Amer. Math. Soc. Transl.
(2) Vol. 189, 1999

On the Scientific Work of Mikhail Shlëmovich Birman

V. Buslaev, M. Solomyak, and D. Yafaev

This volume is dedicated to the seventieth birthday of an outstanding mathematician Mikhail Shlëmovich Birman. The sphere of his scientific interests is wide and diverse. M. Sh. always brings new insights in any field he works in. His most important achievements belong to Hilbert space theory (including scattering theory) and the spectral theory of differential operators. M. Sh. is one of the world's acknowledged leaders in these fields.

M. Sh. was born on January 17, 1928 in Leningrad (now St. Petersburg). His father was Professor of Theoretical Mechanics in one of Leningrad technical institutes, and his mother was a school teacher. During World War II, when Leningrad was besieged by Nazi troops, the family settled in Sverdlovsk (now Ekaterinburg), in the Ural Region, where M. Sh. finished high school. In 1945, after the family returned to Leningrad, he entered Leningrad Electrical Engineering Institute. His professor of Mathematics, highly impressed by the results of the first exams passed by M. Sh., advised him to switch to the Department of Mathematics and Mechanics of Leningrad University, and so he did in 1946.

Being still an undergraduate student, M. Sh. took a part time position at Steklov Mathematical Institute, in the laboratory of L. V. Kantorovich. Very soon Kantorovich distinguished the new young colleague for his strong intellect and independent thought, and he began to give M. Sh. assignments going far beyond standard technical calculations. The years of this work were very important for the mathematical development of M. Sh.

M. Sh. graduated in 1950, under the supervision of M. K. Gavurin. By that time, he proved to be one of the best students of the Department. Antisemitic trends in the Soviet university policy of those days made it impossible for him to study for Ph.D. in the normal way. Instead of being offered support for his research, he had to accept a teaching assistantship at Leningrad Mining Institute.

In the early 1950s the famous Seminar in Mathematical Physics was initiated by V. I. Smirnov, and M. Sh. became one of its most active participants. Practically all his results obtained during the whole of his scientific work, were first reported in the meetings of this seminar. Now M. Sh. (together with O. A. Ladyzhenskaya) is the leader of the seminar.

1991 *Mathematics Subject Classification.* Primary 01A70.

M. Sh. defended his Ph.D. thesis in 1954. In 1956, during "Khrushchev's thaw", on a suggestion by O. A. Ladyzhenskaya and V. I. Smirnov, he joined the Department of Mathematical Physics of the School of Physics of Leningrad University. Since then, all his activity as a scientist and teacher has been inseparable from the activity of this Department and the whole School.

In the fifties the Department of Mathematical Physics was allowed to award M.Sc. degrees. Before, its only task was teaching. M. Sh. became one of the key figures in this conversion. He started giving courses for students specializing in Mathematical and Theoretical Physics, and very soon won the reputation of an outstanding lecturer, since his lectures always elucidated the subject. Especially his graduate courses in spectral theory of differential operators and in scattering theory were not only instructive, but also inspiring.

M. Sh. is one of the founders of the St. Petersburg school in spectral theory. He initiated research in different branches of this theory. Many mathematicians, whose names became well known later, began their scientific career either under direct supervision of M. Sh., or influenced by his lectures and papers.

This short biographic introduction would be incomplete if we did not mention that M. Sh. also has always had an exceptional personal influence on his colleagues and students. This is explained not only by his professional authority, but also by the fact that he is very responsive to the difficulties and problems of other people. Many of M. Sh.'s younger colleagues benefited from his wise influence in their professional lifes. At the same time, he is a man of principle in his attitude to the highest level of the scientific work performed both by others and, especially, by himself. Both sides of his personality, his willingness to support people and his high standards of professional requirements, brilliantly manifested themselves in his editing of a series of scientific publications.

We warmly congratulate Mikhail Shlёmovich on his 70-th birthday and wish his vitality and creative spirit be preserved for a long time.

* * *

Below we give an account of the principal scientific achievements of M. Sh. during almost 50 years of his mathematical activity. The authors of this paper enjoyed the privilege of being in close scientific and personal contact with M. Sh. for many years, and two of us participated in many joint projects with him. Thus we hope that this survey gives some picture, incomplete though it may be, of the wide scale and superior quality of his ideas and results.

Many papers were written by M. Sh. in cooperation with his colleagues and students. When describing the results, we almost never mention coauthors; their names can be recovered from the list of publications (the next item of this collection, pp. 17–26).

1. Early papers

The first publications of M. Sh. were inspired by the ideas of L. V. Kantorovich in numerical analysis. In the papers [**1-3**] the multistep versions of the methods of steepest descent and successive approximations were studied. We describe here the main idea of the paper [**3**], in which an iterative procedure for the equation

$$Ax = \varphi \tag{1}$$

was analyzed. In (1), A is a bounded, selfadjoint and positive definite operator in a Hilbert space. Let m and M stand for the lower and upper bounds of the spectrum of A. In the traditional approach, the equation (1) is replaced by the equivalent equation $x = x - \varepsilon(Ax - \varphi) = (I - \varepsilon A)x + \varepsilon\varphi$, where ε is chosen so that $\|I - \varepsilon A\| < 1$. The new equation can be solved by the standard iterative procedure. In the multistep (p-step) version, one replaces (1) by the equation $x = B_p x - \varphi_p$, where

$$B_p = I + \sum_{k=0}^{p-1} \varepsilon_k^{(p)} A^{k+1} \qquad \text{and} \qquad \varphi_p = -\sum_{k=0}^{p-1} \varepsilon_k^{(p)} A^k \varphi.$$

The problem is to make the appropriate choice of the coefficients $\varepsilon_k^{(p)}$. This choice should minimize the norm of B_p and, as a consequence, optimize the rate of convergence of the successive approximations procedure.

By developing an idea of M. K. Gavurin, M. Sh. takes $B_p = T_p(A)$, where T_p is the Chebyshev polynomial of order p, transferred to the segment $[m, M]$. M. Sh. gives a detailed analysis of the suggested procedure, showing its advantages compared with the standard approach.

