STRUCTURED PERTURBATION THEORY FOR EIGENVALUES OF SYMPLECTIC MATRICES

Fredy Ernesto Sosa Nuñez

en cumplimiento parcial de los requisitos para el grado de
Doctor en

Ingeniería Matemática

Universidad Carlos III de Madrid

Director:
Julio Moro Carreño

Leganés, 10 de mayo de 2019

Este trabajo ha sido desarrollado en el Departamento de Matemáticas de la Universidad Carlos III de Madrid (UC3M) bajo la dirección del profesor Julio Moro Carreño. Se contó con una beca de la UC3M de ayuda al estudio de máster y posteriormente con un contrato predoctoral de la UC3M. Adicionalmente se recibió ayuda parcial del proyecto de investigación: " Matemáticas e Información Cuántica: de las Álgebras de Operadores al Muestreo Cuántico " (Ministerio de Economía y Competitividad de España, Número de proyecto: MTM2014-54692-P).

a mi familia

Agradecimientos

Son muchas las personas que se han visto involucradas tanto en mi formación como profesional, como, en particular, en el desarrollo de este trabajo; llegue a todas ellas mi agradecimiento sincero.

Quiero agradecer en especial:

A mi director de tesis Julio Moro, por su apoyo incondicional desde el primer momento, su paciencia y dedicación para que este trabajo de investigación saliera adelante, pero sobre todo, muchas gracias por ser más amigo que tutor.

A Chistian Mehl por involucrarse activamente en nuestro trabajo con matrices estructuradas, permitiendo el desarrollo de un trabajo conjunto que ha dado lugar a muchos de los resultados que aquí se presentan.

A Volker Mehrmann por recibirme en su grupo de investigación y brindarme todo lo necesario para introducirme en el maravilloso mundo de las matrices estructuradas y que ha supuesto un punto de inflexión en el desarrollo de este trabajo de investigación.

Al departamento de Matemáticas todo el apoyo recibido durante estos años, sin el cuál, nada de esto habría sido posible.

A Alberto Calvo, el imprescindible.

Contenidos Publicados y Presentados

- *First order eigenvalue expansions for matrices under multiplicative perturbation* . Bajo este título se ha presentado el documento correspondiente al Trabajo de Fin de Máster en el programa de Máster en Ingeniería Matemática de la Universidad Carlos III de Madrid. La defensa ha tenido lugar en julio de 2014.

  El material de esta fuente está parcialmente incluido en los capítulos 1, 3 y en los apéndices A y B. Dicho contenido no está señalado por medios tipográficos ni referencias.

- SOSA, F. AND MORO, J., *First order asymptotic expansions for eigenvalues of multiplicatively perturbed matrices*, Publicado electrónicamente en SIAM Journal on Matrix Analysis and Applications, https://doi.org/10.1137/15M1050057 , October 2016, Volumen 37 No. 4, 1478-1504.

  El material de esta fuente está incluido en el Capítulo 3 y en los apéndices A y B. Este contenido no está señalado por medios tipográficos ni referencias.

- SOSA, F., MORO, J., MEHL, C., *First order structure-preserving perturbation theory for eigenvalues of symplectic matrices*, presentado para su publicación en SIAM Journal on Matrix Analysis and Applications.
  URL: https://www3.math.tu-berlin.de/preprints/files/Preprint-04-2018.pdf

  El material de esta fuente está parcialmente incluido en los capítulos 1 y 4, y no está señalado por medios tipográficos ni referencias.

Otros méritos de investigación

- **Contribuciones a Conferencias y Seminarios**

  Los resultados originales contenidos en este documento han sido presentados en las siguientes presentaciones:

  On first order expansions for multiplicative perturbation of eigenvalue. Presented as a contributed talk in the Joint ALAMA-GAMM/ANLA'2014 MEETING, July 14–16, 2014, Barcelona, Spain.

  First order asymptotic expansions for eigenvalues of multiplicatively perturbed matrices. Presented in Master/Phd seminar, Institut für Mathematik, June 11, 2015, Belin, Germany.

  On First Order Expansions for Multiplicative Perturbation of Eigenvalues. Presented as a contributed talk in the SIAM Conference on Applied Linear Algebra, October 26–30, 2015, Atlanta, USA.

  Structured multiplicative perturbation of eigenvalues of J-symplectic matrices. Presented as a contributed talk in the ALAMA 2016 Meeting, June 20–22, 2016, Leon, Spain.

  Structured multiplicative perturbation of eigenvalues of J-symplectic matrices. Presented as invited talk in the Minisymposium: " Perturbation theory and distance problems associated with eigenvalues ", organized by Christian Mehl and Shreemayee Bora in the 20th Conference of the International Linear Algebra Society (ILAS). July 11–15, 2016, Leuven, Belgium.

- **Estancias de Investigación**

  Research visit at Rice University, Houston, USA., October 30 – November 12, 2015.

  Research visit at Institut für Mathematik, TU-Berlin, Berlin, Germany., May 28 – June 30, 2015.

Abstract

The problem of computing eigenvalues, eigenvectors, and invariant subspaces of symplectic matrices plays a major role in many applications, in particular in control theory when the focus is on discrete systems. If standard numerical methods for the solution of the symplectic eigenproblem are applied that do not take into account the special symmetry structure of the problem, then not only the existing symmetry in the spectrum of symplectic matrices may be lost in finite precision arithmetic, but more importantly other relevant intrinsic features or invariants may be ignored although they have a major influence in the corresponding computed eigenvalues. The importance of structure-preservation has been acknowledged in the Numerical Linear Algebra community since several decades, and consequently many algorithms have been developed for the symplectic eigenvalue problem that preserve the given structure at each iteration step. The error analysis for such algorithms requires a corresponding *stucture-preserving* perturbation theory. This is the general framework in which this dissertation can be placed.

In this work, a first order perturbation theory for eigenvalues of real or complex J-symplectic matrices under structure-preserving perturbations is developed. Since the class of symplectic matrices has an underlying multiplicative structure, Lidskii's classical formulas for small additive perturbations of the form $\widehat{A} = A + \varepsilon B$ cannot be applied directly, so a new *multiplicative* perturbation theory is first developed: given an arbitrary square matrix A, we obtain the leading terms of the asymptotic expansions in the small, real parameter ε of multiplicative perturbations $\widehat{A}(\varepsilon) = (I + \varepsilon\, B + \cdots)A(I + \varepsilon\, C + \cdots)$ of A for arbitrary matrices B and C. The analysis is separated in two complementary cases, depending on whether the unperturbed eigenvalue is zero or not. It is shown that in either case the leading exponents are obtained from the partial multiplicities of the eigenvalue of interest, and the leading coefficients generically involve only appropriately normalized left and right eigenvectors of A associated with that eigenvalue, with no need of generalized eigenvectors. It should be noted that, although initially motivated by the needs for the symplectic case, this multiplicative (unstructured) perturbation theory is of independent interest and stands on its own.

After showing that any small structured perturbation $\widehat{S}$ of a symplectic matrix S can be written as $\widehat{S} = \widehat{S}(\varepsilon) = (I + \varepsilon B + \cdots)\, S$ with Hamiltonian first-order coefficient B, we apply the previously obtained Lidskii-like formulas for multiplicative perturbations to the symplectic case by exploiting the particular connections that symplectic structure induces in the Jordan form between normalized left and right eigenvectors. Special attention is given to eigenvalues on the unit circle, particularly to the exceptional eigenvalues ± 1, whose behavior under structure-preserving perturbations is known to differ significantly from the behavior under arbitrary ones. Also, several numerical examples are generated

in order to illustrate the asymptotic expansions and confirm our findings.

Although the approach described above via multiplicative expansions works in most situations, there is a very specific one, the one we call the *nongeneric case*, which requires a separate, completely different analysis. It corresponds to the case in which, in the absence of structure, the rank of the perturbation would break an odd number of odd-sized Jordan blocks corresponding to the eigenvalue either 1 or -1. Since this is not allowed by symplecticity, one among that odd number of Jordan blocks does not break, but increases its size by one becoming an even-sized block. This very special behavior lies outside of the theory developed for what we might call the generic cases, and requires a completely different perturbation analysis, based on Newton diagram techniques like the one performed to obtain the multiplicative expansions. The main difference with the previous expansions is that in this nongeneric case the leading coefficients depend not only on eigenvectors, but also on first generalized Jordan vectors.

Resumen

El problema de calcular autovalores, autovectores y subespacios invariantes de matrices simplécticas juega un papel crucial en muchas aplicaciones, en particular en la Teoría de Control cuando ésta se centra en sistemas discretos. Si para resolver el problema simpléctico de autovalores se emplean métodos numéricos estándar que no tienen en cuenta la simetría especial del problema, entonces no solo se perderá en aritmética finita la simetría natural del espectro de las matrices simplécticas, sino que, aún más importante, podemos estar ignorando otras características o invariantes intrínsecas que tienen una influencia crucial en los correspondientes autovalores calulados. La importancia de preservar la estructura ha sido reconocida por la comunidad del Álgebra Lineal Numérica desde hace varias décadas y, en consecuencia, se han desarrollado diversos algoritmos para el problema simpléctico de autovalores que mantienen la estructura simpléctica en cada paso del proceso iterativo. El análisis de errores de tales algoritmos demanda una teoría de perturbación asociada *que también preserve la estructura*. Este es el marco general en el que se puede inscribir esta tesis doctoral.

En este trabajo se desarrolla una teoría de perturbación de autovalores de matrices J-simplécticas frente a perturbaciones que preservan la simplecticidad de la matriz. Dado que la clase de matrices simplécticas tiene una estructura multiplicativa subyacente, las fórmulas clásicas de Lidskii para perturbaciones aditivas pequeñas de la forma $\widehat{A} = A + \varepsilon B$ no se pueden aplicar de manera directa, de modo que desarrollamos una nueva teoría de perturbación *multiplicativa*: dada cualquier matriz cuadrada A, obtenemos el término director del desarrollo asintótico en el parámetro real (y pequeño) ε de autovalores de perturbaciones multiplicativas $\widehat{A}(\varepsilon) = (I + \varepsilon\, B + \cdots)A(I + \varepsilon\, C + \cdots)$ de A para matrices arbitrarias B y C. El análisis se separa en dos casos complementarios, dependiendo de que el autovalor a perturbar sea nulo o no. Se demuestra que en ambos casos los exponentes directores se obtienen a partir de las multiplicidades parciales del autvalor bajo estudio, y que los coeficientes directores solo involucran genéricamente autovectores derechos e izquierdos adecuadamente normalizados, sin necesidad de autovalor generalizado alguno. Debe señalarse que, aunque inicialmente motivados por la necesidad para el caso simpléctico, esta teoría (no estructurada) de perturbación multiplicativa reviste interés *per se* independientemente de su aplicación al caso simpléctico.

Tras mostrar que cualquier perturbación estructurada peque na $\widehat{S}$ de una matriz simpléctica S puede escribirse como $\widehat{S} = \widehat{S}(\varepsilon) = (I + \varepsilon B + \cdots)\, S$ con coeficiente de primer orden B Hamiltoniano, aplicamos las fórmulas tipo Lidskii obtenidas para perturbaciones multiplicativas al caso simpléctico, explotando la particular conexión que la estructura simpléctica induce entre los autovectores derechos e izquierdos normalizados por la forma de Jordan. Especial atención se le dedica a los autovalores sobre el círculo unidad, particularmente a los autovalores excepcionales ± 1, cuyo comportamiento

frente a perturbaciones estructuradas es sabido que difiere muy significativamente del comportamiento frente a perturbaciones arbitrarias. Además, presentamos varios ejemplos numéricos que ilustran (y confirman) los desarrollos asintóticos obtenidos.

Aunque el enfoque que acabamos de describir via desarrollos multiplicativos funciona en la mayor parte de las situaciones, hay una muy específica, la que llamamos el *caso no-genérico*, que requiere de un análisis por separado, completamete distinto del anterior. Corresponde al caso en que, en ausencia de estructura, el rango de la perturbación rompería un número impar de bloques de Jordan de tamaño impar asociados a uno de los autovalores 1 ó -1. Como esto es incompatible con la simplecticidad, uno de entre los bloques de tamaño impar no se rompe, sino que incrementa en uno su dimensión, convirtiéndose en un bloque de Jordn de tamaño par. Este comportamiento tan especial no está explicado por la teoría de lo que podríamos llamar los casos 'genéricos' , y requiere de un análisis de perturbación completamente distinto, basado en técnicas del Diagrama de Newton, como el llevado a cabo para obtener los desarrollos multiplicativos. La diferencia principal con los desarrollos anteriores es que en el caso no genérico los coeficientes directores dependen no solo de autovectores, sino también de vectores primeros generalizados de Jordan.

Contents

Acknowledgements **v**

Abstract **xi**

Resumen **xiii**

1 Introduction **1**

2 Preliminaries **9**

2.1 Symplectic and Hamiltonian matrices 9

2.1.1 Definitions 9

2.1.2 Canonical forms for symplectic matrices 10

2.1.3 Basic properties of symplectic matrices 13

2.2 First order additive eigenvalue perturbation theory 14

2.2.1 First order eigenvalue perturbation theory: a short historical overview 14

2.2.2 Eigenvalue Perturbation and the Newton Diagram. 15

2.2.3 First order asymptotic expansions: Lidskii's Theorem 19

2.2.4 Proof of Lidskii's Theorem via the Newton Diagram 22

3 First order expansions for eigenvalues of multiplicative perturbations **29**

3.1 Multiplicative perturbation of matrix eigenvalues 29

3.1.1 The case $\lambda \neq 0$ 30

3.1.2 The case $\lambda = 0$ 31

3.1.3 The matrix $\widehat{H}(\varepsilon)$ 33

3.1.4 The principal minors of $\widehat{H}(\varepsilon)$ 36

3.1.5 Proof of Theorem 3.4 38

3.1.6 Proof of Theorem 3.3 39

3.2 Asymptotic singular value expansions for multiplicative perturbations . . 46

4 Structured perturbation of eigenvalues of symplectic matrices **49**

4.1 Preliminaries 49

4.1.1 Small multiplicative symplectic perturbations 49

4.2 Jordan chains for symplectic matrices 52

4.2.1 Extracting Jordan bases from bases associated with symplectic canonical forms 53
4.2.2 Relationships between left and right Jordan chains 56
4.3 Asymptotic expansions for structured symplectic perturbations 61
4.3.1 Asymptotic expansions: the case $\lambda \notin \{-1\,,\,1\}$ 62
4.3.2 Asymptotic expansions: the case $S,\,\widehat{S} \in \mathbb{R}^{2n\times 2n}$ with $|\lambda| = 1$ . . 63
4.3.3 Asymptotic expansions: the case $\lambda \in \{-1\,,\,1\}$. 66
4.4 Rank-one structure-preserving perturbations 68
4.4.1 Observations on coalescing eigenvalues 70
4.5 Numerical examples 71
4.5.1 An auxiliary perturbation matrix 72
4.5.2 Numerical examples 74

5 Structured perturbation of eigenvalues of symplectic matrices: The non-generic case 81
5.1 The nongeneric case: $\lambda \in \{\,-1\,,\,1\}$, n_j odd, ρ odd. 81
5.1.1 The main result 82
5.1.2 An illustrative example 85
5.1.3 Proof of Theorem 5.1 90
5.2 Rank-one structure preserving perturbations: the case when $\mathbf{n_1}$ is odd . . 95

6 Conclusions, publications and open problems 99
6.1 Conclusions and original contributions 99
6.2 Future Work 100

Appendices

Appendix A Proof of (3.21) **105**

Appendix B Proof of Lemma 3.9 109

Bibliography 113

Chapter 1

Introduction

The central question of perturbation theory is: how does a function change when its argument is subject to perturbation, i.e., when the variables the function depends on are slightly changed. The function may be almost anything: the modes of a vibrating system, the solution of a differential equation, or the states of an electron. In this work we will focus on first order perturbation of the eigenvalues of a matrix, a research area whose foundations were laid by Lord Rayleigh [59] at the end of the 19th century. One of his calculations aimed at determining both the eigenfrequencies and eigenmodes of an oscillatory string with constant elasticity and whose density is a small deviation of a constant value. This particular problem illustrates perfectly the typical setting of eigenvalue perturbation theory: the matrix or operator under study is assumed to be a slight deviation from some close, simpler matrix or operator for which the spectral problem is completely (and, in most cases, easily) solved. Consequently, the given operator $\widehat{A}$ is replaced by a neighboring one A whose eigenvalues and eigenvectors are known. Then, the influence on the spectral objects (eigenvalues, eigenvectors, invariant subspaces,...) of the difference between $\widehat{A}$ and A is analyzed by way of an appropriate perturbation theory.

Traditionally, the difference between the perturbed and unperturbed operator is modelled as an *additive* perturbation $E = \widehat{A} - A$ or, if the perturbation is sufficiently small, as

$$\widehat{A} = \widehat{A}(\varepsilon) = A + \varepsilon\, B, \tag{1.1}$$

where B is a matrix or operator, and $\varepsilon > 0$ is a small real number. Obviously, the eigenvalues and eigenvectors of $\widehat{A}$ depend on ε and *are assumed* to converge to the corresponding eigenvalues and eigenvectors of A as ε goes to zero. In a first stage of the analysis the leading terms of the ε-expansions of the spectral objects of $A(\varepsilon)$ are determined (hence the name of *first order* perturbation theory). Two crucial quantities, the leading exponent and the leading coefficient, fully describe the behavior of perturbed spectral objects for small values of ε. In a second stage, the convergence of these expansions is justified up to $\varepsilon = 1$, or some other finite threshold. Here we will focus only on the first stage; the convergence of the eigenvalue and eigenvector expansions follows from the classical theory of analytic functions (see, for instance, [4, 23]).

The framework in which this first order perturbation theory will be analyzed here is that of *structured* eigenvalue perturbation: it is well known that, for certain kinds of matrices, whenever *both* the perturbed and unperturbed matrix belong to that same family of matrices, the behavior of perturbed eigenvalues may be very different from the one when the perturbed matrix is an arbitrary one. The simplest example one can think of is, probably, that of symmetric matrices: if both the unperturbed matrix A and the perturbed one $\widehat{A}(\varepsilon)$ are symmetric, then every perturbed eigenvalue $\lambda(\varepsilon)$ is constrained to lie on the real line, something which is not true for arbitrary $\widehat{A}(\varepsilon)$.

This special behavior is to be expected especially for classes of matrices, like symmetric ones, which impose strong constraints on the location of their eigenvalues. Furthermore, the interest in studying such a special behavior grows with the relevance of the corresponding class of matrices in their applications to other scientific contexts. Several different classes of matrices satisfying both conditions about the existence of spectral constraints and their importance in applications can be found among those satisfying symmetry conditions with respect to indefinite inner products. Symplectic and Hamiltonian matrices, which we shall be using extensively throughout this document, are two of those classes (see §2.1.1 below for the appropriate definitions).

The symplectic eigenvalue problem, i.e., the problem of computing eigenvalues, eigenvectors, and invariant subspaces of symplectic matrices is an important one in linear control theory for discrete-time systems. In particular, the symplectic eigenvalue problem plays a major role in the solution of the linear-quadratic optimal control problem or the solution of discrete-time algebraic Riccati equations, see [14] and references therein. Symplectic matrices play also a crucial role in Quantum Mechanics and Quantum Optics, see [19, 1].

From the computational point of view, several algorithms for the solution of the symplectic eigenvalue problem have been proposed, see, e.g. [6, 5, 8, 32, 39]. All these approaches have in common that they focus on *structure preservation*, i.e., unlike for general-purpose algorithms, like QR or divide-and-conquer, the symplecticity of the underlying matrix is preserved at each and every step of the algorithm. The reason for this is that, for instance, the eigenvalues of real symplectic matrices are symmetric with respect to the unit circle, a symmetry that is sometimes referred to as *symplectic eigenvalue symmetry*. If the eigenvalue algorithm we employ ignores the existing structure, then not only may round-off errors cause the loss of this spectral symmetry, but also other important aspects of the problem may become invisible, since additional invariants, that only exist under structure preserving transformations, may be the origin of unwanted effects if the structure is lost. In the case of J-symplectic matrices, one such invariant having a significant impact on the behavior of the spectrum under structure-preserving perturbations is the so-called *sign characteristic* of unimodular eigenvalues (see (2.9) and (2.16) below for a formal definition).

This preservation of structure in the eigenvalue algorithms is actually one of the main motivations for the *structured spectral perturbation theory* we perform here: if one thinks of ε as the unit roundoff of the computer's finite arithmetic, then one can make use of the theory of backward error analysis and think of the eigenvalues of the perturbed matrix $\widehat{A}(\varepsilon)$ as those computed by the structure-preserving numerical algorithm, round-off errors included. Thus, by understanding the perturbation error we are at the same time understanding better the computational error made by those algorithms.

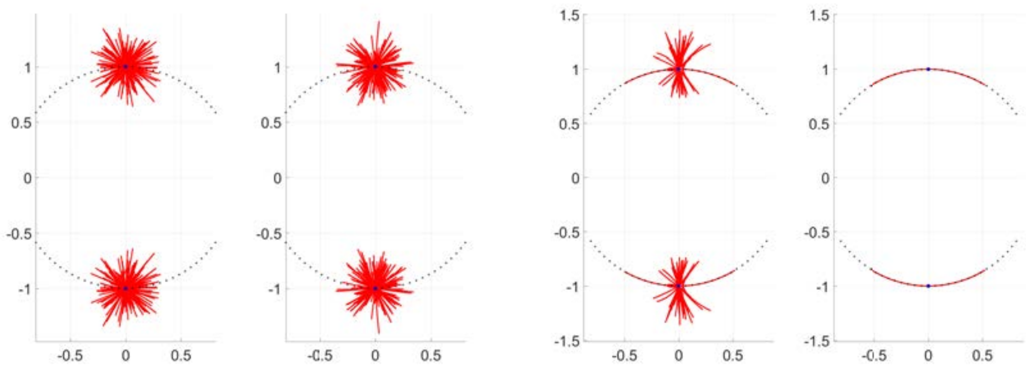
To be more specific, the situation we analyze in this dissertation is the following one: given a (possibly multiple) eigenvalue λ of a J-symplectic matrix S, we consider another J-symplectic matrix $\widehat{S}$ close to S, and develop asymptotic expansions of the eigenvalues of $\widehat{S}$ close to λ by interpreting $\widehat{S} = \widehat{S}(\varepsilon)$ as a particular value of an analytic J-symplectic matrix function $\widehat{S}(\cdot)$ depending on a real parameter ε. The fact that both S and $\widehat{S}$ are J-symplectic is crucial: it is well known (see, for instance, [38, 3]) that the eigenvalues of J-symplectic matrices behave quite differently under structure-preserving perturbations in comparison with arbitrary perturbations. Our goal is to completely describe the local asymptotic behavior of eigenvalues of symplectic matrices under structure-preserving perturbations.

To illustrate the kind of differences we may expect in the perturbation behavior, as well as the important role that the aforementioned sign characteristics play in the perturbation analysis, let us consider the following example. First, recall that stability of discrete systems depends on the location of eigenvalues of the underlying matrix with respect to the unit circle. A real symplectic matrix can be considered to be stable if all its eigenvalues are on the unit circle and semisimple (i.e., their algebraic and geometric multiplicities coincide). If we consider the two symplectic matrices

$$S_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \tag{1.2}$$

then both have the eigenvalues $\pm i$ with algebraic and geometric multiplicity two, and thus correspond to stable systems. In particular, both have the same Jordan canonical form and hence one may expect both matrices to behave similarly under perturbations. Indeed this is the case when perturbation are applied that ignore the symplectic structure, because arbitrary small perturbations can move eigenvalues to the outside of the unit circle, and will thus make the system unstable. This is illustrated Figure 1, which shows the eigenvalues of 100 random small arbitrary perturbations of S_1 (left picture) and S_2 (right picture). These eigenvalues are located in “clouds” around the eigenvalues $\pm i$ of S_1 and S_2, respectively, and may be inside or outside the unit disc. The situation is completely different if the system is subject to structure-preserving transformations. Figure 1 depicts the effect of 100 random small perturbations that result again in a symplectic matrix. While again arbitrarily small perturbations may move eigenvalues of S_1 to the outside of the unit circle (see third picture), the situation is different for the matrix S_2, because the eigenvalues of all perturbed symplectic matrices remain on the unit circle (see fourth picture). From this point of view, the system given by S_2 is robustly stable under structure-preserving perturbations while the system given by S_1 is not. This surprising behavior of unimodular eigenvalues of symplectic matrices has been observed and explained in the literature before, see e.g. [16, 34], and it is caused by the fact that although both matrices have the same Jordan canonical form, they have different sign characteristics (for a more detailed explanation we refer to Section 4 and, especially, to Remark 4.13.)

In particular, this example shows that the perturbation theory can be expected to be significantly different if arbitrary versus structure-preserving perturbations are considered.



(a) ignoring structure (b) preserving structure

Figure 1.0.1: Structure-preserving vs. structure-ignoring perturbations

Driven by the wish for a better understanding of structure-preserving perturbations of symplectic matrices, the effect of generic rank one perturbations on the Jordan structure of J-symplectic matrices was analyzed in [38], and it was discussed in [3] how the results can be extended to perturbations of rank $k > 1$. It was observed that the preservation of structure sometimes has an important impact on how algebraic and partial multiplicities (i.e., sizes of Jordan blocks) of eigenvalues change under low rank perturbations. In particular, it became apparent that the eigenvalues ± 1 show an exceptional behavior due to symmetry restrictions in the Jordan canonical forms of symplectic matrices - an effect that cannot be observed if structure is ignored and arbitrary perturbations are applied.

