Introduction

As an introduction, we first describe the historical background of the spectral theory of linear operators. In the second part, the state of the art of research on the spectral theory of block operator matrices is outlined. The third and last part contains a brief description of the contents of this book.

1. Historical background. The spectral theory of linear, in particular self-adjoint, operators in a Hilbert space is one of the major advances in mathematics of the 20th century (see [Ste73a] for a historical survey). It was initiated by D. Hilbert in his famous six papers on integral equations from 1904 to 1910 (see [Hil53]) which contain all basic ideas for the spectral theorem for bounded self-adjoint operators. Simultaneously, H. Weyl further advanced the theory for singular second order ordinary differential equations in his seminal paper [Wey10]. The next major breakthrough came during the years 1927 to 1929 when J. von Neumann developed the abstract concept of Hilbert space and linear operators therein and initiated the spectral theory for unbounded self-adjoint operators (see [vN30a]). Von Neumann's work was driven by the needs of quantum mechanics (see [vN27], [vN32]), which was created in 1925/26 mainly by W. Heisenberg, E. Schrödinger, and P. Dirac (see [BJ25], [BHJ26], [Sch26a], [Sch26b], [Dir25], [Dir26]). Between 1927 and 1932, this spectral theory was elaborated and extended to unbounded normal operators by F. Riesz (see [Rie30]) and, in great detail and independently of further work by von Neumann (see [vN30b]), by M.H. Stone (see [Sto32]). Further important contributions concerning extensions of semi-bounded symmetric operators and applications to differential operators are due to K.O. Friedrichs in 1934 (see [Fri34]) and M.G. Kreĭn in 1947 (see [Kre47b], [Kre47a]).

Another important field that was stimulated by problems from mathematical physics is the perturbation theory of linear operators (see [Kat95]).

It originates in the works of Lord Rayleigh from 1877 on vibrating systems (see [Ray26]) and of E. Schrödinger from 1926 on eigenvalue problems in quantum mechanics. The first important contribution in this field is due to H. Weyl in 1909 and concerns the stability of the now so-called essential spectrum of a bounded self-adjoint operator (see [Wey09]). A mathematically rigorous perturbation theory for eigenvalues and eigenvectors of self-adjoint operators $T(\kappa)$ depending analytically on a parameter κ was developed by F. Rellich between 1937 and 1942 in a series of fundamental papers (see [Rel42]). Later, in 1948, K.O. Friedrichs created a perturbation theory for the continuous spectrum which proved to be useful in scattering theory and quantum field theory (see [Fri48], [Fri52]). In 1949 T. Kato started his deep investigations on the perturbation theory of linear operators (see [Kat50], [Kat52], [Kat53]) which form one of the bases of his famous so entitled 1966 monograph [Kat95].

The spectral theory of non-self-adjoint linear operators is much more involved than that of self-adjoint operators. Although the first pioneering works of G.D. Birkhoff on eigenfunction expansions for non-self-adjoint differential operators (see [Bir08], [Bir12], [Bir13]) were written almost at the same time as Hilbert's famous six papers, it took about 40 years for abstract results to follow. Since then, a wealth of results has become available in the literature which are not so widely known, in particular, among physicists and engineers. It is impossible to give even a brief account of them and so only a few milestones can be mentioned. In the years 1951/52 B. Sz.-Nagy, F. Wolf and T. Kato generalized Rellich's results and obtained the first theorems on the perturbation theory of closed linear operators (see [SN51], [Wol52], [Kat52]). At the same time, in 1951, M.V. Keldysh's cornerstone work on the completeness of eigenvectors and associated vectors and eigenvalue asymptotics of non-self-adjoint operator polynomials was published (see [Kel51]), which had a great impact on the spectral theory of non-self-adjoint differential operators. A seminal work from 1957 on the stability of various spectral properties, in particular, of the index for non-self-adjoint operators is due to I.C. Gohberg and M.G. Kreĭn (see [GK60]). Almost simultaneously and independently, T. Kato established similar results (see [Kat58]). I.C. Gohberg and M.G. Kreĭn also wrote a comprehensive account of results on non-self-adjoint operators as early as 1965 (see [GK69]). One focus of this book is on completeness results as initiated by V.I. Matsaev (see *e.g.* [Mat61], [Mat64]) and by V.B. Lidskiĭ (see *e.g.* [Lid59b], [Lid59a]); it also contains many other important contributions, *e.g.* on classes of compact linear operators and s-numbers, estimates of the growth of the resolvent of non-self-adjoint linear operators, and perturbation determinants. Another important direction was opened up in 1954 by N. Dunford who developed the theory of spectral operators (see [Dun54]); a detailed presentation of the latter and an enormous collection of results on non-self-adjoint operators is contained in the volume [DS88] by N. Dunford and J.T. Schwartz. A structure theory for non-self-adjoint operators was created in 1954 by M.S. Livšic; in [Liv54] he introduced the notions of operator colligations (or nodes, following the literal translation from Russian) and characteristic functions and employed them intensively with his colleagues, in particular, M.S. Brodskiĭ (see [Bro71], [BL58], and the monograph [LY79] with its review in [Bal81]). In the 1960ies, the notion of characteristic functions appeared also in the work of B. Sz.-Nagy and C. Foiaş in the different context of unitary dilation spaces for contractions (see [SNF70]). One of the main tools in almost all of the above works is the extensive use of the theory of functions, either for studying the behaviour of the resolvents of non-self-adjoint operators or of operator colligations by means of the factorization of characteristic functions (see *e.g.* the monograph [BGK79] by H. Bart, I.C. Gohberg, and M.A. Kaashoek).

Linear operators that are self-adjoint with respect to an indefinite inner product were brought up in quantum field theory, in works from 1942/43 by P. Dirac (see [Dir42]) and W. Pauli (see [Pau43]). The first basic result for operators in infinite dimensional spaces with indefinite inner product is due to L.S. Pontryagin in 1944 (see [Pon44]). Inspired by a mechanical problem considered by S.L. Sobolev (which was, at that time, secret military research and published only in 1960, see [Sob60]), Pontryagin proved the existence of a special invariant subspace for a self-adjoint operator in a space with finite-dimensional positive part (now called Pontryagin space). This result was extended and generalized by I.S. Iohvidov, M.G. Kreĭn, and H. Langer between 1956 and 1962 (see [IK56], [IK59], [Lan62], and the joint monograph [IKL82]). In 1963 M.G. Kreĭn and H. Langer established a spectral function for self-adjoint operators in Pontryagin spaces (see [KL63]). Shortly after, in 1965, a comprehensive spectral theory for definitizable self-adjoint operators in Krein spaces (where positive and negative part may be infinite dimensional) was set up by H. Langer (see [Lan65], [Lan82]). Many of the above results may be found in the monographs [Bog74] by J. Bognár and [AI89] by T.Ya. Azizov and I.S. Iohvidov.

In spite of all its inherent problems, the spectral theory of non-self-adjoint linear operators is of utmost importance for applications: Self-adjointness is an intrinsic property of energy-preserving (so-called closed) systems; however, in practice, many systems *e.g.* from engineering are not energy-preserving and have a non-self-adjoint state operator (so-called open systems). The recent book [TE05] by L.N. Trefethen and M. Embree on the behaviour of non-normal matrices and operators provides striking evidence of this by mentioning 19 fields with more than 8000 representative citations and by its own bibliography comprising 851 references. The significance of the problems arising due to the lack of self-adjointness, especially with regard to numerical calculations, have now become widely accepted, at least among mathematicians. Major contributions to this are due to F. Chatelin who further developed Kato's perturbation theory in view of applications to the numerical spectral analysis in 1983 (see the monograph [Cha83]). Other important concepts are pseudospectra and spectral value sets originating in works of H.J. Landau from 1975 (see [Lan75]), S.K. Godunov *et al.* from 1990 (see [GKK90]), N. Trefethen from 1992 (see [Tre92] and the monograph [TE05]), and of D. Hinrichsen and A. Pritchard in view of uncertain systems (see [HP92] and the monograph [HP05]). Important contributions to pseudospectra of differential operators are due to E.B. Davies (see [Dav99], [Dav00], and the recent survey [Dav07]).