In [**1, 2**] a similar analysis was given for the numerical solution of equation (1) and for computation of the eigenvalues of A by the method of steepest descent.

The qualities, typical for all the scientific production of M. Sh., manifested themselves in these first publications: exhaustive analysis of the problem, extremely transparent exposition, and numerous comments, useful for the reader oriented to applications.

2. Extension theory of positive definite symmetric operators; elliptic boundary value problems

In 1952–54, M. Sh. turned his attention to the variational theory of elliptic boundary value problems. He had intensive discussions with S. G. Mikhlin on the subject. Partially influenced by these discussions, M. Sh. undertook an analysis of Treftz's method, which was suggested in 1926 for the Dirichlet problem in a bounded domain $\Omega \subset \mathbb{R}^d$:

$$(2) \qquad -\Delta u = f \text{ in } \Omega; \qquad u = 0 \text{ on } \partial\Omega.$$

The method consists in minimizing the Dirichlet integral $\int_\Omega |\nabla u|^2 dx$ among all solutions $u \in H^1(\Omega)$ of the equation $-\Delta u = f$ subject to no boundary conditions. The convergence of the method was proved by Mikhlin in 1950. Together with the standard variational procedure, based upon minimizing the functional $\int_\Omega \left(|\nabla u|^2 - 2\,\mathrm{Re}\,(\bar{u} f)\right) dx$, the Treftz method enables one to give two-sided estimates for the value of the Dirichlet integral of the solution of (2).

M. Sh. set up the problem of extending Treftz's method to the case of other boundary conditions. The main difficulty is to construct a quadratic functional ("Treftz's functional"), which attains its minimum value on the solution of a given boundary value problem. For the Dirichlet problem it is just the Dirichlet integral, but it is not so easy to find an appropriate substitute for other boundary conditions.

Approaching this problem, M. Sh. realized the relevance of the extension theory of positive definite symmetric operators. He intensively studied the papers by M. G. Krein and M. I. Vishik on the subject; in his own words, these papers, especially those of Krein, opened a new world for him. Since then M. Sh. regards himself as a "distant student" of Krein.

The basic results of M. Sh. in extension theory are presented in the papers [4] and [10]. Let A_0 be a positive definite, symmetric and closed operator in a Hilbert space. Denote by A_F the Friedrichs extension and by A an arbitrary positive definite selfadjoint extension of A_0. Developing the ideas of Krein and Vishik, M. Sh. considers the so-called parametric representation of A. The role of a parameter is played by a selfadjoint operator T, acting in a subspace of $\operatorname{Ker}(A_0^*)$. Basically,

$$T = (A^{-1} - A_F^{-1}) \restriction \operatorname{Ker} A_0^*. \tag{3}$$

M. Sh. gives a detailed analysis of relations between the spectral properties of A and those of the corresponding "operator parameter" T.

In order to apply the results obtained to second order elliptic boundary value problems, in [5] M. Sh. gives an exhaustive analytic description of the main objects appearing in the extension theory for this case. Then in [6, 11] he applies the results of his papers [4, 5] to the construction of Treftz's functionals for the basic second order elliptic boundary value problems. In [8] the same was done for the biharmonic equation. Incidentally, it is in this paper that various boundary conditions of the theory of thin plates were written down in an invariant form for the first time.

Shortly after that, M. Sh. has found another important application of his results on extension theory. By the classical Weyl theorem, the compactness of $A_1^{-1} - A_2^{-1}$ guarantees the coincidence of the essential spectra (σ_{ess}) of the operators A_1 and A_2. The equality (3) opens a convenient way to study the stability of σ_{ess} for singular elliptic operators.

M. Sh. started this study in [7]. He investigated elliptic operators in domains $\Omega \subset \mathbb{R}^d$ with compact complement. We describe his approach and results for the Laplacian. Consider the subspace $G^1(\Omega)$ of all harmonic functions in the Sobolev space $H^1(\Omega)$. Based upon equality (3), M. Sh. reduces the stability problem for σ_{ess} to the investigation of the embedding properties of $G^1(\Omega)$ in L_2. By proving compactness of the embedding operator, M. Sh. establishes the stability of σ_{ess} of the Laplacian with respect to a change of the domain and of the type of boundary conditions. Later the results of [7] were widely extended in [22]. Namely, M. Sh. found quantitative characteristics of the above embedding and, subsequently, eigenvalue estimates for the difference of the resolvents corresponding to a given elliptic operator with different boundary conditions.

One of the seminal papers of M. Sh. is [19] (a short preliminary version appeared in [15]), where he investigates the stability of σ_{ess} with respect to a change of the coefficients of an operator. A typical example is a comparison of σ_{ess} of the minus Laplacian on $\mathbb{R}^d$ and of the Schrödinger operator with a decreasing potential. The results obtained by M. Sh. in [15, 19], gave new insight into perturbation theory, and served for many followers as a basis for their own work.

Here we list some of the results of these papers.

1. For a semibounded selfadjoint operator A, $\sigma_{\text{ess}}(A)$ is stable with respect to a wide class of perturbations. In particular, it was shown that this class contains all the form-compact perturbations.

2. Suppose that $C \geq 0$ is such a perturbation; then the spectrum of $A - C$, lying to the left of $\sigma(A)$, is discrete. M. Sh. found an equality, connecting the total multiplicity of this spectrum with the eigenvalue behavior of the compact operator

generated by the quadratic form of C in the energy space of A. This important equality was rediscovered by Schwinger (for the particular case of Schrödinger operator) and later was called *the Birman–Schwinger principle*. Until now, this principle remains the basic tool for investigating the eigenvalues of operators with nonempty essential spectrum.

3. Based on this equality, M. Sh. gives finiteness and discreteness criteria for the spectra of all operators $A-\alpha C$, $\alpha > 0$, below $\sigma(A)$. For any "individual" value of the parameter α, these criteria turn into sufficient conditions.

