tr} \left\{ [I \operatorname{ch}(s_k\lambda_B) + \tilde{B} \operatorname{sh}(s_k\lambda_B)] \cdots [I \operatorname{ch}(r_1\lambda_A) + \tilde{A} \operatorname{sh}(r_1\lambda_A)] \right\} \right|\,. \end{aligned} \tag{12}$$

Let us think of the last equation as being multiplied out. Because of $\tilde{A}^2 = \tilde{B}^2 = I$ and the linearity of the trace map, all terms containing an odd number of matrices will produce zero. Thus we get a finite series of the form

$$|\operatorname{tr}(M)| = \left| \operatorname{tr}\left\{ \sum_{\nu=0}^{k} \varphi_\nu\, (\tilde{A}\tilde{B})^\nu \right\} \right| \tag{13}$$

in which all coefficients φ_ν are positive, as they are sums and products of hyperbolic functions of positive arguments. So, in order to estimate (13) we have to investigate φ_ν and $\operatorname{tr}(\tilde{A}\tilde{B})^\nu$ in greater detail. Now one can easily see that $\tilde{A}$ and $\tilde{B}$ are isospectral, which means that they are similar, i.e. there exists a non-degenerate matrix U such that

$$\tilde{B} = U\,\tilde{A}\,U^{-1} \qquad \text{or equivalently} \qquad U\,\tilde{A} = \tilde{B}\,U. \tag{14}$$

Besides, one can choose U to additionally satisfy

$$\operatorname{tr}(U) = 0 \quad \det(U) = -1 \quad U^2 = I \qquad \text{or, identically,} \qquad U^{-1} = U. \tag{15}$$

Taking all this into consideration, U can be written in the form

$$U = \begin{pmatrix} x & y \\ z & -x \end{pmatrix} \qquad \det(U) = -x^2 - yz = -1. \tag{16}$$

Let $(u_A,\ v_A,\ w_A)$ and $(u_B,\ v_B,\ w_B)$ be the coordinates of $\tilde{A}$ and $\tilde{B}$ as specified in (8). Equalization of the matrix elements in (14) leads to a system of homogeneous linear equations for the unknown $(x,\ y,\ z)$ entries of U:

$$\begin{aligned} x\,w_A + y\,v_A &= x\,w_B + z\,u_B \\ x\,u_A - y\,w_A &= y\,w_B - x\,u_B \\ z\,w_A - x\,v_A &= x\,v_B - z\,w_B \\ z\,u_A + x\,w_A &= y\,v_B + x\,w_B. \end{aligned} \tag{17}$$

One can take any solution of this system provided that the matrix U is non-degenerate.

For example, working in the fixed local chart $\Phi^z_{14} = \{w_A \neq w_B,\ v_A \neq -v_B\}$ from the atlas $\mathcal{A} = \left\{\bigcup \Phi^\tau_{\alpha\beta};\ \alpha, \beta = 1, 2, 3, 4,\ \alpha < \beta,\ \tau = x, y, z\right\}$,

$$\begin{array}{lll}
\Phi^x_{12} = \{u_B \neq 0, w_A \neq -w_B\} & \Phi^y_{12} = \{u_B \neq 0, u_A \neq -u_B\} & \Phi^z_{12} = \{u_B \neq 0, v_A \neq -v_B\} \\
\Phi^x_{13} = \{v_B \neq 0, w_A \neq -w_B\} & \Phi^y_{13} = \{v_A \neq 0, u_A \neq -u_B\} & \Phi^z_{13} = \{v_A \neq 0, v_A \neq -v_B\} \\
\Phi^x_{14} = \{u_A v_A \neq u_B v_B\} & \Phi^y_{14} = \{w_A \neq w_B, u_A \neq -u_B\} & \Phi^z_{14} = \{w_A \neq w_B, v_A \neq -v_B\} \\
\Phi^x_{23} = \{w_A \neq -w_B\} & \Phi^y_{23} = \{w_A \neq -w_B, u_A \neq -u_B\} & \Phi^z_{23} = \{w_A \neq -w_B, v_A \neq -v_B\} \\
\Phi^x_{24} = \{u_A \neq 0, w_A \neq -w_B\} & \Phi^y_{24} = \{u_A \neq 0, u_A \neq -u_B\} & \Phi^z_{24} = \{u_A \neq 0, v_A \neq -v_B\} \\
\Phi^x_{34} = \{v_B \neq 0, w_A \neq -w_B\} & \Phi^y_{34} = \{v_B \neq 0, u_A \neq -u_B\} & \Phi^z_{34} = \{v_B \neq 0, v_A \neq -v_B\}
\end{array} \tag{18}$$

which covers the product manifold $H^2_{\tilde{A}} \times H^2_{\tilde{B}}$, one can easily derive from the first and fourth equations in (17) that

$$x = \frac{w_A + w_B}{v_A + v_B}\, z \qquad \text{and} \qquad y = \frac{u_A + u_B}{v_A + v_B}\, z. \tag{19}$$