2. Spectral theory of block operator matrices. If $\mathcal{A}$ is a bounded linear operator in a Hilbert space $\mathcal{H}$ and a decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ into two Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ is given, then $\mathcal{A}$ always admits a block operator matrix representation

$$\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{I}$$

with linear operators A, B, C, and D acting in or between the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. For an unbounded linear operator $\mathcal{A}$ in $\mathcal{H}$, its domain $\mathcal{D}(\mathcal{A})$ need not be decomposable as $\mathcal{D}_1 \oplus \mathcal{D}_2$ with $\mathcal{D}_1 \subset \mathcal{H}_1$, $\mathcal{D}_2 \subset \mathcal{H}_2$ and hence it is an additional assumption that $\mathcal{A}$ has a representation (I) such that $\mathcal{D}(\mathcal{A}) = \big(\mathcal{D}(A) \cap \mathcal{D}(C)\big) \oplus \big(\mathcal{D}(B) \cap \mathcal{D}(D)\big)$. In this case, we call $\mathcal{A}$ an unbounded block operator matrix.

Our aim is to use information about the entries A, B, C, and D to investigate various spectral properties of the block operator matrix $\mathcal{A}$. If $\mathcal{A}$ is not self-adjoint or symmetric and/or all entries of $\mathcal{A}$ are unbounded, we face a number of difficulties:

a) The results on self-adjoint operators and perturbations of them rely heavily on the following properties: if $\mathcal{A}$ with domain $\mathcal{D}(\mathcal{A})$ is self-adjoint in a Hilbert space $\mathcal{H}$ with scalar product $(\cdot,\cdot)$, then its numerical range $W(\mathcal{A}) = \big\{(\mathcal{A}x, x) : x \in \mathcal{D}(\mathcal{A}), \|x\| = 1\big\}$ and its spectrum $\sigma(\mathcal{A})$ are real, the norm of the resolvent $(\mathcal{A} - \lambda)^{-1}$ is bounded by $1/\mathrm{dist}\big(\lambda, \sigma(\mathcal{A})\big)$ for $\lambda \notin \sigma(\mathcal{A})$, all eigenvalues are algebraically simple (*i.e.* there are no Jordan chains), and, if the spectrum of $\mathcal{A}$ is discrete (*i.e.* if it consists only of isolated eigenvalues with finite multiplicities), then $\mathcal{H}$ has an orthonormal basis of eigenvectors. For a non-self-adjoint linear operator $\mathcal{A}$ all these properties may fail: the numerical range and the spectrum need no longer be real, there are no estimates of the resolvent in terms of the spectrum (which may result in an extreme sensitivity of the spectrum to perturbations), the eigenvalues need not be algebraically simple, and, if the spectrum is discrete, then the system of eigenvectors and associated vectors may not be complete and may fail to be "linearly independent".

b) For the spectral theory of unbounded linear operators, it has to be required that the operator is closed or at least closable. However, the sum of closed or closable operators need not be closed or closable, respectively; similarly, the sum of self-adjoint operators need not be self-adjoint. As a consequence, an unbounded block operator matrix need not be closed even if so are its entries; if it is closable, then the closure need not have a block operator matrix representation. We identify classes of closable block operator matrices and describe their closures. This classification is based on inclusion relations between the domains of the entries: diagonally dominant, off-diagonally dominant, and upper dominant block operator matrices. It turns out that, in many respects, these classes require different approaches.

c) The most powerful tool to investigate all kinds of spectral data of a self-adjoint operator is the spectral function. Nothing comparable exists in the non-self-adjoint case. For bounded isolated parts of the spectrum, a contour integral over the resolvent, the so-called Riesz projection, can be used to identify a corresponding spectral subspace. However, if the spectrum consists of two unbounded isolated parts it is not clear if the spectrum splits at infinity, *i.e.* if there exist corresponding spectral subspaces which reduce the operator. Moreover, even for self-adjoint operators that are non-semi-bounded and hence have spectrum tending to ∞ and $-\infty$, the classical variational principles do not apply to eigenvalues in gaps of the spectrum.

d) For 2×2 matrices, the eigenvalues can be characterized as the zeroes of a quadratic polynomial, the characteristic determinant. Since, in general, the entries A, B, C, and D of a block operator matrix (I) act between different spaces, it is not possible to multiply them in a way resembling the scalar 2×2 case; *e.g.* if $\mathcal{H}_1 = \mathbb{C}^k$ and $\mathcal{H}_2 = \mathbb{C}^l$ with $k \neq l$, then the product of matrices $AD - BC$ cannot be formed. To study the spectrum of 2×2 block operator matrices, one therefore has to find other functions encoding the spectral data (*e.g.* Schur complements). These functions depend on the complex spectral parameter and their values are linear operators; how to build them, may depend on the invertibility of the entries of the block operator matrix and on inclusions between their domains.

The above mentioned problems have been attacked in the theory of block operator matrices by various different methods:

2.1. *Quadratic numerical range and block numerical ranges.* The numerical range $W(\mathcal{A}) = \big\{(\mathcal{A}x, x) : x \in \mathcal{D}(\mathcal{A}),\ \|x\| = 1\big\}$ of a linear operator $\mathcal{A}$ in a Hilbert space $\mathcal{H}$ is a reliable method to localize its spectrum. It was first studied by O. Toeplitz in 1918 (see [Toe18]); he proved that the numerical range of a matrix contains all its eigenvalues and that its boundary is a convex curve. In 1919 F. Hausdorff showed that indeed the set $W(\mathcal{A})$ is convex (see [Hau19]). In fact, it turned out that this continues to hold for general bounded linear operators and that the spectrum is contained in the closure $\overline{W(\mathcal{A})}$ (see [Win29]). For unbounded linear operators $\mathcal{A}$, the inclusion of the spectrum prevails if every component of $\mathbb{C} \setminus \overline{W(\mathcal{A})}$ contains at least one point of the resolvent set of $\mathcal{A}$; moreover, the resolvent estimate $\|(\mathcal{A}-\lambda)^{-1}\| \le 1/\mathrm{dist}\big(\lambda, W(\mathcal{A})\big)$ holds for $\lambda \notin \overline{W(\mathcal{A})}$ (see [Kat95]). Another interesting property is that every corner of $\overline{W(\mathcal{A})}$ belongs to the spectrum and is an eigenvalue if it lies in $W(\mathcal{A})$ (see [Don57], [Hil66]).

At first sight, the convexity of the numerical range seems to be a useful property, *e.g.* to show that the spectrum of an operator lies in a half plane. However, the numerical range often gives a poor localization of the spectrum and it cannot capture finer structures such as the separation of the spectrum in two parts. In view of these shortcomings, the new concept of quadratic numerical range was introduced in 1998 in [LT98] and further studied in [LMMT01], [LMT01]. If $\mathcal{A}$ is a bounded linear operator, a decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ of the Hilbert space is given, and (I) is the corresponding block operator matrix representation of $\mathcal{A}$, then the quadratic numerical range $W^2(\mathcal{A})$ of $\mathcal{A}$ is the set of all eigenvalues of the 2×2 matrices

$$\mathcal{A}_{x,y} = \begin{pmatrix} (Ax, x) & (By, x) \\ (Cx, y) & (Dy, y) \end{pmatrix}, \quad x \in \mathcal{H}_1,\ y \in \mathcal{H}_2,\ \|x\| = \|y\| = 1. \qquad \text{(II)}$$

The obvious generalization to $n \times n$ block operator matrices is called block numerical range of $\mathcal{A}$ (see [Wag00], [TW03]); for unbounded block operator matrices $\mathcal{A}$, only matrices $\mathcal{A}_{x,y}$ with normed elements $x \in \mathcal{D}(A) \cap \mathcal{D}(C)$ and $y \in \mathcal{D}(B) \cap \mathcal{D}(D)$ are considered (see [LT98], [Tre08]).

The quadratic numerical range is always contained in the numerical range: $W^2(\mathcal{A}) \subset W(\mathcal{A})$. However, unlike the numerical range, the quadratic numerical range is no longer convex; it consists of at most two components which need not be convex either. On the other hand, the quadratic numerical range shares many other properties with the numerical range: it enjoys the spectral inclusion property; it furnishes a resolvent estimate of the form $\|(\mathcal{A}-\lambda)^{-1}\| \le 1/\operatorname{dist}\big(\lambda, W^2(\mathcal{A})\big)^2$ for $\lambda \notin \overline{W^2(\mathcal{A})}$; a corner of $\overline{W^2(\mathcal{A})}$ belongs to the spectrum of $\mathcal{A}$ or one of its diagonal entries A, D. Analogous results hold for the block numerical range and some, *e.g.* the spectral inclusion, also for certain classes of unbounded block operator matrices. Compared to other spectral enclosures of Gershgorin or Brauer type (see [Ger31], [Bra58]), the quadratic numerical range has the advantage of not requiring any norms of inverses.

Besides the spectral inclusion, the most important feature of the quadratic numerical range is that it yields a criterion for block diagonalizability. The corresponding theorem generalizes the well-known fact that every 2×2 matrix with two distinct eigenvalues can be diagonalized: If the closure of $W^2(\mathcal{A})$ consists of two disjoint components, then $\mathcal{A}$ can be block diagonalized. An analogue for the block numerical range was proved recently under some additional conditions (see the PhD thesis [Wag07]).

