

SOME REMARKS ON REFLECTION POSITIVITY*)

A. UHLMANN

Karl-Marx University, Leipzig, GDR

The requirement of reflection positivity is investigated and its general applicability to different physical theories is pointed out. Its role is illustrated on an example from electrostatics and on several simple examples of field theories. Then, after presenting an abstract construction of the concept, the role of reflection positivity in classical lattice systems is discussed.

Reflection positivity has appeared in Euclidean quantum field theory and in lattice theory. It has been used in constructing the Hamiltonian in quantum field theory, the transfer operators in spin lattice systems, and in formalizing part of the Peierls argument in existence proofs for phase transitions.

What will be said below to this topic is not only incomplete but also highly subjective. Therefore I have to point out some very important aspects which are not considered here, and which have been handled much better than I can do in textbooks [1] and review articles [2] already: As a constructive tool RP is one of the Osterwalder and Schrader axioms [3]. Here RP reflects the positivity condition of the Wightman axioms as analytically continued to the Schwinger points. How is it possible to continue a positivity condition? For the complicated case of QFT this is well described in [1, 3], in GLASER [4], and other papers. Here, to see the flavour of the argument, let us notice one version of a theorem due to Fitz-Gerald: Assume $f(z, w)$ to be analytic in $|z| < 1$, $|w| < 1$, and choose $0 < \varepsilon < 1$. If then for every natural m and every choice of real $s_1, \dots, s_m$, with $0 \leq s_k \leq \varepsilon$ the matrix $a_{ij} = f(s_i, s_j)$ turns out to be positive definite it follows the positive definiteness of the matrix $b_{jk} = f(z_j, \bar{z}_k)$, no matter how the complex numbers $z_1, z_2, \dots$ have been chosen out of the interior of the unit circle [5]. Considering f along the imaginary axis and redefining $g(s_j, s_k) = f(is_j, is_k)$ one arrives at the positive definiteness of the matrix $g(s_j, -s_k)$ for real numbers s_j bounded by one. Though this example is widely oversimplified it shows how to obtain RP by analytic continuation.

Quite another way RP is entering in Nelson's approach, [7, 6], to Euclidean QFT. The centre of this theory is the famous generalizing the concept of Markov random processes to that of Markov random fields [1, 2]. RP shows up as a consequence of the Markov property. There are some indications, [8, 9], that with the aid of reflection positivity a "good" class of stochastic processes can be defined in which the Markovian behaviour is "weakened". The derivation of RP from Markov properties was presented by MACK [10] last year at Primorsko. His lecture further includes an application to lattice gauge theory found also in [11].

*) Invited talk at the "Symposium on Mathematical Methods in the Theory of Elementary Particles", Liblice castle, Czechoslovakia, June 18–23, 1978.

In preparing this talk I was influenced by the work of DYSON, LIEB, and SIMON on the Heisenberg ferromagnet [12] and by the Rome talk of LIEB [13] where also more references can be found.

An example from electrostatics.

Let us start with an example, due to LIEB [13], that nicely explains some important concepts. We consider the set L of all electric charge distributions $q = q(x, y, z)$ with finite electrostatic self energy. L is a real linear space. In L we introduce a positive definite scalar product $\langle, \rangle$ which, up to an unimportant for us factor, gives the interaction energy of two charge distributions

$$\langle q', q \rangle = \int d^3r d^3r' q'(r') q(r) |r - r'|^{-1}.$$

Thus L becomes pre-hilbertian. Now define the reflection operator

$$(\theta q)(x, y, z) = q(-x, y, z).$$

The square of θ is the identity operator. θ is an isometric map of L onto L . This we see either by calculation or by recalling that q and θq should have the same electrostatic energy. A one-parameter group $s \rightarrow T_s$ of translations

$$T_s q = q_s \quad \text{with} \quad q_s(x, y, z) = q(x - s, y, z)$$

is naturally defined in L . The reflection plane $x = 0$ divides the Euclidean 3-space into two parts, and we shall now consider such charge distributions which are in one of these half-spaces. They form a linear subspace L_+ of L