While the change in the Jordan canonical form for a given eigenvalue of a symplectic matrix is now well understood, not much has been said about *new* eigenvalues, i.e., eigenvalues of the perturbed matrix which were not eigenvalues of the original unperturbed one. Based on the techniques used in [38], it was possible to show that generically these new eigenvalues will be simple, but only very limited statements on their location in the complex plane could be made. It is the aim of our work to fill this gap and to investigate the movement of eigenvalues under structure-preserving perturbations depending on a small real parameter ε.

The formulas we obtain rely on two fundamental ingredients, namely (i) the multiplicative spectral perturbation theory described in Chapter 3 below and (ii) the detailed analysis performed in §4.2.2 of the connection between left and right eigenvectors of J-symplectic matrices. It may be worth at this point to highlight the importance of each of these two items:

(i) the choice of a *multiplicative* approach(see, for instance, Chapter 3) for the perturbation analysis, instead of the usual additive one (see, for instance, §2), is more natural in this context due to the underlying multiplicativity of symplectic structure : let S be a J-symplectic matrix and let $\widehat{S}(\varepsilon)$ be an analytic J-symplectic matrix function in the real parameter ε with $\widehat{S}(0) = S$. If we write $\widehat{S}(\varepsilon)$ as a multiplicative perturbation

$$\widehat{S}(\varepsilon) = (I + \varepsilon B + O(\varepsilon^2))S, \tag{1.3}$$

one can easily check (see § 4.1.1 below) that the matrix B must be J-Hamiltonian, a fact that will be exploited extensively throughout our analysis. This crucial property would be lost in an additive representation

$$\widehat{S}(\varepsilon) = S + \varepsilon E + O(\varepsilon^2),$$

where the perturbation matrix $E = BS$ is the product of a J-Hamiltonian and a J-symplectic matrix, and has therefore no recognizable structure;

(ii) while multiplicative perturbation theory quantitatively accounts for the size of the leading asymptotic terms, it does not provide much geometric information as to the symmetry constraints the perturbed eigenvalues must satisfy (recall that for each value of ε the perturbed matrix $\widehat{S}(\varepsilon)$ is still J-symplectic). Such symmetries are not apparent right away from the expansions, unless one digs deeper into the formulas for the coefficients: the key observation here is that, as shown in Theorem 3.2, those formulas involve left and right eigenvectors of the unperturbed matrix, *and these vectors are connected in a special way when the matrix is J-symplectic*. This, *together with the special properties of B in* (1.3) *as a J-Hamiltonian matrix*, will allow us to make the symmetries explicit in the formulas.

The main tool to reveal the connections between left and right eigenvectors are *structured symplectic canonical forms*, described in § 2.1 : working out how they relate with the Jordan canonical form leads to explicit relationships between appropriately chosen left and right eigenvectors which, in turn, allow further refinement of the formulas obtained via Theorem 3.2, so that spectral symmetries are explicitly shown in the expansions. This is especially important when the unperturbed eigenvalue is either of the critical eigenvalues ± 1, since we will show that any other eigenvalues behave similarly under structure-preserving and under general perturbations. The exceptional behavior of ± 1 under structure-preserving perturbations is what will make them the main object of our focus in § 4.3.3.

The rest of this dissertation is organized as follows:

Chapter 2 lays out all the preliminaries required to perform our subsequent analysis: first, symplectic and Hamiltonian matrices are defined and the most relevant properties of symplectic matrices are recalled, especially the structured canonical forms which will be instrumental in Chapter 4 in order to establish a crucial connection between left and right eigenvectors of symplectic matrices. The second section of Chapter 2 presents classic results in first order eigenvalue perturbation theory for *additive* perturbations. Although the most natural setting for structure-preserving symplectic perturbations is multiplicative, the additive context will serve us as a model in Chapter 3 for our analysis of the multiplicative one. The main tool in both cases is the so-called Newton Diagram, a geometric construction dating back to Sir Isaac Newton, whose particulars are briefly described in §2.2.2. Sections 2.2.3 and 2.2.4 are devoted to stating and proving, respectively, Lidskii's Theorem, which is our model in the additive case for the asymptotic expansions we will be obtaining later on in the multiplicative setting.

All original results in this dissertation are presented in Chapters 3, 4 and 5. These are structured as follows:

Chapter 3 contains all results for arbitrary (unstructured) multiplicative perturbations. These are especially suited to being applied later to the symplectic case due to its natural multiplicative underlying structure. There is a sharp distinction here between the expansions for zero and nonzero unperturbed eigenvalue λ due to the rank preservation property of multiplicative perturbations: Section 3.1.1 deals with the expansions for nonzero unperturbed eigenvalues, which are quite similar to the ones for the additive case. The expansions for unperturbed zero eigenvalues, whose proof will turn out to be quite more involved, are derived in § 3.1.2. After stating the main Theorem of § 3.1.2, Theorem 3.3, we state and prove several auxiliary results in § 3.1.4 and § 3.1.5, which will be needed in the proof of Theorem 3.3, which in turn is presented in § 3.1.6. Two of these auxiliary results are of an especially technical nature, and their proofs are somewhat independent of the main ideas in Chapter 3. This is why the corresponding proofs are deferred to Appendices A and B. As in the additive case, the leading term of the asymptotic expansions we obtain will depend only on the first order perturbation matrices, on the sizes of the Jordan blocks for the unperturbed eigenvalue λ, and on left and right eigenvectors associated with these Jordan blocks. Section § 3.2 closes this chapter with a *multiplicative perturbation theory* for singular values, which can be easily obtained from the theory for eigenvalues via the so-called Jordan-Wielandt form.

As already mentioned, Chapter 4 is devoted to asymptotic expansions for eigenvalues of structure-preserving *symplectic* perturbations. We begin by showing in § 4.1 that any small structured perturbation of a symplectic matrix can be modeled multiplicatively. This, together with the fact that the set of symplectic matrices is not a linear subspace, but a nonlinear manifold, makes it natural to use the multiplicative theory developed in Chapter 3. Furthermore, the first order coefficient matrix in the multiplicative perturbation is identified as being Hamiltonian (see Definition 2.2 below), a fact which will be crucial in our analysis. Using this, we propose replacing the original perturbing matrix with an alternative one (4.29), which has the advantage that it can be understood as a sequence of rank one perturbations, a special kind of perturbations that has been extensively studied in the literature and is, in general, better understood.

Although the multiplicative theory in Chapter 3 can, in principle, be applied right away to the symplectic case, this is not enough to account for the special spectral behaviors induced by symplectic structure. To do so, one has to incorporate into the analysis the special relationships existing between left and right eigenvectors for symplectic matrices. This is the goal of § 4.2, where explicit relationships between left and right Jordan chains (hence, also between left and right eigenvectors) are established. Once these eigenvector relationships are obtained, they are used in § 4.3 to refine the leading coefficient formulas in the unstructured asymptotic expansions derived in Chapter 3. The most interesting cases (namely, those when the unperturbed eigenvalue is unimodular, especially for the critical eigenvalues ± 1) are fully analyzed, and one very special non-generic situation is identified where the multiplicative theory of Chapter 3 cannot be applied, namely the case when an odd number of odd-sized Jordan blocks associated with an unperturbed eigenvalue $\lambda = \pm 1$ has to be broken. This special situation, in which symplectic spectral symmetry interferes with the usual eigenvalue perturbation behavior, is not covered by the previous theory, so the corresponding expansions have to be worked out from scratch, and are therefore deferred to Chapter 5, where the Newton diagram is again the main tool as in § 3.1.5.

Section §4.4 is devoted to specializing the most relevant expansions obtained in § 4.3 to the case of perturbations of rank 1: such low-rank perturbations are not only most relevant in several applications, but the corresponding asymptotic formulas are simple enough to display in a very clear way the crucial role played in the perturbation behavior of unimodular eigenvalues by certain intrinsic symplectic invariants, the so-called *sign characteristics*. In connection to this, §4.4.1 contains some observations on how these rank-1 expansions can be used in certain situations to describe the coalescence of eigenvalues on the unit circle. Finally, $ 4.5 illustrates with several specific examples, generated via MATLAB, the different asymptotic expansions derived in § 4.3.

As previously announced, Chapter 5 is devoted to a specific, separate analysis of the non-generic particular case identified in §4.3.3. Theorem 5.1 is the main result here: in order to prove it, we introduce in § 5.1.2 an example to illustrate some crucial ideas for the proof, which is then rigorously presented in § 5.1.3. Finally, in § 5.2 our results are specialized again to the case of rank-one perturbations.

Conclusions and some possible ideas for future work are included in Chapter 6, and two appendices include, as already mentioned, the proofs of two auxiliary results of an especially technical nature employed in Chapter 3.

Chapter 2

Preliminaries

2.1 Symplectic and Hamiltonian matrices

In this first section we collect some basic information on symplectic and Hamiltonian matrices, with special emphasis on symplectic canonical forms, which will be crucial in Chapter 4 below to describe the relationship between left and right eigenvectors of symplectic matrices.

2.1.1 Definitions

Let J be an invertible and skew-symmetric matrix. Then J induces a nondegenerate skew-symmetric bilinear form in $\mathbb{R}^m$ (or $\mathbb{C}^m$) as follows

$$\langle\, x \,,\, y \,\rangle := y^T J x \,, \quad x, y \in \mathbb{R}^m \,(\text{or } \mathbb{C}^m)\,,$$

where ^{T} stands for transposition. Now, given a matrix A, the J*-adjoint* of A is defined as the unique matrix $A^{\langle T \rangle}$ such that

$$\langle\, x \,,\, Ay \,\rangle = \left\langle\, A^{\langle T \rangle} x \,,\, y \,\right\rangle \,, \quad \forall x, y \in \mathbb{R}^m \,(\text{or } \mathbb{C}^m)\,.$$

Then, it follows that

$$A^T J = J\, A^{\langle T \rangle} \quad \Rightarrow \quad A^{\langle T \rangle} = J^{-1} A^T J.$$

Finally, depending on whether the matrix $A^{\langle T \rangle}$ is A, $-A$ or A^{-1}, the matrix A is called J-selfadjoint, J-skew-adjoint, or J-unitary, respectively.
Symplectic and Hamiltonian matrices are both particular cases of this special kind of matrices:

Definition 2.1. *Let $J \in \mathbb{C}^{2n \times 2n}$ be an invertible skew-symmetric matrix (i.e., $J^T = -J$). Then a matrix $S \in \mathbb{C}^{2n \times 2n}$ is called J*-symplectic *if it satisfies*

$$S^T\, J\, S = J.$$

Definition 2.2. *Let $J \in \mathbb{C}^{2n\times 2n}$ be an invertible skew-symmetric matrix. Then a matrix $H \in \mathbb{C}^{2n\times 2n}$ is called J-*Hamiltonian *if it satisfies*

$$H^T J + J H = 0.$$

From previous definitions we easily obtain that

$$\begin{aligned} S^{\langle T\rangle} &= J^{-1}S^T J = S^{-1}, \\ H^{\langle T\rangle} &= J^{-1}H^T J = -H. \end{aligned}$$

Hence, a J-symplectic matrix S is also J-unitary and a J-Hamiltonian matrix is also J-skew-adjoint. The case most relevant for applications is when S (or H) is real and when J takes the special form

$$J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \tag{2.1}$$

where I_n denotes the $n \times n$ identity matrix. Note that this J is not only skew-symmetric, but also orthogonal, i.e., $J^{-1} = J^T = -J$.
In this special case, the matrices S and H are simply called *symplectic* and *Hamiltonian* respectively, rather than J*-symplectic* or J*-Hamiltonian*.

Notice that, even though they are not the central object of this memoir, we still need to define Hamiltonian matrices together with symplectic ones. The reason for this is that, as we will see in Chapter 3 below, the multiplicative representation of symplectic perturbations leads naturally to a Hamiltonian first-order coefficient matrix.

2.1.2 Canonical forms for symplectic matrices

For every $n \in \mathbb{N}$ we define the auxiliary matrices

$$\Sigma_n = \begin{bmatrix} 0 & & (-1)^0 \\ & \iddots & \\ (-1)^{n-1} & & 0 \end{bmatrix}, \qquad R_n = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}. \tag{2.2}$$

Furthermore, if $a = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}^T \in \mathbb{C}^n$, we denote by Toep$(a_1, \dots, a_n)$ the upper triangular Toeplitz $n \times n$ matrix that has a^T as its first row. As before, J is a nonsingular skew-symmetric matrix, i.e., $J^T = -J$.

Theorem 2.3 (Theorem 8.5 in [36])**.** *Let $S \in \mathbb{C}^{2n\times 2n}$ be a J-unitary matrix. Then there exists a nonsingular matrix $\mathcal{T}$ such that*

$$\mathcal{T}^{-1} S \mathcal{T} = S_1 \oplus \ldots \oplus S_p, \quad \mathcal{T}^T J \mathcal{T} = H_1 \oplus \ldots \oplus H_p, \tag{2.3}$$

where S_j and H_j have one of the following forms:

i) even-sized blocks associated with $\lambda_j = \pm 1$*, where* $n_j \in \mathbb{N}$ *is even:*

$$S_j = \textit{Toep}\left(\lambda_j,\, 1,\, t_2,\, \dots,\, t_{n_j-1}\right)\,, \quad H_j = \Sigma_{n_j}, \tag{2.4}$$

where the parameters t_k *are real and uniquely determined by the recursive formula*

$$t_2 = \frac{1}{2}\lambda_j, \quad t_{2\ell+1} = 0, \quad t_{2\ell+2} = -\frac{1}{2}\lambda_j \sum_{\nu=1}^{\ell} t_{2\nu} t_{2(\ell+1-\nu)}, \quad \ell \geq 1; \tag{2.5}$$

ii) paired blocks associated with $\lambda_j = \pm 1$*, where* $n_j \in \mathbb{N}$ *is odd:*

$$S_j = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda_j) & \\ \hline & \mathcal{J}_{n_j}(\lambda_j)^{-T} \end{array}\right], \quad H_j = \left[\begin{array}{c|c} 0 & I_{n_j} \\ \hline -I_{n_j} & 0 \end{array}\right]; \tag{2.6}$$

iii) blocks associated with the pair $\left(\lambda_j\,,\, \lambda_j^{-1}\right) \in \mathbb{C}\times\mathbb{C}$*, satisfying* $\mathbf{Re}(\lambda_j) > \mathbf{Re}(\lambda_j^{-1})$ *or* $\mathbf{Im}(\lambda_j) > \mathbf{Im}(\lambda_j^{-1})$ *if* $\mathbf{Re}(\lambda_j) = \mathbf{Re}(\lambda_j^{-1})$*, where* $n_j \in \mathbb{N}$*:*

$$S_j = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda_j) & \\ \hline & \mathcal{J}_{n_j}(\lambda_j)^{-T} \end{array}\right], \quad H_j = \left[\begin{array}{c|c} 0 & I_{n_j} \\ \hline -I_{n_j} & 0 \end{array}\right]. \tag{2.7}$$

Moreover, the form (2.3) *is unique up to the permutation of blocks.*

Theorem 2.4 (Theorem 5.5 in [35])**.** *Let* $S \in \mathbb{R}^{2n\times 2n}$ *be a* J*-unitary matrix. Then there exists a nonsingular matrix* $\mathcal{T} \in \mathbb{R}^{2n\times 2n}$ *such that*

$$\mathcal{T}^{-1}\, S\, \mathcal{T} = S_1 \oplus \dots \oplus S_p\,, \quad \mathcal{T}^T\, J\, \mathcal{T} = H_1 \oplus \dots \oplus H_p, \tag{2.8}$$

where S_j *and* H_j *have one of the following forms:*

i) even-sized blocks associated with $\lambda_j = \pm 1$*, where* $n_j \in \mathbb{N}$ *is even:*

$$S_j = \textit{Toep}\left(\lambda_j,\, 1,\, t_2,\, \dots,\, t_{n_j-1}\right)\,, \quad H_j = \varsigma_j \Sigma_{n_j}, \tag{2.9}$$

where $\varsigma_j = \pm 1$*, and the parameters* t_k *are real and uniquely determined by the recursive formula*

$$t_2 = \frac{1}{2}\lambda_j, \quad t_{2\ell+1} = 0, \quad t_{2\ell+2} = -\frac{1}{2}\lambda_j \sum_{\nu=1}^{\ell} t_{2\nu} t_{2(\ell+1-\nu)}, \quad \ell \geq 1; \tag{2.10}$$

ii) paired blocks associated with $\lambda_j = \pm 1$*, where* $n_j \in \mathbb{N}$ *is odd:*

$$S_j = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda_j) & \\ \hline & \mathcal{J}_{n_j}(\lambda_j)^{-T} \end{array}\right], \quad H_j = \left[\begin{array}{c|c} 0 & I_{n_j} \\ \hline -I_{n_j} & 0 \end{array}\right]; \tag{2.11}$$

iii) blocks associated with the pair $\left(\lambda_j\,,\,\lambda_j^{-1}\right)\in\mathbb{R}\times\mathbb{R}$*, where* $\lambda_j>\lambda_j^{-1}$ *and* $n_j\in\mathbb{N}$*:*

$$S_j=\left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda_j) & \\ \hline & \mathcal{J}_{n_j}(\lambda_j)^{-T} \end{array}\right], \quad H_j=\left[\begin{array}{c|c} 0 & I_{n_j} \\ \hline -I_{n_j} & 0 \end{array}\right]. \tag{2.12}$$

iv) blocks associated with the pair $\left(\lambda_j\,,\,\overline{\lambda_j}\right)$ *of nonreal and unimodular eigenvalues:*

$$S_j=\begin{bmatrix} \Lambda_j & \Theta_j & -t_2\Lambda_j & \cdots & -t_{n_j-1}\lambda_j \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -t_2\Lambda_j \\ \vdots & & \ddots & \ddots & \Theta_j \\ 0 & \cdots & \cdots & 0 & \Lambda_j \end{bmatrix}, \quad H_j=\varsigma_j R_j\oplus\Sigma_2 \tag{2.13}$$

where $|\lambda_j|=1$, $Im(\lambda_j)>0$, $\varsigma_j=\pm1$*, and*

$$\Lambda_j=\begin{bmatrix} Re(\lambda_j) & Im(\lambda_j) \\ -Im(\lambda_j) & Re(\lambda_j) \end{bmatrix}, \quad \Theta_j=\begin{bmatrix} Im(\lambda_j) & -Re(\lambda_j) \\ Re(\lambda_j) & Im(\lambda_j) \end{bmatrix}.$$

Similar than in (2.10)*, parameters* t_k *are real and uniquely determined by the formula*

$$t_2=\frac{1}{2}, \quad t_{2\ell+1}=0, \quad t_{2\ell+2}=-\frac{1}{2}\sum_{\nu=1}^{\ell}t_{2\nu}t_{2(\ell+1-\nu)}, \quad \ell\geq1;$$

iv) blocks associated with a quadruplet $\left(\lambda_j\,,\,\overline{\lambda_j}\,,\,\lambda_j^{-1}\,,\,\left(\overline{\lambda_j}\right)^{-1}\right)$ *of non-real and non-unimodular eigenvalues:*

$$S_j=\begin{bmatrix} \mathcal{J}_{n_j}(\alpha,\beta) & 0 \\ 0 & \mathcal{J}_{n_j}(\alpha,\beta)^{-T} \end{bmatrix}, \quad H_j=\begin{bmatrix} 0 & I_{2n_j} \\ -I_{2n_j} & 0 \end{bmatrix}, \tag{2.14}$$

where $\alpha=Re(\lambda_j)$, $\beta=Im(\lambda_j)>0$, $\alpha^2+\beta^2>1$ *and* $n_j\in\mathbb{N}$.

Moreover, the form (2.8) *is unique up to the permutation of blocks.*

The case when $S\in\mathbb{R}^{2n\times2n}$ and the eigenvalue λ lies on the unit circle is special: in this case it will be advantageous to consider the real matrix S as a complex matrix which turns out to be H-unitary with $H=iJ$ Hermitian. Next theorem provides us from another canonical form for these special matrices.

Theorem 2.5 (Theorem 6.6 in [36]). *Let H be a Hermitian matrix and $S \in \mathbb{C}^{2n\times 2n}$ be H-unitary. Then there exists a nonsingular matrix $\mathcal{T} \in \mathbb{C}^{2n\times 2n}$ such that*

$$\mathcal{T}^{-1}\, S\, \mathcal{T} = S_1 \oplus \ldots \oplus S_p\,, \quad \mathcal{T}^*\, J\, \mathcal{T} = H_1 \oplus \ldots \oplus H_p, \tag{2.15}$$

where S_j and H_j have one of the following forms:

i) blocks associated with unimodular eigenvalues λ_j, i.e. $|\lambda_j| = 1$:

$$S_j = \mathit{Toep}\left(\lambda_j,\, i\,\lambda_j,\, -i\lambda_j t_2,\, \ldots,\, -i\lambda_j t_{n_j-1}\right), \quad H_j = \varsigma_j\, R_{n_j}, \tag{2.16}$$

where $\varsigma_j = \pm 1$, and the parameters t_k are real and uniquely determined by the formula

$$t_2 = \frac{1}{2}, \quad t_{2\ell+1} = 0, \quad t_{2\ell+2} = -\frac{1}{2}\sum_{\nu=1}^{\ell} t_{2\nu} t_{2(\ell+1-\nu)}, \quad \ell \geq 1;$$

ii) blocks associated with a pair $\left(\lambda_j\,,\, \left(\overline{\lambda_j}\right)^{-1}\right)$ of nonunimodular eigenvalues:

$$S_j = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda_j) & \\ \hline & \mathcal{J}_{n_j}(\lambda_j)^{-*} \end{array}\right], \quad H_j = \left[\begin{array}{c|c} 0 & I_{n_j} \\ \hline I_{n_j} & 0 \end{array}\right]; \tag{2.17}$$

where $|\lambda_j| < 1$ and $n_j \in \mathbb{N}$.

Moreover, the form (2.15) *is unique up to the permutation of blocks.*

2.1.3 Basic properties of symplectic matrices

It is well known that both symplectic and Hamiltonian matrices satisfy several different special properties derived from their particular structure. In what follows, we list those ones for symplectic matrices that will be used later in this work. Most of them can be immediately read from the canonical forms §2.1.2 above. In all statements, we assume that S is a symplectic matrix as defined in §2.1.1:

i) If a matrix is symplectic, then its determinant is either 1 or -1. Hence, every symplectic matrix is nonsingular.

ii) If a matrix S is symplectic, then S^{-1}, S^T and $-S$ are also symplectic.

iii) If S_1 and S_2 are symplectic matrices with the same sizes then $S_1 + S_2$ in general is not symplectic but $S_1 S_2$ is always symplectic.

iv) Let λ be an eigenvalue of S, then $\frac{1}{\lambda}$ is also an eigenvalue of S. If S is a real matrix, then $\overline{\lambda}$ and $\frac{1}{\overline{\lambda}}$ are also in the spectrum of S.

v) Both eigenvalues λ and $\frac{1}{\lambda}$ of S have the same Jordan structure, i.e. the same number of Jordan blocks with the same sizes. If S is a real matrix, then the quadruplet $\{\lambda\,,\, \frac{1}{\lambda}\,,\, \overline{\lambda}\,,\, \frac{1}{\overline{\lambda}}\,\}$ have the same Jordan structure associated in S.

vi) If $\lambda = \pm 1$ is an eigenvalue of S, $k \in \mathbb{N}$ and $\mathcal{J}_k(\lambda)$ is a $k \times k$ Jordan block of S associated with λ, then

1. If k is even, then S may have any number of blocks $\mathcal{J}_k(\lambda)$ in its Jordan structure.
2. If k is odd, then the number of blocks $\mathcal{J}_k(\lambda)$ in the Jordan structure of S must be even.

vi) If S is real and λ is a unimodular eigenvalue of S (i.e., $|\lambda| = 1$), then there is a signature associated with the real Jordan form of S for the Jordan blocks associated with λ. This is usually called a *sign characteristic*.

2.2 First order additive eigenvalue perturbation theory

In this second section we aim at an overview of some classic results in first order eigenvalue perturbation theory for arbitrary *additive perturbations*, which we aim to extend to multiplicative ones. These results, embodied primarily by Lidskii's Theorem, are in the spirit of several previous ones, briefly described in §2.2.1, closely following the summary made in [41]. Since our approach to the multiplicative case in Chapter 3 will mirror the proof of Lidskii's Theorem using the so-called *Newton diagram* technique, we briefly introduced this technique in §2.2.2. Then we fully state Lidskii's Theorem in §2.2.3, and recall its proof via the Newton Diagram in §2.2.4 as the model we will follow to prove Theorem 3.3 in Chapter 3. Furthermore, the Newton diagram will be also the main tool in Chapter 5 to obtain the appropriate asymptotic expansions for the nongeneric case.