Although the quadratic numerical range is a relatively recent concept, applications of it have already appeared in the literature. V.V. Kostrykin, K.A. Makarov, A.K. Motovilov used it to prove perturbation results for spectra and spectral subspaces of bounded self-adjoint operators (see [KMM07]); some of their results are presented in Section 1.3 (see Theorems 1.3.6, 1.3.7). In [Lin03] H. Linden applied the quadratic numerical range to derive enclosures for the zeroes of monic polynomials. K.-H. Förster and N. Hartanto developed a Perron-Frobenius theory for the block numerical range of (entrywise) nonnegative matrices in [FH08], thus generalizing corresponding results for the spectrum and the numerical range.

2.2. *Schur complements and factorization.* Schur complements were first used in the theory of matrices. The idea to associate the matrix $D-CA^{-1}B$ with a block matrix $\mathcal{A}$ as in (I) (with non-singular A) may be traced back at least to Schur (see [Sch17], and maybe even to earlier work by J. Sylvester). The name “Schur complement” was created by E. Haynsworth in 1968 when she began to study partitioned matrices (see [Hay68]). In Hilbert spaces, Schur complements may be found first in M.G. Kreĭn’s famous paper [Kre47b] on the extension of self-adjoint operators, under the name “shorted operators”. Apart from their numerous applications in matrix theory and numerical linear algebra, Schur complements are used in many other areas including statistics, electrical engineering, C^*-algebras (see *e.g.* [Zha05] with its exhaustive bibliography, [Bha07], and [CIDR05]), and in mathematical systems theory, where they appear as transfer functions of state space realizations of linear time invariant systems (see [BGKR05]).

In the theory of bounded and unbounded block operator matrices, Schur complements are powerful tools to study the spectrum and various spectral properties. This was first recognized by R. Nagel in a series of papers starting in 1985 (see [Nag85], [Nag89], [Nag90], [Nag97]). He began to develop a "matrix theory" for unbounded operator matrices with "diagonal domain" (block operator matrices in our terminology) and with "non-diagonal domain" (allowing for some coupling between the two components of elements of the domain). The intimate relation between the spectral properties of the block operator matrix $\mathcal{A}$ and those of its Schur complements

$$S_1(\lambda) = A - \lambda - B(D-\lambda)^{-1}C, \quad \lambda \notin \sigma(D),$$
$$S_2(\lambda) = D - \lambda - C(A-\lambda)^{-1}B, \quad \lambda \notin \sigma(A),$$

is obvious from the so-called (formal) Frobenius-Schur factorizations, *e.g.*

$$\mathcal{A} - \lambda = \begin{pmatrix} I & 0 \\ C(A-\lambda)^{-1} & I \end{pmatrix} \begin{pmatrix} A-\lambda & 0 \\ 0 & S_2(\lambda) \end{pmatrix} \begin{pmatrix} I & (A-\lambda)^{-1}B \\ 0 & I \end{pmatrix}, \quad \lambda \notin \sigma(A),$$

and the corresponding factorization for the resolvent $(\mathcal{A}-\lambda)^{-1}$. Important milestones in this direction are: the paper [ALMS94] by F.V. Atkinson, H. Langer, R. Mennicken, and A.A. Shkalikov from 1994 where the essential spectrum of a upper dominant block operator matrix was determined by means of Schur complements; the paper [AL95] by V.M. Adamjan and H. Langer from 1995 which contains the key ideas for the block diagonalization of operator matrices with self-adjoint separated diagonal entries $D \ll 0 \ll A$ and bounded corners B, C that are self-adjoint ($C = B^*$) or $\mathcal{J}$-self-adjoint ($C = -B^*$). In the subsequent papers [ALMS96], [Shk95], [AMS98], the approach of [AL95] was extended to self-adjoint block operator matrices with unbounded entries, and in [LT98] to non-self-adjoint diagonal entries A, D with spectra separated by a vertical strip and $C = B^*$. The Schur complement approach of [ALMS94] to determine the essential spectrum was further developed by A.A. Shkalikov (see [Shk95]) and A. Jeribi *et al.* (see [MDJ06], [DJ07]) in the Banach space case, and, in the Hilbert space case, by A.Yu. Konstantinov *et al.* (see [Kon96], [Kon97], [Kon98], [Kon02], [KM02], [AK05]) who also studied the absolutely continuous spectrum and gave applications to upper dominant singular matrix differential operators.

Like Livšic' characteristic functions, Schur complements are operator-valued analytic functions reflecting the spectral properties of the associated block operator matrix $\mathcal{A}$. Also here, methods of complex analysis such as the factorization theorems by A.I. Virozub and V.I. Matsaev for the self-adjoint case (see [VM74]) and by A.S. Markus and V.I. Matsaev for the general case (see [MM75]) may be employed. The former was used by R. Mennicken and A.A. Shkalikov (see [MS96]) who generalized the results of [AL95], [ALMS96] for self-adjoint block operator matrices with $D \leq A$ to the case where A and the Schur complement S_2 satisfy a certain separation condition. The factorization theorem by Markus and Matsaev, together with Brouwer's fixed point theorem, was used in [LMMT01] to prove the theorem on block diagonalization for bounded non-self-adjoint block operator matrices. It was shown that if the closure of the quadratic numerical range consists of two disjoint components, $\overline{W^2(\mathcal{A})} = \mathcal{F}_1 \dot\cup \mathcal{F}_2$, then the Schur complements admit factorizations

$$S_j(\lambda) = M_j(\lambda)(Z_j - \lambda), \quad j = 1, 2, \tag{III}$$

with operator functions M_j that are holomorphic in $\mathcal{F}_j$ and have boundedly invertible values and linear operators Z_j such that $\sigma(Z_j) = \sigma(\mathcal{A}) \cap \mathcal{F}_j$. As a consequence of the factorization (III), the block operator matrix $\mathcal{A}$ turns out to be similar to the block diagonal matrix diag (Z_1, Z_2).

Interestingly, in 1990/1991, before the abstract methods described above were developed, eigenvalue problems for second order differential expressions depending on the spectral parameter rationally (with so-called "floating singularity", see [Bog85]) were studied (see [LMM90], [FM91], [ALM93]). In fact, the corresponding linear operators are the Schur complements of upper dominant matrix differential operators. These investigations, especially on λ-rational Sturm-Liouville problems on compact intervals were further elaborated to study also eigenvalue accumulation and embedded eigenvalues by means of Sturm's comparison and oscillation theories (see [ALM93], [MSS98], [Lan00], [ALL01], [Lan01]). The case of singular intervals was treated by J. Lutgen in [Lut99]; the case of the whole axis with periodic boundary conditions was considered by R.O. Hryniv, A.A. Shkalikov, and A.A. Vladimirov in [HSV00], [HSV02]. It is impossible to give an account of the vast literature on matrix differential operators studied purely by means of techniques from the theory of differential equations; we only mention Dirac systems or Schrödinger type operator matrices.

In mathematical physics, Schur complements were first used by H. Feshbach in 1958 (see [Fes58]) and have since become valuable tools under the name Feshbach maps or decimation maps (see *e.g.* [Bac01, Section 7], [Lut04], [BCFS03], [GH08]). Here the entry A corresponds *e.g.* to low energy states of the system without interaction and the operator $A - BD^{-1}C$ is called decimated Hamiltonian. The fact that the spectrum and eigenvalues of a block operator matrix outside the spectrum of the diagonal entry A, say, coincide with the spectrum and eigenvalues, respectively, of its first Schur complement S_1 and the relation $P_1(\mathcal{A}-\lambda)^{-1}P_1 = S_1(\lambda)^{-1}$ are referred to as the Feshbach projection method or Grushin problem in the physics literature (see [BFS98, Chapter II] and [SZ03]).

2.3. *Algebraic Riccati equations.* There are two algebraic Riccati equations formally associated with a block operator matrix $\mathcal{A}$ as in (I):

$$K_1BK_1 + K_1A - DK_1 - C = 0, \quad K_2CK_2 + K_2D - AK_2 - B = 0. \quad \text{(IV)}$$

Even in the matrix case, the existence of solutions to such quadratic operator equations is a non-trivial problem. In the following we describe a purely analytical approach which relies on rewriting the Riccati equations as operator Sylvester equations (sometimes also called Kreĭn-Rosenblum equations) and using a fixed point argument; subsequently, we present methods that are based on the close relation of solutions of Riccati equations and invariant graph subspaces of the block operator matrix.