4. Sometimes, the bottom of $\sigma(A)$ (for definiteness, suppose that this is $\lambda = 0$) is a resonance point. This means that the negative spectrum of $A - C$ is nonempty for an arbitrarily small negative form-compact perturbation $-C$. This phenomenon was explained in [**19**] as a consequence of the noncompatibility of two topologies on the energy space of A: one of them is the topology of the underlying Hilbert space, and the other is the topology generated by the quadratic form of A.

In order to apply these abstract results to differential operators, one needs to answer the following question from the theory of function spaces: for which weight-functions V is the Sobolev space $H^l(\mathbb{R}^d)$, or the "homogeneous" Sobolev space $\mathcal{H}^l(\mathbb{R}^d)$, compactly embedded into the weighted space $L_{2,V}(\mathbb{R}^d)$? The first embedding corresponds to the discreteness and the second to the finiteness of the negative spectra of the operators $(-\Delta)^l - \alpha V$ in $L_2(\mathbb{R}^d)$ (for all $\alpha > 0$ at once).

M. Sh. addresses this problem in the same papers [**15, 19**] and in [**18**]. For $2l > d$, a complete description of admissible V was found. For $l = 1$, $d \geq 2$, some necessary and (separately) sufficient compactness conditions were obtained in [**19**]. Later, a compactness criterion for this case was found by Maz'ya; it involves capacity.

As one application of these general results to the Schrödinger operator, we present here the famous estimate (first published in [**15**]) for the number of negative eigenvalues of the operator $-\Delta - V$ on $\mathbb{R}^3$:

$$N_-(-\Delta - V) \leq \frac{1}{16\pi^2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{V_+(x)V_+(y)}{|x-y|^2} dxdy. \tag{4}$$

This was the first quantitative estimate for the negative spectrum in the multidimensional case. It was independently found by Schwinger and is usually called Birman–Schwinger estimate.

3. Scattering theory: trace class approach

The trace class approach in scattering theory originated in papers by Kato and M. Rosenblum (1957) where the following fundamental result was proven. Let A and B be selfadjoint operators in a Hilbert space $\mathcal{H}$ and let $P_{\rm ac}(A)$ be the orthogonal projection on the absolutely continuous subspace $\mathcal{H}_{\rm ac}(A)$ of A. Suppose that the difference $B - A$ belongs to the trace class $\mathfrak{S}_1$. Then the strong limits

$$\operatorname*{s-lim}_{t\to\pm\infty} \exp(iBt)\exp(-iAt)P_{\rm ac}(A) =: W_\pm(B,A) \tag{5}$$

(called *the wave operators*) exist. The operators $W_\pm(B,A)$ are automatically isometric on $\mathcal{H}_{\rm ac}(A)$ and $BW_\pm(B,A) = W_\pm(B,A)A$. Since the assumptions of the Kato–Rosenblum theorem are symmetric, the wave operators $W_\pm(A,B)$ also exist and hence the absolutely continuous parts of the operators A and B are unitarily equivalent.

If the limits (5) do exist, then the scattering operator $\mathcal{S} := W_+^*(B,A)W_-(B,A)$ commutes with A and thus acts as multiplication by an operator-function $S(\lambda)$ in the diagonal for A representation of the space $\mathcal{H}$. The scattering operator $\mathcal{S}$ and the scattering matrix $S(\lambda)$ first appeared in the theory of scattering of quantum particles, which explains the title "scattering theory" for the perturbation theory of the absolutely continuous spectrum.

The simplest quantum mechanical system is described by the Schrödinger operator

$$B = -\Delta + V(x),$$

where $V(x)$ is a real function (the potential energy) which decays sufficiently rapidly at infinity. Thus B is a perturbation of the kinetic energy operator $A = -\Delta$ by a multiplication operator which is never compact. Therefore the Kato–Rosenblum theorem cannot be directly applied to this important case. Naturally, after the papers by Kato and Rosenblum the problem of applications of their theorem to differential operators was immediately posed. This problem was studied by Kato himself, Kuroda, and many other mathematicians. The contribution of M. Sh. to this highly competitive domain was crucial.

The study of the absolutely continuous spectrum $\sigma_{\rm ac}$ was for M. Sh. a natural continuation of his analysis of the essential spectrum. The connecting point is the paper [**22**] where the invariance of $\sigma_{\rm ac}$ was verified for perturbations of the boundary or of the type of boundary condition for elliptic operators in unbounded domains. The initial, and as it turned out later very fruitful, idea of M. Sh. was to consider suitable functions φ (for example, inverse powers) of these operators and to apply the Kato–Rosenblum theorem to the pair $\varphi(A)$, $\varphi(B)$.

The invariance of the absolutely continuous spectrum allowed M. Sh. to conjecture that under the assumption

$$\varphi(B) - \varphi(A) \in \mathfrak{S}_1 \tag{6}$$

not only $\sigma_{\rm ac}(B) = \sigma_{\rm ac}(A)$ but also the wave operators $W_\pm(\varphi(B), \varphi(A))$ exist and

$$W_\pm(\varphi(B), \varphi(A)) = W_\pm(B, A). \tag{7}$$

This result, proven by M. Sh. in [**27**] for a wide class of functions φ, was later called *the invariance principle.* At the same period, in the joint paper [**24**] of M. G. Krein and M. Sh., the Kato–Rosenblum theorem was carried over to unitary operators. This corresponds to the invariance principle for a fractional-linear function φ when $\varphi(A)$ and $\varphi(B)$ are the Cayley transforms of the operators A and B. The Birman–Krein theorem implies that for a pair of selfadjoint operators A, B the wave operators $W_\pm(B, A)$ exist if the difference of their resolvents belongs to the trace class. This is an important generalization of the Kato–Rosenblum theorem which can be directly applied to the Schrödinger operator.

The invariance principle is a concept, introduced by M. Sh. in the trace class framework, but it preserves its meaning in a much more general framework. Other important examples of this type are the local wave operators [**41**], related to some given interval of the spectral axis, and wave operators for pairs of selfadjoint operators acting in different Hilbert spaces [**42**]. These concepts, which were introduced in the context of abstract operator theory, play a decisive role in scattering problems for differential operators.