2.2.1 First order eigenvalue perturbation theory: a short historical overview

The origins of first order eigenvalue perturbation theory have a strong operator-theoretic flavor: after Lord Rayleigh's initial research, the first steps in first order eigenvalue perturbation theory owe much to the formalism proposed by Schrödinger [54, 55] in his approach to Quantum Mechanics. In his theory any observable of a quantum mechanical system is described by a selfadjoint operator H on a certain Hilbert space. An eigenvector x_0 associated with an isolated eigenvalue λ_0 is interpreted as a bound state with energy level λ_0. Any external influence on the system, or any neglected interaction which

should be taken into account, are modelled as an *additive* perturbation H_1 to the operator H. The problem now is to find whether or not there is at least one bound state of the perturbed operator in the neighborhood of x_0. Schrödinger was one of the first to use the additive approach (1.1), writing the perturbed operator as $H + \varepsilon H_1$. He obtained explicit formulas for both eigenvalues and eigenvectors of the perturbed operator $H + \varepsilon H_1$ in a neighborhood of λ_0 and x_0, respectively. The explicit formulas he obtained for eigenvalues and eigenvectors are known as perturbation series in Quantum Mechanics. Even the splitting of multiple eigenvalues was studied in the first order approximation, although Schrödinger did not prove the convergence of such expansions.
It was Rellich who, in a series of papers [47, 48, 49, 50, 51], solved the convergence problem for the isolated eigenvalues of selfadjoint operators in Hilbert space. These results can also be found in Rellich's monograph: *Perturbation theory of eigenvalue problems*[52]. Although convergence was not proved in the non-selfadjoint case, his results stimulated further advances in similar problems. Sz. Nagy [60] extended Rellich's results to complex non-selfadjoint operators by using the Cauchy-Riesz integral method. Other advances in this area were the perturbation theory of continuous spectra, developed by Friedrichs [15], or the perturbation theory for one-parameter semigroups by Hille and Phillips [20]. We finally mention Kato's [23] and Baumgärtel's [4] monographies, which established a general framework for the perturbation theory of linear operators.

Other important results in first order perturbation theory were obtained by Vishik and Lyusternik [62]. They first obtained first order expansions for additive perturbations in the context of differential operators. These results were specialized by Lidskii [31] for the finite dimensional case. He obtained simple explicit formulas for the perturbation coefficients and provided, at the same time, a much more elementary proof. Results in both [62] and [31] were refined later by Baumgärtel [4], in the sense of dealing not only with perturbation series for eigenvalues and eigenvectors, but also with the corresponding eigenprojections as functions of ε. Later on, Moro, Burke & Overton [40] provided an alternative proof for Lidskii's results, using the Newton diagram technique, and extended them to certain nongeneric perturbations. This kind of Newton diagram techniques are the ones we will use in this work to obtain similar results for multiplicative perturbations.

2.2.2 Eigenvalue Perturbation and the Newton Diagram.

As stated at the beginning of this chapter, our first step in this work will be to find the leading terms of asymptotic expansions in the parameter ε of the eigenvalues of multiplicative perturbations of a matrix A. As will be seen in Chapter 3, for nonzero eigenvalues we can use Lisdkii's theory to find those leading terms, but, the case of the eigenvalue $\lambda = 0$ is highly non-generic from the additive point of view. Thus we need to find the leading terms starting from scratch, and, a way to do that, is to analyze the characteristic polynomial of the perturbed matrix.

Obviously, such perturbed eigenvalues are roots of

$$p(z,\varepsilon) = \det(z\,I - (I + \varepsilon\,B)A(I + \varepsilon\,C)),$$

which is a polynomial in z with ε-dependent coefficients. The classical tool to find the leading term in the asymptotic expansions of roots of such polynomials is the so-called *Puiseux-Newton Diagram* (in short, Newton Diagram, or also Newton Polygon), an elementary geometrical construction going back to Sir Isaac Newton (but only rigorously founded by Puiseux [44]), which provides us with both leading powers and leading coefficients of the expansions (see [4, Appendix A7], [7, §8.3] or [40, 41] for more details).

2.2.2.1 The Newton Diagram

The Newton Diagram technique applies to any complex polynomial[1]

$$P(z\,,\,\varepsilon) = z^n + \alpha_1(\varepsilon)z^{n-1} + \ldots + \alpha_{n-1}(\varepsilon)z + \alpha_n(\varepsilon). \tag{2.18}$$

in a variable z with coefficients depending analytically on a parameter ε. In order to simplify the exposition, we assume there is only one zero root of multiplicity n for $\varepsilon = 0$, i.e., the coefficients $\alpha_k(\varepsilon)$ satisfy

$$\alpha_k(\varepsilon) = \widehat{\alpha}_k\varepsilon^{a_k} + o(\varepsilon^{a_k})\;, \qquad k = 1,\,2,\,\ldots,\,n, \tag{2.19}$$

with $\widehat{\alpha}_k \neq 0$, and no term of order lower than a_k appears in the expansion of $\alpha_k(\varepsilon)$. Otherwise, we just shift $z \mapsto z - \lambda$ for any nonzero root λ of $P(z, 0)$.

It is well known [4, 23] that in this situation the roots of equation (2.18) can be written as a series in fractional powers of ε, and we are interested in finding the leading term (i.e., both the leading exponent and leading coefficient) of these series.
Now, let us introduce an example to see how to use the Newton diagram technique and then, a formal proof of why it works will be given.

The Newton Diagram associated with equation (2.18) is obtained as follows: let $I_d = \{k \in \{0,\,\ldots,\,n\} \,:\, \alpha_k(\varepsilon) \neq 0\}$ and $k_{max} = \max I_d$; notice that $\alpha_0(\varepsilon) = 1$ and, as a consequence, $a_0 = 0$. Now we plot the set of points $\{(k\,,\,a_k) \,:\, k \in I_d\} \subset \mathbb{Z}^2$ on a Cartesian grid, and draw the segments on the lower boundary of the convex hull of the plotted points. These segments constitute the so-called *Newton Diagram* associated with the polynomial $P(z\,,\,\varepsilon)$ in (2.18). For instance, the diagram associated with the polynomial

$$P(z,\varepsilon) = z^5 + (2\varepsilon^2 - \varepsilon^3)z^4 - \varepsilon z^3 + (-6\varepsilon^2 + 3\varepsilon^5)z + \varepsilon^3 - \varepsilon^4$$

is as illustrated in Fig. 2.2.1.

It turns out that *the leading exponents of the asymptotic expansions of the different roots of P are just the slopes of the different segments in the Newton Diagram.* More

[1] The Newton Diagram technique applies, in fact, to more general analytic functions, but we restrict ourselves to the special case of polynomials in z with coefficients analytic in a parameter ε.

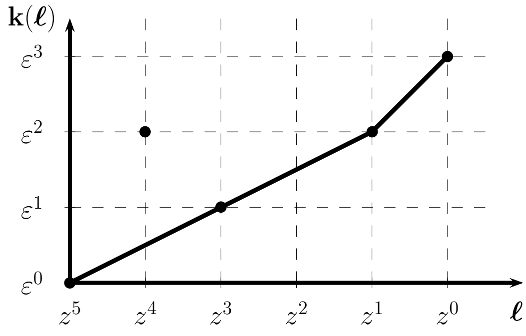


Figure 2.2.1: *Newton diagram for* $P(z,\varepsilon) = z^5 + (2\varepsilon^2 - \varepsilon^3)z^4 - \varepsilon z^3 + (-6\varepsilon^2 + 3\varepsilon^5)z + \varepsilon^3 - \varepsilon^4$.

specifically, let S be an arbitrary segment in the diagram, and $I_S = \{k \in I_d \,:\, (k\,,\, a_k) \in S\}$. If we denote by η the slope of S, $k_{\min} = \min I_S$ and $k_{\max} = \max I_S$, then there are $k_{\max} - k_{\min}$ nonzero roots of $P(z\,,\,\varepsilon)$ with asymptotic expansions.

$$\lambda_j(\varepsilon) = \mu_j \varepsilon^{\eta} + \sum_{s=2}^{\infty} a_{js}\, \varepsilon^{s\eta}\,, \qquad j = 1,\, \ldots,\, k_{\max} - k_{\min}. \tag{2.20}$$

Futhermore, *the leading coefficients* $\{\mu_j\}_{j=1}^{k_{max}-k_{min}}$ *are just the roots of the polynomial*

$$\sum_{k \in I_S} \widehat{\alpha}_k\, \mu^{k_{\max}-k}, \tag{2.21}$$

which is, in general, of a much lower order than $P(z\,,\,\varepsilon)$.

Summarizing, to obtain both the leading exponent η and the leading coefficient μ_j in the asymptotic expansions (2.20), all we have to do is

1. Draw the associated Newton Diagram;
2. Compute the different slopes η of the segments on the Newton Diagram. These are the leading exponents of the different roots of (2.18);
3. For each slope η, find the length of the projection on the horizontal axis of the segment with slope η. This is the number of roots of the order of ε^{η};
4. The leading coefficient μ_j for each root of order ε^{η} is each of the roots of equation (2.21), where S is the segment of the Newton diagram with slope η.

2.2.2.2 Why this works

Without lose of generality, we may assume that the unique eigenvalue of the unperturbed matrix A is zero, otherwise we can shift and project to obtain a reference matrix $\widehat{H}(\varepsilon)$

whose eigenvalues have the same leading terms than the perturbed eigenvalues we are interesting in and such that $\lambda_0 = 0$ is the unique eigenvalue of $\widehat{H}(0)$ having the same Jordan structure than the one associated with the unperturbed eigenvalue λ. As it should be expected, the Newton Diagram technique will be applied to the characteristic polynomial of $\widehat{H}(\varepsilon)$

$$\pi(z, \varepsilon) = det(z\, I - \widehat{H}(\varepsilon)),$$

which clearly satisfies the conditions needed, posed at the beginning of this section. Notice that if we write $\pi(z, \varepsilon)$ using the notation in (2.18) and (2.19), it is well known [22] that the coefficients α_k of $\pi(z, \varepsilon)$, except for a sign, are sum of all the main minors of $\widehat{H}(\varepsilon)$ with size k, but any of them is $O(\varepsilon)$ because of the last property. Finally, $\pi(z, 0) = z^n$, so, we are able to use the Newton diagram for our purposes.

Next, we proceed to proof why this technique give us exactly what we want. This proof can be found in [41] but in seek of completeness we include it here.
Recall that we are interested in the leading term of the roots of the polynomial $P(z, \varepsilon)$ in (2.18) and such that $P(z, 0) = z^n$. Suppose that

$$\lambda(\varepsilon) = \mu \varepsilon^\beta + o\left(\varepsilon^\beta\right)$$

is a root of $P(z, \varepsilon)$, then $P(\lambda(\varepsilon), \varepsilon) = 0$ for all ε, that is

$$\sum_{k=0}^{n} \left[\mu^{n-k} \widehat{\alpha}_k \varepsilon^{a_k + (n-k)\beta} + o\left(\varepsilon^{a_k+(n-k)\beta}\right)\right] = 0\,, \tag{2.22}$$

where $\alpha_0(\varepsilon) = 1$, so, $\widehat{\alpha}_0 = 1$ and $a_0 = 0$. Now, the left part of (2.22), is an infinite polynomial in ε and all the coefficient most be zero. Recall that for all k, $\widehat{\alpha}_k \neq 0$, then, the lowest power of ε in sum (2.22) need to appear at least twice, let denote it by η_β. Therefore, there exist at least two different indexes $i\,, j \in \{0,\, 1,\, \ldots,\, n\}$ such that

$$\eta_\beta = a_i + (n-i)\beta = a_j + (n-j)\beta \leq a_k + (n-k)\beta\,, \quad k = 0,\, 1,\, ,\ldots,\, n.$$

From the equality we have that

$$\beta = \frac{a_i - a_j}{i - j}\,,$$

which is the slope of the segment joining points $(\, i\,,\, a_i\,)$ and $(\, j\,,\, a_j\,)$. We denote by S_β the smallest segment containing all points $(\, k\,,\, a_k\,)$ such that

$$\beta = \frac{a_i - a_k}{i - k}$$

and by I_β the index set of all those values of k. Suppose now that there exist $k_0 \in \{\, 0\,,\, \ldots\,,\, n\,\}$ such that $(\, k_0\,,\, a_{k_0}\,) \notin S_\beta$, then

$$a_i + (n-i)\beta < a_{k_0} + (n-k_0)\beta$$

from where we have that point $(k_0\,,\,a_{k_0})$ is over the line containing segment S_β. Geometrically, this means that if we plot k versus a_k in the Cartesian grid for $k = 0,\,1,\,\ldots,\,n$, and we draw the lower boundary of the convex hull, the so called Newton diagram, then segment S belongs to it. Therefore, the different values of β are de slopes of the segments in the Newton diagram.

Now, we just need to care about the leading coefficient μ. Notice that a segment S_β in the diagram is associated with the lowest power of ε in (2.22), assumed to be $\varepsilon^{a_i+(n-i)\beta}$, for some i, and that we have denoted by I_β the index set for all points belonging to S_β, this is

$$I_\beta = \{k \,|\, (\,k\,,\,a_k\,) \in S_\beta\}\ .$$

So, (2.22) can be written as

$$\left(\sum_{k\in I_\beta} \mu^{n-k}\widehat{\alpha}_k\right)\varepsilon^{\eta_\beta} + o\left(\varepsilon^{\eta_\beta}\right) = 0\,.$$

It follows that

$$\sum_{k\in I_\beta} \mu^{n-k}\widehat{\alpha}_k = 0.$$

The roots of this equation, will give us the coefficient of ε^β for those roots $\lambda(\varepsilon)$ of $P(\lambda)$ with leading exponent larger or equal to β. Then, if we denote by $k_{min} = \min I_\beta$ and $k_{max} = \max I_\beta$, the equation up, can be factorized as follows

$$\mu^{n-k_{max}} \sum_{k\in I_\beta} \mu^{k_{max}-k}\widehat{\alpha}_k = 0\,.$$

The $n - k_{max}$ zero roots of first factor corresponds to the roots $\lambda(\varepsilon)$ of $P(\lambda)$ associated with the segments of slope higher than β(at right of S_β in the diagram) which are $o\left(\varepsilon^\beta\right)$ thus $\mu = 0$. The roots of the second factor, give us the leading coefficient for those $\lambda(\varepsilon)$ that have β as leading exponent. Finally, associated with each segment in the Newton diagram, there are as many roots $\lambda(\varepsilon)$ of $P(z,\varepsilon)$ as it's horizontal projection, with leading exponent β equal to the slope of the segment and leading coefficients μ as the complex solutions of the polynomial equation

$$\sum_{k\in I_\beta} \mu^{k_{max}-k}\widehat{\alpha}_k = 0\,.$$

2.2.3 First order asymptotic expansions: Lidskii's Theorem

In this section we briefly describe some classical first order additive perturbation results which will be employed later in the multiplicative case. They go back to Vishik and Lyusternik [62], who first obtained them in the context of differential operators, but we will be using mostly the expansions obtained later by Lidskii [31], who specialized them to the finite-dimensional case.

In order to state Lidskii's result, we first need to introduce some notation: let A be an arbitrary complex $n \times n$ matrix, and consider an additive perturbation

$$\widetilde{A}(\varepsilon) = A + \varepsilon B \tag{2.23}$$

for arbitrary $B \in \mathbb{C}^{n\times n}$ and small, real $\varepsilon > 0$. Suppose that the unperturbed matrix A in (2.23) has Jordan structure

Now, let the unperturbed matrix A in (2.23) have Jordan structure

$$\left[\begin{array}{c|c} \mathcal{J} & \\ \hline & \widehat{\mathcal{J}} \end{array}\right] = \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] A \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] \tag{2.24}$$

with

$$\left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = I, \tag{2.25}$$

where $\mathcal{J}$ corresponds to a (possibly multiple) eigenvalue λ, and $\widehat{\mathcal{J}}$ is the part of the Jordan form containing the other eigenvalues of A. Moreover, we take $\mathcal{J}$ to be partitioned in the form

$$\mathcal{J} = \text{Diag}(\Gamma_1^1, \ldots, \Gamma_1^{r_1}, \ldots, \Gamma_q^1, \ldots, \Gamma_q^{r_q}), \tag{2.26}$$

where, for $j = 1, \ldots, q$,

$$\Gamma_j^1 = \ldots = \Gamma_j^{r_j} = \mathcal{J}_{n_j}(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda \end{bmatrix}$$

is a Jordan block of dimension n_j, repeated r_j times, and ordered so that

$$n_1 > n_2 > \ldots > n_q.$$

The matrices P and Q are further partitioned as

$$P = \left[\begin{array}{c|c|c|c|c|c|c} P_1^1 & \ldots & P_1^{r_1} & \ldots & P_q^1 & \ldots & P_q^{r_q} \end{array}\right],$$

conformally with (2.26). Notice that the columns of each P_j^k form a right Jordan chain of A with length n_j corresponding to λ. If we denote by x_j^k the first column of P_j^k, each x_j^k is a right eigenvector of A associated with λ. Analogously, if we split

$$Q = \left[\begin{array}{c|c|c|c|c|c|c} Q_1^1 & \dots & Q_1^{r_1} & \dots & Q_q^1 & \dots & Q_q^{r_q} \end{array} \right]^T$$

also according to (2.26), the rows of each Q_j^k form a left Jordan chain of A of length n_j corresponding to λ. Hence, if we denote by y_j^k the last (i.e. n_j−th) row of Q_j^k, each y_j^k is a left eigenvector corresponding to λ. With these eigenvectors we build up the following matrices

$$Y_j = \begin{bmatrix} y_j^1 \\ \vdots \\ y_j^{r_j} \end{bmatrix}, \qquad X_j = [x_j^1, \dots, x_j^{r_j}],$$

for $j = 1, \dots, q$,

$$W_s = \begin{bmatrix} Y_1 \\ \vdots \\ Y_s \end{bmatrix}, \qquad Z_s = [X_1, \dots, X_s], \tag{2.27}$$

for $s = 1, \dots, q$.

Finally, given any arbitrary matrix $K \in \mathbb{C}^{n \times n}$, we define associated square matrices $\Phi_s(K)$ and E_s of dimension

$$f_s = \sum_{j=1}^{s} r_j \tag{2.28}$$

by

$$\begin{aligned} &\Phi_s(K) = W_s K Z_s, \qquad s = 1, \dots, q, \\ &E_1 = I, \qquad E_s = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \quad \text{for } s = 2, \dots, q, \end{aligned} \tag{2.29}$$

where the identity block in E_s has dimension r_s. Note that, due to the cumulative definitions of W_s and Z_s, every $\Phi_{s-1}(K)$, is the upper left block of $\Phi_s(K)$ for $s = 2, \dots, q$.This nested structure allows us to define the **Schur complement** $(\Phi_s/\Phi_{s-1})(B)$ of $\Phi_{s-1}(B)$ in $\Phi_s(B)$ whenever $\Phi_{s-1}(B)$ is nonsingular. For each $s > 1$, we have

$$\Phi_s(B) = \left[\begin{array}{c|c} \Phi_{s-1}(B) & W_{s-1}\, B\, X_s \\ \hline Y_s\, B\, Z_{s-1} & Y_s\, B\, X_s \end{array} \right]$$

and, consequently, the Schur complement is defined as

$$(\Phi_s/\Phi_{s-1})(B) = Y_s \big(B - B\, Z_{s-1}\, (\Phi_{s-1}(B))^{-1}\, W_{s-1}\, B\big)\, X_s. \tag{2.30}$$

Remark 2.6. *As a consequence of the partitions above, we have*

$$Q_j^k\, S\, P_j^k = \Gamma_j^k = \mathcal{J}_{n_j}(\lambda)$$

for every $j = 1 \ldots, q,\ k = 1, \ldots, r_j$.

We are now in the position to state Lidskii's Theorem [31]:

Theorem 2.7. *(Lidskii [31]) Let A be a complex $n \times n$ matrix with an eigenvalue λ and Jordan structure (2.24). Let B be any complex $n \times n$ matrix, and let $j \in \{1, \ldots, q\}$ be such that, if $j > 1$, $\Phi_{j-1}(B)$ is nonsingular. Then there are $r_j n_j$ eigenvalues of the perturbed matrix $A + \varepsilon B$ admitting first-order expansions*

$$\widehat{\lambda}_{j,k,l} = \lambda + (\xi_{j,k})^{1/n_j}\, \varepsilon^{1/n_j} + o\left(\varepsilon^{1/n_j}\right) \tag{2.31}$$

for $k = 1, \ldots, r_j,\ l = 1, \ldots, n_j$, where

(i) *the* $\xi_{j,k},\ k = 1, \ldots, r_j$, *are the roots of equation*

$$\det\left(\Phi_j(B) - \xi\, E_j\right) = 0.$$

where Φ_j and E_j are as in (2.29). Equivalently, the $\xi_{j,k}$ are the eigenvalues of the Schur complement of $\Phi_{j-1}(B)$ in $\Phi_j(B)$; as in (2.30) *(if $j = 1$, the $\xi_{1,k}$ are just the r_1 eigenvalues of $\Phi_1(B)$);*

(ii) *the different values $\widehat{\lambda}_{j,k,l}(\varepsilon)$ for $l = 1, \ldots, n_j$ are obtained by taking the n_j distinct n_j-th roots of $\xi_{j,k}$.*

Notice that Theorem 2.7 applies to any perturbation matrix B except those for which some $\Phi_s(B)$ is singular. The singularity of any such matrix amounts to a polynomial condition on the entries of B, so the set of matrices B for which some $\Phi_s(B)$ is singular has zero measure within the set $\mathbb{C}^{n\times n}$ of complex $n\times n$ matrices. In other words, Theorem 2.7 describes the *generic behavior* of λ under additive matrix perturbations.

Theorem 2.7 can be proved in different ways, but the one most relevant to our purposes and included next is the one making use of the Newton Diagram (see [40, 41]).

2.2.4 Proof of Lidskii's Theorem via the Newton Diagram

Before we apply the Newton Diagram, we need to make some transformations in the characteristic polynomial of the perturbed matrix $A(\varepsilon)$ (recall that we defined the Newton

Diagram in §2.2.2.1 for polynomials $P(z,\varepsilon)$ such that $P(z,0)$ has zero as its only root). Let B be the perturbation matrix and P, Q as defined in (2.24). If we set

$$\widetilde{B} = \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] B \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = \left[\begin{array}{c|c} QBP & QB\widehat{P} \\ \hline \widehat{Q}BP & \widehat{Q}B\widehat{P} \end{array}\right] = \left[\begin{array}{c|c} \widetilde{B}_{11} & \widetilde{B}_{12} \\ \hline \widetilde{B}_{21} & \widetilde{B}_{22} \end{array}\right], \qquad (2.32)$$

then the characteristic polynomial of the perturbed matrix $A(\varepsilon) = A + \varepsilon B$ can be written in the form

$$\begin{aligned} P(z\,,\,\varepsilon) &= \det\left(zI - A(\varepsilon)\right) = \\ &= \det\left(\left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right](zI - A - \varepsilon B)\left[\begin{array}{c|c} P & \widehat{P} \end{array}\right]\right) = \\ &= \det\left(zI - \left[\begin{array}{c|c} \mathcal{J} & \\ \hline & \widehat{\mathcal{J}} \end{array}\right] - \varepsilon\widetilde{B}\right) = \\ &= \det\left(\left[\begin{array}{c|c} zI - \left(\mathcal{J} + \varepsilon\widetilde{B}_{11}\right) & \varepsilon\widetilde{B}_{12} \\ \hline \varepsilon\widetilde{B}_{21} & zI - \left(\widehat{\mathcal{J}} + \varepsilon\widetilde{B}_{22}\right) \end{array}\right]\right). \end{aligned}$$

For sufficiently small ε, no eigenvalue of $\widehat{\mathcal{J}} + \varepsilon\widetilde{B}_{22}$ will be close to λ, so if $\widehat{\lambda}$ is an eigenvalue of the perturbed matrix $A(\varepsilon)$ close to λ, the matrix $\widehat{\lambda}\,I - \widehat{\mathcal{J}} - \varepsilon\widetilde{B}_{22}$ is nonsingular. Therefore, we may factorize the last determinant using Schur's formula

$$P(z\,,\,\varepsilon) = \widehat{\pi}(z\,,\,\varepsilon)\pi(z\,,\,\varepsilon),$$

where

$$\widehat{\pi}(z\,,\,\varepsilon) = \det\left(zI - \widehat{\mathcal{J}} - \varepsilon\widetilde{B}_{22}\right) \text{ and} \qquad (2.33)$$

$$\pi(z\,,\,\varepsilon) = \det\left(zI - \mathcal{J} - \varepsilon\widetilde{B}_{11} - \varepsilon^2\left(\widetilde{B}_{12}\left(zI - \widehat{\mathcal{J}} - \varepsilon\widetilde{B}_{22}\right)^{-1}\widetilde{B}_{21}\right)\right). \qquad (2.34)$$

Hence, the eigenvalues of $A(\varepsilon)$ in which we are interested satisfy the equation $\pi(z\,,\,\varepsilon) = 0$. It only remains to shift λ to the origin: write $\mathcal{J} = \lambda I + \mathcal{J}_0$, i.e., $\mathcal{J}_0$ is the same Jordan matrix as $\mathcal{J}$ replacing λ with zero. Substituting this into equation (2.33) we can write π as a characteristic polynomial

$$\pi(\widetilde{z}\,,\,\varepsilon) = \det\left(\tilde{z}I - H\left(\varepsilon\right)\right),$$

where $\tilde{z} = z - \lambda$ and the matrix $H(\varepsilon)$ is

$$H(\varepsilon) = \mathcal{J}_0 + \varepsilon\widetilde{B}_{11} + O\left(\varepsilon^2\right). \qquad (2.35)$$

Notice that the roots of $\pi(z, \varepsilon)$ are just the eigenvalues of $A(\varepsilon)$ close to λ, and if we use the shifted variable $\widetilde{z}$ then $\pi(\widetilde{z}, \varepsilon)$ satisfies the conditions allowing application of the Newton Diagram. Hence, the key to proving Theorem 2.7 will be to find the correlations between powers of z and powers of ε in the coefficients of the characteristic polynomial of $H(\varepsilon)$.