The starting point for the fixed point approach is to write *e.g.* the first Riccati equation in the equivalent form $KA - DK = Y$ with $Y := C - KBK$. Solutions K to such operator equations in integral form seem to have been found first by M.G. Kreĭn in 1948 (see [Phó91]), and later independently by Yu. Daleckiĭ (see [Dal53]) and M. Rosenblum (see [Ros56]). The key condition is that the spectra of the operator coefficients A and D on the left hand side have to be disjoint; then the solution is given by the so-called Daleckiĭ-Kreĭn formula (see [DK74a, Theorem I.3.2] or [GGK90, Theorem I.4.1])

$$K = -\frac{1}{2\pi \mathrm{i}} \oint_{\Gamma_D} (D-z)^{-1} Y (A-z)^{-1} =: \Phi(K);$$

here Γ_D is a Cauchy contour around $\sigma(D)$ separating it from $\sigma(A)$. Since $Y = C - KBK$, this formula may be viewed as a fixed point equation for K. To ensure that the mapping Φ so defined is a contraction, smallness conditions have to be imposed on the coefficients B and C. As a result of Banach's fixed point theorem, we obtain the existence and uniqueness of contractive solutions of the Riccati equation and a fixed point iteration converging to the solution in the operator norm.

The first to use a fixed point argument in connection with a Kreĭn-Rosenblum equation was G.W. Stewart in a series of papers between 1971 and 1973 (see [Ste71], [Ste72], [Ste73b]). He introduced a special implicit measure $\delta(A,D)$ for the separation of the spectra of A and D, which is defined *e.g.* if one of them is bounded and, in the self-adjoint case, reduces to the usual distance. Assuming $\sqrt{\|B\|\,\|C\|} < \delta(A,D)/2$, Stewart proved the existence of a bounded solution of the associated Riccati equation, which guarantees the existence of an invariant subspace of a closed linear operator (see below). Motivated by applications to two-channel Hamiltonians from elementary particle physics, A.K. Motovilov applied the fixed point method to self-adjoint diagonally dominant block operator matrices in 1991/1995 (see [Mot91], [Mot95]); he proved existence and uniqueness of solutions of Riccati equations if $\operatorname{dist}\big(\sigma(A),\sigma(D)\big) > 0$ and the Hilbert-Schmidt norm of B satisfies $\|B\|_2 < \operatorname{dist}\big(\sigma(A),\sigma(D)\big)/2$. This approach was further advanced, and more general conditions were found, for the non-self-adjoint case in [ALT01] with bounded B, C, D, in [AMM03] with bounded B, C, and in [AM05] for the case of one normal diagonal entry. The fixed point method also applies if the spectra of A and D are not disjoint: In a series of papers, R. Mennicken, A.K. Motovilov, and V. Hardt (see [MM98], [MM99], [HMM02], [HMM03]) used it for upper dominant block operator matrices for which $\sigma(D)$ is partly or entirely embedded in the continuous spectrum $\sigma_c(A)$ of A; they assumed that the second Schur complement admits an analytic continuation under the cuts along the branches of the absolutely continuous spectrum of A, which is ensured by conditions on B. In numerical linear algebra, iterative schemes for solving Riccati equations were used by K. Veselič and E. Kovač Striko in [KSV01] in order to establish an algorithm for block diagonalizing non-self-adjoint matrices (see below).

Riccati equations play an important role in mathematical systems theory (see the monographs [LR95, Chapter IV] by L. Rodman and P. Lancaster, [CZ95, Chapter 6] by R. Curtain and H. Zwart and the bibliographies therein). They arise *e.g.* in linear quadratic (LQ) optimal control on an infinite time interval and have been the subject of intense research since Kalman's seminal paper from 1960 (see [Kal60]). There, and in other areas like computational physics and chemistry (see [Ben00]), the operator coefficients in (IV) have the special properties that $D = -A^*$ and B, C are non-negative; the corresponding block operator matrices are called Hamiltonians. In systems theory results on existence and uniqueness of non-negative Hermitian solutions of Riccati equations are sought as well as iterative schemes to approximate solutions for infinite time intervals by solutions for finite time intervals (see *e.g.* [Mår71], [CZ95], [KSV01]). An idea of the vast literature on Riccati equations for Hamiltonians with bounded coefficients is given in [LR95]; for unbounded coefficients, important contributions are due to A. Pritchard and D. Salamon (see [PS84], [PS87]), and to I. Lasiecka and R. Triggiani (see [LT91], [LT00b], [LT00a]) who gave applications to linear partial differential equations with boundary and point controls. Although some of the methods presented here apply to Riccati equations from systems theory (see [LRT97], [LRvdR02], Remark 2.7.26, and Corollary 2.9.23), this wide area is out of the scope of this book.

2.4. *Invariant graph subspaces and block diagonalization.* Invariant subspaces of matrices and linear operators are a key tool in many areas of mathematics and its applications (see the monograph [GLR86] by I.C. Gohberg, P. Lancaster, and L. Rodman for an exhaustive account). The existence of solutions to the Riccati equations (IV) implies that the subspaces

$$\mathcal{L}_1 = \left\{ \begin{pmatrix} x \\ K_1 x \end{pmatrix} : x \in \mathcal{H}_1 \right\}, \quad \mathcal{L}_2 = \left\{ \begin{pmatrix} K_2 y \\ y \end{pmatrix} : y \in \mathcal{H}_2 \right\} \tag{V}$$

are invariant for the block operator matrix $\mathcal{A}$ formed from their coefficients. As a consequence, the operator $\mathcal{A}$ formally admits the block diagonalization

$$\begin{pmatrix} I & K_2 \\ K_1 & I \end{pmatrix}^{-1} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & K_2 \\ K_1 & I \end{pmatrix} = \begin{pmatrix} A + BK_1 & 0 \\ 0 & D + CK_2 \end{pmatrix}. \tag{VI}$$

Vice versa, if a block operator matrix $\mathcal{A}$ possesses invariant subspaces of the form (V), then the operators K_1, K_2 therein are solutions of the Riccati equations (IV). Hence the existence of an invariant graph subspace of a block operator matrix is equivalent to the existence of a solution of a corresponding Riccati equation.

The operators K_1, K_2 in (V) are called angular operators since they provide a measure for the perturbation of the invariant subspaces $\mathcal{H}_1 \oplus \{0\}$, $\{0\} \oplus \mathcal{H}_2$ of the block diagonal operator $\operatorname{diag}(A, D)$ if B and C are turned on. As an instructive example, we consider a real 2×2 matrix $\mathcal{A}$ (see [Hal69], [KMM05]). If $\mathcal{A}$ has two different real eigenvalues λ_1, λ_2 with eigenvectors $(x_1\ y_1)^{\mathrm{t}}$, $(x_2\ y_2)^{\mathrm{t}}$, say $x_1 \neq 0$, then *e.g.*

$$\mathcal{L}_1 = \left\{ \begin{pmatrix} x_1 \\ k_1 x_1 \end{pmatrix} : x_1 \in \mathcal{H}_1 \right\}, \quad k_1 = \frac{y_1}{x_1} =: \tan\theta;$$

here $\theta \in [0, \pi/2)$ is the angle between the axis $\mathbb{R} \oplus \{0\}$ and the eigenspace $\mathcal{L}_1$. The orthogonal projection P of $\mathbb{C}^2$ onto $\mathcal{L}_1$ is given by

$$P = \begin{pmatrix} \cos^2\theta & \sin\theta\,\cos\theta \\ \sin\theta\,\cos\theta & \sin^2\theta \end{pmatrix}. \tag{VII}$$

If P_1 is the orthogonal projection of $\mathbb{C}^2$ onto $\mathbb{C} \oplus \{0\}$, then it is not difficult to check that $\|P - P_1\| = \sin\theta$. If $\mathcal{A}$ is a bounded block operator matrix that can be block diagonalized and $\mathcal{L}_1$ is an invariant subspace as in (V), then the orthogonal projection P of $\mathcal{H}_1 \oplus \mathcal{H}_2$ on $\mathcal{L}_1$ is given by

$$P = \begin{pmatrix} (I + K_1^* K_1)^{-1} & (I + K_1^* K_1)^{-1} K_1^* \\ K_1 (I + K_1^* K_1)^{-1} & K_1 (I + K_1^* K_1)^{-1} K_1^* \end{pmatrix}.$$

Comparing with formula (VII), we recognize the formal correspondence

$$I + K_1^* K_1 \;\longleftrightarrow\; \frac{1}{\cos^2\theta} = 1 + \tan^2\theta \quad \text{or} \quad \sqrt{K_1^* K_1} \;\longleftrightarrow\; \tan\theta,$$

which is the identity $k_1 = \tan\theta$ in the matrix case. Using the notion of an operator angle Θ of a pair of subspaces (see *e.g.* [KMM03a, Section 2]), one arrives at the rigorous identities $\sqrt{K_1^* K_1} = \tan\Theta$ and $\|P - P_1\| = \sin\Theta$ where P_1 is the orthogonal projection of $\mathcal{H}_1 \oplus \mathcal{H}_2$ on $\mathcal{H}_1 \oplus \{0\}$.