$$L_+ = \{q \in L: q(x, y, z) = 0 \quad \text{if} \quad x \leq 0\}.$$

Let us imagine a charge distribution $q \in L_+$ and its reflected counterpart θq . It is suggestive to think that the interaction energy of q and θq is always positive. Another version of this assertion arises as follows: If the reflecting plane were ideally conducting, then q would be attracted by the plane. The potential and the electric field of this situation is generated by the joint distribution $q - \theta q$. Setting $q_s = T_s q$ with $s \geq 0$ the electric energy

$$\langle q_s - \theta q_s, q_s - \theta q_s \rangle$$

has to increase with increasing s . This is equivalent with the decreasing of $\langle q_s, \theta q_s \rangle$. Hence the last quality has to be non-negative because it decreases to zero with $s \rightarrow \infty$. The conclusion is

$$\langle q, \theta q \rangle \geq 0 \quad \text{for all} \quad q \in L_+.$$

Let us now prove this inequality rigorously.

Using continuity arguments we can restrict ourselves to sums of Dirac measures though these distributions are not in L_+ . Let us, therefore, start with the distribution

$$\varrho(\mathbf{r}) = \sum q_j \delta(\mathbf{r} - \mathbf{r}_j), \quad x_j \geq 0.$$

We thus have to prove the inequality

$$\sum q_n q_m ((x_n + x_m)^2 + (y_n - y_m)^2 + (z_n - z_m)^2)^{-1/2} \geq 0.$$

Replacing $1/r$ by its Fourier transform $1/k^2$ and introducing

$$a_n = q_n \exp i(k_y y_n + k_z z_n)$$

our inequality is converted into

$$\int \sum a_n \bar{a}_m k^{-2} (\exp i(x_n + x_m) k_x) d^3 k \geq 0.$$

Let us now carry out the integration with respect to k_x only. With $t^2 = (k_y)^2 + (k_z)^2$ and noticing $x_j \geq 0$ the integral

$$\int |k|^{-2} \exp i(x_n + x_m) k_x dk_x$$

becomes a product of two real factors b_n and b_m with

$$b_j = \pi^{1/2} t^{-1/2} \exp(t^{-1} x_j).$$

Hence our left-hand side becomes the non-negative expression

$$\int a_n \bar{a}_m b_n b_m dk_y dk_z$$

and the assertion is proved.

From electrostatics to other examples.

At first we mention the one-particle Hilbert space of Euclidean quantum field theory. One defines $S(x)$, $x \in \mathbb{R}^d$, $d = 1, 2, 3, \dots$ as the solution of

$$(-\Delta + m^2) S(x) = \delta(x)$$

which vanishes at infinity. (If $d = 1, 2$ one has to have $m \neq 0$. For $d = 2$ RP can be obtained under the additional assumption that the total charges are zero.) In a suitable function space the scalar product reads

$$\langle \psi, \psi' \rangle = \int \psi(x) \psi'(x') S(x - x') dx dx'.$$

With $x = (x^1, \dots, x^d)$ we define as above the reflection on the hyperplane $x^1 = 0$

and the translation operator T_s for a shift of amount s . L_+ denotes again the subspace of such functions of L which vanish for $x^1 \leq 0$. Then we have reflection positivity

$$\langle \psi, \theta \psi \rangle \geq 0, \quad \psi \in L_+.$$

The proof can be done literally as in the electrostatic case. The usual next step [1] is to construct the Fock space over the completion of L and to extend RP to the Fock space.

There are not only scalar theories showing up reflection positivity – though, to my knowledge, this point has not been discussed to its end yet. One of the best known examples is the massless free vector field, or, magnetism with stationary currents. Indeed, let L be the set of all divergence-free vector fields, $j(r)$, in Euclidean 3-space with (say) compact supports. (The dimension can be altered as in the scalar theory but the vanishing of the divergence is important.) The translations T_s in x -direction are defined obviously. The reflection is

$$(\theta j_x)(x, y, z) = -j_x(-x, y, z), \quad (\theta j_y)(x, y, z) = +j_y(-x, y, z)$$

and j_z behaves as j_y . The scalar product is defined by the expression for the magnetic energy

$$\langle j, j \rangle = \int j(r) j(r') |r - r'|^{-1} d^3r d^3r'.$$

With this definition it is again possible to show RP. If L_+ is the space of currents situated right off the reflection plane,

$$\langle j, \theta j \rangle \geq 0 \quad \text{for all } j \in L_+.$$

There is an interesting observation to be made: Also in the last example we can refer to a problem in magnetostatics: Here we have to choose the reflection plane to have infinite permeability. Then every system of currents, placed on one side of this plane, will be attracted by the plane. The magnetic energy decreases during a move of the plane towards the electric current system. It seems so that the same effect can be reached with other hypersurfaces, especially if they are convexly curved. Thus it looks like the same phenomena as in RP, or, as a formulation of “RP without symmetry”.