In addition to the time-dependent formulation, scattering theory admits also a stationary formulation, where the unitary groups are replaced by the resolvents of the operators A and B. In this formulation instead of limits (5) for large t one has to study the limits of resolvents as the complex spectral parameter approaches the real axis. M. Sh. developed the consistent stationary scheme in the trace class framework [**36**]. From the analytic point of view, the approach of this paper relies on the existence almost everywhere of boundary values (on $\mathbb{R}$) of the resolvent of any selfadjoint operator sandwiched between Hilbert–Schmidt operators. This result, proved in [**36**], is important in its own sake. Much later, M. Sh. returned to this topic in [**94**] where the stationary scheme was carried out in a very general framework.

Thus, M. Sh., partially together with his students, developed abstract scattering theory to the level where it could be directly applied to differential operators. These applications are summarized in the article [**46**], where a wide class of equations of mathematical physics is considered.

4. Spectral shift function

The spectral shift function $\xi(\lambda) = \xi(\lambda; B, A)$ may be introduced by the relation

$$\operatorname{Tr}(\varphi(B) - \varphi(A)) = \int_{-\infty}^{\infty} \varphi'(\lambda)\xi(\lambda)d\lambda, \tag{8}$$

called the trace formula. The concept of the spectral shift function in perturbation theory appeared at the beginning of fifties in the physics literature in the papers of I. M. Lifshitz. Its mathematical theory was created shortly after by M. G. Krein who proved relation (8) for a pair of selfadjoint operators A and B with a trace class difference $C = B - A$ and a wide class of functions φ.

A link between the spectral shift function and the scattering matrix $S(\lambda)$, associated with A and B, was found by M. Sh. and M. G. Krein in their joint note [**24**], which was already cited in the previous section. Actually, it was shown in [**24**] that

$$S(\lambda) - I \in \mathfrak{S}_1 \tag{9}$$

(so that the determinant of $S(\lambda)$ is well defined) and

$$\operatorname{Det} S(\lambda) = \exp(-2\pi i \xi(\lambda)) \tag{10}$$

for almost all $\lambda \in \sigma_{\rm ac}(A)$. This elegant relation is often used as the definition of the spectral shift function on the absolutely continuous spectrum. Actually, $\xi(\lambda)$ is also well defined on the discrete spectrum, where it depends on the shift of the eigenvalues of the operator B relative to the eigenvalues of A. This explains the term "spectral shift function."

It follows from (9) that the spectrum of a unitary operator $S(\lambda)$ consists of eigenvalues lying on the unit circle and accumulating at the point 1 only. This assertion was made more precise for perturbations C of definite sign. Actually, it was shown in [**24**] that there is no accumulation from above (from below) if $C \geq 0$ (if $C \leq 0$). In the abstract framework the eigenvalues of $S(\lambda)$ play the role of phase shifts for the Schrödinger operator with a radial potential.

The short note [**24**] (and report [**28**]) laid foundations for further studies of the function $\xi(\lambda)$ and of the spectral properties of $S(\lambda)$, both by students of M. Sh.

and by researchers from quite different mathematical schools. M. Sh. himself also maintained interest in that field for a long time. Thus, in [**80**] the asymptotics of the spectrum of the scattering matrix was found for the Schrödinger operator with a general (not radial) potential. The concept of the spectral shift function also turned out to be quite useful for the study of discrete spectrum; in particular, some of the results presented in Section 10 of this article were guessed by M. Sh. on the basis of this concept.

5. Double operator integrals

In 1965 M. Sh. started a systematic study of double operator integrals. His interest in the area originated in his work on scattering theory, where some technical problems reduce to the consideration of these integrals. As the theory developed it became clear that double operator integrals appear in many applications of quite different nature. Thus this formalism is more than just a technical trick; the development of the theory even stimulated new problems in Real Analysis. An original approach to approximation of functions in Sobolev spaces was invented to attack these problems. Moreover, this approach proved effective for many problems, some of which have nothing in common with operator integrals. The corresponding results are described in Sections 6 and 7.

A double operator integral is an expression of the form

$$Q = \int\int \psi(\lambda, \mu) dF(\mu) T dE(\lambda), \tag{11}$$

where dE and dF are spectral measures in a Hilbert space, T is a bounded operator and ψ is a scalar function. For given ψ, dE, and dF, (11) defines a linear mapping $\Psi : T \mapsto Q$. Such mappings, considered between appropriate spaces of operators, are usually called "transformers". Double operator integrals were introduced by Daletskii and S. G. Krein in 1956 to study the differentiation of operator-valued functions depending on a scalar parameter. In their work, some rather crude convergence conditions for the integrals (11) were found.

In [**32, 33, 35, 37, 59**] the theory of double operator integrals was extensively developed. In particular, the theory of them as transformers in the von Neumann–Schatten classes $\mathfrak{S}_p$ was established. It was shown that any bounded function ψ generates a bounded transformer in $\mathfrak{S}_2$. Boundedness criteria in $\mathfrak{S}_1$, $\mathfrak{S}_\infty$ and in $\mathfrak{B}$, and sufficient boundedness conditions in $\mathfrak{S}_p$ for other values of p were obtained. It was realized that many quite different objects appearing in Analysis (say, pseudodifferential operators on the one hand and the "triangle truncation transformer" in the theory of Volterra operators on the other) can be treated in the framework of double operator integrals.

In order to understand the idea of the double operator integral better, let us discuss one of its realizations. Consider the set of all integral operators acting from $L_2(\Lambda, d\lambda)$ to $L_2(M, d\mu)$, where $(\Lambda, d\lambda)$ and $(M, d\mu)$ are two measure spaces. Given a function $\psi(\lambda, \mu)$, consider the linear mapping which sends an integral operator T with kernel $T(\lambda, \mu)$ into the operator Q, with the kernel $\psi(\lambda, \mu)T(\lambda, \mu)$. It turns out that this mapping can be interpreted as a double operator integral. Moreover, any double operator integral can be represented as such a "multiplier transformation," if one admits integral operators with operator-valued kernels. Consequently, the results on double operator integrals can be interpreted as results on "multiplier problem for kernels of integral operators," which is a far reaching generalization of the classical problem of Schur multipliers.