Now we use the property that the coefficients of a characteristic polynomial of a matrix are just sums of principal minors of the matrix (see, for instance, [22, §1.2]), i.e.,

$$\pi(\widetilde{z}\,,\,\varepsilon) = \widetilde{z}^m - \mathbb{E}_1[H(\varepsilon)]\widetilde{z}^{m-1} + \ldots + (-1)^m \mathbb{E}_m[H(\varepsilon)]$$

where for $k = 1, \ldots, n$, $\mathbb{E}_k[H(\varepsilon)]$ denotes the sum of all k-by-k principal minors of $H(\varepsilon)$. Relating this with equation (2.18) we get

$$\alpha_k(\varepsilon) = (-1)^k \mathbb{E}_k[H(\varepsilon)].$$

To identify the Newton Diagram associated with π one can do the following: given a power ε^l, $l \in \{1, \ldots, f_q\}$, with f_q as in (2.28), we will find the largest possible $k = k(l)$ such that there exists a perturbation matrix B for which $a_{k(l)} = l$ (i.e., $\alpha_{k(l)}(\varepsilon) = \widehat{\alpha}_{k(l)}\varepsilon^l + \ldots$). This amounts to fixing a vertical height l in the Newton Diagram, and finding the rightmost possible point on the diagram at that specific height (equivalently, finding the largest size k such that there exists a $k \times k$ principal minor of $H(\varepsilon)$ of order ε^l).

An specific example may help to better understand the situation: suppose, for instance, that

$$\mathcal{J}_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then $H(\varepsilon)$ can be written schematically as

$$H(\varepsilon) = \left[\begin{array}{ccc|ccc|cc} * & +1 & * & * & * & * & * & * \\ * & * & +1 & * & * & * & * & * \\ \square_{11} & * & * & \square_{12} & * & * & \diamond_{13} & * \\ \hline * & * & * & * & +1 & * & * & * \\ * & * & * & * & * & +1 & * & * \\ \square_{21} & * & * & \square_{22} & * & * & \diamond_{23} & * \\ \hline * & * & * & * & * & * & * & +1 \\ \diamond_{31} & * & * & \diamond_{32} & * & * & \diamond_{33} & * \end{array}\right],$$

where we represent with "+1" the entries of type $1 + O(\varepsilon)$ and the stars, the boxes and the diamonds denote entries of order ε. Boxes and diamonds have been singled out because the matrices Φ_1 and Φ_2 defined in (2.29) are in this case

$$\begin{bmatrix} \square_{11} & \square_{12} \\ \square_{21} & \square_{22} \end{bmatrix} = \varepsilon\Phi_1(B) + O\left(\varepsilon^2\right) \qquad \text{and} \qquad \left[\begin{array}{cc|c} \square_{11} & \square_{12} & \diamond_{13} \\ \square_{21} & \square_{22} & \diamond_{23} \\ \hline \diamond_{31} & \diamond_{32} & \diamond_{33} \end{array}\right] = \varepsilon\Phi_2(B) + O\left(\varepsilon^2\right).$$

Notice that if we take $l = 1$, of course any 1×1 principal minor of $H(\varepsilon)$ is of order ε, but we may find 2×2, and even 3×3 principal minors which are still of order ε: take, for instance, the upper left 3×3 principal minor. This is because the specific position of the "1+" entries allows us to include in the principal minor all consecutive rows and columns of any Jordan block increasing the order in ε by just one unit. Another 3 by 3 minor of $O(\varepsilon)$ is the one formed by rows and columns $4, 5, 6$ of the second 3 by 3 Jordan block. One can check that any 4 by 4 principal minor is of order ε^2, so in this case $k(1) = 3$. For $l = 2$, a similar argument shows that we may find a 6×6 principal minor (the upper left one) of $H(\varepsilon)$ of order ε^2, so $k(2) = 6$. Finally, one can easily check that $k(3) = 8$, so the Newton Diagram in this situation is as shown in Figure 2.2.2 below.

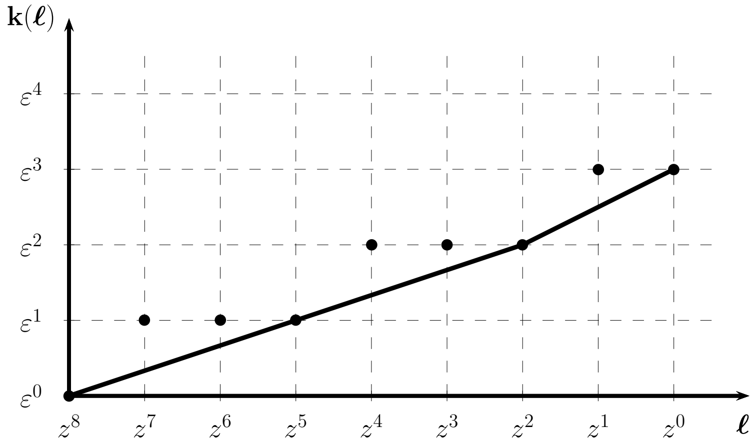


Figure 2.2.2: *Newton diagrams for the example.*

In general, given an arbitrary Jordan form (2.24), one can prove (see [40, Theorem 3.1]) that if $l = f_{j-1} + \rho,\ 0 < \rho \le r_j$ for some $j \in \{1, \ldots, q\}$, with f_{j-1} as defined in (2.28), then the largest $k = k(l)$ we are seeking is

$$k(l) = r_1 n_1 + \ldots + r_{j-1} n_{j-1} + \rho\, n_j.$$

As a consequence of this the Newton Diagram is generically the concatenation of the q segments connecting the points $(k_j, f_j),\ j = 0, 1, \ldots, q$, where f_j is as defined in (2.28), the k_j are defined as

$$k_j = r_1 n_1 + \ldots + r_j n_j,$$

and by convention $f_0 = 0,\ k_0 = 0$. This implies that the leading exponents of the eigenvalue expansions (i.e., the slope of the segments above) are $1/n_j,\ j = 1, \ldots, q$. As to the leading coefficients, let us consider the first segment of the Newton Diagram: it has slope $1/n_1$, it begins at the origin and ends at the point $(k_1,\ f_1)$. It turns out that the only k_1 by k_1 principal minor of $H(\varepsilon)$ of order $O(\varepsilon^{f_1})$ is the upper left k_1 by k_1 minor. In the example above $f_1 = 2$ and $k_1 = 6$, so

$$\begin{aligned}\alpha_6(\varepsilon) &= \det \left[\begin{array}{ccc|ccc} * & +1 & * & * & * & * \\ * & * & +1 & * & * & * \\ \square_{11} & * & * & \square_{12} & * & * \\ \hline * & * & * & * & +1 & * \\ * & * & * & * & * & +1 \\ \square_{21} & * & * & \square_{22} & * & \end{array}\right] = \det \begin{bmatrix} \square_{11} & \square_{12} \\ \square_{21} & \square_{22} \end{bmatrix} + o(\varepsilon^2) = \\ &= \varepsilon^2 \det \Phi_1(B) + o(\varepsilon^2).\end{aligned}$$

In the general case we obtain

$$\alpha_{k_1}(\varepsilon) = (-1)^{f_1} \det\, \Phi_1(B)\varepsilon^{f_1} + o\left(\varepsilon^{f_1}\right).$$

Now take another $j \in \{2, \ldots, q\}$: the j-th segment S_j of the Newton Diagram has slope $1/n_j$, begins at the point $(k_{j-1}\,,\, f_{j-1})$ and ends at $(k_j\,,\, f_j)$. As before, it turns out that the only k_j-by-k_j principal minor of $H(\varepsilon)$ of order $O(\varepsilon^{f_j})$ is the upper left k_j by k_j one, so

$$\alpha_{k_j}(\varepsilon) = (-1)^{f_j} \det\, \Phi_j(B)\varepsilon^{f_j} + o\left(\varepsilon^{f_j}\right).$$

In the example above, $f_2 = 3,\ k_2 = 8$, so

$$\alpha_8(\varepsilon) = \det\, H(\varepsilon) = \det \left[\begin{array}{cc|c} \square_{11} & \square_{12} & \diamond_{13} \\ \square_{21} & \square_{22} & \diamond_{23} \\ \hline \diamond_{31} & \diamond_{32} & \diamond_{33} \end{array}\right] + o(\varepsilon^3) = \varepsilon^3 \det \Phi_2(B) + o(\varepsilon^3).$$

Notice that the nonsingularity of the matrices $\Phi_j(B)$ is just the genericity condition for the 'bending points' (f_j, k_j) to appear on the Newton Diagram. These may not be, however, the only points on the Diagram: there may be intermediate points $(f_{j-1} + \rho, k_{j-1} + \rho n_j)$ for $\rho = 0, \ldots, r_j$ on the segment S_j. The corresponding coefficient in π, which must be nonzero for this point to lie on the Newton Diagram, is

$$\alpha_{k_{j-1}+\rho n_j}(\varepsilon) = (-1)^{f_{j-1}+\rho}\mathbb{E}^*_\rho[\Phi_j(B)]\varepsilon^{f_{j-1}+\rho} + o\left(\varepsilon^{f_{j-1}+\rho}\right), \tag{2.36}$$

where $\mathbb{E}^*_\rho[\Phi_j(B)]$ is the sum of all principal minors of $\Phi_j(B)$ taking the first f_{j-1} rows, together with ρ other rows among the last r_j ones of $\Phi_j(B)$.

Let's finally prove the formula (1) for the leading coefficients of the perturbed eigenvalue expansions. Given $j \in \{1, \ldots, q\}$, we consider the $r_j n_j$ zeros of π admitting first-order expansion

$$\lambda_j = \lambda + \xi_{j,k}\varepsilon^{\frac{1}{n_j}} + o\left(\varepsilon^{\frac{1}{n_j}}\right), \qquad k = 1, \ldots, r_j n_j.$$

We know, from the Newton Diagram theory, that the $\xi_{j,k}$ are the solutions of a polynomial equation

$$\sum_{k \in I_{S_j}} \mu^{k_{max}-k}\hat{\alpha}_k = 0, \tag{2.37}$$

where $k_{max} = k_j$ and $I_{S_j} = \{k_{j-1} + \rho n_j \ : \ \rho = 0, \ldots, r_j\}$ is the set of possible abscissae for the points on S_j.

If we multiply equation (2.37) by $\frac{(-1)^{f_{j-1}}}{\det(\Phi_{j-1}(B))}$ and make the change of variable $\xi = \mu^{n_1}$, we get the equivalent equation

$$\xi^{r_j} - \frac{\mathbb{E}_1^*[\Phi_j(B)]}{\det(\Phi_{j-1}(B))}\xi^{r_j-1} + \ldots + (-1)^{r_j}\frac{\mathbb{E}_{r_j}^*[\Phi_j(B)]}{\det(\Phi_{j-1}(B))} = 0 \tag{2.38}$$

Finally, consider a principal minor ϖ of $\Phi_j(B)$ including its first f_{j-1} rows and ρ other rows among the r_j remaining ones. Then, $\varpi = \det(M)$ for a matrix

$$M = \left[\begin{array}{c|c} \Phi_{j-1}(B) & M_{12} \\ \hline M_{21} & M_{22} \end{array}\right].$$

The basic properties of Schur complements imply that

$$\det\left(M_{22} - M_{21}\Phi_{j-1}^{-1}(B)M_{12}\right) = \frac{\det(M)}{\det(\Phi_{j-1}(B))}.$$

Now, if we denote by $(\Phi_j/\Phi_{j-1})(B)$ the Schur complement of $\Phi_{j-1}(B)$ in $\Phi_j(B)$, one can easily prove that for $\rho = 1, \ldots, r_j$

$$\frac{\mathbb{E}_\rho^*[\Phi_j(B)]}{\det(\Phi_{j-1}(B))} = \mathbb{E}_\rho\left[(\Phi_j/\Phi_{j-1})(B)\right],$$

so equation (2.38) may be rewritten as

$$\xi^{r_j} - \mathbb{E}_1\left[(\Phi_j/\Phi_{j-1})(B)\right]\xi^{r_j-1} + \ldots + (-1)^{r_j}\mathbb{E}_{r_j}\left[(\Phi_j/\Phi_{j-1})(B)\right] = 0.$$

But this is just the characteristic equation of $(\Phi_j/\Phi_{j-1})(B)$, so the solutions of equation (2.37) are just the eigenvalues of $(\Phi_j/\Phi_{j-1})(B)$.

Chapter 3

First order expansions for eigenvalues of multiplicative perturbations

In this Chapter we obtain a result similar to Theorem 2.7 above, but for multiplicative, instead of additive perturbations. Let $A \in \mathbb{C}^{n\times n}$ be a matrix with arbitrary Jordan form (2.24), and consider a multiplicative perturbation

$$\widehat{A} = \widehat{A}(\varepsilon) = (I + \varepsilon B)\, A\, (I + \varepsilon C). \tag{3.1}$$

for arbitrary matrices B, $C \in \mathbb{C}^{n\times n}$.

We begin by noting that multiplicative perturbations are, in a way, less powerful than additive ones, since *a multiplicative perturbation always preserves rank*. This will induce a very sharp distinction in our subsequent analysis between the case when $\lambda \neq 0$ (which is essentially the same as the additive one), and the case $\lambda = 0$, which will require a completely new analysis.

Remark 3.1. *In general, the matrices B and C are assumed to not depend on ε. However, all results in this chapter remain true for perturbations of the form*

$$\widehat{A}(\varepsilon) = \left(I + \varepsilon B + O(\varepsilon^2)\right)\; A\; \left(I + \varepsilon C + O(\varepsilon^2)\right).$$

3.1 Multiplicative perturbation of matrix eigenvalues

We begin with the simpler case of nonzero eigenvalues, which will turn out to be a straightforward consequence of the additive case:

3.1.1 The case $\lambda \neq 0$

Suppose the unperturbed eigenvalue λ of A is different from zero. We may rewrite the perturbation (3.1) additively as

$$\widehat{A}(\varepsilon) = A + \varepsilon\,(BA + AC) + O\left(\varepsilon^2\right),$$

and apply Lisdkii's Theorem 2.7 for additive perturbations to conclude that the asymptotic behavior of λ under perturbation depends on the matrices $\Phi_s(BA + AC)$, as defined in (2.29). Now, recall that the rows of W_s (respectively, columns of Z_s) are left eigenvectors (respectively., right eigenvectors) of A associated with λ, i.e.

$$\begin{aligned} W_s A &= \lambda W_s \\ A Z_s &= \lambda Z_s\,, \qquad s = 1,\,\ldots,\,q \end{aligned}$$

and, therefore,

$$\Phi_s(BA + AC) = W_s\,(BA + AC)\,Z_s = \lambda W_s\,(B + C)\,Z_s = \lambda \Phi_s\,(B + C)$$

Hence, although the additive perturbation matrix $BA + AC$ does depend on A, the corresponding $\Phi_s(\cdot)$ does not. This leads directly to the following result:

Theorem 3.2 ([57]). *Let $\lambda \neq 0$ be an eigenvalue of a complex $n \times n$ matrix A with Jordan structure (2.24), and let B, C be arbitrary $n \times n$ complex matrices. Let $j \in \{1,\,\ldots,\,q\}$ be given and assume that if $j > 1$, $\Phi_{j-1}(B+C)$ is nonsingular, where $\Phi_{j-1}(\cdot)$ is defined as in (2.29). Then there are $r_j n_j$ eigenvalues of the perturbed matrix $\widehat{A}(\varepsilon) = (I + \varepsilon B)\,A\,(I + \varepsilon C)$ admitting first order expansions*

$$\widehat{\lambda}_{j,k,l} = \lambda + \left(\lambda \xi_{j,k}\right)^{1/n_j} \varepsilon^{1/n_j} + o\left(\varepsilon^{1/n_j}\right) \tag{3.2}$$

where

(i) *the $\xi_{j,k},\ k = 1, \ldots, r_j,$ are the roots of the equation*

$$\det\left(\Phi_j(B + C) - \xi\, E_j\right) = 0 \tag{3.3}$$

where Φ_j and E_j are as in (2.29). Equivalently, the $\xi_{j,k},\ k = 1, \ldots, r_j,$ are the eigenvalues of the Schur complement of $\Phi_{j-1}(B + C)$ in $\Phi_j(B + C)$; as in (2.30) *(if $j = 1$, the $\xi_{1,k}$ are just the r_1 eigenvalues of $\Phi_1(B + C)$);*

(ii) *the different values $\widehat{\lambda}_{j,k,l}(\varepsilon)$ for $l = 1, \ldots, n_j$ are defined by taking the n_j distinct n_j-th roots of $\xi_{j,k}$.*

3.1.2 The case $\lambda = 0$

Let us now consider the case when the eigenvalue under examination is zero. Notice that the argument in section 3.1.1 above gives no information whatsoever, since now both AZ_s and W_sA are zero, so $\Phi_s(BA + AC) = 0$. Furthermore, the fact that A and $\widehat{A}(\varepsilon)$ have the same rank forces $\lambda = 0$ to be an eigenvalue of both matrices, and with the same geometric multiplicity. Hence, both matrices A and $\widehat{A}$ have the same number of Jordan blocks associated with $\lambda = 0$. The algebraic multiplicity, however, will generically decrease: it is well known (see, for instance, [21, 42, 53]) that the larger the dimension of a Jordan block, the more unstable it is under perturbation. Hence, the most likely behavior of the zero eigenvalue under multiplicative perturbations is that any of its 1×1 Jordan blocks in A is preserved in $\widehat{A}$, while any Jordan block of A of dimension larger than one becomes a 1×1 Jordan block of $\widehat{A}$. In other words, the algebraic multiplicity of $\lambda = 0$ is expected to go from its initial value of $n_1r_1 + \ldots + n_qr_q$ down to $f_q = r_1 + \ldots + r_q$, creating $r_1(n_1 - 1) + \ldots + r_q(n_q - 1)$ nonzero eigenvalues in the process.

In terms of the Newton Diagram, this amounts to the diagram being formed by q segments with slopes $1/(n_j - 1),\ j = 1, \ldots, q$, instead of q segments of slope $1/n_j$ as in the additive case. The length of their horizontal projections is therefore smaller, $r_j(n_j - 1)$ instead of r_jn_j. Loosely speaking, the Newton Diagram for the multiplicative case should be obtained by moving the first point $(n_1, 1)$ on the additive Newton Diagram one unit to the left, the second point on the additive diagram two units to the left, an so on. Consider, for instance the following 8×8 example,

$$\mathcal{J}_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{3.4}$$

In this case, for instance, the two Newton Diagrams, additive and multiplicative, will be the ones in Figure 3.1.1 below. The one for additive perturbations is the dashed one on the bottom, while the diagram for the multiplicative ones is the solid line on top.

More specifically, the expected behavior under multiplicative perturbations of a zero eigenvalue is described by our main Theorem in this section:

Theorem 3.3. *Let A be any complex $n \times n$ matrix with Jordan structure (2.24) and $\lambda = 0$. Let B, C be arbitrary $n \times n$ complex matrices, let $j \in \{1, \ldots, q\}$ be given and assume that if $j > 1$, $\Phi_{j-1}(B + C)$ is nonsingular. Then, there are $r_j\,(n_j - 1)$ eigenvalues of the perturbed matrix $\widehat{A}(\varepsilon) = (I + \varepsilon B)\,A\,(I + \varepsilon C)$ admitting first order expansions*

$$\widehat{\lambda}_{j,k,l} = (\xi_{j,k})^{\frac{1}{n_j - 1}}\,\varepsilon^{\frac{1}{n_j - 1}} + o\left(\varepsilon^{\frac{1}{n_j - 1}}\right). \tag{3.5}$$

Moreover, if $n_j \geq 2$, then

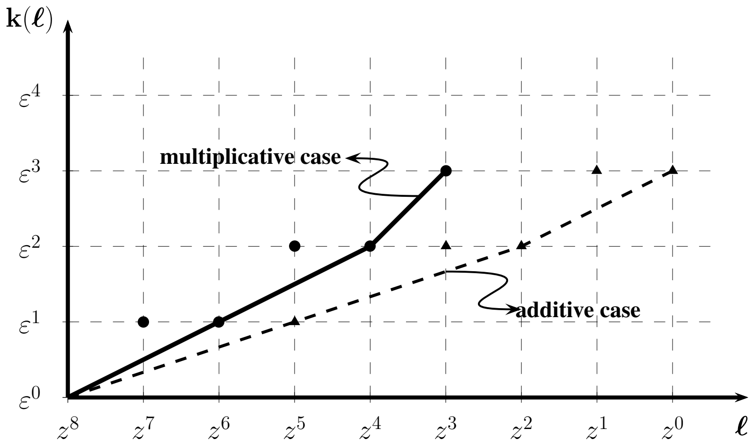


Figure 3.1.1: *Newton diagrams for example (3.4). The Newton diagram corresponding to the multiplicative case is depicted with a solid line, the one with the dashed line is associated with additive perturbations.*

(i) *the* $\xi_{j,k}$, $k = 1, \ldots, r_j$, *are the roots of equation*

$$\det\left(\Phi_j(B+C) - \xi\, E_j\right) = 0 \tag{3.6}$$

where Φ_j *and* E_j *are as in (2.29) or, equivalently, the eigenvalues of the Schur complement of* $\Phi_{j-1}(B+C)$ *in* $\Phi_j(B+C)$ *(if* $j = 1$, *the* $\xi_{1,k}$ *are just the* r_1 *eigenvalues of* $\Phi_1(B+C)$*);*

(ii) *the different values* $\widehat{\lambda}_{j,k,l}(\varepsilon)$ *for* $l = 1, \ldots, n_j - 1$ *are defined by taking the* $(n_j - 1)$ *distinct* (n_j-1)*-th roots of* ξ_j^k.

(iii) *the remaining* r_j *eigenvalues are zero, i.e.,*

$$\widehat{\lambda}_{j,k,n_j} = 0, \;\; k = 1, \ldots, r_j \tag{3.7}$$

Notice that one consequence of Theorem 3.3 (and of Theorem 3.2 as well, for that matter) is that the leading coefficients depend on the perturbation matrices B an C only through their sum, so the perturbed eigenvalues of (3.1) have the same leading term as those of the matrix $(I + \varepsilon(B+C))A$, where A is perturbed only from the left, or those of $A(I + \varepsilon(B+C))$, since all three perturbations give rise to the same matrices $\Phi_j(B+C)$. Of course the two latter matrices have the same eigenvalues, since they are products of the same two matrices in reverse order, but this is not true for (3.1). One can easily check that the eigenvalues of (3.1) are close to those of the one-sided perturbations only for small values of ε.

Now, in order to prove Theorem 3.3 we will separately prove the validity of the expansions (3.5) and of the formula (3.6) for the leading coefficients.

Mirroring the idea of the proof of Lidskii's theorem in [40], we first prove the validity of (3.5) by finding, for every height $l \in \{1, \ldots, f_q\}$, the rightmost possible point on the Newton Diagram at that specific height. This amounts to identifying the lowest possible Newton Diagram in this situation (the *Newton envelope* in the terminology of [40]). The corresponding result is as follows:

Theorem 3.4. *Let $A \in \mathbb{C}^{m\times m}$ be a complex matrix with Jordan form (2.24) and $\lambda = 0$. For every $l \in \{1, \ldots, f_q\}$, let $k(l)$ be the largest possible integer such that there exist perturbation matrices B, C for which $a_{k(l)} = l$. If $l = f_{j-1} + \rho$ for some $j \in \{1, \ldots, q\}$ with $n_j \geq 2$ and $1 \leq \rho \leq r_j$, then*

$$k(l) = r_1\,(n_1 - 1) + \ldots + r_{j-1}\,(n_{j-1} - 1) + \rho\,(n_j - 1)\,. \tag{3.8}$$

(the case $n_j = 1$ is left out of both Theorems 3.3 and 3.4, since we know that 1 by 1 Jordan blocks corresponding to a zero eigenvalue are preserved by multiplicative perturbations). In order to prove Theorem 3.4, we will need to carefully analyze all possible ways of constructing principal minors of order ε^l of an appropriately chosen ε-dependent matrix $\widehat{H}(\varepsilon)$, to be defined in the next subsection. We first identify the matrix $\widehat{H}(\varepsilon)$ in section 3.1.3, and describe basic properties of its principal minors in section 3.1.4 below. Section 3.1.5 contains the proof of Theorem 3.4, while the proof of Theorem 3.3 is included in § 3.1.6.