An abstract construction.

Our aim is to abstract some general structure from the foregoing examples. We have had, (a), a real or complex linear space L equipped with a hermitian form $\langle, \rangle$. In the examples, this form had turned out to be positive definite. This, however, is not the general case. Further we had, (b), an isometric reflection θ with

$$\theta^2 = \text{identity}, \quad \langle \theta \xi, \eta \rangle = \langle \xi, \theta \eta \rangle.$$

There was given, (c), a subspace L_+ of L with RP property

$$\langle \xi, \theta \xi \rangle \geq 0 \quad \text{for all } \xi \in L_+.$$

Then there appeared, (d), an isometric 1-parameter group $s \rightarrow T_s$, $T_0 = \text{identity}$, satisfying for all s

$$\theta T_s \theta = T_{-s} \quad \text{and, for } s \geq 0, \quad T_s L_+ \subseteq L_+.$$

We have to have two further conditions of technical character. (e) Continuity condition: For all $\eta \in L$

$$s \rightarrow \langle \eta, T_s \eta \rangle \quad \text{is continuous in } s.$$

(f) Growth condition: For all $\eta \in L$ and every $\varepsilon > 0$

$$\lim_{s \rightarrow \infty} (\exp - \varepsilon s) \langle \eta, T_s \eta \rangle = 0.$$

The growth condition is satisfied automatically for positive definite $\langle, \rangle$. But it is just the example of axiomatic Euclidean QFT where we generally do not have positive definiteness of $\langle, \rangle$ but only polynomial boundedness of $\langle \eta, T_s \eta \rangle$ in s and hence condition (f), see [1, 3].

Lemma: Given L, L_+, θ, T_s , satisfying the conditions (a) to (f). Then there is a Hilbert space $\mathcal{H}$ with scalar product $(,)$, a self-adjoint operator H with $H \geq 0$, and a mapping τ from L_+ onto a dense subset of $\mathcal{H}$ with

$$\begin{aligned} \langle \eta, \theta \eta' \rangle &= (\tau \eta, \tau \eta'), \quad \eta, \eta' \in L_+ \\ \tau T_s \eta &= (\exp(-sH)) \tau \eta, \quad s \geq 0. \end{aligned}$$

To prove this one first define on L_+ the scalar product $\langle \eta, \eta' \rangle_\theta = \langle \eta, \theta \eta' \rangle$ so that by RP L_+ becomes pre-hilbertian. With this new scalar product $T_s, s \geq 0$, restricted to L_+ becomes hermitian. We choose a $\langle, \rangle_\theta$ -normed $\eta \in L_+$ and examine $v(s) = \langle \eta, T_s \eta \rangle_\theta$. One gets $v(0) = 1, v(s) \geq 0$, and, by Schaarz' inequality, $v(s) v(t) \geq v(s/2 + t/2)^2$. Because $w(s) = \ln v(s)$ is continuous this means convexity of $s \rightarrow w(s)$ for $s \geq 0$. Using $w(0) = 0$ we have

$$w(sp) \leq p w(s) + (1-p) w(0) = p w(s).$$

Therefore, setting $p = s_0/s$,

$$s w(s_0) \leq s_0 w(s), \quad s \geq s_0 \geq 0.$$

Now (f) tells us

$$0 = \lim_{s \rightarrow \infty} (\exp - \varepsilon s) v(s) \geq \lim_{s \rightarrow \infty} \exp(-\varepsilon s + s/s_0 w(s_0)).$$

For all $\varepsilon > 0$ this can be true if and only if $w(s_0) \leq 0$ whatsoever $s_0 \geq 0$ was. Hence

space of all observables concentrated on X^+ . The reflection operator is given by $(\theta f)(A) = f(\vartheta A)$. We are now concerned with rather simple in form hermitian scalar products: We choose $d \in L$ with $d = \theta d$ and write

$$\langle f, g \rangle = \langle fgd \rangle_0$$

so that θ becomes an isometric reflection. We further specify by a Gibbsian ansatz $d = \exp(-h)$ with $h = \sum J_A s_A$.