Even in the case of matrices or bounded linear operators, there are no simple answers neither to the problem of existence of invariant graph subspaces nor to the problem of existence of solutions of Riccati equations; in the unbounded case, additional problems with domains and closures arise. However, the equivalence of the problems widens the range of methods available for their solution. Besides the fixed point methods for Riccati equations described before, we distinguish four main directions:

For self-adjoint operators $\mathcal{A}$, a geometric approach was initiated by works of C. Davis and W.M. Kahan in the 1960ies (see [Dav63], [Dav65], [DK70]); they studied the perturbation of spectral subspaces of a self-adjoint operator $\mathcal{A}_0$ the spectrum of which has two disjoint components. Note that this induces a decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ in which $\mathcal{A}_0$ is block diagonal, say $\mathcal{A}_0 = \operatorname{diag}(A, D)$ with $\sigma(A) \cap \sigma(D) = \emptyset$. Their main results are four different types of theorems, called $\sin\theta$ theorem, $\tan\theta$ theorem, $\sin 2\theta$ theorem, and $\tan 2\theta$ theorem, which give the best possible bound on the angle between the perturbed and unperturbed spectral subspaces.

Independently, and only in the 1990ies, V.M. Adamjan and H. Langer developed a different analytic approach in [AL95] for self-adjoint and $\mathcal{J}$-self-adjoint operators of the form

$$\mathcal{A} = \begin{pmatrix} A & B \\ \pm B^* & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} + \begin{pmatrix} 0 & B \\ \pm B^* & 0 \end{pmatrix} = \mathcal{A}_0 + \mathcal{V} \qquad \text{(VIII)}$$

with bounded off-diagonal perturbation $\mathcal{V}$. Under the stronger assumption $\sigma(D) < \sigma(A)$, but without bounds on the norm of $\mathcal{V}$ in the self-adjoint case, they proved that the interval $\big(\max\sigma(D), \min\sigma(A)\big)$ remains free of spectrum for the perturbed operator $\mathcal{A} = \mathcal{A}_0 + \mathcal{V}$ and that the spectral subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$ corresponding to the intervals $\big[\min\sigma(A), \infty\big)$ and $\big(-\infty, \max\sigma(D)\big]$ admit angular operator representations (V) with a uniform contraction K_1 and $K_2 = -K_1^*$ (the latter being a consequence of the orthogonality of $\mathcal{L}_1$ and $\mathcal{L}_2$); their method is based on analytic estimates of integrals over the inverse of the Schur complement.

In 2001 a novel approach involving indefinite inner products was established in [LT01] for block operator matrices (VIII) with non-self-adjoint $\mathcal{A}_0$ and self-adjoint $\mathcal{V}$ (*i.e.* $C = B^*$). It is based on a theorem on accretive linear operators in Krein spaces which applies if we assume that $\operatorname{Re} W(D) \le 0 \le \operatorname{Re} W(A)$. In fact, if $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ is equipped with the indefinite inner product $[\cdot,\cdot] = (\mathcal{J}\cdot,\cdot)$ with $\mathcal{J} = \operatorname{diag}(I, -I)$, then

$$\begin{aligned}\operatorname{Re}\left[\begin{pmatrix} A & B \\ B^* & D\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix}, \begin{pmatrix} x \\ y\end{pmatrix}\right] &= \operatorname{Re}\left(\begin{pmatrix} A & B \\ -B^* & -D\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix}, \begin{pmatrix} x \\ y\end{pmatrix}\right) \\ &= \operatorname{Re}(Ax,x) - \operatorname{Re}(Dx,x) \;\ge\; 0\end{aligned} \tag{IX}$$

for $(x\ y)^{\mathrm{t}} \in \mathcal{D}(\mathcal{A})$. This theorem yields the existence of invariant subspaces of $\mathcal{A}$ that are maximal non-positive and maximal non-negative with respect to $[\cdot,\cdot]$, provided $\mathcal{A}$ is exponentially dichotomous (see the next subsection). As a consequence of the definiteness of these invariant subspaces, we obtain angular operator representations (V) with contractions K_1, K_2. This approach does not only furnish a new and more elegant proof for the self-adjoint case treated in [AL95], it also covers non-self-adjoint diagonally dominant and off-diagonally dominant block operator matrices. Note that, for the latter, the off-diagonal part $\mathcal{V}$ can no longer be regarded as a perturbation of the diagonal part $\mathcal{A}_0$ and so, even in the self-adjoint case, none of the previous results applies.

Finally, in parallel, a fourth method was developed in [LMMT01] which relies on the factorization theorems by Markus and Matsaev used for the Schur complements. It applies to bounded linear operators $\mathcal{A}$ and decompositions $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ such that the closure of the quadratic numerical range consists of two disjoint components. In this case, *e.g.* the angular operator K_1 and the operator Z_1 in the linear factor of $S_1(\lambda) = M_1(\lambda)(Z_1 - \lambda)$ are related by the formulae

$$Z_1 = A + BK_1, \quad K_1 = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_1} (D-\lambda)^{-1} C (Z_1 - \lambda)^{-1}\,\mathrm{d}\lambda;$$

here Γ_1 is a Cauchy contour separating the two components of the quadratic numerical range. Similarly, we have $Z_2 = D + CK_2$ and a corresponding integral formula for K_2. This shows that, indeed, the operators Z_1, Z_2 in the factorizations (III) of the Schur complements are the diagonal entries in the block diagonalization (VI).

Numerous papers were published following one or two of the previous approaches. The analytic approach by Adamjan and Langer was further pursued for unbounded upper dominant self-adjoint block operator matrices with $\max\sigma(D) \le \min\sigma(A)$ in [ALMS96] and, for the first time, for non-self-adjoint A and D with $\operatorname{Re} W(D) \leq 0 \leq \operatorname{Re} W(A)$, bounded D, and $C = B^*$ in [LT98]. The method of factorizing the Schur complements was first applied in [MS96] to allow for a certain overlapping of the spectra of A and D. The geometric approach of Davis and Kahan was further elaborated in a series of papers by A.K. Motovilov *et al.* (see [KMM03b] for general bounded $\mathcal{V}$; [KMM03a], [KMM07] for bounded off-diagonal $\mathcal{V}$; [KMM04] for bounded off-diagonal $\mathcal{V}$ and the case $\max \sigma(D) \leq \min \sigma(A)$; [KMM05] for off-diagonal $\mathcal{V}$ and the case that A is bounded and $\sigma(A)$ lies in a finite gap of $\sigma(D)$; [AMS07] for bounded off-diagonal $\mathcal{V}$ and [MS06] for unbounded off-diagonal $\mathcal{V}$ and, in both cases, $\sigma(A) \cap \operatorname{conv} \sigma(D) = \emptyset$). As in [KMM07] and [AMM03], optimal bounds on the norm of $\mathcal{V}$ guaranteeing that the perturbed spectrum remains separated and on the angle between the perturbed and the unperturbed spectral subspaces are given.

2.5. *Dichotomous block operator matrices.* A linear operator is called dichotomous if its spectrum does not intersect the imaginary axis $\mathrm{i}\mathbb{R}$; in this case, the spectrum of $\mathcal{A}$ splits into two parts $\sigma_1 \subset \mathbb{C}_+$, $\sigma_2 \subset \mathbb{C}_-$ in the open right and left half plane, respectively. The notion of dichotomous operators is closely related to the notion of dichotomy for differential equations. In the most classical case, it means that the solution of a Sturm-Liouville equation on $L_2(\mathbb{R})$ is the sum of two solutions from $L_2(0,\infty)$ and $L_2(-\infty,0)$, respectively. For evolution equations $u'(t) = \mathcal{A}u(t)$, $t \in [0,\infty)$, in an abstract Banach or Hilbert space $\mathcal{H}$, the concept of exponential dichotomy was considered by S.G. Kreĭn and Ju.B. Savčenko (see [KS72]); essentially, it means that there exist two invariant subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$ of $\mathcal{A}$ such that $\mathcal{H} = \mathcal{L}_1 \oplus \mathcal{L}_2$ and $\|u(t)\|$ decays (increases, respectively) exponentially for $t \to \infty$ if $u(t) \in \mathcal{L}_2$ ($u(t) \notin \mathcal{L}_2$, respectively). If $\mathcal{A}$ is a bounded dichotomous operator, then $\mathcal{L}_1$, $\mathcal{L}_2$ can be chosen to be the spectral subspaces of $\mathcal{A}$ corresponding to σ_1, σ_2, *i.e.* the ranges of the corresponding Riesz projections; if $\mathcal{A}$ is self-adjoint and unbounded, then $\mathcal{L}_1$, $\mathcal{L}_2$ can be chosen to be the ranges of the corresponding spectral projections. If, however, $\mathcal{A}$ is unbounded and not self-adjoint, then both σ_1 and σ_2 may be unbounded and the problem of "separating the spectrum at infinity" (see [GGK90, Section XV.3]) arises.