3.1.3 The matrix $\widehat{H}(\varepsilon)$

We begin by transforming the characteristic polynomial of the perturbed matrix $\widehat{A}(\varepsilon)$ in order to apply the Newton Diagram technique:

$$\begin{aligned} P\,(z\,,\,\varepsilon) &= \det\left(zI - \widehat{A}(\varepsilon)\right) = \det\left(\,zI - (I+\varepsilon B)\,A\,(I+\varepsilon C)\,\right) = \\ &= \det\left(zI - \left(I+\varepsilon\widetilde{B}\right)\left[\begin{array}{c|c} \mathcal{J}_0 & \\ \hline & \widehat{\mathcal{J}} \end{array}\right]\left(I+\varepsilon\widetilde{C}\right)\right), \end{aligned}$$

where $\mathcal{J}_0$ contains the Jordan blocks associated with $\lambda = 0$, $\widehat{\mathcal{J}}$ contains the Jordan blocks corresponding to nonzero eigenvalues, and

$$\begin{aligned} \widetilde{B} &= \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] B \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = \left[\begin{array}{c|c} QBP & QB\widehat{P} \\ \hline \widehat{Q}BP & \widehat{Q}B\widehat{P} \end{array}\right] = \left[\begin{array}{c|c} \widetilde{B}_{11} & \widetilde{B}_{12} \\ \hline \widetilde{B}_{21} & \widetilde{B}_{22} \end{array}\right], \\ \widetilde{C} &= \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] C \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = \left[\begin{array}{c|c} QCP & QC\widehat{P} \\ \hline \widehat{Q}CP & \widehat{Q}C\widehat{P} \end{array}\right] = \left[\begin{array}{c|c} \widetilde{C}_{11} & \widetilde{C}_{12} \\ \hline \widetilde{C}_{21} & \widetilde{C}_{22} \end{array}\right]. \end{aligned}$$

Partitioning the matrix zI, and using the properties of the Schur complement, we may factorize $P\,(z\,,\,\varepsilon)$ as

$$P(z\,,\,\varepsilon) \;=\; \widehat{\pi}(z\,,\,\varepsilon)\pi(z\,,\,\varepsilon),$$

for

$$\begin{aligned}
\widehat{\pi}(z\,,\,\varepsilon) &= \det\left(M\right), \\
\pi(z\,,\,\varepsilon) &= \det\left(zI - \left(I + \varepsilon\widetilde{B}_{11}\right)\mathcal{J}_0\left(I + \varepsilon\widetilde{C}_{11}\right) - \varepsilon^2\left(\widetilde{B}_{12}\widehat{\mathcal{J}}\widetilde{C}_{21} + \widehat{S}(z\,,\,\varepsilon)\right)\right),
\end{aligned}$$

where

$$M = zI - \left(I + \varepsilon\widetilde{B}_{22}\right)\widehat{\mathcal{J}}\left(I + \varepsilon\widetilde{C}_{22}\right) - \varepsilon^2\widetilde{B}_{21}\mathcal{J}_0\widetilde{C}_{12}$$

$$\widehat{S}(z\,,\,\varepsilon) =$$
$$\left(\left(I + \varepsilon\widetilde{B}_{11}\right)\mathcal{J}_0\widetilde{C}_{12} + \widetilde{B}_{12}\widehat{\mathcal{J}}\left(I + \varepsilon\widetilde{C}_{22}\right)\right)M^{-1}\left(\widetilde{B}_{21}\mathcal{J}_0\left(I + \varepsilon\widetilde{C}_{11}\right) + \left(I + \varepsilon\widetilde{B}_{22}\right)\widehat{\mathcal{J}}\widetilde{C}_{21}\right).$$

For small ε, the matrices $I + \varepsilon B$, $I + \varepsilon C$ and $\widehat{\mathcal{J}}$ are nonsingular. Hence, if $\widehat{\lambda}$ is an eigenvalue of the perturbed matrix $\widehat{A}(\varepsilon)$ close to zero, it cannot be a root of $\widehat{\pi}(z\,,\,\varepsilon)$, so it must be a root of the polynomial $\pi(z\,,\,\varepsilon)$. Hence, we write $\pi(z\,,\,\varepsilon) = \det\left(zI - \widehat{H}_1(\varepsilon)\right)$ for

$$\widehat{H}_1(\varepsilon) = \mathcal{J}_0 + \varepsilon\left(\widetilde{B}_{11}\mathcal{J}_0 + \mathcal{J}_0\widetilde{C}_{11}\right) + \varepsilon^2\left(\widetilde{B}_{12}\widehat{\mathcal{J}}\widetilde{C}_{21} + \widetilde{B}_{11}\mathcal{J}_0\widetilde{C}_{11} + \widehat{S}(z\,,\,\varepsilon)\right). \tag{3.9}$$

One can easily check that the sum $\widetilde{B}_{11}\mathcal{J}_0 + \mathcal{J}_0\widetilde{C}_{11}$ has zero entries on the lower left corner of every submatrix resulting from the partition conformal with $\mathcal{J}_0$. Consequently, the entries of $\widehat{H}_1(\varepsilon)$ in those positions are of order $O(\varepsilon^2)$. The remaining entries $\widehat{H}_1(\varepsilon)$ are either of order $O(1)$ (the ones coming from the superdiagonal 1s in $\mathcal{J}_0$) or of order $O(\varepsilon)$ (all the rest). To illustrate this, consider again the 8×8 example $\mathcal{J}_0 = \mathcal{J}_3(0) \oplus \mathcal{J}_3(0) \oplus \mathcal{J}_2(0)$ in (3.4). In that case, the sum $\widetilde{B}_{11}\mathcal{J}_0 + \mathcal{J}_0\widetilde{C}_{11}$ has a zero on the lower left corner of each block in the 3×3 block partition conformal with $\mathcal{J}_0$, and the corresponding entries of $\widehat{H}(\varepsilon)$ will be of order $O(\varepsilon^2)$. We may schematically write

$$\widehat{H}_1(\varepsilon) = \left[\begin{array}{ccc|ccc|cc}
* & +1 & * & * & * & * & * & * \\
* & * & +1 & * & * & * & * & * \\
\bullet & * & * & \bullet & * & * & \bullet & * \\
\hline
* & * & * & * & +1 & * & * & * \\
* & * & * & * & * & +1 & * & * \\
\bullet & * & * & \bullet & * & * & \bullet & * \\
\hline
* & * & * & * & * & * & * & +1 \\
\bullet & * & * & \bullet & * & * & \bullet & *
\end{array}\right]$$

where the bullets denote $O(\varepsilon^2)$ entries, the asterisks denote $O(\varepsilon)$ entries, and the '$+1$' denote entries of type $1+O(\varepsilon)$. This is roughly the general form of $\widehat{H}_1(\varepsilon)$ for any dimension.

We now concentrate on the matrix $\widetilde{B}_{12}\widehat{\mathcal{J}}\widetilde{C}_{21} + \widehat{S}(z\,,\,\varepsilon)$ in the second order term of $\widehat{H}_1(\varepsilon)$. We shall prove that the entries of this matrix lying in the same positions as the bullet entries in $\widehat{H}_1(\varepsilon)$ are $O\left(\varepsilon^\eta\right)$ for some $\eta > 0$. As a consequence, we will write

$$\widehat{H}_1(\varepsilon) = \mathcal{J}_0 + \varepsilon\left(\widetilde{B}_{11}\mathcal{J}_0 + \mathcal{J}_0\widetilde{C}_{11}\right) + \varepsilon^2\left(\widetilde{B}_{11}\mathcal{J}_0\widetilde{C}_{11} + R\right) + O\left(\varepsilon^{2+\eta}\right) \tag{3.10}$$

for some matrix R whose entries in the bullet positions of $\widehat{H}_1(\varepsilon)$ are all zero.

In order to prove this, we begin by noting that any principal minor of $\widehat{H}_1(\varepsilon)$ is of order at least ε. Hence, the Newton diagram applied to the polynomial $\pi(z\,,\,\varepsilon)$ implies that any of its nonzero roots is $O\left(\varepsilon^\eta\right)$ for some $\eta,\ 0 < \eta \leq 1$. Taking $z = O\left(\varepsilon^\eta\right)$ in the formula for M leads to $M = -\widehat{\mathcal{J}} + \phi(\varepsilon)$ for some matrix $\phi(\varepsilon)$ of order $O\left(\varepsilon^\eta\right)$ and, consequently, $-\widehat{\mathcal{J}}^{-1}M = I - \widehat{\mathcal{J}}^{-1}\phi(\varepsilon)$. Since $\widehat{\mathcal{J}}^{-1}\phi(\varepsilon)$ can be made arbitrarily small by taking ε small enough, we obtain

$$M^{-1} = -\left(I - \widehat{\mathcal{J}}^{-1}\phi(\varepsilon)\right)^{-1}\widehat{\mathcal{J}}^{-1} = -\widehat{\mathcal{J}}^{-1} + O\left(\varepsilon^\eta\right).$$

Replacing this in the expression for $\widehat{S}(z\,,\,\varepsilon)$, we obtain

$$\begin{aligned}\widehat{S}(z\,,\,\varepsilon) &= -\left(\widetilde{B}_{12}\widehat{\mathcal{J}} + \mathcal{J}_0\widetilde{C}_{12}\right)\widehat{\mathcal{J}}^{-1}\left(\widetilde{B}_{21}\mathcal{J}_0 + \widehat{\mathcal{J}}\widetilde{C}_{21}\right) + O\left(\varepsilon^\eta\right) = \\ &= -\widetilde{B}_{12}\widehat{\mathcal{J}}\widetilde{C}_{21} + R + O\left(\varepsilon^\eta\right)\end{aligned}$$

where

$$R = -\mathcal{J}_0\widetilde{C}_{12}\left(\widehat{\mathcal{J}}^{-1}\widetilde{B}_{21}\mathcal{J}_0 + \widetilde{C}_{21}\right) - \widetilde{B}_{12}\widetilde{B}_{21}\mathcal{J}_0$$

is the matrix we announced in (3.10). This implies

$$\widehat{H}_1(\varepsilon) = \mathcal{J}_0 + \varepsilon\left(\widetilde{B}_{11}\mathcal{J}_0 + \mathcal{J}_0\widetilde{C}_{11}\right) + \varepsilon^2\left(\widetilde{B}_{11}\mathcal{J}_0\widetilde{C}_{11} + R\right) + O\left(\varepsilon^{2+\eta}\right),$$

so the leading matrix for $\widehat{H}_1(\varepsilon)$ is

$$\widehat{H}(\varepsilon) = \left(I + \varepsilon\widetilde{B}_{11}\right)\mathcal{J}_0\left(I + \varepsilon\widetilde{C}_{11}\right)$$

and, therefore, the leading terms in the coefficients of the polynomial π will be sums of principal minors of $\widehat{H}(\varepsilon)$. Thus, the crucial question from now on is to identify which principal minors of $\widehat{H}(\varepsilon)$ give rise to terms of a given order in ε. Notice that the $O(\varepsilon^2)$ entries of $\widehat{H}(\varepsilon)$ are placed precisely in those positions which were most important in the expansions for the additive case. This will somewhat complicate the analysis.

If we write

$$\pi(z,\varepsilon) = \det(z\,I - \widehat{H}_1(\varepsilon)) = z^m + \alpha_1(\varepsilon)z^{m-1} + \ldots + \alpha_{m-1}(\varepsilon)z + \alpha_m(\varepsilon) \qquad (3.11)$$

with coefficients

$$\alpha_k(\varepsilon) = \widehat{\alpha}_k \varepsilon^{a_k} + o\left(\varepsilon^{a_k}\right), \qquad k = 1, \ldots, m,$$

we now know that under the conditions of Theorem 3.3 the polynomial π has a root $\lambda = 0$ with multiplicity

$$f_q = r_1 + \ldots + r_q.$$

If we denote $\widetilde{m} = m - f_q$, then $\alpha_k(\varepsilon) = 0$ for $k = \widetilde{m}+1, \ldots, m$. On the other hand, it is well known [22, §1.2] that each coefficient of a characteristic polynomial, except for a sign, is a sum of principal minors of the matrix. Hence, every coefficient α_k can be written as

$$\alpha_k(\varepsilon) = (-1)^k \mathbb{E}_k\left[\widehat{H}(\varepsilon)\right], \qquad k = 1, \ldots, \widetilde{m}, \qquad (3.12)$$

where $\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right]$ denotes the sum of all k-by-k principal minors of $\widehat{H}(\varepsilon)$.

3.1.4 The principal minors of $\widehat{H}(\varepsilon)$

In this subsection we will identify which principal minors of $\widehat{H}(\varepsilon)$ give rise to terms of a given order in ε. We begin by fixing the notation for principal minors: let $M \in \mathbb{C}^{m\times m}$ be an arbitrary matrix, and, for each $k \in \{1, \ldots, m\}$, let Ξ_k be the family of all increasingly ordered lists γ of length k with entries taken from $\{1, \ldots, m\}$. For each $\gamma \in \Xi_k$ we denote by $M[\gamma]$ the k-by-k principal sub-matrix of M whose entries are those lying on the rows and columns of M with indices in γ. In a more general context, and for $\gamma,\, \theta \in \Xi_k$, we denote by $M[\gamma\,|\,\theta]$ the sub-matrix of M whose entries are those lying on the rows with indices in γ and on the columns with indices in θ. With this definition,

$$\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right] = \sum_{\beta\in\Xi_k} \det\, \widehat{H}(\varepsilon)[\beta]$$

For simplicity, set $M_1 = I + \varepsilon B_{11}$ and $M_2 = I + \varepsilon C_{11}$, so

$$\widehat{H}(\varepsilon) = M_1 \mathcal{J}_0 M_2.$$

Then, the Cauchy-Binet formula [22] applied to $\widehat{H}(\varepsilon)[\beta]$ leads to

$$\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right] = \sum_{\beta\in\Xi_k}\sum_{\gamma\in\Xi_k}\sum_{\theta\in\Xi_k} \det\left(M_1[\beta\,|\,\gamma]\right)\det\left(\mathcal{J}_0[\gamma\,|\,\theta]\right)\det\left(M_2[\theta\,|\,\beta]\right).$$

One can easily check that $\det\left(\mathcal{J}_0[\gamma\,|\,\theta]\right) \in \{0\,,\,1\}\ \ \forall \gamma,\, \theta \in \Xi_k$.
More precisely, $\det(\mathcal{J}_0[\gamma\,|\,\theta]) = 1$ if and only if

$$\gamma(i) \in \{1, \ldots, m\} \setminus \Omega \qquad \wedge \quad \theta(i) = \gamma(i) + 1, \qquad i = 1, \ldots, k,$$

where

$$\Omega = \left\{ \sum_{i=1}^{j-1} n_i r_i + \rho n_j \;\middle|\; \begin{array}{ccc} j & = & 1, \ldots, q \\ \rho & = & 1, \ldots, r_j \end{array} \right\} \tag{3.13}$$

is the set of indices corresponding to the last row of each Jordan block in $\mathcal{J}_0$, and we denote by $\xi(j)$ the index placed in the j-th position in a list ξ. Therefore, $\mathbb{E}_k \left[\widehat{H}(\varepsilon)\right]$ can be written as

$$\mathbb{E}_k \left[\widehat{H}(\varepsilon)\right] = \sum_{\gamma \in \widehat{\Xi}_k} \sum_{\beta \in \Xi_k} \det\left(M_1[\beta \,|\, \gamma]\right) \det\left(M_2[\theta \,|\, \beta]\right) \quad , \qquad \theta(i) = \gamma(i)+1,\ i = 1, \ldots, l, \tag{3.14}$$

where $\widehat{\Xi}_k$ denotes the family of all increasingly ordered lists of length k and entries in $\{1, \ldots, m\} \setminus \Omega$.

We now highlight some entries of M_1 and M_2 which will play a crucial role in our analysis: consider the lower left entry of each block in the Jordan partition (2.26), and denote with a club $\clubsuit$ (respectively, a spade $\spadesuit$) the entry of M_1 (resp. of M_2) in that specific position. In our 8 by 8 previous example, the highlighted positions are as follows:

$$M_1 = \left[\begin{array}{ccc|ccc|cc}
+1 & * & * & * & * & * & * & * \\
* & +1 & * & * & * & * & * & * \\
\clubsuit_{11}\,\varepsilon & * & +1 & \clubsuit_{12}\,\varepsilon & * & * & \clubsuit_{13}\,\varepsilon & * \\
\hline
* & * & * & +1 & * & * & * & * \\
* & * & * & * & +1 & * & * & * \\
\clubsuit_{21}\,\varepsilon & * & * & \clubsuit_{22}\,\varepsilon & * & +1 & \clubsuit_{23}\,\varepsilon & * \\
\hline
* & * & * & * & * & * & +1 & * \\
\clubsuit_{31}\,\varepsilon & * & * & \clubsuit_{32}\,\varepsilon & * & * & \clubsuit_{33}\,\varepsilon & +1
\end{array}\right],$$

$$M_2 = \left[\begin{array}{ccc|ccc|cc}
+1 & * & * & * & * & * & * & * \\
* & +1 & * & * & * & * & * & * \\
\spadesuit_{11}\,\varepsilon & * & +1 & \spadesuit_{12}\,\varepsilon & * & * & \spadesuit_{13}\,\varepsilon & * \\
\hline
* & * & * & +1 & * & * & * & * \\
* & * & * & * & +1 & * & * & * \\
\spadesuit_{21}\,\varepsilon & * & * & \spadesuit_{22}\,\varepsilon & * & +1 & \spadesuit_{23}\,\varepsilon & * \\
\hline
* & * & * & * & * & * & +1 & * \\
\spadesuit_{31}\,\varepsilon & * & * & \spadesuit_{32}\,\varepsilon & * & * & \spadesuit_{33}\,\varepsilon & +1
\end{array}\right].$$

The reason these entries are singled out is that the club (resp., the spade) entries are precisely the entries of the nested matrices $\Phi_j(B)$ (resp., $\Phi_j(C)$) defined in (2.29).

We are now in the position of introducing the following auxiliary result, which will be the basis of the proof of Theorem 3.4.

Lemma 3.5. *Let $k \in \{1, \ldots, m\}$ and $\gamma \in \widehat{\Xi}_k$, where $\widehat{\Xi}_k$ is as defined in (3.14). Let $\theta(i) = \gamma(i) + 1, i = 1, \ldots, k$, and let η be the cardinal of the set*[1] *$\gamma \bigcap \theta$. Then the lowest possible order in ε of*

$$\det\left(M_1[\beta \,|\, \gamma]\right) \det\left(M_2[\theta \,|\, \beta]\right) \tag{3.15}$$

for any $\beta \in \Xi_k$ is $O\left(\varepsilon^{k-\eta}\right)$, and the order $k - \eta$ is actually attained if and only if β satisfies the two following properties:

$$\begin{aligned} \beta &\supset \gamma \bigcap \theta, \\ \beta &\subset \gamma \bigcup \theta. \end{aligned} \tag{3.16}$$

Proof: our goal is to make the exponent of ε as small as possible by choosing an appropriate β. Let β be any set of indices in Ξ_k and let $\eta_1 = \operatorname{card}(\beta \bigcap \gamma)$, $\eta_2 = \operatorname{card}(\beta \bigcap \theta)$. Then it is easy to check that

$$\begin{aligned} \det\left(M_1[\beta \,|\, \gamma]\right) &= O\left(\varepsilon^{k-\eta_1}\right) \text{ and} \\ \det\left(M_2[\theta \,|\, \beta]\right) &= O\left(\varepsilon^{k-\eta_2}\right), \end{aligned}$$

since all entries in both M_1 and M_2 are $O(\varepsilon)$, except diagonal ones, which are $O(1)$. Hence, the number of $O(1)$ entries in $M_1[\beta \,|\, \gamma]$ (resp. in $M_2[\theta \,|\, \beta]$) is precisely the cardinal of the intersection between β and γ (resp. β and θ). Therefore,

$$\det\left(M_1[\beta \,|\, \gamma]\right) \det\left(M_2[\theta \,|\, \beta]\right) = O\left(\varepsilon^{2k-\eta_1-\eta_2}\right)$$

Now, to make $\eta_1 + \eta_2$ as large as possible , we need the index set β to have as much overlapping as possible with both γ and θ , in order to include as many 1 entries as possible in the minors $M_1[\beta \,|\, \gamma]$ and $M_2[\theta, \beta]$. One can easily check that this is equivalent to both conditions in (3.16).

□

The importance of Lemma 3.5 lies in identifying the situations which produce the lowest possible order in ε as those when η takes its maximum possible value. Obviously, the more consecutive elements γ has, the more common elements the sets γ and $\theta = \{i + 1 \,|\, i \in \gamma\}$ will have. These are the ideas we will be using in the following subsection to characterize the principal minors of $\widehat{H}(\varepsilon)$ with lower order in ε for a given size k.

3.1.5 Proof of Theorem 3.4

Let $l = f_{j-1} + \rho$ for some index $j \in \{1, \ldots, q\}$ with $n_j \geq 2$ and $1 \leq \rho \leq r_j$. We must prove two things:

[1] Strictly speaking, γ and θ are not sets, but lists. However, we have chosen not to make this distinction explicit, since that would complicate the notation even more. In other words, from now on we identify, when needed, each increasingly ordered list with the corresponding set of indices.

(i) there exist k-by-k principal minors of $\widehat{H}(\varepsilon)$ of order $O(\varepsilon^l)$, where

$$k = r_1(n_1 - 1) + \ldots + r_{j-1}(n_{j-1} - 1) + \rho(n_j - 1);$$

(ii) any $(k+1)$-by-$(k+1)$ principal minor of $\widehat{H}(\varepsilon)$ is $O(\varepsilon^{l+1})$.

The first statement is easy to prove: take γ as the set of indices of all rows, except the last one, in each of the first (i.e., largest) l Jordan blocks in $\mathcal{J}_0$. Then, the cardinal of γ is k and there are exactly l indices in $\{i+1 \mid i \in \gamma\}$ which are not in γ, so

$$\mathrm{card}\left(\gamma \bigcap \theta\right) = k - l.$$

Hence, Lemma 3.5 shows the existence of index sets β with

$$\det\left(M_1[\beta \mid \gamma]\right) \det\left(M_2[\theta \mid \beta]\right) = O\left(\varepsilon^l\right).$$

As to statement (ii), notice that, since the choice of γ excludes the last row of each Jordan block (recall that $\gamma \subset \{1, \ldots, m\} \setminus \Omega$ for Ω given by (3.13)), any choice of $k+1$ rows must be taken from at least $l+1$ Jordan blocks. As a consequence, the cardinal of $\gamma \cap \theta$ becomes $k - (l+1)$, and the corresponding principal minor of $\widehat{H}(\varepsilon)$ is $O\left(\varepsilon^{l+1}\right)$ as claimed.

□

It should be noted that, in fact, the restriction $\gamma \subset \{1, \ldots, m\} \setminus \Omega$ implies that the only possible choices of $k(l)$ rows for γ in order to produce principal minors of order $O(\varepsilon^l)$ are those described when proving statement (i) above, i.e., to choose all rows, except the last one, taken from l among the largest Jordan blocks (notice that if $\rho < f_j$, the choice of l largest Jordan blocks is not unique).

We are finally in the position of proving the main result in this paper.

3.1.6 Proof of Theorem 3.3

Up to this point we have already established that under the conditions in the statement of Theorem 3.3 the Newton Diagram associated with the polynomial

$$\pi(z\,,\,\varepsilon) = \det\left(zI - \widehat{H}_1(\varepsilon)\right) \tag{3.17}$$

is generically the concatenation of the q segments of slopes $1/(n_j - 1),\ j = 1, \ldots, q$, joining the $q+1$ points $(k(f_j), f_j),\ j = 0, 1, \ldots, q$, where $k(\cdot)$ is given by (3.8) and we make the convention $f_0 = k(f_0) = 0$. Hence, the leading exponents in the asymptotic expansion (3.5) are correct and it only remains to show that the leading coefficients ξ_j^k are given by formula (3.6).

In order to do that, we will find formulas for the coefficients $\widehat{\alpha}_{k(l)}$ associated with each point $(k(l), l)$ in the Newton Diagram (recall that $(k(l), l)$ lies on the diagram if and only if $\widehat{\alpha}_{k(l)} \neq 0$). Such $\widehat{\alpha}_{k(l)}$ are (up to a sign) just the leading coefficients in the sum (3.12) of $k(l) \times k(l)$ principal minors of $\widehat{H}(\varepsilon)$.

Let $l = f_{j-1} + \rho$ for some index $j \in \{1, \ldots, q\}$ with $n_j \geq 2$ and $1 \leq \rho \leq r_j$. Our first goal is to show that the coefficient of $\varepsilon^l z^{m-k(l)}$ in (3.17) is

$$\widehat{\alpha}_{k(l)} = (-1)^l \sum_{\varpi} \det\left(\Phi_j(B+C)[\varpi]\right) \tag{3.18}$$

where $k(l)$ is given by (3.8) and the sum on ϖ is over all index sets with cardinality l whose first f_{j-1} entries are $1, \ldots, f_{j-1}$ and whose remaining ρ entries are taken from the set $\{f_{j-1}+1, \ldots, f_j\}$.