We shall distinguish two cases. They may be illustrated by a 2-dimensional cubic lattice with nearest neighbour interaction. If the lattice points then are denoted by (n, m) with integers n, m running, say, from $-N$ to $+N$, we can choose ϑ to be the reflection $(n, m) \rightarrow (-n, m)$. Another possibility is to choose ϑ to be the reflection $(n, m) \rightarrow (m, n)$. For both choices we have the general situation referred to further as

Case A: The intersection $X^0 = X^+ \cap X^-$ consists of fix-points of ϑ only. $J_A \neq 0$ should imply either $A \subseteq X^+$, or $A \subseteq X^-$, or $A \subseteq X^0$.

Obviously, in case A, every bond connecting X^+ and X^- is entirely in X^0 .

Lemma: In case A we have reflection positivity.

Indeed, we may write $h = h_+ + h_- + h_0$ with $\theta h_+ = h_-$ and h_0 is the sum of all such $J_A s_A$ having $A \subseteq X^0$. Introducing the exponentials $d_+ = \exp(-h_+)$ and so on, we get

$$\langle df\theta f \rangle = \langle d_0(fd_+) \theta(fd_-) \rangle_0.$$

Assuming f to be supported in X_+ then the same is true for $g = fd_+$. In $g = \sum g^B s_B$ we collect all terms with $B \subseteq X^0$ and denote their sum by g_0 . Then we have

$$\langle df\theta f \rangle^0 = \langle d_0 g_0 \theta g_0 \rangle_0 \geq 0$$

because $\theta g_0 = g_0$ and $d_0(g_0)^2$ is non-negative. Thus the lemma is proved.

Returning to our 2-dimensional model, there is a further possibility. Let the integers in (n, m) run from $-N$ to $N + 1$. Then $(n, m) \rightarrow (-n + 1, m)$ is a reflection. With this example is connected the following

Case B: The intersection $X^+ \cap X^-$ is empty. Rewriting

$$I(A, B) = -J_C \quad \text{if} \quad C = A \cup \vartheta B \quad \text{with} \quad A, B \subseteq X^+$$

we demand the matrix $A, B \rightarrow I(A, B)$, indexed by the subsets of X^+ , to be positive semidefinite.

Lemma: In case B we have reflection positivity.

To prove this we first write $h = h_+ + h_- + h_0$ where h_+ resp. h_- is concentrated

sign of m_1 . We get a "case A" problem. The hamiltonian is of the form $h = \sum h_{xy}$ with $h_{xy} \in \mathfrak{B}_{\{x,y\}}$ and x, y nearest neighb. We call X^+ and X^- the sets of lattice sites with $m_1 \geq 0$ and $m_1 \leq 0$. Reflection symmetry of h enables us to decompose h in a sum

$$h = h_+ + h_-, \quad \theta h_+ \theta = h_-, \quad h_+ \in \mathfrak{B}_{X^+}.$$

Because the intersection of X^+ and X^- is not empty, the two matrices h_+ and h_- do not commute in general. However, if b is an operator supported in $X^+ \setminus (X^+ \cap X^-)$, b will commute with h_- . One conjecture of Dyson, Lieb, and Simon then asserts

$$\text{Tr.} (b\theta b\theta \exp h) \geq 0.$$

A little bit more general this reads [12]

DLS-conjecture: Assume h_+ to be a hermitian operator supported in X^+ and define $h = h_+ + \theta h_+ \theta$. Assume further the hermitian operator b to be supported in $X^+ \setminus (X^+ \cap X^-)$. Then $\text{Tr.} (b\theta b\theta \exp h) \geq 0$.