If an unbounded non-self-adjoint block operator matrix $\mathcal{A}$ as in (VIII) is exponentially dichotomous, then it can be transformed into block diagonal form. The spectral inclusion theorem for the quadratic numerical range yields a criterion for dichotomy. If we assume that the numerical ranges of A and D are separated by a strip around $\mathrm{i}\mathbb{R}$, *i.e.* for some α, $\delta > 0$

$$W(A) \subset \{z \in \mathbb{C} : \operatorname{Re} z \geq \alpha\}, \quad W(D) \subset \{z \in \mathbb{C} : \operatorname{Re} z \leq -\delta\},$$

that the strip $S := \{z \in \mathbb{C} : -\delta < \operatorname{Re} z < \alpha\}$ contains at least one point of $\rho(A) \cap \rho(D)$, and that certain relative boundedness assumptions are satisfied for the entries of $\mathcal{A}$, then $S \subset \rho(\mathcal{A})$ and hence $\mathcal{A}$ is dichotomous. Note that this implies that the block operator matrix $\mathcal{A}$ is $\mathcal{J}$-accretive (see (IX)). If we require, in addition, that $W(A)$ and $W(D)$ lie in certain sectors of angle less than π in the right and left half plane, respectively, then $\mathcal{A}$ is exponentially dichotomous. This follows from a deep theorem proved by H. Bart, I.C. Gohberg, and M.A. Kaashoek (see [BGK86, Theorem 3.1], [GGK90, Theorem XV.3.1]); they studied exponentially dichotomous operators intensively in relation with Wiener-Hopf factorization. Equivalent conditions for the separation of the spectrum at infinity were given by G. Dore and A. Venni in [DV89] in terms of powers of the operator in question. In [RvdM04], [vdMR05], A.C.M. Ran and C. van der Mee considered additive and multiplicative perturbations of exponentially dichotomous operators; they used methods different from those presented here, *e.g.* the Bochner-Phillips theorem. A survey of this area and of applications of exponentially dichotomous operators, *e.g.* to transport equations, diffusion equations of indefinite Sturm-Liouville type, noncausal infinite-dimensional linear continuous-time systems, and functional differential equations of mixed type are given in the recent monograph [vdM08] by C. van der Mee.

2.6. *Variational principles and eigenvalue estimates.* The variational characterization of eigenvalues goes back well into the 19th century, to H. Weber (see [Web69]) and Lord Rayleigh (see [Ray26]). If $\mathcal{A}$ is a self-adjoint operator that is semi-bounded, say from below, then the eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots$ of $\mathcal{A}$ below its essential spectrum can be characterized by means of the classical min-max principle

$$\lambda_n = \min_{\substack{\mathcal{L} \subset \mathcal{D}(\mathcal{A}) \\ \dim \mathcal{L} = n}} \max_{\substack{x \in \mathcal{L} \\ x \neq 0}} p(x), \quad p(x) := \frac{(\mathcal{A}x, x)}{\|x\|^2}, \tag{X}$$

(see *e.g.* [RS78, Chapter XIII], [WS72], [Gou57], [Ste70]). Here p is the so-called Rayleigh functional defining the numerical range of $\mathcal{A}$. Besides the min-max principle, which is based on the inequalities of H. Poincaré (see [Poi90]), there also exists a max-min characterization which relies on the inequalities of H. Weyl (see [Wey12]). Min-max and max-min principles are effective tools in the qualitative and quantitative analysis of eigenvalues of self-adjoint operators for several reasons. They do not require any knowledge about eigenvectors and can be used for comparing eigenvalues of operators, deriving eigenvalue estimates, locating the bottom of the essential spectrum (or showing it is empty), proving the existence of eigenvalues below the essential spectrum, and for numerical approximations of eigenvalues; corresponding algorithms have been used in countless applications from physics and engineering sciences for decades, *e.g.* in elasticity theory for calculating buckling loads of beams or plates (see *e.g.* [Mik64]).

Due to the convexity of the numerical range, the classical variational principles only apply to eigenvalues to the left (or to the right) of the essential spectrum, but not to eigenvalues in gaps of the essential spectrum. This excludes eigenvalues of some important operators from mathematical physics like Dirac operators, Klein-Gordon operators, and Schrödinger operators with periodic potentials. The first abstract min-max principles for block operator matrices with spectral gap were proved by M. Griesemer and H. Siedentop in 1999 and generalized in 2000 jointly with R.T. Lewis (see [GS99], [GLS99]), followed by work of M.J. Esteban, J. Dolbeault, and E. Séré beginning in 2000 (see [DES00b]), [DES00c], [DES06]). The main motivation of these authors was to characterize the eigenvalues of Dirac operators with Coulomb potential, as suggested in earlier work of J. Talman and of S.N. Datta and G. Deviah (see below). While the assumptions and methods of proof are different, the common idea in these papers is to use the given decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ and to impose the dimension restriction only in one component, *e.g.* in the bounded case

$$\lambda_n = \min_{\substack{\mathcal{L}_1 \subset \mathcal{H}_1 \\ \dim \mathcal{L}_1 = n}} \max_{\substack{x \in \mathcal{L}_1 \oplus \mathcal{H}_2 \\ x \neq 0}} p(x). \tag{XI}$$

Much earlier, beginning in the 1950ies, min-max principles were proved for operator functions depending non-linearly on the spectral parameter. In 1955 R. Duffin was the first to consider the case of self-adjoint quadratic operator polynomials $T(\lambda)$ (see [Duf55] and the later work [Bar74] of E.M. Barston); self-adjoint continuously differentiable functions of matrices and of bounded linear operators defined on some real interval I were studied in 1964 by E.H. Rogers (see [Rog64], [Rog68]) and by B. Werner in 1971 (see [Wer71]), respectively. Here a generalized Rayleigh functional p was introduced which, *e.g.* if $T(\cdot)$ is strictly monotonically increasing, is defined to be the unique zero $p(x)$ of $\big(T(\lambda)x, x\big)$ (or $\pm\infty$ if no zero exists). This definition is closely related to the numerical range of the operator function T, which is given by $W(T) := \big\{\lambda \in I : \big(T(\lambda)x, x\big) = 0 \text{ for some } x \in \mathcal{H}, x \neq 0\big\}$. In the particular case $T(\lambda) = \mathcal{A} - \lambda$, the unique zero of $\big(T(\lambda)x, x\big) = 0$ is just the classical Rayleigh functional $p(x)$ (see (X)). Under much weaker assumptions on the operator functions, these min-max principles were generalized by P. Binding, H. Langer, and D. Eschwé in [BEL00]) for the case of bounded values $T(\lambda)$, and by D. Eschwé, and M. Langer in [EL04] for the case of unbounded values $T(\lambda)$. In both papers, the variational principles were applied to the Schur complements to characterize eigenvalues of self-adjoint and even skew-self-adjoint block operator matrices.

A new type of variational principles for eigenvalues in spectral gaps appeared in 2002 (see [LLT02], [KLT04]). They apply to block operator matrices having real quadratic numerical range $W^2(\mathcal{A})$, *i.e.* to self-adjoint and certain skew-self-adjoint block operator matrices. Here the role of the classical Rayleigh functional is played by the functionals $\lambda_\pm$ induced by the quadratic numerical range; they are defined as the zeroes $\lambda_\pm\binom{x}{y}$ of the quadratic polynomial $\det(\mathcal{A}_{x,y}-\lambda)$ (*i.e.* the eigenvalues of the matrix $\mathcal{A}_{x,y}$ given by (II)). The proof of these novel min-max and max-min principles uses the variational principle of [EL04] and the inclusion of the numerical range of the Schur complement in the quadratic numerical range. As a corollary, we obtain the min-max principle (XI) with the classical Rayleigh functional p. In the off-diagonally dominant case, the results of [KLT04] are restricted to bounded diagonal entries; a generalization to one relatively bounded diagonal entry was given in [Win05] (see also [Win08]).