We have seen in §3.1.4 and §3.1.5 that a term of order $O(\varepsilon^l)$ can only be obtained in the sum (3.14) if the corresponding principal minor is associated with sets γ, θ and β, each with cardinal $k(l)$, satisfying the following conditions:

$$\begin{array}{ll} 1. & \gamma \text{ contains the indices of all rows, except the last one, taken among } l \text{ of the largest blocks} \\ & \text{in } \mathcal{J}_0 \text{ (if } \rho < f_j\text{, the choice of } l \text{ largest Jordan blocks is not unique) ;} \\ 2. & \theta = \{i+1 \mid i \in \gamma\}\,; \\ 3. & \beta \subset \gamma \bigcup \theta; \\ 4. & \beta \supset \gamma \bigcap \theta. \end{array} \tag{3.19}$$

Since we are only interested in the leading terms of products of the form (3.15), we may replace minors of the ε-dependent matrices $M_1 = I + \varepsilon B_{11}$ and $M_2 = I + \varepsilon C_{11}$ with 'simplified minors' of constant matrices as follows:

- Replace every entry of type $1 + O(\varepsilon)$ in the appropriate submatrix of M_1 or M_2 by 1, and set all entries in the same row or column of the submatrix to zero;

- For each of the remaining entries in the submatrix, if any, replace the entry by its coefficient in ε (recall that if any entry remains, it must be $O(\varepsilon)$).

If we denote the simplified minors with $\widetilde{\ }$, it is obvious that if γ, θ and β satisfy conditions 1.– 4. in (3.19) above, then

$$\det\left(M_1[\beta \mid \gamma]\right) \det\left(M_2[\theta \mid \beta]\right) = \det\left(\widetilde{M_1}[\beta \mid \gamma]\right) \det\left(\widetilde{M_2}[\theta \mid \beta]\right) \varepsilon^l + O\left(\varepsilon^{l+1}\right) \tag{3.20}$$

In order not to complicate the proof unnecessarily, we will only analyze in full the case when $\gamma = \{1, \ldots, k(l) + l\} \setminus \Omega$. Any other choice for γ can be analyzed analogously.

The way we will identify the coefficients of all $O(\varepsilon^l)$ terms of the form (3.20) is by

(i) first locating all the $O(1)$ entries in both $\widetilde{M_1}[\beta \mid \gamma]$ and $\widetilde{M_2}[\theta \mid \beta]$, and then

(ii) expanding the corresponding minors along the rows where those $O(1)$ entries lie.

Of course, such locations will very much depend on the set β. Table 3.1.6 below summarizes the relevant information on all possible positions of the $O(1)$ entries, as well as the size of the diagonal blocks (according to the Jordan partition induced by $\mathcal{J}_0$) including them. We distinguish four cases, depending on whether the index of the first row, say r_F, and the index of the last row, say r_L, of the diagonal block is included in β or not.
In order to illustrate this and simplify subsequent proofs let us introduce some examples based on the 8×8 example introduced in (3.4) with $l = 3$. Here $k(3) = 5$ and we have only one possibility for γ:

$$\Omega = \{3\,,\,6\,,\,8\},\ \gamma = \{1\,,\,2\,,\,4\,,\,5\,,\,7\},\ \theta = \{2\,,\,3\,,\,5\,,\,6\,,\,8\}$$

Now, depending on the choice of the index set β satisfying conditions (3.16), there are different possibilities:

Example 3.6. $4, 7, 8 \in \beta \quad \wedge \quad 1, 3, 6 \notin \beta$
Here we have chosen the indices in β in such a way that the first diagonal block is in Case 4, as described in Table 3.1.6, the second block is in Case 2, and the third block in Case 1. Hence, $\beta = \{1\,,\,2\,,\,3\,,\,5\,,\,7\}$ and

$$\widetilde{M_1}[\beta\,|\,\gamma] = \left[\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \clubsuit_{31} & 0 & 0 & 0 & 0 \end{array}\right], \qquad \widetilde{M_2}[\theta\,|\,\beta] = \left[\begin{array}{c|cc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & \spadesuit_{12} & 0 & \spadesuit_{13} & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & \spadesuit_{22} & 0 & \spadesuit_{23} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Example 3.7. $1, 4, 7 \in \beta \quad \wedge \quad 3, 6, 8 \notin \beta$
In this case we have chosen the indices in β so that all three diagonal blocks are in Case 2, so $\beta = \{1\,,\,2\,,\,4\,,\,5\,,\,7\}$ and

$$\widetilde{M_1}[\beta\,|\,\gamma] = \left[\begin{array}{cc|cc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right], \qquad \widetilde{M_2}[\theta\,|\,\beta] = \left[\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & 0 \\ \spadesuit_{11} & 0 & \spadesuit_{12} & 0 & \spadesuit_{13} \\ \hline 0 & 0 & 0 & 1 & 0 \\ \spadesuit_{21} & 0 & \spadesuit_{22} & 0 & \spadesuit_{23} \\ \hline \spadesuit_{31} & 0 & \spadesuit_{32} & 0 & \spadesuit_{33} \end{array}\right]$$

Example 3.8. $3, 4, 8 \in \beta \quad \wedge \quad 1, 6, 7 \notin \beta$
Finally, we consider β such that the first and third diagonal blocks are in Case 3 and the second one is in Case 2. Now, $\beta = \{2\,,\,3\,,\,4\,,\,5\,,\,8\}$ and

$$\widetilde{M_1}[\beta\,|\,\gamma] = \left[\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & 0 \\ \clubsuit_{11} & 0 & 0 & 0 & \clubsuit_{13} \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline \clubsuit_{31} & 0 & 0 & 0 & \clubsuit_{33} \end{array}\right], \qquad \widetilde{M_2}[\theta\,|\,\beta] = \left[\begin{array}{cc|cc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \spadesuit_{22} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Cases	Diagonal block of $\widetilde{M_1}[\beta \mid \gamma]$	Diagonal block of $\widetilde{M_2}[\theta \mid \beta]$
$1) \to \begin{cases} r_F \in \beta \\ r_L \in \beta \end{cases}$	$n_i \times (n_i - 1)$ block with the 1s on the main diagonal. The last row does not contain a 1.	$(n_i - 1) \times n_i$ block with the 1s on the superdiagonal. The first column does not contain a 1.
$2) \to \begin{cases} r_F \in \beta \\ r_L \notin \beta \end{cases}$	The block is the identity matrix of order $(n_i - 1)$.	An $(n_i - 1) \times (n_i - 1)$ block with the 1s on the superdiagonal. Neither the first column nor the last row contain a 1.
$3) \to \begin{cases} r_F \notin \beta \\ r_L \in \beta \end{cases}$	$(n_i - 1) \times (n_i - 1)$ block with the 1s on the superdiagonal. Neither the first column nor the last row contain a 1.	The block is the identity matrix of order $(n_i - 1)$.
$4) \to \begin{cases} r_F \notin \beta \\ r_L \notin \beta \end{cases}$	$(n_i - 2) \times (n_i - 1)$ block with the 1s on the superdiagonal. The first column does not contain a 1.	$(n_i - 1) \times (n_i - 2)$ block with the 1s on the superdiagonal. The last row does not contain a 1.

Table 3.1.6

One can check that in this configuration there are twenty different possible choices for β. If we compute $\mathbb{E}_5\left[\widehat{H}(\varepsilon)\right]$ via (3.14) and (3.20) we obtain

$$\begin{aligned}
\mathbb{E}_5\left[\widehat{H}(\varepsilon)\right] &= \Bigg((\clubsuit_{31}) \det \begin{bmatrix} \spadesuit_{12} & \spadesuit_{13} \\ \spadesuit_{22} & \spadesuit_{23} \end{bmatrix} + \det \begin{bmatrix} \spadesuit_{11} & \spadesuit_{12} & \spadesuit_{13} \\ \spadesuit_{21} & \spadesuit_{22} & \spadesuit_{23} \\ \spadesuit_{31} & \spadesuit_{32} & \spadesuit_{33} \end{bmatrix} + \\
&\qquad - \det \begin{bmatrix} \clubsuit_{11} & \clubsuit_{13} \\ \clubsuit_{31} & \clubsuit_{33} \end{bmatrix} (-\spadesuit_{22}) + \cdots \Bigg) \varepsilon^3 + O(\varepsilon^4) \\
&= \det\left(\Phi_2(B) + \Phi_2(C)\right) \varepsilon^3 + O(\varepsilon^4).
\end{aligned}$$

In general by construction, the only entries in the modified submatrix $\widetilde{M_1}[\beta \mid \gamma]$ (resp., $\widetilde{M_2}[\theta \mid \beta]$) which are not zero or 1 are placed in those rows and columns not containing a 1, and are just the coefficients in ε of entries of $\Phi_j(B)$ (resp. $\Phi_j(C)$). Thus, if we expand the determinants in (3.15) along the rows containing 1s, then for each choice of β the leading coefficient in the product (3.15) will be, up to a sign, just a product of two appropriately chosen minors of $\Phi_j(B)$ and $\Phi_j(C)$.

The sign, of course, will depend on the positions of the 1s in both $\widetilde{M_1}[\beta\,|\,\gamma]$ and $\widetilde{M_2}[\theta\,|\,\beta]$. In order to identify both the sign and the minors we need to further specify, for each diagonal block, to which of the four cases in Table 3.1.6 it belongs: for each $i \in \{1,\dots,l\}$ we include the index i in either of four increasingly ordered lists $\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4$ depending on which, among the four cases in Table 3.1.6, is the one corresponding to the i-th diagonal block.

Since we will need to concatenate some of these lists, for each pair ξ, χ of disjoint lists of indices with lengths a and b, we denote by (ξ,χ) the list of length $a+b$ obtained from concatenating ξ and χ in that order, i.e.,

$$(\xi,\chi) = \xi(1),\,\dots,\,\xi(a),\,\chi(1),\,\dots,\,\chi(b).$$

Furthermore, we denote by $\xi + \chi$ the increasingly ordered list obtained from reordering the concatenation (ξ,χ) and by $\operatorname{sgn}(\xi,\chi)$ the sign of the permutation $\begin{pmatrix} \xi+\chi \\ \xi,\chi \end{pmatrix}$ which transforms the reordered list $\xi+\chi$ into the concatenation (ξ,χ). Finally, for each increasingly ordered list υ with a indices taken from $\{1,\dots,l\}$, we denote by υ^c its complementary list, i.e., the list of length $l-a$ containing those indices in $\{1,\dots,l\}$ which are not in υ, increasingly ordered.

With this notation, one can prove that, if we denote

$$\vartheta = \upsilon_1 + \upsilon_3, \qquad \zeta = \upsilon_3 + \upsilon_4,$$

then

$$\det\left(\widetilde{M_1}[\beta\,|\,\gamma]\right)\det\left(\widetilde{M_2}[\theta\,|\,\beta]\right) = S\,\det\left(\Phi_j(B)[\vartheta\,|\,\zeta]\right)\det\left(\Phi_j(C)[\vartheta^c\,|\,\zeta^c]\right) \tag{3.21}$$

where S is the sign

$$S = (-1)^{k(l)-l}\operatorname{sgn}(\vartheta,\vartheta^c)\operatorname{sgn}(\zeta,\zeta^c) \tag{3.22}$$

for $k(l)$ given by (3.8). Since the proof of (3.21) is not central to our argument, we defer it to the Appendix.

Obviously, the four lists υ_1, υ_2, υ_3, υ_4 amount to a partition of $\{1,\dots,m\}$. Furthermore, notice that, according to Table 3.1.6,

1. each block whose index is in υ_1 contributes n_i indices to the set β;

2. each block whose index is either in υ_2 or in υ_3 contributes $n_i - 1$ indices to the set β; and

3. each block whose index is in υ_4 contributes $n_i - 2$ indices to the set β.

Since we know that the cardinal of β is

$$k(l) = r_1(n_1 - 1) + \ldots + r_{j-1}(n_{j-1} - 1) + \rho(n_j - 1),$$

we conclude that *the lengths of* υ_1 *and* υ_4 *coincide*.

Now that we have a formula for each product (3.15), we focus on the inner sum in (3.14), which runs over all index sets β for a fixed γ. Recall that we have fixed γ by taking rows from only the first l Jordan blocks, and that each choice of β induces a choice of υ_1, υ_2, υ_3, υ_4 and, therefore, of ϑ, ζ, in such a way that υ_1 and υ_4 have the same length. Let us now show that taking all possible choices for β is equivalent to making all possible choices of pairs (ϑ, ζ) in the cartesian product $\Lambda_i \times \Lambda_i$, where Λ_i denotes the family of all index sets with entries in $\{1, \ldots, l\}$ and cardinality

$$i = \mathrm{card}(\vartheta) = \mathrm{card}(\upsilon_3) + \mathrm{card}(\upsilon_1) = \mathrm{card}(\upsilon_3) + \mathrm{card}(\upsilon_4) = \mathrm{card}(\zeta),$$

with i varying from 0 to l: on one hand, we have already seen that every choice of β produces two sets ϑ, ζ of the same cardinal i for some appropriate i, satisfying the constraint on υ_1 and υ_4. We only need to show that, given $i \in \{0, \ldots, l\}$, and given any two sets $\vartheta, \zeta \in \Lambda_i$, we can uniquely define four lists υ_1, υ_2, υ_3, υ_4 covering all indices in $\{1, \ldots, l\}$ and such that the lengths of υ_1 and υ_4 coincide. One can easily check that

$$\begin{aligned} \upsilon_1 &= \vartheta \setminus \zeta \\ \upsilon_2 &= \{1, \ldots, l\} \setminus \left(\vartheta \bigcap \zeta\right) \\ \upsilon_3 &= \vartheta \bigcap \zeta \\ \upsilon_4 &= \zeta \setminus \vartheta \end{aligned}$$

is such a choice, and it is the only one satisfying the required conditions. Hence, we may use (3.21) to rewrite the coefficient in ε^l of

$$\sum_{\beta} \det\left(M_1[\beta \,|\, \gamma]\right) \det\left(M_2[\theta \,|\, \beta]\right)$$

in formula (3.14) as

$$\sum_{i=0}^{l} \sum_{\vartheta, \in \Lambda_i} \sum_{\zeta \in \Lambda_i} S \, \det\left(\Phi_j(B)[\vartheta \,|\, \zeta]\right) \det\left(\Phi_j(C)[\widetilde{\vartheta} \,|\, \widetilde{\zeta}]\right) \tag{3.23}$$

with S given by (3.22).

We now make use of the following technical lemma, whose proof is also deferred to the Appendix.

Lemma 3.9. *Let $M,\ N \in \mathbb{C}^{l\times l}$ and, for $i = 0,\ \ldots,\ l$, let Λ_i be the family of increasingly ordered lists of indices, taken from $\{1,\ \ldots,\ l\}$, and with length i. Then*

$$\det(M+N) = \sum_{i=0}^{l} \sum_{\vartheta\in\Lambda_i} \sum_{\zeta\in\Lambda_i} \operatorname{sgn}(\vartheta,\vartheta^c)\operatorname{sgn}(\zeta,\zeta^c)\det\left(M\left(\vartheta\,|\,\zeta\right)\right)\det\left(N\left(\vartheta^c\,|\,\zeta^c\right)\right), \tag{3.24}$$

where ζ^c (resp. ϑ^c) denotes the complementary list of ζ (resp. ϑ) in $\{1,\ldots,l\}$.

This Lemma implies that, since $\Phi_j(B+C) = \Phi_j(B)+\Phi_j(C)$, if we take $M = \Phi_j(B)[1:l]$ and $N = \Phi_j(C)[1:l]$, then (3.14) is equal to

$$(-1)^{k(l)-l}\det\left(\Phi_j(B+C)[1:l]\right).$$

This, of course, corresponds to our initial simplifying choice $\gamma = \{1,\ \ldots,\ k(l)+l\}\setminus\Omega$. If we repeat the same argument for any other admissible choice of γ in the outer sum of (3.14), we obtain

$$\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right] = (-1)^{k(l)-l}\varepsilon^l\sum_{\varpi}\det\left(\Phi_j(B+C)[\varpi]\right)+O\left(\varepsilon^{l+1}\right) \tag{3.25}$$

where the sum on ϖ is over all index sets with cardinality l whose first f_{j-1} entries are $1,\ldots,f_{j-1}$ and whose remaining ρ entries are taken from the set $\{f_{j-1}+1,\ \ldots,\ f_j\}$. Thus, using (3.12) we obtain

$$\widehat{\alpha}_k = (-1)^l\sum_{\varpi}\det\left(\Phi_j(B+C)[\varpi]\right).$$

In other words, except for the sign, $\widehat{\alpha}_k$ is the sum of all $k\times k$ principal minors of $\Phi_j(B+C)$ which include the first f_{j-1} rows, together with ρ rows chosen among the last r_j ones.

Now, we know from (2.21) that the leading coefficients $\xi_{j,k}$ in Theorem 3.3 are solutions of a polynomial equation

$$\sum_{k\in I_{S_j}}\widehat{\alpha}_k\,\mu^{k_{max}-k} = 0, \tag{3.26}$$

where $k_{max} = k(f_j)$ and $I_{S_j} = \{k(f_{j-1})+\rho(n_j-1)\ :\ \rho = 0,\ \ldots,\ r_j\}$ is the set of possible abscissae for points on the segment S_j of the Newton diagram joining $(k(f_{j-1}), f_{j-1})$ and $(k(f_j), f_j)$. Of course, only those terms with nonzero $\widehat{\alpha}_k$ appear on the equation. If we multiply (3.26) by $(-1)^{f_{j-1}}(\det\left(\Phi_{j-1}(B+C)\right))^{-1}$ and make the change of variables $\xi = \mu^{n_j-1}$, we get the equivalent equation

$$\xi^{r_j} - \frac{\mathbb{E}_1^*[\Phi_j(B+C)]}{\det\left(\Phi_{j-1}(B+C)\right)}\xi^{r_j-1}+\ldots+(-1)^{r_j}\frac{\mathbb{E}_{r_j}^*[\Phi_j(B+C)]}{\det\left(\Phi_{j-1}(B+C)\right)} = 0 \tag{3.27}$$

where $E_\rho^*[\Phi_j(B+C)]$, $\rho = 1, \ldots, r_j$, stand for the sum of all principal minors of $\Phi_j(B+C)$ including the first f_{j-1} rows together with ρ rows chosen from the last r_j ones.

Finally, consider a principal minor of $\Phi_j(B+C)$ including its first f_{j-1} rows and ρ other rows among the r_j last ones. Then such a minor is just the determinant of a matrix

$$M = \left[\begin{array}{c|c} \Phi_{j-1}(B+C) & M_{12} \\ \hline M_{21} & M_{22} \end{array}\right],$$

and the basic properties of Schur complements imply that

$$\det\left(M_{22} - M_{21}\Phi_{j-1}^{-1}(B+C)M_{12}\right) = \frac{\det(M)}{\det\left(\Phi_{j-1}(B+C)\right)}.$$

Hence, if we denote by Ξ_j the Schur complement of $\Phi_{j-1}(B+C)$ in $\Phi_j(B+C)$, one can easily prove that, for $\rho = 1, \ldots, r_j$,

$$\frac{\mathbb{E}_\rho^*[\Phi_j(B+C)]}{\det\left(\Phi_{j-1}(B+C)\right)} = \mathbb{E}_\rho[\Xi_j],$$

so equation (3.27) may be rewritten as

$$\xi^{r_j} - \mathbb{E}_1\left[\Xi_j\right]\xi^{r_j-1} + \ldots + (-1)^{r_j}\mathbb{E}_{r_j}\left[\Xi_j\right] = 0.$$

But this is just the characteristic equation of Ξ_j, so the solutions of equation (3.26) or, equivalently, of (3.27), are just the eigenvalues of the Schur complement Ξ_j of $\Phi_{j-1}(B)$ in $\Phi_j(B)$. This completes the proof of Theorem 3.3.

3.2 Asymptotic singular value expansions for multiplicative perturbations

All the ideas above can be easily translated into the context of multiplicative perturbation of singular values: let

$$A = U\,\Sigma\,V^*, \quad \Sigma = \begin{bmatrix} \Sigma_n \\ 0 \end{bmatrix}, \quad \Sigma_n = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}$$

be a singular value decomposition of $A \in \mathbb{C}^{m\times n}$ with $m \geq n$. It is well known [58] that the Hermitian $(m+n)\times(m+n)$ matrix

$$M = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \tag{3.28}$$

has $2n$ eigenvalues $\pm\sigma_i$, $i = 1, \ldots, n$, plus $m-n$ zero eigenvalues. Furthermore, if we partition $U = [\,U_1\,|\,U_2]$, with $U_1 \in \mathbb{C}^{m\times n}$, then M can be unitarily diagonalized as

$$\begin{bmatrix} \Sigma_n & & \\ & -\Sigma_n & \\ & & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} U_1^* & V^* \\ U_1^* & -V^* \\ \sqrt{2}U_2^* & 0 \end{bmatrix}\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}\begin{bmatrix} U_1 & U_1 & \sqrt{2}U_2 \\ V & -V & 0 \end{bmatrix}$$

In this setting, it is straightforward to prove the following result on multiplicative perturbation of nonzero singular values (recall that zero singular values of A correspond to zero eigenvalues of M, which are unchanged by multiplicative perturbations):

Corollary 3.10. *Let $A \in \mathbb{C}^{m\times n}$, $m \geq n$, and let σ_0 be a nonzero singular value of A with multiplicity k. Let $U_0 \in \mathbb{C}^{m\times k}$ and $V_0 \in \mathbb{C}^{n\times k}$ be matrices whose columns span simultaneous bases of the respective left and right singular subspaces of A associated with σ_0. Then, for any $B \in \mathbb{C}^{m\times m}$ and $C \in \mathbb{C}^{n\times n}$, the matrix $\widehat{A}(\varepsilon) = (I + \varepsilon B)A(I + \varepsilon\, C)$ has k singular values analytic in ε which can be expanded as*

$$\sigma_j(\varepsilon) = \sigma_0\left(1 + \xi_j\ \varepsilon + o(\varepsilon)\right), \tag{3.29}$$

where the $\xi_j,\ j = 1, \ldots, k$ are the eigenvalues of the $k \times k$ Hermitian matrix

$$\Phi = \frac{1}{2}\left(U_0^*\left(B + B^*\right)U_0 + V_0^*\left(C + C^*\right)V_0\right).$$

Proof. We view the nonzero singular values of $\widehat{A}(\varepsilon) = (I+\varepsilon B)A(I+\varepsilon C)$ as the positive eigenvalues of $\widehat{M} = (I + \varepsilon\widetilde{B})M(I + \varepsilon\widetilde{C})$ for

$$\widetilde{B} = \begin{bmatrix} B & 0 \\ 0 & C^* \end{bmatrix},\quad \widetilde{C} = \begin{bmatrix} B^* & 0 \\ 0 & C \end{bmatrix},$$

and σ_0 as the unperturbed eigenvalue of the matrix M in (3.28) with (algebraic and geometric) multiplicity k. Hence, we are in the simplest case

$$q = n_1 = 1, \qquad r_1 = k$$

of Theorem 3.3. Since the columns of

$$X_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} U_0 \\ V_0 \end{bmatrix} \in \mathbb{C}^{(m+n)\times k}$$

form an orthonormal basis of the right eigenspace associated with the semisimple eigenvalue σ_0 of M, then the correspondent left eigenvectors belonging to the same Jordan form are $Y_1 = X_1^*$. Hence, an straightforward application of Theorem 3.3 leads to the expansions (3.29).

□

Chapter 4

Structured perturbation of eigenvalues of symplectic matrices

As announced in the introduction, an important application of the multiplicative perturbation theory developed in the previous Chapter is to obtain asymptotic expansions for eigenvalues of small structured perturbations of symplectic matrices. Our goal in this chapter is to develop a first order perturbation theory under structure-preserving perturbations for eigenvalues of J-symplectic matrices. To be more precise, given a (possibly multiple) eigenvalue λ of a J-symplectic matrix S, we consider another J-symplectic matrix $\widehat{S}$ close to S, and find asymptotic expansions of the eigenvalues of $\widehat{S}$ close to λ by interpreting $\widehat{S} = \widehat{S}(\varepsilon)$ as a particular value of an analytic J-symplectic matrix function $\widehat{S}(\cdot)$ depending on a real parameter ε.

We shall see that, although the theory we develop in this Chapter covers most possible situations, there is a very specific one, when the unperturbed eigenvalue is one of the critical eigenvalues ± 1, which is not covered by it. This highly nongeneric situation, already identified in [38], will require a completely different approach, which will be presented in Chapter 5 below.

4.1 Preliminaries

4.1.1 Small multiplicative symplectic perturbations

We start our analysis by showing that, as announced above, any analytic J-symplectic matrix function $\widehat{S}(\varepsilon)$ in a real parameter ε with $\widehat{S}(0) = S$ can be written multiplicatively in the form (1.3) for some J-Hamiltonian matrix B: if $\|\cdot\|$ is any matrix norm then we can write

$$\frac{\|\widehat{S}(\varepsilon) - S\|}{\|S\|} = O(\varepsilon).$$

Since S is J-symplectic, and therefore invertible, we have

$$\frac{\|\widehat{S}(\varepsilon) - S\|}{\|S\|} = \frac{\|\widehat{S}(\varepsilon) - S\|\|S^{-1}\|}{\kappa(S)} \geq \frac{\|\widehat{S}(\varepsilon)\, S^{-1} - I_{2n}\|}{\kappa(S)} \,,$$

where $\kappa(S) = \|S\|\|S^{-1}\|$ is the condition number of S in the norm of choice. Hence,

$$\|\widehat{S}(\varepsilon)\, S^{-1} - I_{2n}\| \leq \kappa(S)\, O(\varepsilon)$$

and consequently the matrix function $\widehat{I}(\varepsilon) := \widehat{S}(\varepsilon)\, S^{-1}$ satisfies $\widehat{I}(0) = I_{2n}$ and can be expanded as a power series

$$\widehat{I}(\varepsilon) = I_{2n} + \varepsilon B + O(\varepsilon^2) \tag{4.1}$$

for some matrix $B \in \mathbb{C}^{2n \times 2n}$. Now, recall that the set of J-symplectic matrices forms a (Lie) group with respect to matrix multiplication. Hence, $\widehat{I}(\varepsilon)$ is J-symplectic which implies

$$J = \widehat{I}(\varepsilon)^T J\, \widehat{I}(\varepsilon) = J + \varepsilon(B^T J + JB) + O(\varepsilon^2).$$

Comparing the coefficients on both sides yields $B^T J + JB = 0$, i.e., B must be J-Hamiltonian. Therefore, any small structure-preserving perturbation $\widehat{S}$ of a J-symplectic matrix, can be modelled as

$$\widehat{S} = \widehat{S}(\varepsilon) = \widehat{I}(\varepsilon)\, S = \big(I_{2n} + \varepsilon B + O(\varepsilon^2)\big)\, S \,, \tag{4.2}$$

where B is J-Hamiltonian. Notice that parametrizing the perturbation in terms of ε as in (1.3) opens the way for describing the perturbed eigenvalues $\lambda(\varepsilon)$ of $\widehat{S}(\varepsilon)$ through fractional power expansions in ε (see [23, 4]).