Now, the determinant condition on U implies that

$$\left\{ \left[\frac{w_A + w_B}{v_A + v_B} \right]^2 + \frac{u_A + u_B}{v_A + v_B} \right\} z^2 = 1 \tag{20}$$

and it can be solved for z if and only if

$$u_A\, v_B + u_B\, v_A + 2\, w_A\, w_B + 2 > 0. \tag{21}$$

Similarly, the same term comes out in any other local chart, which actually means that equation (21) is a global condition. As our goal is to estimate $\sum_{\nu=0}^{k} \varphi_\nu\ \mathrm{tr}(\tilde{A}\tilde{B})^\nu$ efficiently, let us recall other important properties of the trace map besides the linearity which has been used to derive (13). First of all these are invariances under permutation and conjugations with non-degenerate matrices, i.e.

$$\mathrm{tr}(X\,Y) = \mathrm{tr}(Y\,X) \quad \mathrm{tr}(g\,X\,g^{-1}) = \mathrm{tr}(X) \quad \forall\, g, X, Y \in Mat(2, R) \qquad \det(g) \neq 0. \tag{22}$$

Furthermore, if S and T are unimodular matrices we can apply the Hamilton–Cayley theorem to find

$$\mathrm{tr}(S\,T) = \mathrm{tr}(S)\ \mathrm{tr}(T) - \mathrm{tr}(T\,S^{-1}). \tag{23}$$

So, let us take the simplest case of the $A^{r_1} B^{s_1}$ section for which we have

$$\mathrm{tr}(A^{r_1} B^{s_1}) = \varphi_0\ \mathrm{tr}(I) + \varphi_1\ \mathrm{tr}(\tilde{A}\tilde{B}) \tag{24}$$

where $\varphi_0 = \mathrm{ch}(r_1\lambda_A)\ \mathrm{ch}(s_1\lambda_B)$ and $\varphi_1 = \mathrm{sh}(r_1\lambda_A)\ \mathrm{sh}(s_1\lambda_B)$. Since by hypothesis the arguments of the hyperbolic functions are positive, the hyperbolic cosines are strongly greater than unity, i.e. $\varphi_0 > 1$. We will also have $|\,\mathrm{tr}(M)| > 2$ provided that the second term is positive. On the same footing the hyperbolic sines are strongly positive and therefore we end with the condition

$$\mathrm{tr}(\tilde{A}\tilde{B}) = u_A\, v_B + u_B\, v_A + 2w_A\, w_B \geqslant 0 \tag{25}$$

which is suficient for the validity of the Saxon–Hutner theorem in this case. More complex cases of two, three and an arbitrary number of segments $A^{r_1} B^{s_1} \ldots A^{r_i} B^{s_i}$, $i = 2, 3, \ldots, k$ are dealt with by combining (23) and (24), which by an induction process produce

$$\mathrm{tr}(M) = \sum_{\nu=0}^{k} \psi_\nu \left[\mathrm{tr}(\tilde{A}\tilde{B})\right]^\nu. \tag{26}$$

However, it is trivial to observe that ψ_0 is of the form

$$\mathrm{ch}(m_1\lambda_A)\,\mathrm{ch}(n_1\lambda_B)\cdots\mathrm{ch}(m_k\lambda_A)\,\mathrm{ch}(n_k\lambda_B) \qquad m_i,\ n_i \in \mathbb{Z}^+ \quad i = 1,\ 2, \ldots, k \tag{27}$$

plus other non-negative terms which appear as cross multiplication terms in the expansion of $A^{r_i} B^{s_i}$, $i = 1, 2, \ldots, k$ and consecutive simplifications. Relying as before on the fundamental properties of the hyperbolic functions, we can conclude that $\psi_0 > 1$ holds generally and $\psi_\nu > 0$ is fulfilled for all $\nu \in \{1, 2, 3, \ldots, k\}$. Hence equation (25) is just the condition we seek. It might be appropriate to point out that this condition is consistent with (21), because the Saxon–Hutner theorem can also be viewed as a part of the long-standing problem of classifying (orbits of) pairs of matrices under simultaneous similarity transformations:

$$(H, K) \quad \rightarrow \quad (H', K') = (g\,H\,g^{-1},\ g\,K\,g^{-1}) \qquad H, K \in \mathfrak{sl}(2, R). \tag{28}$$

Taken together, equations (21) and (25) separate the class (stratum) of orbits for which the Saxon–Hutner theorem is satisfied (for more details about orbits see [18]).