The problem of eigenvalue accumulation in gaps of the essential spectrum of self-adjoint block operator matrices was investigated also in [AMS98]; some of the results therein follow from more general considerations of V.A. Derkach and M.M. Malamud on self-adjoint extensions of symmetric operators with spectral gaps (see [DM91]).

2.7. *Motivation by applications.* Many physical systems are described by systems of partial or ordinary differential equations or linearizations thereof. The corresponding spectral problems tend to be challenging; profound physical intuition and advanced techniques from the analysis of differential equations are common ways to address them. The theory of block operator matrices opens up a new line of attack.

Magnetohydrodynamics and fluid mechanics. The study of block operator matrices occurring in magnetohydrodynamics and fluid mechanics was initiated by several papers of G. Grubb, G. Geymonat (see [GG74], [GG77], [GG79]) and of J. Descloux, G. Geymonat (see [DG79], [DG80]). Using pseudo-differential calculus, they developed methods to determine the essential spectrum of so-called Douglis-Nirenberg elliptic systems. It seems that these authors were the first to observe that, unlike regular differential operators, regular *matrix* differential operators may have non-empty essential spectrum. Applications of the results were given *e.g.* to the linearized Navier-Stokes operator from fluid mechanics.

In magnetohydrodynamics, knowledge of the essential spectrum, especially of its so-called Alfvén range part, may be used to heat up the plasma inexpensively (see [Lif89, Section 7.9]). A variety of singular and regular spectral problems from magnetohydrodynamics were studied by T. Kako (see [Kak84], [Kak85], [Kak87], [Kak88a], [Kak88b], [KD91], [Kak94]) and by G. Raĭkov by means of variational techniques (see [Raĭ85], [Raĭ86a], [Raĭ86b], [Raĭ87], [Raĭ88], [Raĭ90], [Raĭ91], [Raĭ97]); G. Raĭkov also investigated the discrete eigenvalues and proved asymptotic formulae for them.

The methods developed in [ALMS94] for upper dominant block operator matrices were employed to determine the essential spectrum for several problems in magnetohydrodynamics, fluid mechanics, and astrophysics: regular two-dimensional problems from magnetohydrodynamics were studied by M. Faierman, A. Yu. Konstantinov, H. Langer, R. Mennicken, and M. Möller (see [FLMM94], [FMM95], [LM96], [Kon02]); singular problems were considered in [FMM99], [FMM00], and by S.N. Naboko *et al.* in [HMN99], [MNT02], [KN02], [KN03]. In the more involved singular case considered in [DG79], Descloux and Geymonat's conjecture on the essential spectrum is still waiting to be proved; by now, it was only confirmed for a simpler model problem in [FMM04]. The result for the essential spectrum of the linearized Navier-Stokes operator obtained in [GG79] was reproved in [FFMM00]. The methods of [ALMS94] apply also to problems from the stability analysis of stellar oscillations, *e.g.* to a model arising in the Cowling approximation, which was considered by H. Beyer directly by Sturm-Liouville techniques (see [Bey00]), or to differentially rotating stars (see [FLMM99]).

Quantum mechanics. Another field where block operator matrices naturally occur is quantum mechanics, the most prominent example being the Dirac operator (see [Tha92]):

$$\mathbf{H}_{\mathbf{\Phi}} := \begin{pmatrix} (mc^2 + e\Phi)\, I & c\,\vec{\sigma}\cdot\big(-\mathrm{i}\hbar\nabla - \frac{e}{c}\vec{A}\big) \\ c\,\vec{\sigma}\cdot\big(-\mathrm{i}\hbar\nabla - \frac{e}{c}\vec{A}\big) & (-mc^2 + e\Phi)\, I \end{pmatrix}.$$

Here the problem of block diagonalization amounts to a decoupling of positive- and negative-energy states (electrons and positrons). For the free Dirac operator in $\mathbb{R}^3$ (*i.e.* $\vec{A} = 0$, $\Phi = 0$), an exact unitary transformation into block diagonal form was found by L.L. Foldy and S.A. Wouthuysen in 1950 (see [FW50] and the review articles [DV70], [CM95]). For non-vanishing electric potential Φ, the Foldy-Wouthuysen method suggests a sequence of unitary transformations eliminating the lowest order off-diagonal term in $1/c$ in each step. However, the resulting series expansion in $1/c$ for the successively transformed Dirac operator is ill-behaved (see [GNP89], [RW04a]). In 1974, another iterative diagonalization scheme was proposed by M. Douglas and N.M. Kroll who used an expansion in terms of orders in the potential Φ rather than in $1/c$ (see [DK74b]); its efficiency for quantum chemical implementations was first noticed by B.A. Hess in [Heß86] and investigated in detail in a recent series of papers by M. Reiher and A. Wolf (see [RW04a], [RW04b], [WR06a], [WR06b]). Unlike the Foldy-Wouthuysen approximation, this method yields a well-behaved series with a limiting block diagonal operator having the same spectrum as the original Dirac operator (see [RW04a]). For Coulomb potentials up to atomic number 51, convergence of this series in norm resolvent sense was proved recently by H. Siedentop and E. Stockmeyer in [SS06]; numerically, Reiher and Wolf have tested convergence for the ground state energies of one-electron ions over the whole periodic table in [RW04b]. In 1988, B. Thaller generalized the Foldy-Wouthuysen transformation for abstract supersymmetric Dirac operators, which include the Dirac operator with vector potential $\vec{A}$, but not with electric potential (see [Tha88], [Tha91], and the monograph [Tha92]).

The idea of achieving a block diagonalization for various Hamiltonians in quantum mechanics by means of invariant graph subspaces may be traced back to 1954 to S. Ôkubo (see [Ôku54] and also the survey paper [SS01] for an introduction to Ôkubo's method). His approach includes free Dirac operators as a special case, but it also has applications in elementary particle physics (see e.g. [GM81], [KS93], and [SS01]); there the two components of the Hilbert space correspond, for example, to a space of external (*e.g.* hadronic) degrees of freedom and a space of internal (*e.g.* quark) degrees of freedom (see [DHM76], [Mot95]). In a mathematically rigorous way, Ôkubo's idea was used by V.A. Malyshev and R.A. Minlos for the construction of invariant subspaces for a class of self-adjoint block operator matrices in statistical physics, so-called clustering operators (see [MM79], [MM82]); their assumptions amount to separated spectra of the diagonal entries and sufficiently small off-diagonal terms.

The first exact Foldy-Wouthuysen transformation for Dirac operators with electric potentials was proved in 2001 in [LT01]. This result is a consequence of the abstract theorem on block diagonalization described above; at the moment, it is restricted to bounded potentials, but we conjecture that the method continues to be applicable as long as the potential does not mix the positive and negative part of the spectrum too badly. This conjecture is supported by the recent paper [Cor04] of H.O. Cordes. Independently of [LT01], in 2004, he proved the existence of an exact Foldy-

Wouthuysen transformation for Dirac operators with unbounded smooth electromagnetic fields; in [Cor04], the latter is ensured by the assumption that all derivatives of order k decay like $(1+|x|)^{-k-1}$ (see also [Nen76]). Such fields do not alter the essential spectrum, but may produce a finite mixing of eigenvalues emerging from the left and the right end-point of the essential spectrum. Cordes' method relies on his earlier works on pseudo-differential calculus and its application to Dirac operators, begun already in the 1980ies (see [Cor79], [Cor83b], [Cor83a], [Cor95], [Cor00], [Cor01]).

Variational principles for Dirac operators of the form (XI) were first suggested in the 1980ies by J.D. Talman (see [Tal86]) and by S.N. Datta and G. Deviah (see [DD88]). The main difficulty, compared to Schrödinger operators, is that the Dirac operator is not semi-bounded and eigenvalues appear in a gap of the essential spectrum. Remarkably, Talman already used the Schur complement implicitly for his heuristic arguments (see [Tal86, (4)]). The first rigorous min-max result for Dirac operators with Coulomb potential was considered in 1997 by M. Esteban *et al.* (see [ES97] and also [DES00b]). Later the abstract min-max principles by H. Siedentop *et al.* in [GS99], [GLS99] and by M. Esteban *et al.* in [DES06] were applied to Dirac operators with Coulomb-like potentials. The variational principles in [KLT04] were applied to Dirac operators with bounded potentials and to Dirac operators on Riemannian spin manifolds with warped product metric. Recently, M. Winklmeier used block operator methods to establish variational principles for eigenvalues of the angular part of the Dirac operator in curved spacetime (see the PhD thesis [Win05] and [Win08]). The resulting eigenvalue estimates are the first analytic results, preceded only by numerical calculations of K.G. Suffern, E.D. Fackerell, C.M. Cosgrove and of S.K. Chakrabarti in the 1980ies (see [SFC83], [Cha84]).