One of important applications of double operator integrals concerns perturbation theory. If A and B are selfadjoint operators and dE, dF are their spectral measures, then for any "nice" function $\varphi(s)$ on $\mathbb{R}$ one has

$$\varphi(B) - \varphi(A) = \int\int \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} dE_\lambda (B - A) dF_\mu. \tag{12}$$

Wide conditions on φ, under which this representation (and a similar relation for unitary operators) can be justified, were obtained in [**33, 59**]. The Daletskii–S. G. Krein differentiation formula was also justified there, under rather mild assumptions on φ. If $B - A \in \mathfrak{S}_p$, the derivative exists in $\mathfrak{S}_p$-norm.

Comparison with the equality (8) allows one to guess that there should be a direct relation between the Daletskii–S. G. Krein differentiation formula and the spectral shift function. Such a relation was established in [**55**], where a new representation for the "integrated" spectral shift function $\Xi(\Delta) = \int_\Delta \xi(\lambda) d\lambda$ was obtained.

One more application of the same differentiation formula was found in the paper [**64**]. Namely, if A and B are nonnegative selfadjoint operators, then the following elegant inequality holds:

$$\|(B^\alpha - A^\alpha)_\pm\|_{\mathfrak{S}_p} \le \|(B - A)^\alpha_\pm\|_{\mathfrak{S}_p}, \qquad 0 < \alpha < 1,\ 1 \le p \le \infty.$$

6. Function spaces and piecewise-polynomial approximation. Singular number estimates for integral operators

The interest of M. Sh. in the theory of function spaces (especially in the embedding properties of Sobolev spaces) originates from his work in spectral theory. His first results in this field were briefly mentioned in Section 2. They were considerably refined in the mid-sixties, when M. Sh. started investigating quantitative characteristics of such embeddings. This study was based on a new approach to approximation of functions in Sobolev spaces $W^{l,p}$, which was suggested and developed in the papers [**34, 38**]; see also the book [**62**]. This approach can be regarded as a many-dimensional analog of spline approximation with nonfixed nodes.

Below we give a short description of this approach, for $l = 1$. Let X be a space of functions on the unit cube $Q^d \subset \mathbb{R}^d$, and suppose that the Sobolev space $W^{1,p}(Q^d)$ is compactly embedded in X. We want to approximate a given $u \in W^{1,p}(Q^d)$ by a suitable piecewise-constant function f, in the norm $\|\cdot\|_X$. The construction of f depends on the choice of a partition of Q^d into a finite family of smaller cubes, whose sizes are not prescribed. Only the number of cubes involved is under control: it should not exceed a given integer n. We take f equal to the mean value of u on each cube of the partition. Thus, the choice of f is completely determined by the choice of the partition. The next step is crucial: for a given n, we minimize the error $\|f-u\|_X$ among all partitions subject to the above restriction. An algorithmic partitioning procedure, suggested in [**34, 38**], gives an optimal approximation for a wide class of metrics $\|\cdot\|_X$.

This approach turned out to be quite flexible and efficient. The results are of two types. In the first type, the choice of the partition depends on the function to be approximated; this makes the procedure nonlinear. The main result, obtained by means of this version of the algorithm, is the sharp (up to the order) estimate of the epsilon-entropy $\mathcal{H}(\varepsilon)$ (in Kolmogorov's sense) of the unit ball of $W^{l,p}(Q^d)$ as a compact set in L_q, provided that the corresponding embedding is compact. For simplicity, we shall discuss this result for $d = 1$, $l = 1$, and $q = \infty$.

Let us recall the definition of the epsilon-entropy of a metric compact K. For a given $\varepsilon > 0$, let $\mathcal{N}(\varepsilon)$ be the least cardinality of epsilon-nets for K. The epsilon-entropy $\mathcal{H}(\varepsilon)$ of K is defined as $\log_2 \mathcal{N}(\varepsilon)$. The behavior of $\mathcal{H}(\varepsilon)$ as $\varepsilon \to 0$ characterizes, roughly speaking, the best possible rate of approximation of K by any approximating procedure, depending on ε^{-1} parameters.

It was well known by the end of fifties that for the unit ball of $C^\alpha[0,1]$, viewed as a compact set in $C[0,1]$, one has $\mathcal{H}(\varepsilon) = O(\varepsilon^{-1/\alpha})$. This order can be achieved by a simple linear procedure (say, for $\alpha \le 1$ by piecewise-linear approximation with equidistant nodes). The unit ball of $W^{\alpha,p}(0,1)$, regarded as a compact set in $L_p(0,1)$, has the same order of $\mathcal{H}(\varepsilon)$, namely $O(\varepsilon^{-1/\alpha})$. However, nothing was known on the epsilon-entropy of Sobolev embeddings with different exponents, including the simplest case $W^{1,p}(0,1) \subset C[0,1]$. This problem was solved in full generality in [**38**].

For the particular case discussed, the estimate obtained is $\mathcal{H}(\varepsilon) \asymp \varepsilon^{-1}$. This shows that the order of $\mathcal{H}(\varepsilon)$ for Sobolev classes is "nonsensitive" to the character of the metric in which the smoothness is measured, as well as to the metric of approximation. Together with the results on the d-widths of the same embedding (obtained later, mainly by Kashin), the estimate shows that the order $\mathcal{H}(\varepsilon) \asymp \varepsilon^{-1}$ cannot be achieved by any linear approximation method.

In the second type, the approximation is considered in a weighted space $L_{q,V}$, and the partitions are chosen independently of the function u to be approximated. Instead, the partitioning algorithm depends on V, and the estimates obtained are uniform with respect to a wide class of weights. Independence of u makes the approximation operators linear, which is important for applications to spectral theory.

The results on the weighted approximation were immediately applied to estimates of the singular numbers of integral operators. This problem is closely connected with double operator integrals: the boundedness criterium of the corresponding transformer in the classes $\mathfrak{S}_1$, $\mathfrak{S}_\infty$ and $\mathfrak{B}$, mentioned in Section 5, is formulated in terms of $\mathfrak{S}_1$-norm estimates for a family of integral operators. All these operators have the same kernel $\psi(\lambda, \mu)$, but act between different weighted L_2-spaces. Thus, in order to apply double operator integrals to concrete problems, one needs singular number estimates for such operators. No results on this subject existed in the mid-sixties, because integral operators in the "usual" L_2 setting were considered to be the only interesting case.