In finite dimension, and with the case we are concerned here, one may disentangle this a little bit further: There are hermitian operators $b_1, b_2, \dots$ supported in $X^+ \setminus (X^+ \cap X^-)$ and $c_1, c_2, \dots$ supported in $X^+ \cap X^-$ such that

$$h_+ = b_1 \otimes c_1 + b_2 \otimes c_2 + \dots$$

$$h_- = \tilde{b}_1 \otimes c_1 + \tilde{b}_2 \otimes c_2 + \dots$$

with $\tilde{b}_j = \theta b_j \theta$. One sees every "b" commuting with every " $\tilde{b}$ ", and every "b" and every " $\tilde{b}$ " commuting with every "c".

The conjecture is not only of very general nature but also highly non-trivial. Hence, if someone were clever enough to prove (or disprove?) this wonderful assertion, we surely would gain important new insight in quantum lattices.

Let me remark the triviality of the conjecture in the case of mutually commuting operators $c_1, c_2, \dots$ because then h_+ commutes with $h_- = \theta h_+ \theta$. Writing

$$2 \exp h = (\exp h_+) (\exp h_-) + (\exp h^+) (\exp h_+) + R$$

one sees that the conjecture is "good up to a remainder R of order three in the norm of h ". However, the remainder R is of awfully complex algebraic structure.

A by far less trivial, and up to now also unproved consequence is the conjecture of BESSIS, MOUSSA, and VILLANI. It asserts, [14], that the matrix $a_{ij} = j(s_i + s_j)$ is positive definite for arbitrary real $s_1, s_2, \dots$ and with j given by $j(s) = \text{Tr.} \exp . (h + sk)$ with any choice of the hermitian matrices h and k . The DSL-conjecture gives even a slightly sharper conjecture: For every finite set $d_1, d_2, \dots$ of hermitian matrices the matrix $a_{ij} = \text{Tr.} \exp (d_i + d_j)$ should be positive definite. This, indeed, is equivalent to the DSL-conjecture in case of mutually commuting operators $a_1, a_2, \dots$.

Applying Lie-Trotter decomposition of the exponential the last problem is connected with one concerning block matrices. Give natural numbers n and m . Let $d = d = (d_{ij})$ denote a nm -dimensional matrix in "block form", the blocks may be assumed of dimension m , so that every d_{ij} is itself a matrix of dimension m and $i, j = 1, 2, \dots, n$. Assume d to be positive definite. Construct the n -dimensional matrix (b_{ij}) , $b_{ij} = \text{Tr. } (d_{ij})^k$. Is (b_{ij}) positive definite? For $n = 2$ the answer is "yes". It is unknown what happens for $n > 2$.

For valuable discussions I should like to thank J. LÖFFENHOLZ.

Received 26. 6. 1978.

References

- [1] SIMON B., The P_2 Euclidean (Quantum) Field Theory, Princeton University Press, Princeton N. J. 1974.
- [2] GUERRA F., ROSEN L., SIMON B., Math. 101 (1975), 111 and 191.
- [3] OSTERWALDER K., SCHRADER R., Comm. Math. Phys. 31 (1973), 83.
- [4] GLASER V., in "Problems of Theoretical Physics", Moscow 1961.
- [5] FITZGERALD C., Proc. Am. Math. Soc. 18 (1967), 788.
- [6] NELSON E., J. Funct. Anal. 12 (1973), 97 and 211.
- [7] DOBRUSHIN R. L., MINLOS R. A., in "Functional and Probabilistic Methods in Quantum Field Theory", Acta Univ. Wratislaviensis, No 368, Wrocław 1976.
- [8] HEGERFELD G. C., Comm. Math. Phys. 35 (1974), 155.
- [9] KLEIN A., Bull. Am. Math. Soc. 82 (1976), 762.
- [10] MACK G., in "Proceedings of the IIIrd School of Elementary Particle Physics", Primorsko (Bulgaria) 1977.
- [11] OSTERWALDER K., SEILER E., Princeton Univ. preprint (1977).
- [12] DYSON J. F., LIEB E. H., SIMON B., Princeton Univ. preprint (1976).
- [13] LIEB E. H., "New proofs of long range order", to appear.
- [14] BESSIS D., MOUSSA P., VILLANI M., J. Math. Phys. 16 (1975), 2318.