As an illustration, let us apply our sign test in the most classical situation, that of symmetric δ-potentials of different amplitudes, as treated by Luttinger [2]. Curiously enough, he was also able to show that the Saxon–Hutner conjecture has an analogy in the case of transmission of a wave down a line loaded with two terminals. In both cases the parameters of the transfer matrices are of the form

$$u_C = \sqrt{\left|\frac{C_{11}-1}{C_{11}+1}\right|}\,\mu_C, \quad v_C = \sqrt{\left|\frac{C_{11}+1}{C_{11}-1}\right|}\,\mu_C^{-1}, \quad w_C = 0 \qquad C \equiv A \text{ or } B \tag{29}$$

so that

$$u_A\,v_B + u_B\,v_A + 2\,w_A\,w_B = \sqrt{\left|\frac{(A_{11}-1)\,(B_{11}+1)}{(A_{11}+1)\,(B_{11}-1)}\right|}\,\mu_A\,\mu_B^{-1} + \sqrt{\left|\frac{(A_{11}+1)\,(B_{11}-1)}{(A_{11}-1)\,(B_{11}+1)}\right|}\,\mu_A^{-1}\,\mu_B$$

and we have finished, because $\mu_A \equiv \mu_B$ (this follows either by the properties of even and odd solutions of the Schrödinger equation or the nature of the transmission line), and makes it possible for the last expression to be rewritten (with the clear orbit interpretation) as

$$\omega + \frac{1}{\omega} \geqslant 2 \qquad \omega \in \mathbb{R}^+. \tag{30}$$

It also seems worthwhile to mention that these results are applicable to any one-dimensional or quasi-one-dimensional systems, such as isotopically or non-isotopically polyatomic disordered chains or transmission lines with varying physical parameters composed of inequivalent electrical, mechanical or optical filters (cf, e.g., [6, 7, 19, 20]).

Our final comment is to mention that the very basic Lie group theory has once again been found to be useful in a problem with no apparent symmetry. Another discrete, and in some sense 'hidden', symmetry which stems from the time-reversal invariance of the fundamental time-independent Schrödinger equation left implicit in our considerations manifests itself by the possibility of forgetting about the dichotomy that appears in (8), and indicates that the relevant group in this case is $SO(2, 1)$.

References

[1] Saxon D and Hutner R 1949 *Philips Res. Rep.* **4** 81
[2] Luttinger J 1951 *Philips Res. Rep.* **6** 303
[3] Landauer R and Helland J 1954 *J. Chem. Phys.* **22** 1655
[4] Kerner E 1954 *Phys. Rev.* **95** 687
[5] Sah P and Srivastava K 1969 *Physica* **43** 528
[6] Matsuda H 1962 *Progr. Theor. Phys.* **27** 81
[7] Hori J 1964 *Progr. Theor. Phys.* **32** 371
[8] Dworin L 1965 *Phys. Rev.* A **138** 1121
[9] Tong B and Tong S 1969 *Phys. Rev.* **180** 739
[10] Mladenov I 1985 *CR Acad. Sci. Sofia* **38** 993
[11] Subramanian R and Bhagwat K 1971 *Phys. Stat. Solidi* b **48** 399
[12] Mladenov I 1989 *Phys. Lett.* **137A** 313
[13] Hermann M 1989 *Lett. Math. Phys.* **18** 19
[14] Erdös P and Herndon R 1982 *Adv. Phys.* **31** 65
[15] Fässler A and Stiefel E 1992 *Group Theoretical Methods and Their Applications* (Basel: Birkhäuser)
[16] Peres A 1983 *J. Math. Phys.* **24** 1110
[17] Kerner E 1956 *Proc. Phys. Soc.* **69** 234
[18] Friedland S 1983 *Adv. Math.* **50** 189
[19] Brillouin L 1953 *Wave Propagation in Periodic Structures* (New York: Dover)
[20] Mladenov I 1987 *CR Acad. Sci. Sofia* **40** 35