The problem was investigated in detail in [**39, 40, 43**], where singular number estimates, uniform with respect to weights, were obtained. This was the main novelty of these papers. However, many results were new even for the "usual" L_2, especially for integral operators on the whole of $\mathbb{R}^d$. In [**43, 68**] the estimates obtained using piecewise-polynomial approximation, were then extended and sometimes improved using interpolation. One more way to extend the results to wider classes of integral operators relies on the concept of multiplier transformations described in Section 5.

As was mentioned above, the original purpose of this study was to obtain convenient tests for integral operators to be of trace class $\mathfrak{S}_1$. Another approach to establish this property for integral operators between weighted L_2-spaces, which does not rely on piecewise-polynomial approximations, was found in [**45**].

Main results on the subject were summarized in [**69**]. Besides results of a general character, this paper contains material on some special classes of integral operators, including the class of operators in $L_2(\mathbb{R}^d)$ with kernels $b(x)e^{ix\xi}a(\xi)$. This class appears in many applications, including the study of the Schrödinger operator and the theory of pseudodifferential operators. Later, M. Sh. returned to this subject; this class of integral operators was studied in detail in the paper [**112**].

7. Spectral asymptotics for nonsmooth elliptic problems

Results on piecewise-polynomial approximation lead directly to eigenvalue estimates for the boundary value problems with weights, such as

$$-\Delta u = \lambda V u, \qquad u\,|_{\partial\Omega} = 0, \qquad \Omega \text{ is an open set in } \mathbb{R}^d, \tag{13}$$

and its higher order analogs. Let $N_\pm(\lambda)$ be the distribution functions of the positive and negative eigenvalues for the problem (13); here we admit indefinite weights, that is, weights which may change sign. The asymptotic behavior of $N_\pm(\lambda)$ is given by the Weyl-type formula

$$\lim_{\lambda\to\infty} \lambda^{-d/2} N_\pm(\lambda) = c_0(d)\int_\Omega V_\pm^{d/2}\,dx, \qquad c_0(d)^{-1} = \Gamma(d/2+1)(2\sqrt{\pi})^d. \tag{14}$$

For $V = 1$, (14) was obtained by Weyl in 1912 by the variational approach. The Tauberian technique, suggested by Carleman in 1936, turned out to be more flexible, and since then most of the results on the subject were obtained using this approach. However, it always required some smoothness restrictions on $\partial\Omega$ and V. For the nonsmooth case nothing was known by the end of sixties, except for the early papers of Weyl and Courant, where the formula (14) was justified for continuous weights bounded away from zero, and for domains Ω with "not too bad" boundary.

The estimates for $N_\pm(\lambda)$ found using piecewise-polynomial approximation techniques, involve the L_p-norm of V for an appropriate value of p. Such estimates prove to be a powerful tool for the study of spectral asymptotics. The way to apply them is based on an important (though, rather elementary) statement from abstract perturbation theory proved in [**49**]. When combined with the above-mentioned estimates, this statement shows that the (nonlinear) functionals

$$V \mapsto \limsup_{\lambda\to\infty} \lambda^{-d/2} N_\pm(\lambda), \qquad V \mapsto \liminf_{\lambda\to\infty} \lambda^{-d/2} N_\pm(\lambda),$$

for the problem (13) are continuous in L_p. It is important that this is an *a priori* fact, independent of the analytic form of the asymptotic coefficients, which is given by (14). The application of this general fact and its analogs lies at the heart of a renewed version of the variational approach to spectral asymptotics, which was suggested by M. Sh. in [**49**], developed in [**56, 57, 60**], and later summarized in [**62**]. It was shown in these papers that this approach is adequate for a wide class of nonsmooth eigenvalue problems.

Here are some of the results, obtained by this approach.

1. The asymptotic formula (14) is valid for the problem (13) in an arbitrary bounded open set $\Omega \subset \mathbb{R}^d$ for any real $V \in L_1(\Omega)$ $(d = 1)$, $V \in \bigcup_{p>1} L_p(\Omega)$ $(d = 2)$, and $V \in \bigcup_{p>d/2} L_p(\Omega)$ $(d \geq 3)$. In particular, the original Weyl asymptotic formula for the Dirichlet Laplacian (that is, (14) for $V = 1$) holds for any bounded domain in

any dimension. For $d \geq 3$, the conditions on V were later relaxed by G. Rozenblum, to become $V \in L_{d/2}(\Omega)$. Moreover, the result was extended by him to arbitrary (not necessarily bounded) open sets Ω.

Note that in the framework of the approach suggested, one encounters no additional difficulties when considering indefinite weights. Before, only some particular results (by Plejel) were known for the indefinite case.

2. The Weyl–Carleman eigenvalue asymptotic formula was justified in [**57**] for arbitrary uniformly elliptic second order operators in the divergent form, with measurable bounded coefficients. All the previous results of this sort required some regularity (at least, continuity) of the leading coefficients. Similar results were obtained in [**57, 58**] for more general problems, such as $Au = \lambda Bu$, where A is a selfadjoint elliptic operator, corresponding to a positive quadratic form of differential order l, and B is a symmetric operator, corresponding to a quadratic form of differential order $r < l$. The results apply also to elliptic systems. The assumptions about the coefficients of both forms are formulated in terms of appropriate L_p-spaces. In particular, this admits degeneration of ellipticity for A. Moreover, the form (Bu, u) can be indefinite.

3. The eigenvalue asymptotics for a wide class of integral operators with "weakly polar" kernels, acting in $L_2(\Omega)$, $\Omega \subset \mathbb{R}^d$, was established in [**48**]. Another, more general approach to this class of problems was suggested later in [**67, 73**]. An asymptotic formula of the Weyl type was obtained there for pseudodifferential operators of negative order, under minimal assumptions on the smoothness of the symbol.