Other areas. There are many other applications that we do not touch in this book. They include linear evolution problems that can be written formally as first order Cauchy problems

$$\dot{u}(t) = \mathcal{A}\,u(t), \quad u(0) = u_0, \tag{XII}$$

with a block operator matrix $\mathcal{A}$ in a product space $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ and a function $u : \mathbb{R}^+ \to \mathcal{H}$ (see *e.g.* [Nag89]). Often these problems arise as linearizations of second order Cauchy problems; examples include the wave equation (see *e.g.* [Gol85]) and the Klein-Gordon equation from quantum mechanics (see *e.g.* [Lun73]) for which

$$\mathcal{A} = \begin{pmatrix} 0 & I \\ B & D \end{pmatrix}, \quad B = c^2\big(-\mathrm{i}\hbar\nabla - \tfrac{e}{c}\vec{A}\big)^2 + m^2c^4 - e^2\Phi^2, \quad D = 2e\Phi;$$

other linearizations lead to different coefficients $\mathcal{A}$ in (XII) (see *e.g.* [Ves91]). Here the crucial problem is to show that the operator $\mathcal{A}$ is the generator of a semi-group. For wave equations, the operator matrix approach initiated by Nagel was applied recently by D. Mugnolo (see [Mug06a], [Mug06c], [Mug06b]). Other classes of block operator matrices, such as one-sided coupled operator matrices or polynomial operator matrices, were studied intensively by K.-J. Engel *et al.* in numerous papers (see [Eng89a], [Eng89b], [EN90], [Eng90], [Eng91a], [Eng91b], [Eng93], [Eng96], [Eng98], [Eng99], [BE04]); applications range from epidemiology to elasticity theory and reaction-diffusion systems. Operator matrices with non-diagonal domain in the sense of Nagel and applications to differential equations with delay were considered in [BBD$^+$05]. For abstract operator matrices of Klein-Gordon type, the block operator matrix approach presented in this book and indefinite inner product methods were applied by various authors, including K. Veselič, B. Najman, and P. Jonas (see [Ves91], [Naj83], [Jon93], [Jon00], [LNT06], [LNT08]); variational principles for the Klein-Gordon equation, even in the presence of complex eigenvalues, were derived in [LT06]. Along similar lines, Cauchy problems arising in hydrodynamics describing small oscillations of a fluid in a partially filled container were studied in [AHKM03], [KMPT04]).

3. Outline of contents. The three chapters of this book are organised in the following way:

Chapter 1 deals with *bounded block operator matrices.* In Section 1.1 the quadratic numerical range $W^2(\mathcal{A})$ of a bounded block operator matrix $\mathcal{A}$ is introduced and some elementary properties are proved. In Section 1.2 we study the quadratic numerical range of several special classes of block operator matrices. Section 1.3 contains the most important property of the quadratic numerical range, the spectral inclusion. In Section 1.4 we establish an estimate of the resolvent in terms of the quadratic numerical range. In Section 1.5 the corners of the quadratic numerical range are studied. In Section 1.6 we analyse the Schur complements of the block operator matrix and the relation of their numerical ranges to the quadratic numerical range. We show that if the closure of the quadratic numerical range consists of two disjoint components, then the Schur complements admit factorizations. This result is used in Section 1.7 to prove one of the main results of this chapter, the theorem on block diagonalization (see Theorem 1.7.1). A consequence of this theorem is the existence of bounded solutions of the corresponding Riccati equations. In Section 1.8 so-called spectral supporting subspaces $\mathcal{H}_1^{\Delta}$ of self-adjoint block operator matrices are studied; they are characterized by the property that the spectral subspace $\mathcal{L}_\Delta(\mathcal{A})$ is the graph of a linear operator defined on $\mathcal{H}_1^\Delta$. In Section 1.9 variational principles based on the functionals $\lambda_\pm$ defining the quadratic numerical range are established; they apply to eigenvalues in gaps of the essential spectrum of bounded linear operators with real quadratic numerical range. Section 1.10 deals with so-called $\mathcal{J}$-self-adjoint block operator matrices (*i.e.* $A = A^*$, $D = D^*$, $C = -B^*$). The functionals $\lambda_\pm$ are used to classify the spectral points of $\mathcal{A}$ and to identify an interval $[\nu, \mu]$ outside of which the spectrum of $\mathcal{A}$ is of definite type; the latter implies that $\mathcal{A}$ possesses a local spectral function on $\mathbb{R} \setminus [\nu, \mu]$. In addition, the results on spectral supporting subspaces are extended to the case $C = -B^*$. In Section 1.11 the quadratic numerical range is generalized to $n \times n$ block operator matrices; all results of Sections 1.1 to 1.4 carry over to this so-called block numerical range. Moreover, we prove that a refinement of the decomposition of the Hilbert space leads to an inclusion of the corresponding block numerical ranges. Together with the spectral inclusion property, this yields a successively improved localization of the spectrum. In Section 1.12 we show that the block numerical range of the companion matrix of an operator polynomial contains its numerical range. Finally, in Section 1.13, Gershgorin's theorem for $n \times n$ block operator matrices is presented and compared with the spectral inclusion by means of the block numerical range.

Chapter 2 focuses on *unbounded block operator matrices*. Section 2.1 contains some basic properties of closed linear operators; in particular, the notions of relative boundedness and relative compactness are presented. In Section 2.2 we establish criteria for closedness and closability of block operator matrices. To this end, three classes of block operator matrices are distinguished, depending on the positions of the "strongest" operators in each column (*i.e.* the entries with smallest domains): diagonally dominant, off-diagonally dominant, and upper dominant block operator matrices. In Section 2.3 and Section 2.4 we investigate the spectrum and the essential spectrum, respectively, of block operator matrices by means of Schur complements and quadratic complements. In Section 2.5 we introduce the quadratic numerical range for unbounded block operator matrices and prove the spectral inclusion theorem for diagonally dominant and off-diagonally dominant block operator matrices. In Section 2.6 some properties of symmetric and $\mathcal{J}$-symmetric block operator matrices are studied; in particular, a classification of eigenvalues in terms of the quadratic numerical range is given. Section 2.7 contains some of the key results of Chapter 2 (see Theorems 2.7.7, 2.7.21, and 2.7.24); they concern the existence of invariant graph subspaces and of solutions of Riccati equations for dichotomous block operator matrices. These theorems apply to essentially self-adjoint block operator matrices and, in the non-self-adjoint case, to diagonally dominant and off-diagonally dominant block operator matrices with certain relative boundedness assumptions between the entries (ensuring exponential dichotomy). In Section 2.8 we exploit the theorems of Section 2.7 further to obtain results on block diagonalization as well as on half range completeness and half range basis properties of eigenfunctions and associated functions. In Section 2.9 we reconsider the existence problem for solutions of Riccati equations by means of fixed point methods, which also provide uniqueness of the solution and a convergent iteration scheme. In Section 2.10 we derive variational principles for eigenvalues in gaps of the essential spectrum for self-adjoint and $\mathcal{J}$-self-adjoint block operator matrices. In Section 2.11 we use the variational principles to derive two-sided eigenvalue estimates. Many of the results of Chapter 2 are illustrated by matrix differential operators.

Chapter 3 concentrates on a selection of *applications in mathematical physics* to which the results of Chapter 2 are applied. In Section 3.1 we consider a stability problem for small oscillations of a magnetized gravitating plane equilibrium layer of a hot compressible ideal plasma, which is described by an upper dominant essentially self-adjoint block operator matrix. In particular, we derive a formula for the essential spectrum, which consists of two intervals called slow magnetosonic and Alfvén spectrum, and we give estimates for the eigenvalues if there is a gap between these two intervals. In Section 3.2 we study the stability of the two-dimensional Ekman boundary layer flow which is produced in a rotating tank with small inflow; the corresponding block operator matrix is diagonally dominant and non-self-adjoint. We calculate the essential spectrum which consists of a curve in the complex plane confined to a semi-strip. Section 3.3 is dedicated to Dirac operators in $\mathbb{R}^3$ and in curved spacetime; in both cases, the corresponding block operator matrices are self-adjoint and off-diagonally dominant. We show that, for the free Dirac operator in $\mathbb{R}^3$, our diagonalization theorem reproduces the well-known Foldy-Wouthuysen transformation. In contrast to the latter, our method also applies to Dirac operators in $\mathbb{R}^3$ with electric potential under some boundedness condition. Finally, we study the angular part of the Dirac equation in the spacetime generated by an electrically charged rotating massive black hole. We use the variational principles of Section 2.10 to derive upper and lower bounds for the eigenvalues and compare them with numerical values from the physics literature.