We conclude by observing that *any* arbitrary J-Hamiltonian matrix can take the role of the matrix B in the multiplicative perturbation as in (1.3). To show this, recall that the matrix $\widehat{I}(\varepsilon)$ in (4.1) is J-symplectic, and its upper left $n \times n$ block is nonsingular for ε small enough. Thus, we can make use of a characterization due to Dopico & Johnson of the subset of $2n \times 2n$ symplectic matrices with regular upper left $n \times n$ block, namely Theorem 3.1 in [12]. Adapting that result to the particular structure (1.3) of our perturbations, it is straightforward to prove the following:

Lemma 4.1. *Let $S \in \mathbb{C}^{2n \times 2n}$ be J-symplectic. Then a complex $2n \times 2n$ matrix*

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

partitioned into $n \times n$ blocks, is the first-order coefficient of a J-symplectic analytic matrix function $\widehat{S}(\varepsilon)$ of the form (1.3) *if and only if $B_{22} = -B_{11}^T$ and B_{12}, B_{21} are complex symmetric.*

This result will also be useful in Section 4.5 below when generating numerical examples. Before we start the structured perturbation analysis, we highlight some preliminary observations one can easily make about Theorem 3.2, the multiplicative result from Chapter 3 we will be using in this chapter, since symplectic matrices are never singular:

asymptotic expansions:

Remark 4.2. *Theorem 3.2 only speaks of the perturbed eigenvalues with lowest possible leading exponent, i.e., the ones moving fastest away from λ, which are the ones depending only on the first-order coefficient matrix B. Of course the $O(\varepsilon^2)$ terms in* (1.3) *may increase the rank of the perturbation $\widehat{S}(\varepsilon)-S$, and consequently destroy further Jordan blocks of S, which gives rise to other perturbed eigenvalues, not covered by Theorem 3.2. But these eigenvalues will have leading exponents larger than $1/n_j$, and will be left out of our analysis. Section 4.5 below provides many instances when this happens.*

Remark 4.3. *Theorem 3.2 does not say that $\widehat{S}$ has exactly $r_j n_j$ eigenvalues of order ε^{1/n_j}. This only happens if all r_j eigenvalues of the Schur complement $\left(\Phi_j/\Phi_{j-1}\right)(B)$ are nonzero. To be more precise, let $r=$rank(B). Then there exist two positive integers $j\in\{1,\,...\,,q\}$ and $\rho\in\{1,\,...\,,r_j\}$ such that*

$$r=\sum_{k=1}^{j-1} r_k+\rho=f_{j-1}+\rho \tag{4.3}$$

where f_{j-1} is as defined in (2.28). *Then, for every index $i\in\{1,\,...\,,q\}$, $i<j$, Theorem 3.2 implies that there will be generically $r_i n_i$ eigenvalues of order ε^{1/n_i} generated by the destruction of all r_i Jordan blocks of size n_i. For the index j, however, at most ρn_j eigenvalues of order ε^{1/n_j} may exist in $\widehat{S}$ if $\Phi_{j-1}(B)$ is nonsingular, since rank $\Phi_{j-1}(B)=f_{j-1}=r-\rho$ and, therefore,*

$$\text{rank } \left(\Phi_j/\Phi_{j-1}\right)(B)=\text{rank}\,\Phi_j(B)-\text{rank}\,\Phi_{j-1}(B)\le\rho\,, \tag{4.4}$$

i.e., the Schur complement $\left(\Phi_j/\Phi_{j-1}\right)(B)$ may have at most ρ nonzero eigenvalues. For each of these nonzero eigenvalues, formula (3.2) *provides n_j different expansions, one for each n_j-th root. Notice that the expansions* (3.2) *still formally hold for all perturbed eigenvalues $\widehat{\lambda}_{j,k,l}(\varepsilon)$, but they provide no information whatsoever for the $(r_j-\rho)n_j$ perturbed eigenvalues corresponding to the values $\xi_{j,k}$ that are zero (other than the fact that the corresponding perturbed eigenvalues are of order ε^q for some q larger than $1/n_j$). As to the indices $i>j$, Theorem 3.2 does not apply, since every Φ_{i-1} is singular.*

Remark 4.4. *As long as it corresponds to a nonzero eigenvalue $\xi_{j,k}$ of the Schur complement, the leading term in the expansion* (3.2) *is the same* for all perturbations (1.3) having the same first order matrix coefficient B. *More specifically, let $\widehat{S}(\varepsilon)$ and $\widetilde{S}(\varepsilon)$ be two different J-symplectic perturbations of the form* (1.3) *with the same first order J-Hamiltonian coefficient B, and let the common Schur complement $\left(\Phi_j/\Phi_{j-1}\right)(B)$ have*

ρ nonzero eigenvalues. Then, Theorem 3.2 ensures that both $\widehat{S}(\varepsilon)$ and $\widetilde{S}(\varepsilon)$ have ρn_j eigenvalues of order ε^{1/n_j} with identical leading coefficients. The remaining $(r_j - \rho)n_j$ eigenvalues of $\widehat{S}(\varepsilon)$ and $\widetilde{S}(\varepsilon)$ may differ in their leading coefficients, or even in their leading exponents. It may be even that the number of Jordan blocks destroyed by $\widehat{S}(\varepsilon)$ and by $\widetilde{S}(\varepsilon)$ is different, since the ranks of $\widehat{S} - S$ and of $\widetilde{S} - S$ may be different. For both perturbations, however, at least ρ Jordan blocks of size n_j will be destroyed, and the ρn_j corresponding eigenvalues will display the very same first order behavior, described by Theorem 3.2, for both perturbations. This observation will be relevant in Section 4.5 below, where we will introduce such an auxiliary perturbation $\widetilde{S}$ matrix in order to single out from the numerical examples those eigenvalues which conform to the expansions obtained in Section 4.3.

4.2 Jordan chains for symplectic matrices

Theorem 3.2 shows that the leading coefficients we are looking for are determined by the matrices $\Phi_j(B)$ defined in (2.29), and these largely depend on left and right eigenvectors of S associated with λ. Notice that the choice of these left and all right eigenvectors is not arbitrary: first, they are implicitly normalized by the fact that they show up in the same Jordan canonical form. Our goal in this section is to show that we may further narrow the choice of the Jordan vectors in such a way that a special connection, induced by symplecticity, is revealed between left and right eigenvectors. Our main auxiliary tools to do so will be structured canonical forms for J-symplectic matrices introduced in §2.1.

In the most general case, when all we know is that the unperturbed J-symplectic matrix S is complex, we shall use the complex canonical form in Theorem 2.3. It is important to remark here that the second equation in (2.3) can also be written as

$$\mathcal{T}^T J = (H_1 \oplus \ldots \oplus H_p)\, \mathcal{T}^{-1}, \tag{4.5}$$

which is sometimes more convenient, as we will see later.

If S is a *real* symplectic matrix, then we may also consider the canonical form in Theorem 2.4 to get the Jordan chains associated with real eigenvalues. The only difference with the canonical form in Theorem 2.3 is that now $\mathcal{T} \in \mathbb{R}^{2n\times 2n}$, and thus Jordan chains associated with real eigenvalues obtained from the symplectic canonical form have real entries. For each even-sized block corresponding to $\lambda_j = \pm 1$, a sign ς_j appears in equation (2.9) for H_j,

Finally, the case when $S \in \mathbb{R}^{2n\times 2n}$ and the eigenvalue λ lies on the unit circle is special: here it will be advantageous to consider the real matrix S as a complex matrix which turns out to be (iJ)-unitary, i.e., which satisfies $S^*(iJ)S = iJ$. This allows us to make use of the structured canonical form for H-unitary matrices for Hermitian H in Theorem 2.5. This canonical form distinguishes two cases, just like Theorem 2.3, depending on whether the eigenvalue lies on the unit circle or not. Since we will be only interested in unimodular eigenvalues (for the other ones, either of the two previous canonical forms mentioned above[1] already give better results), we need only describe the blocks S_j and H_j associated with λ_j such that $|\lambda_j| = 1$, namely,

$$S_j = \text{Toep}\left(\lambda_j,\, i\,\lambda_j,\, -i\lambda_j t_2,\, \dots,\, -i\lambda_j t_{n_j-1}\right), \quad H_j = \varsigma_j\, R_{n_j}, \quad \varsigma_j = \pm 1, \tag{4.6}$$

where the parameters t_k are as in (2.5). It should be noted that even though this is also true for $\lambda = \pm 1$, for these two particular eigenvalues it is more advantageous to use the real canonical form, since it provides more detailed information for odd-sized blocks, and it also takes into account the fact that they appear in pairs in the canonical form.

4.2.1 Extracting Jordan bases from bases associated with symplectic canonical forms

Although the structured canonical forms from §2.1 are closely related to the Jordan form, they do not directly display the left and right eigenvectors needed to apply Theorem 3.2. This is why we need to find appropriate changes of basis which similarity-transform the structured symplectic form into the Jordan form. These transformations will reveal the connections induced by the symplectic structure between certain left and right eigenvectors, which will in turn allow us to write more detailed formulas for the leading terms in the asymptotic expansions.

- Let $\lambda \in \mathbb{C} \setminus \{0\}$, let p be a positive integer and let $\mathcal{J}_p(\lambda)$ be an upper triangular Jordan block of size p with eigenvalue λ. Then there exists a matrix D_p such that

  $$D_p^{-1} \mathcal{J}_p(\lambda)^{-T} D_p = \mathcal{J}_p\left(\lambda^{-1}\right) \tag{4.7}$$

  For $p = 2,\, 3,\, 4$ the matrix D_p can be chosen as

  $$D_2 = \begin{bmatrix} 0 & -\lambda^2 \\ 1 & 0 \end{bmatrix}, \quad D_3 = \begin{bmatrix} 0 & 0 & \lambda^4 \\ \mathbf{0} & -\boldsymbol{\lambda}^{\mathbf{2}} & \lambda^3 \\ \mathbf{1} & \mathbf{0} & 0 \end{bmatrix}, \quad D_4 = \begin{bmatrix} 0 & 0 & 0 & -\lambda^6 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\lambda}^{\mathbf{4}} & -2\lambda^5 \\ \mathbf{0} & -\boldsymbol{\lambda}^{\mathbf{2}} & \boldsymbol{\lambda}^{\mathbf{3}} & \lambda^4 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & 0 \end{bmatrix}.$$

  In general, one can prove that the following recursive sequence provides us with the required change of basis:

  $$D_1 = [1], \qquad D_{p+1} = \left[\begin{array}{c|c} 0_{1\times p} & (-\lambda^2)^p \\ \hline D_p & d_p \end{array}\right], \qquad p = 1,\, 2,\, 3,\, \dots, \tag{4.8}$$

[1] The complex canonical form in Theorem 2.3 or the real one in Theorem 5.5 of [35].

with $d_p \in \mathbb{C}^p$ being an appropriate vector of the form $d_p = \begin{bmatrix} * & \dots & * & 0 \end{bmatrix}^T$. In order to prove this by induction, note first that (4.7) is equivalent to

$$\mathcal{J}_p(\lambda)^T D_p \mathcal{J}_p\left(\lambda^{-1}\right) = D_p, \tag{4.9}$$

which clearly holds for $p = 1$. We now state the induction hypothesis: assume that (4.9) holds for some $p \in \mathbb{N}$ and D_p as in (4.8). We will show that this implies

$$\mathcal{J}_{p+1}(\lambda)^T D_{p+1} \mathcal{J}_{p+1}\left(\lambda^{-1}\right) = D_{p+1}. \tag{4.10}$$

with D_{p+1} as in (4.8). First, partition the matrices above as

$$\mathcal{J}_{p+1}(\lambda)^T = \left[\begin{array}{c|c} \lambda & 0_{1\times p} \\ \hline e_1 & \mathcal{J}_p(\lambda)^T \end{array}\right] \qquad \mathcal{J}_{p+1}(\lambda^{-1}) = \left[\begin{array}{c|c} \mathcal{J}_p(\lambda^{-1}) & e_p \\ \hline 0_{1\times p} & \lambda^{-1} \end{array}\right],$$

where e_1 and e_p are the first and last column, respectively, of the $p \times p$ identity matrix. If we plug these block matrices into (4.10), we get

$$\mathcal{J}_{p+1}(\lambda)^T D_{p+1} \mathcal{J}_{p+1}\left(\lambda^{-1}\right) = \left[\begin{array}{c|c} 0_{1\times p} & (-1)^p \lambda^{2p} \\ \hline \mathcal{J}_p(\lambda)^T D_p\, \mathcal{J}_p\left(\lambda^{-1}\right) & \mathcal{J}_p(\lambda)^T \left(D_p e_p + \lambda^{-1} d_p\right) + (-1)^p \lambda^{2p-1} e_1 \end{array}\right].$$

Now, since $\mathcal{J}_p(\lambda)^T D_p\, \mathcal{J}_p\left(\lambda^{-1}\right) = D_p$ by the induction hypothesis, we only need to prove the existence of a vector d_p satisfying

$$\mathcal{J}_p(\lambda)^T \left(D_p e_p + \lambda^{-1} d_p\right) + (-1)^p \lambda^{2p-1} e_1 = d_p,$$

or, equivalently,

$$\left(\lambda^{-1} \mathcal{J}_p(\lambda)^T - I_p\right) d_p = (-1)^{p-1} \lambda^{2p-1} e_1 - \mathcal{J}_p(\lambda)^T D_p e_p. \tag{4.11}$$

However, since $D_p e_p = [(-\lambda^2)^p, *, \cdots, *]^T$, we have

$$\mathcal{J}_p(\lambda)^T D_p e_p = \begin{bmatrix} \lambda\,(-\lambda^2)^p \\ * \\ \vdots \\ * \end{bmatrix} \quad \Rightarrow \quad (-1)^{p-1} \lambda^{p-1} e_1 - \mathcal{J}_p(\lambda)^T D_p e_p = \begin{bmatrix} 0 \\ * \\ \vdots \\ * \end{bmatrix}.$$

On the other hand, the first row of $\lambda^{-1}\mathcal{J}_p(\lambda)^T - I_p$ is identically zero, so the subspace generated by its columns is $(p-1)$-dimensional. Therefore, the vector $(-1)^{p-1}\lambda^{p-1}e_1 - \mathcal{J}_p(\lambda)^T D_p e_p$ is in the column space of $\lambda^{-1}\mathcal{J}_p(\lambda)^T - I_p$, which proves the existence of the vector solution d_p of (4.11). It only remains to prove that the last entry of d_p can be chosen to be zero, but since the last column of $\lambda^{-1}\mathcal{J}_p(\lambda)^T - I_p$ is identically zero, the last entry of d_p is a free variable, and can thus be chosen to be zero.

- Another type of blocks showing up in the symplectic canonical forms are upper triangular Toeplitz matrices of the form

$$T_p = \text{Toep}(\lambda\,,\, t_1\,,\, t_2\,,\, \dots\,,\, t_{p-1}), \qquad t_1 \neq 0.$$

As before, we need a matrix A_p such that

$$A_p^{-1}\, T_p\, A_p = \mathcal{J}_p(\lambda).$$

One can check that, for $p = 2,\, 3,\, 4$ the matrices A_p

$$A_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{t_1} \end{bmatrix}, \quad A_3 = \begin{bmatrix} \mathbf{1} & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{\frac{1}{t_1}} & -\frac{t_2}{(t_1)^3} \\ 0 & 0 & \frac{1}{(t_1)^2} \end{bmatrix}, \quad A_4 = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{\frac{1}{t_1}} & -\mathbf{\frac{t_2}{(t_1)^3}} & \frac{2(t_2)^2 - t_1 t_3}{(t_1)^5} \\ \mathbf{0} & \mathbf{0} & \mathbf{\frac{1}{(t_1)^2}} & -\frac{2t_2}{(t_1)^4} \\ 0 & 0 & 0 & \frac{1}{(t_1)^3} \end{bmatrix}$$

satisfy the equation above. Again, the nested structure in these matrices suggests the recursive formula

$$\begin{aligned} A_1 &= [1] \\ A_{p+1} &= \left[\begin{array}{c|c} A_p & a_p \\ \hline 0_{1\times p} & (t_1)^{-p} \end{array}\right], \qquad p = 1,\, 2,\, 3,\, \dots \end{aligned} \tag{4.12}$$

for an appropriate vector $a_p \in \mathbb{C}^p$. Let us prove, again by induction, that these matrices do the job: obviously, A_1 satisfies $A_1^{-1}\, T_1\, A_1 = \mathcal{J}_1(\lambda)$. Next, assume that

$$A_p^{-1} T_p\, A_p = \mathcal{J}_p(\lambda)$$

for some $p \in \mathbb{N}$ and A_p as defined in (4.12). We will prove that for some appropriate a_p the matrix A_{p+1} satisfies the corresponding equation with p replaced by $p+1$. If we partition conformally all matrices involved, we obtain using the abbreviation $t = \begin{bmatrix} t_{p-1} & \dots & t_1 \end{bmatrix}^T$ that

$$\begin{aligned} T_{p+1} A_{p+1} &= \left[\begin{array}{c|c} T_p & t \\ \hline 0_{1\times p} & \lambda \end{array}\right] \left[\begin{array}{c|c} A_p & a_p \\ \hline 0_{1\times p} & (t_1)^{-p} \end{array}\right] = \left[\begin{array}{c|c} T_p A_p & T_p a_p + (t_1)^{-p} t \\ \hline 0_{1\times p} & \lambda (t_1)^{-p} \end{array}\right] \\ A_{p+1}\mathcal{J}_{p+1}(\lambda) &= \left[\begin{array}{c|c} A_p & a_p \\ \hline 0_{1\times p} & (t_1)^{-p} \end{array}\right] \left[\begin{array}{c|c} \mathcal{J}_p(\lambda) & e_p \\ \hline 0_{1\times p} & \lambda \end{array}\right] = \left[\begin{array}{c|c} A_p \mathcal{J}_p(\lambda) & A_p e_p + \lambda a_p \\ \hline 0_{1\times p} & \lambda (t_1)^{-p} \end{array}\right] \end{aligned}$$

The induction hypothesis leads to $T_p A_p = A_p \mathcal{J}_p(\lambda)$, so we only need to prove that there exists a vector $a_p \in \mathbb{C}^p$ such that

$$T_p a_p + (t_1)^{-p} t = A_p e_p + \lambda a_p\,, \qquad t = [\,t_p\; t_{p-1}\; \dots\; t_1]^T$$

or, equivalently,

$$(T_p - \lambda I_p)\, a_p = A_p e_p - (t_1)^{-p} t.$$

Now, since $t_1 \neq 0$ and the last row of $T_p - \lambda I_p$ is identically zero, the vector $A_p e_p - (t_1)^{-p} t$ is in the column space of $T_p - \lambda I_p$ if and only if its last entry is zero. But

$$A_p e_p - (t_1)^{-p} t = \begin{bmatrix} * \\ \vdots \\ * \\ (t_1)^{-p+1} \end{bmatrix} - \begin{bmatrix} t_p (t_1)^{-p} \\ \vdots \\ t_2 (t_1)^{-p} \\ t_1 (t_1)^{-p} \end{bmatrix} = \begin{bmatrix} * \\ \vdots \\ * \\ 0 \end{bmatrix},$$

which completes the proof.

4.2.2 Relationships between left and right Jordan chains

It is well known that if $x \in \mathbb{C}^{2n}$ is a right eigenvector associated with the eigenvalue $\lambda \in \mathbb{C} \setminus \{0\}$ of the J-symplectic matrix S, then $y = x^T J$ is a left eigenvector of S associated with the eigenvalue λ^{-1}. Our aim in this subsection is to extend this result to more general contexts: in §4.2.1 we have identified changes of basis allowing us to extract Jordan vectors from vectors in the structured symplectic canonical form. We will now exploit such transformations in order to establish the connections induced by symplectic structure between those left and right eigenvectors constructed starting from the structured canonical form.

To begin with, we partition the structured symplectic canonical form

$$\left[\begin{array}{c|c} \mathcal{C} & \\ \hline & \widehat{\mathcal{C}} \end{array}\right] = \left[\begin{array}{c} U \\ \hline \widehat{U} \end{array}\right] S \left[\begin{array}{c|c} T & \widehat{T} \end{array}\right]$$

in a way similar to (2.24)–(2.26), i.e., $\mathcal{C}$ contains all blocks in the canonical form associated with the eigenvalue λ, while $\widehat{\mathcal{C}}$ contains those corresponding to the remaining eigenvalues. As above, T and U are partitioned into column blocks and row blocks, respectively, conformally with the block structure of $\mathcal{C}$. We denote by T_j^k the k-th block of columns taken from T, associated with the k-th block of size n_j in $\mathcal{C}$ (resp., by U_j^k the k-th block of rows taken from U associated with the k-th block of size n_j in $\mathcal{C}$). Then we have for each $j = 1, \ldots, q$ and $k = 1, \ldots, r_j$ that

$$U_j^k \, S \, T_j^k = C_j^k$$

where C_j^k is a block of size n_j associated with λ in the canonical form $\mathcal{C}$. We distinguish two cases in the analysis, depending on whether $\lambda = \pm 1$ or not. Since the former case is by far the most interesting one, we shall mostly focus on it, although for the sake of completeness we will also report the relationships for eigenvalues other than ± 1.

4.2.2.1 Case 1: $\lambda \in \{-1\,,\,1\}$

Again, we split the analysis into two different cases, depending on the parity of the block size n_j of the blocks C_j^k in the canonical form:

- **Case 1.a:** $\lambda \in \{-1\,,\,1\}$, n_j even.

  If the size n_j of C_j^k is even, then from the first relation in (2.4) we have that C_j^k is an upper triangular Toeplitz matrix of the form

  $$\text{Toep}\left(\lambda,\, 1,\, t_2,\, \cdots,\, t_{n_j-1}\right).$$

  Although this matrix is not a Jordan block, it is similar to $\mathcal{J}_{n_j}(\lambda)$. Moreover, its structure allows us to conclude that the first column of T_j^k is a right eigenvector associated with λ, and the last row of U_j^k is a corresponding left eigenvector. Furthermore, the nonsingular matrix A_{n_j}, as defined in (4.12), is such that

  $$A_{n_j}^{-1}\,\text{Toep}\left(\lambda,\, 1,\, t_2,\, \cdots,\, t_{n_j-1}\right)\, A_{n_j} = \mathcal{J}_{n_j}(\lambda),$$

  so the left and right Jordan chains associated with this block are, respectively, the columns of

  $$\begin{array}{rcl} P_j^k & = & U_j^k A_{n_j}, \\ Q_j^k & = & (A_{n_j})^{-1} T_j^k. \end{array}$$

  Now, recall that the first column of A_{n_j} (resp., the last row of $A_{n_j}^{-1}$) is the first column $e_1 \in \mathbb{R}^{n_j}$ (resp., the last row $e_{n_j}^T$) of the identity matrix. Hence, the left and right eigenvectors associated with λ for this Jordan block are the first column of U_j^k and the last row of T_j^k, respectively.

  Next, recall that we have denoted $\mathcal{T} = \left[\, T \;\middle|\; \widehat{T}\, \right]$, and its inverse by

  $$\mathcal{T}^{-1} = \left[\begin{array}{c} U \\ \hline \widehat{U} \end{array}\right].$$

  Making use of the congruence relationship (4.5) and the second equation in (2.4), we get

  $$(T_j^k)^T J = \Sigma_{n_j} U_j^k, \qquad j = 1, \ldots, q, \quad k = 1, \ldots, r_j,$$

  with Σ_{n_j} given by (2.2). If, as before, we denote by y_j^k the last row of U_j^k, and by x_j^k the first column of T_j^k (recall that we have established that these are, respectively, a left and a corresponding right eigenvector associated with λ), then multiplying on the left by the row vector $e_1^T \in \mathbb{R}^{1\times n_j}$ we obtain

  $$e_1^T \left(T_j^k\right)^T J = \left(T_j^k\, e_1\right)^T J = e_1^T \Sigma_{n_j} U_j^k = e_{n_j}^T U_j^k,$$

which, read column-wise, leads to

$$y_j^k = \left(x_j^k\right)^T J\,,$$

which, in turn, induces the relationships

$$\boxed{Y_j = X_j^T\, J} \tag{4.13}$$

between the left and right eigenvector matrices corresponding to blocks of size n_j associated with λ.