4. In [**70, 75, 76, 77, 82, 83**] estimates and asymptotics were established for operators associated with a variational quotient, considered on a space of solutions of a given differential equation. As one of the consequences, the spectral asymptotics for operators of the type $A_{\mathcal{N}}^{-1} - A_{\mathcal{D}}^{-1}$ was obtained; here $A_{\mathcal{N}}$ and $A_{\mathcal{D}}$ are Neumann and Dirichlet realizations for a given elliptic differential operator A in a (bounded or unbounded) domain $\Omega \subset \mathbb{R}^d$. These results can be viewed as a refinement of the estimates found by M. Sh. much earlier in [**22**]; see Section 2 of the present paper.

5. A way to obtain eigenvalue estimates and asymptotics for problems on the whole of $\mathbb{R}^d$, starting from the ones for bounded domains, was suggested in [**54**]. Not only the results, but also the approach of [**54**] turned out to be important for many applications. It was widely extended in the subsequent papers [**109, 113**].

8. Laplace and Maxwell operators in domains with nonsmooth boundary

This is one more field in which M. Sh. has been interested since the early stages of his scientific career.

It was well known by the mid-fifties that for any bounded region Ω with smooth boundary the domain of the selfadjoint Dirichlet Laplacian $\Delta_{\mathcal{D}}$ is $H^2(\Omega) \cap H^{1,0}(\Omega)$. It was also known that in general this is not so if $\partial\Omega$ is nonsmooth. However, nothing was known about the possibility of obtaining an analytic description of $\mathrm{Dom}(\Delta_{\mathcal{D}})$ for this case. This problem was stated and solved for plane regions with corners in the pioneering paper [**25**]. In particular, it was shown there that the image of $\Delta_{\mathcal{D}} \restriction H^2(\Omega) \cap H^{1,0}(\Omega)$ is a closed subspace in $L_2(\Omega)$, whose codimension is equal to the number of inward corners of Ω.

Later M. Sh. returned to this class of problems in connection with the theory of Maxwell operator (in anysotropic media). The Maxwell operator is not semibounded. Therefore the quadratic form approach does not apply. Thus the primary problem for nonsmooth regions $\Omega \subset \mathbb{R}^3$ is to find an appropriate selfadjoint realization of the operator. Only then can the smoothness properties of the solutions be investigated.

This program was realized in [**88, 89, 90, 93, 95**]. As a result, a description of the singularities of the electric component of an electro-magnetic field was obtained for regions with Lipshitz boundaries. It was given in terms of singularities of the solutions of an appropriate scalar elliptic equation for which such a description was well known. For the magnetic component, the results were less exhaustive. However, by exploiting the description of the electric component, M. Sh. succeeded in justifying the Weyl eigenvalue asymptotic formula for the Maxwell operator in an arbitrary bounded region with Lipschitz boundary [**91**]. Before this paper, the result was known only for regions with smooth boundary.

It is impossible not to mention the program paper [**114**], where M. Sh. compares the solvability properties of three important problems in 3-dimensional polyhedra: the systems of Maxwell, Stokes, and Lamé. Under appropriate boundary conditions, coming from physics, all three of these problems reduce formally to the same boundary value problem. However, for nonconvex polyhedra the physical origin of these problems dictates a different choice of the selfadjoint realization for each of them. The paper contains a detailed discussion of this striking effect.

9. Estimates on the number of negative eigenvalues of the Schrödinger operator and its analogs

The first such estimate for the multidimensional case, namely the Birman–Schwinger inequality (4), was found by M. Sh. in [**15**]. Another important estimate was obtained in 1972 by G. Rozenblum by refining the piecewise-polynomial approximation approach described in Section 6. The estimate for operators involving a large parameter α (the coupling constant) states that

$$N_-(-\Delta - \alpha V) \le C(d)\alpha^{d/2} \int_{\mathbb{R}^d} V_+^{d/2} dx, \qquad d \ge 3. \tag{15}$$

In addition to (15), the asymptotic formula

$$\lim_{\alpha\to\infty} \alpha^{-d/2} N_-(-\Delta - \alpha V) = c_0(d) \int_{\mathbb{R}^d} V_+^{d/2} dx, \tag{16}$$

is valid, where $c_0(d)$ is the same constant as in (14). The relations (15) and (16) show that the semiclassical behavior of $N_-(-\Delta - \alpha V)$ is typical for the problem considered.

Let us call the estimate (15) "regular". One encounters an "irregular" growth of N_-, say, when a bounded potential V decays at infinity (and therefore, the negative spectrum of $-\Delta - \alpha V$ is still discrete), but $V_+ \notin L_{d/2}$. One example of an irregular estimate is given by (4). Indeed, the right-hand side of (4) acquires the factor α^2 (instead of the regular $\alpha^{3/2}$), if we replace V by αV. The finiteness of the integral in (4) does not imply $V_+ \in L_{d/2}$, so (4) is independent of the regular estimate (15).

Effects appearing in the case of irregular behavior of $N_-(-\Delta - \alpha V)$ were investigated in detail in [**98, 102, 109, 113**]. The main tool was linear interpolation.

An important observation of a rather general nature was made in [**102**]: in contrast with the regular case, irregular estimates are unstable with respect to the additional spectral parameter. Namely, in many cases the behavior of $N_-(-\Delta - \alpha V)$ is irregular, whereas the behavior of $N_-(-\Delta + hI - \alpha V)$ is regular for any fixed $h > 0$. It was also shown that the sharp irregular estimates always involve the (quasi)-norm of V_+ in an appropriate nonseparable Banach or quasi-Banach function space. Consequently, the asymptotic behavior of $N_-(-\Delta - \alpha V)$ is determined by the values of V in a neighbourhood of its main singularities (or for bounded V, by its behavior at infinity). This is also in contrast with the regular case: formula (16) shows that any subset of $\mathbb{R}^d$ with nonzero Lebesgue measure contributes to the asymptotics. The sharpness of the estimates obtained in [**109**] was confirmed by a series of explicit examples where the product $\alpha^{-q}N_-(-\Delta - \alpha V)$ with a prescribed $q > d/2$ has a nonzero limit as $\alpha \to \infty$.

To be more precise, the results of [**109, 113**] concern the "generalized" Schrödinger operator $A_{\alpha V} = (-\Delta)^l - \alpha V$ on $\mathbb{R}^d$. The estimates for $2l < d$ and for $2l \geq d$ look different. It was shown that for $2l > d$, d odd, the irregular estimates obtained are invertible: the same quasi-norms of V are involved both in the upper and lower estimates for $N_-(A_{\alpha V})$. In particular, this occurs for $d = 1$, when the results apply to the operator on the positive semiaxis $A_{\alpha V}y = (-1)^l y^{(2l)} - \alpha V y$, with Dirichlet boundary conditions at $x = 0$. Namely, necessary and sufficient conditions on V were found, guaranteeing that $N_-(A_{\alpha V}) = O(\alpha^q)$, with a prescribed value of $q > (2l)^{-1}$. The importance of this result was realized later, when addressing the operator $(-\Delta)^l - \alpha V$ on $\mathbb{R}^d$ for $2l \geq d$ and d even. It was understood that the "unpleasantness" appearing in this case is caused by some auxiliary differential operator on the semiaxis. The above results make it possible to carry out a detailed analysis of the multidimensional problem in question. This was done in [**130**] for the Schrödinger operator on $\mathbb{R}^2$, and in [**132**] for the higher order case. The results obtained show, in particular, that there are potentials V such that the function $N_-((-\Delta)^l - \alpha V)$ has the semiclassical order $O(\alpha^{d/2l})$ as $\alpha \to \infty$, but non-qWeyl type asymptotics. Basically, the results of [**109, 113, 130, 132**] exhaust the problem of irregular behavior of N_- for the Schrödinger operator and its higher order analogs.

10. Discrete spectrum of a perturbed operator in the gaps of the unperturbed one

Suppose that the essential spectrum of a given selfadjoint operator A is a disconnected set. Typical examples are the Dirac operator, whose spectrum is $(-\infty, -1] \cup [1, \infty)$, and the Schrödinger operator with periodic potential, whose spectrum has a band structure. For an operator A with such a spectrum, the complementary intervals are called gaps. If A is perturbed by a relatively compact operator, then a discrete spectrum may appear in the gaps. Investigation of this spectrum is an important problem. The variational technique does not apply to it in a direct way. It was shown by M. Sh., as far back as in [**16, 19**], how to overcome this difficulty for the Dirac operator. The case of periodic Schrödinger operator also was always in the sphere of his interests. The study of eigenvalues in the gaps of the spectrum, when this operator is perturbed by a decaying potential, was one of the important problems mentioned in the survey talk [**29**]. Some of the students of M. Sh. made contributions towards the solution of this problem, but he himself addressed it again only since 1990.

Let (λ_-, λ_+) be a gap for the unperturbed operator A, and let V be a perturbation. For simplicity, we suppose here that $V \geq 0$ and that the perturbed operator is $A(t) = A - tV$, with $t > 0$. When t grows, the eigenvalues $\lambda_k(t)$ of $A(t)$ spring up at the point λ_+ and move to the left. Fix a point $\lambda \in [\lambda_-, \lambda_+]$ and an $\alpha > 0$, and let $N(A, V, \alpha, \lambda)$ denote the number of eigenvalues $\lambda_k(t)$ which cross the level λ as t grows from 0 to α. The problem of interest is the behavior of $N(A, V, \alpha, \lambda)$ for fixed λ as $\alpha \to \infty$.

M. Sh. investigated this problem in detail. There is a big difference between the cases $\lambda \in [\lambda_-, \lambda_+)$ and $\lambda = \lambda_+$ (behavior at the upper edge of the gap). For the first case, a general "comparison theorem" was established in [**110**]. Let A be bounded from below, and let B be obtained from A by a form-bounded perturbation with zero bound. The theorem says that, under some assumptions on V, the quantities $N(A, V, \alpha, \lambda)$ and $N(B, V, \alpha, \mu)$ have the same asymptotic behavior as $\alpha \to \infty$, for arbitrary real $\lambda \in \rho(A)$ and $\mu \in \rho(B)$; here $\rho(\cdot)$ stands for the set of all regular points of the operator under consideration. In applications to the periodic Schrödinger operator (both in one and many dimensions) this allows one to apply known results on the negative spectrum of the "usual" Schrödinger operator (see Section 9) to the problem discussed. In [**111**] the same general theorem was applied to the magnetic Schrödinger operator.

The case $\lambda = \lambda_+$ is much subtler; here one encounters one more manifestation of the "instability effect" mentioned in Section 9. Irregular behavior of $N(A, V, \alpha, \lambda)$ is possible, and in such a case the eigenvalues and eigenfunctions of the corresponding "quasi-periodic" operators $A(\xi)$, depending on the quasi-momentum ξ, are involved in the asymptotic formulas in an explicit way. These formulas were obtained in [**118, 122, 128**].

The paper [**123**] concerns the Dirac operator on $\mathbb{R}^3$. Both the regular and irregular asymptotic behavior of $N(A, V, \alpha, \lambda)$ was analyzed and some unexpected phenomena, caused by the same instability effect, were discovered.

In conclusion, we mention the remarkable result in the recent paper [**131**]. It was shown there that the periodic magnetic Schrödinger operator on $\mathbb{R}^2$ has absolutely continuous spectrum. This important problem seemed to be unassailable for many years.

* * *

Approaching his seventieth birthday, Mikhail Shlëmovich Birman did not scale down his scientific activity. While this article was under preparation, his publication list became two papers longer. We leave their description, as well as the description of other results not touched upon in this paper, until another jubilee date.

Department of Physics, St. Petersburg State University, Stary Peterhof, St. Petersburg 198904, Russia
E-mail address: `buslaev@mph.niif.spb.su`

Department of Theoretical Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: `solom@wisdom.weizmann.ac.il`

Department of Mathematics, Université de Rennes, IRMAR, Campus de Beaulieu, 35042 Rennes CEDEX, France
E-mail address: `yafaev@maths.univ-rennes1.fr`