If $S \in \mathbb{R}^{2n\times 2n}$, we obtain a similar result, but including the signs for the blocks of size n_j, namely

$$y_j^k = \varsigma_j^k \left(x_j^k\right)^T J\,, \qquad \varsigma_j^k = \pm 1,$$

which leads to

$$\boxed{Y_j = \Upsilon_j\, X_j^T\, J}\,, \qquad \Upsilon_j = \text{Diag}\left(\varsigma_j^1\,,\,\cdots\,,\,\varsigma_j^{r_j}\right)\,, \quad X_j \in \mathbb{R}^{2n\times r_j}. \tag{4.14}$$

- **Case 1.b:** $\lambda \in \{-1\,,\,1\}$, n_j odd

  In this case the procedure is similar to the one above: when n_j is odd, the blocks in the canonical form are paired as

$$\left[\begin{array}{c} U_j^k \\ \hline U_j^{k+1} \end{array}\right] S \left[\begin{array}{c|c} T_j^k & T_j^{k+1} \end{array}\right] = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda) & \\ \hline & \mathcal{J}_{n_j}(\lambda)^{-T} \end{array}\right]. \tag{4.15}$$

  To transform the structured canonical form into the Jordan form, we can use the similarity matrices D_{n_j} satisfying (4.7) or, equivalently,

$$\mathcal{J}_{n_j}\left(\lambda^{-1}\right) = \left(D_{n_j}\right)^{-1} \mathcal{J}_{n_j}(\lambda)^{-T} D_{n_j},$$

  which can also be written as

$$\mathcal{J}_{n_j}(\lambda)^T \left(-D_{n_j}\right) \mathcal{J}_{n_j}\left(\lambda^{-1}\right) = -D_{n_j},$$

  where

$$D_{n_j} = \begin{bmatrix} 0 & \vec{0}^T & 1 \\ \vec{0} & * & * \\ 1 & * & * \end{bmatrix}, \tag{4.16}$$

as in § 4.2.1 above, where $\vec{0}$ represents the zero vector of $\mathbb{R}^{n_j-2}$. Now, D_n satisfies

$$\left[\begin{array}{c|c} I_{n_j} & \\ \hline & (-D_{n_j})^{-1} \end{array}\right] \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda) & \\ \hline & \mathcal{J}_{n_j}(\lambda)^{-T} \end{array}\right] \left[\begin{array}{c|c} I_{n_j} & \\ \hline & -D_{n_j} \end{array}\right] = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda) & \\ \hline & \mathcal{J}_{n_j}(\lambda) \end{array}\right].$$

Replacing this into (4.15) leads to

$$\left[\begin{array}{c} U_j^k \\ \hline (-D_{n_j})^{-1} U_j^{k+1} \end{array}\right] S \left[\begin{array}{c|c} T_j^k & T_j^{k+1}(-D_{n_j}) \end{array}\right] = \left[\begin{array}{c|c} \mathcal{J}_{n_j}(\lambda) & \\ \hline & \mathcal{J}_{n_j}(\lambda) \end{array}\right],$$

but we know from (2.24) that

$$\left[\begin{array}{c} Q_j^k \\ \hline Q_j^{k+1} \end{array}\right] S \left[\begin{array}{c|c} P_j^k & P_j^{k+1} \end{array}\right] = \mathrm{Diag}(\mathcal{J}_{n_j}, \mathcal{J}_{n_j}),$$

so we may identify

$$\begin{aligned} P_j^k &= T_j^k, & P_j^{k+1} &= T_j^{k+1}(-D_{n_j}), \\ Q_j^k &= U_j^k, & Q_j^{k+1} &= (-D_{n_j})^{-1} U_j^{k+1}. \end{aligned} \tag{4.17}$$

Now, recall that the blocks T_j^k and U_j^k in the symplectic canonical form are linked via

$$\left[\begin{array}{c} (T_j^k)^T \\ \hline (T_j^{k+1})^T \end{array}\right] J = \left[\begin{array}{c|c} & I_{n_j} \\ \hline -I_{n_j} & \end{array}\right] \left[\begin{array}{c} U_j^k \\ \hline U_j^{k+1} \end{array}\right],$$

or, in short, $(T_j^k)^T J = U_j^{k+1}$, $(T_j^{k+1})^T J = -U_j^k$. Substituting this into (4.17), we obtain for every $j \in \{1, \ldots, q\}$ and every $k = 1, 3, 5, \cdots, r_j - 1$ the identities

$$(P_j^k)^T J = -D_{n_j} Q_j^{k+1}, \qquad \left(P_j^{k+1}\right)^T J = D_{n_j} Q_j^k \tag{4.18}$$

between the P_j^k and Q_j^k blocks. The special form (4.16) of D_{n_j} implies that $e_1^T D_{n_j} = e_{n_j}^T$, so if we denote right and left eigenvectors by x_j^k and y_j^k, respectively, as before, then

$$(x_j^k)^T J = -e_1^T D_{n_j} Q_j^{k+1} = -y_j^{k+1}$$

and

$$\left(x_j^{k+1}\right)^T J = \left(e_1^T D_{n_j}\right) Q_j^k = y_j^k,$$

which can be summarized in matrix form as

$$\boxed{Y_j = \Theta_j\, X_j^T\, J}\,, \qquad \Theta_j = (\Sigma_2 \oplus \cdots \oplus \Sigma_2) \in \mathbb{R}^{r_j \times r_j}\ , \qquad X_j \in \mathbb{C}^{2n \times r_j}. \tag{4.19}$$

Since there is no difference here between the real and complex canonical forms, the only difference in the equality above whenever $S \in \mathbb{R}^{2n\times 2n}$ is that $X_j \in \mathbb{R}^{2n\times r_j}$.

Finally, we summarize the relationship between the matrices W_j and Z_j in (2.27), i.e., between the matrices of left and right eigenvectors for all Jordan blocks associated with 1 or -1 with size at least n_j as follows:

$$\boxed{W_j = (\Lambda_1 \oplus \Lambda_2 \oplus \cdots \oplus \Lambda_j)\, Z_j^T\, J}\,, \tag{4.20}$$

where, for each $k = 1, \ldots, j$,

- $\Lambda_k = \Theta_k$ if n_k is odd,
- $\Lambda_k = \Upsilon_k$ if n_k is even and $S \in \mathbb{R}^{2n\times 2n}$, and
- $\Lambda_k = I_k$ if n_k is even and $S \in \mathbb{C}^{2n\times 2n}$.

4.2.2.2 Case 2: $\lambda \in \mathbb{C} \setminus \{-1\,,\, 1\}$

This case is treated similarly to the case $\lambda = \pm 1$ with odd n_j described in **Case 1.b** of § 4.2.2.1: the transformations leading to Jordan chains from canonical vectors are the same, with the only difference that now $\lambda^{-1} \neq \lambda$, so we have two identical Jordan structures, each one associated with either eigenvalue λ or λ^{-1}. Recall that, for each $j \in \{1, \ldots, q\}$, the $2n \times r_j$ matrix X_j (resp., the $r_j \times 2n$ matrix Y_j) collects all right (resp., left) eigenvectors of S corresponding to Jordan blocks of size n_j associated with λ. Let $\widetilde{X}_j$ (resp., $\widetilde{Y}_j$) be the analogous matrix associated with the reciprocal eigenvalue $1/\lambda$. Then one can prove that

$$Y_j = -\widetilde{X}_j^T\, J\,, \qquad \widetilde{Y}_j = \left(\frac{-1}{\lambda^2}\right)^{n_j-1} X_j^T\, J\,,$$

or, if we lump together the right (resp., left) eigenvector matrices into Z_s, $\widetilde{Z}_s$ (resp., W_s, $\widetilde{W}_s$), that

$$\boxed{W_s = -\widetilde{Z}_s^T\, J}\,, \qquad \boxed{\widetilde{W}_s = \Theta_s Z_s^T\, J}\,, \tag{4.21}$$

where

$$\Theta_s = \left(\left(\frac{-1}{\lambda^2}\right)^{n_1-1} I_{r_1}\right) \oplus \ldots \oplus \left(\left(\frac{-1}{\lambda^2}\right)^{n_s-1} I_{r_s}\right)\,.$$

We conclude by observing that if $\lambda \in \mathbb{R}$ and $S \in \mathbb{R}^{2n \times 2n}$, then the eigenvectors in the formulas above can all be chosen to be real.

Special Case: $S \in \mathbb{R}^{2n \times 2n} \ \wedge \ \lambda \in \mathbb{C}, \, |\lambda| = 1$**.**

The relationships we have just obtained when $\lambda \neq \pm 1$ are not direct relationships between left and right eigenvectors associated with λ, but crossed relationships between left eigenvectors associated with λ and right eigenvectors associated with λ^{-1} (and viceversa). In the special case when S is real and λ lies on the unit circle we can do better, actually finding direct relationships for left and right λ-eigenvectors: we proceed as before, using the transformations in (4.12) to obtain the corresponding Jordan chains for the blocks in (2.16). This leads to the relationships

$$\left(x_j^k\right)^* (i\,J) = \left(\overline{i\,\lambda}\right)^{n_j - 1} \varsigma_j^k y_j^k$$

between the left eigenvectors y_j^k and the right eigenvectors x_j^k associated with λ in the k-th Jordan block of size n_j (recall that ς_j^k is the sign of the k-th Jordan block of size n_j).

Thus, the relationship between the matrices W_s and Z_s of left and right eigenvectors in (2.27) is

$$\boxed{(Z_s)^* \, (i\,J) = \mathrm{Diag}\{\Lambda_1 \, , \, \ldots \, \Lambda_s\} W_s} \, , \tag{4.22}$$

where

$$\Lambda_j = \left(\overline{i\,\lambda}\right)^{n_j - 1} \mathrm{diag}\{\varsigma_j^1 \, , \, \ldots \, , \, \varsigma_j^{r_j}\} \, , \qquad j = 1 \, , \, \ldots \, , \, s.$$

Note that in this case the upper triangular Toeplitz matrix we need to transform into a Jordan block has $t_1 = i\lambda$ on the superdiagonal, so in this case the matrix corresponding to the change of basis has the form

$$A_{n_j} = \left[\begin{array}{c|c} A_{n_j - 1} & * \\ \hline 0_{1 \times (n_j - 1)} & (i\lambda)^{-(n_j - 1)} \end{array}\right].$$

We conclude by observing that this result is valid also when $S \in \mathbb{C}^{2n \times 2n}$ and $S^* \, J \, S = J$, where $*$ denotes the conjugate transpose, since in this case the matrix S is also (iJ)-unitary.

4.3 Asymptotic expansions for structured symplectic perturbations

Once the relations are known between left and right eigenvectors in the Jordan form of S, we incorporate that information into the formulas in Theorem 3.2 in order to refine them. These formulas depend, in principle, on the eigenvalues of the Schur complement $(\Phi_j / \Phi_{j-1}) \, (B)$ in (2.30), but we shall see that, once the relations between left and right eigenvectors are incorporated, the relevant Schur complements can be written in terms of a slightly different family of nested matrices:

Definition 4.5. *Let* λ *be an eigenvalue of a symplectic matrix* S *with Jordan structure* (2.24)*, and let* $j \in \{1, \ldots, q\}$*. Then, for every matrix* K *with the same dimensions as* S*, we define*

$$\Xi_j(K) = Z_j^* \, K \, Z_j \,,$$

where the columns of Z_j *are right eigenvectors of* S*, associated with Jordan blocks corresponding to* λ *of size at least* n_j*, which have been extracted from one of the symplectic canonical forms in § 2.1 as explained in § 4.2.1.*

Notice that this definition imposes a nested structure on the matrices $\Xi_s(K)$, just as on the matrices $\Phi_s(K)$ before. Hence, for each s we can define the Schur complement of $\Xi_{s-1}(K)$ in $\Xi_s(K)$, and denote it by $(\Xi_s/\Xi_{s-1})(K)$. One can easily check that

$$(\Xi_s/\Xi_{s-1})(K) = X_s^* \left(K - K \, Z_{s-1} \, \Xi_{s-1}(K)^{-1} Z_{s-1}^* \, K\right) X_s,$$

where X_s is the submatrix of Z_s containing all eigenvectors associated with Jordan blocks of size exactly n_s.

As before, we distinguish several cases, depending on whether the unperturbed eigenvalue λ belongs to $\{1, -1\}$ or not, and on whether S is real or complex. Each of these cases will be illustrated with one or more numerical examples in Section 4.5 below. Also, for each case we shall specialize in Section 4.4 the results for the simplest case when the perturbation $\widehat{S}(\varepsilon) - S$ has rank one, not only because it is the situation most commonly studied in the literature, but because the asymptotic formulas come out especially simple, revealing quite clearly the influence of the sign characteristics on the behavior of perturbed eigenvalues.

4.3.1 Asymptotic expansions: the case $\lambda \notin \{-1 \,, \, 1\}$

In this case each eigenvalue λ is paired with its reciprocal λ^{-1}, which is also an eigenvalue of S with the same Jordan structure as λ. The relations (4.21) between normalized eigenvectors corresponding to λ and λ^{-1} lead to the connection

$$\left(\widetilde{\Phi}_j/\widetilde{\Phi}_{j-1}\right)(B) = -\left(\frac{-1}{\lambda^2}\right)^{n_j-1} \left(\Phi_j/\Phi_{j-1}\right)(B)^T$$

between the Schur complements associated with λ and $1/\lambda$. This induces a one-to-one relationship between each eigenvalue ξ of $\big(\Phi_j/\Phi_{j-1}\big)(B)$ and the corresponding eigenvalue $\widetilde{\xi} = -\left(\frac{-1}{\lambda^2}\right)^{n_j-1} \xi$ of $\big(\widetilde{\Phi}_j/\widetilde{\Phi}_{j-1}\big)(B)$. Replacing these values in formula (3.2) of Theorem 3.2 does little more than confirming the symplectic symmetry of perturbed eigenvalues (which was known in advance, since $\widehat{S}$ is assumed to be symplectic).

Thus, the most relevant conclusion here is that, in the absence of further geometric constraints on the reciprocal, *if the unperturbed eigenvalue* λ *is not on the unit circle (i.e.,* $|\lambda| \neq 1$*) then the corresponding perturbed eigenvalues are not forced by symplectic structure to move in any particular direction* (see, for instance, Example 4.11 below, which confirms this statement). The case when $\lambda = \pm 1$ is obviously special, since λ and its reciprocal coincide. Moreover, when λ is on the unit circle *and both* S *and* $\widehat{S}$ *are real*, we have already seen in § 4.2.2.2 that left and right eigenvectors satisfy additional constraints, which in turn restricts the possible behaviors of perturbed eigenvalues. This is the case we study next:

4.3.2 Asymptotic expansions: the case $S,\, \widehat{S} \in \mathbb{R}^{2n\times 2n}$ with $|\lambda| = 1$

We address this situation separately, since, as we have seen in §4.2.2.2, more specific information is available in this special case to refine the asymptotic expansions: if $S,\, \widehat{S} \in \mathbb{R}^{2n\times 2n}$ and $|\lambda| = 1$, we can make use of the left-right eigenvector relationships (4.22) obtained in § 4.2.2.2. Since these connect left and right eigenvectors *both* associated with λ, we may further simplify the Schur complements $(\Phi_j/\Phi_{j-1})\,(B)$: one can easily check that if $\Phi_{j-1}(B)$ is nonsingular, then

$$(\Phi_j/\Phi_{j-1})\,(B) = \mathrm{i}^{n_j}\,\lambda^{n_j-1} H_j\,, \tag{4.23}$$

where $H_j = \mathrm{Diag}\{\,\varsigma_j^1\,,\,\ldots\,,\,\varsigma_j^{r_j}\,\}(\Xi_j/\Xi_{j-1})(J\,B)$, where the ς_j^k are the signs associated with $n_j \times n_j$ Jordan blocks, and Ξ_{j-1} and Ξ_j are as in Definition 4.5. This proves the following result, which describes the structured asymptotic expansions obtained from Theorem 3.2 in this special case.

Theorem 4.6. *Let* λ *be an eigenvalue of a real symplectic matrix* S *with* $|\lambda| = 1$ *and Jordan structure* (2.24)*. Let* $\widehat{S}(\varepsilon) = (\,I + \varepsilon B + O(\varepsilon^2)\,)\,S$ *be an arbitrary real symplectic structured perturbation of* S *as in* (1.3)*. Let* $j \in \{1,\,\ldots,\,q\}$ *be given and assume that* $\Xi_{j-1}(J\,B)$ *is nonsingular if* $j > 1$*, where* $\Xi_{j-1}(\cdot)$ *is given by Definition* 4.5*. Then there are* $r_j n_j$ *eigenvalues of the perturbed matrix* $\widehat{S}(\varepsilon)$ *admitting first order expansions*

$$\widehat{\lambda}_{j,k,l}(\varepsilon) = \lambda\left(1 + i\;(\xi_{j,k})^{\frac{1}{n_j}}\,\varepsilon^{\frac{1}{n_j}}\right) + o\left(\varepsilon^{\frac{1}{n_j}}\right)\,, \tag{4.24}$$

where $\xi_{j,k},\ k = 1, \ldots, r_j$*, is any of the eigenvalues of the matrix*

$$H_j = \mathit{Diag}\{\,\varsigma_j^1\,,\,\ldots\,,\,\varsigma_j^{r_j}\,\}(\Xi_j/\Xi_{j-1})(J\,B)$$

if $j > 1$*, or of the matrix* $H_1 = \mathit{Diag}\{\,\varsigma_1^1\,,\,\ldots\,,\,\varsigma_1^{r_1}\,\}\Xi_1(J\,B)$ *if* $j = 1$*. The values* $\widehat{\lambda}_{j,k,l}(\varepsilon)$ *for* $l = 1, \ldots, n_j$ *are defined by taking the* n_j*-th roots of* $\xi_{j,k}$ *(which are pairwise distinct unless* $\xi_{j,k} = 0$*).*

Notice that, since we are considering the real case, the matrix B is *real* Hamiltonian and, consequently, $(\Xi_j/\Xi_{j-1})(J\,B)$ is Hermitian. Although H_j is a complex matrix and, in fact, may have complex eigenvalues depending on the signs ς_j^k, its characteristic polynomial is always real. Thus, the spectrum of H_j is symmetric with respect to the real line, so the perturbed asymptotic expansions are consistent with the real symmetry of the spectrum of $\widehat{S}$.

Now, we focus on the special case when $\xi_{j,k}$ is a simple real eigenvalue of H_j. This happens, for instance, for rank-one perturbations, or when S has a single Jordan block of size n_j associated with λ. In this case, more specific information can be gathered about the behavior of perturbed eigenvalues: we assume $\xi_{j,k} \neq 0$ and, as usual, distinguish two cases:

- n_j is odd
  In this case, one of the n_j-th roots of $\xi_{j,k}$ must be real, so one of the perturbed eigenvalues moves tangentially to the unit circle, according to Theorem 4.6. Since $\xi_{j,k}$ is a simple eigenvalue of H_j, such a single perturbed eigenvalue cannot actually leave the unit circle, or else symplectic spectral symmetry would be broken (see Remark 4.7 below). Hence, one of the perturbed eigenvalues moves on the unit circle, and as to the remaining eigenvalues, half of them move inwards, and the other half outwards with respect to the unit circle.

- n_j is even
  In this case we need to take into account not only the sign of $\xi_{j,k}$, but the remainder of the integer division of n_j by 4 as well. The following table summarizes all possibilities, depending on the nature of the n_j-th roots of $\xi_{j,k}$:

	$n_j = 4p$	$n_j = 4p + 2$
$\xi_{j,k} > 0$	Two eigenvalues move on the unit circle in opposite directions, another two move away from the unit circle orthogonally to it, one inwards and the other one outwards. The remaining $4(p-1)$ eigenvalues move away from the unit circle, half inwards and half outwards.	Two eigenvalues move on the unit circle in opposite directions, and the remaining $4p$ move away from the circle, half inwards and half outwards, none of them orthogonally to the circle.
$\xi_{j,k} < 0$	$2p$ eigenvalues move inside the unit circle and the remaining $2p$ eigenvalues move outside. None of them moves orthogonally to the circle, none of them stays on the unit circle.	Two perturbed eigenvalues move away from the unit circle orthogonally to it, one inwards and the other one outwards. The remaining $4p$ eigenvalues move away from the unit circle, half inwards and half outwards.

Table 4.3.2

Remark 4.7. *The claims in Table 4.3.2 are just simple consequences of (i) the asymptotic expansions in Theorem 4.6; (ii) the fact that* $\widehat{S}(\varepsilon)$ *is still symplectic; and (iii) the fact that* $\xi_{j,k}$ *in (4.24) is assumed to be a* simple *real eigenvalue of* H_j.
First, it is obvious that any non-real n_j*-th root of* $\xi_{j,k}$ *produces an expansion* (4.24) *whose tangent direction is* not *tangent to the unit circle. Hence, the corresponding perturbed eigenvalue* must *leave the unit circle following that tangent direction. Purely imaginary* n_j*-th roots of* $\xi_{j,k}$*, for instance, correspond to perturbed eigenvalues which leave the unit circle orthogonally to it.*
Of course, the condition of having a trajectory tangent to the unit circle is, in general, necessary but not sufficient for the perturbed eigenvalue to stay on the unit circle. If the real eigenvalue $\xi_{j,k}$ *is* simple*, however, the symplecticity of* $\widehat{S}(\varepsilon)$ *makes that necessary condition also sufficient, because if the perturbed tangent eigenvalue, say* $\widehat{\lambda}$*, leaves the unit circle, then also the eigenvalue* $\overline{(1/\widehat{\lambda})}$*, one of the three symplectic counterparts to the conjugate of* $\widehat{\lambda}$ *according to symplectic spectral symmetry, should leave it, and this can only happen when* $\xi_{j,k}$ *is a* multiple *eigenvalue of* H_j*: consider, for instance, the case of a real symplectic matrix* S *with one or more Jordan blocks of size* 4 *associated with an eigenvalue* λ *on the unit circle. Let* $\widehat{S}(\varepsilon)$ *be a structured perturbation of* S *as in (1.3), with* B *real and* $r = rank(B)$*, such that the matrix* H_j *corresponding to* 4 *by* 4 *Jordan blocks has a simple eigenvalue* $\xi_{1,1} = 1$*. Hence, the asymptotic expansions (4.24) for the four corresponding perturbed eigenvalues are*

$$\begin{aligned} \widehat{\lambda}_1 &= \lambda\left(1 + i\varepsilon^{\frac{1}{4}}\right) + o\left(\varepsilon^{\frac{1}{4}}\right), \\ \widehat{\lambda}_2 &= \lambda\left(1 - i\varepsilon^{\frac{1}{4}}\right) + o\left(\varepsilon^{\frac{1}{4}}\right), \\ \widehat{\lambda}_{3,4} &= \lambda\left(1 \pm \varepsilon^{\frac{1}{4}}\right) + o\left(\varepsilon^{\frac{1}{4}}\right), \end{aligned}$$

i.e., $\widehat{\lambda}_1$ *and* $\widehat{\lambda}_2$ *move tangentially to the unit circle, while* $\widehat{\lambda}_3$ *and* $\widehat{\lambda}_4$ *move orthogonally away from it. Now, let us assume that, say,* $\widehat{\lambda}_1$ *leaves the unit circle, moving outside of it. We shall see this is not possible in this setting: symplectic symmetry implies that*

(i) the complex conjugate $\overline{\widehat{\lambda}_1}$ *is an eigenvalue of* $\widehat{S}(\varepsilon)$ *lying outside the unit circle, close to* $\overline{\lambda}$*;*

(ii) its reciprocal $\widehat{\mu} = \left(\overline{\widehat{\lambda}_1}\right)^{-1}$ *is also an eigenvalue of* $\widehat{S}(\varepsilon)$*, lying* inside *the unit circle, close to* $1/\overline{\lambda} = \lambda$*.*

Furthermore, a simple calculation shows that

$$\widehat{\mu} = \frac{1}{\overline{\widehat{\lambda}_1}} = \frac{1}{\overline{\lambda}(1 - i\varepsilon^{1/4}) + o(\varepsilon^{1/4})} = \lambda\left(1 + i\varepsilon^{\frac{1}{4}}\right) + o\left(\varepsilon^{\frac{1}{4}}\right),$$

i.e., $\widehat{\mu}$ is different from $\widehat{\lambda}_1$, since it lies inside *the unit circle, but it should have the same first order asymptotic expansion. This is impossible, since $\xi_{1,1} = 1$ is a simple eigenvalue of H_j, and has already produced the four perturbed eigenvalues $\widehat{\lambda}_j,\ j = 1, \ldots, 4$, none of which can possibly be $\widehat{\mu}$. Thus in this case both $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ must remain on the unit circle for sufficiently small values of ε.*

4.3.3 Asymptotic expansions: the case $\lambda \in \{-1\,,\,1\}$.

In this situation we expect quite different behaviors depending on the parity of n_j, since the formulas in § 4.2.2.1 are different depending on whether n_j is even or odd. Moreover, recall from §2.1.3 that if n_j is even, there can be any number of $n_j \times n_j$ blocks in the Jordan form of S, whereas if n_j is odd, there can only be an even number of blocks of size n_j. This distinction is important, as we shall soon see.

Since the leading coefficients of the asymptotic expansions are given by the eigenvalues of the Schur complements $(\Phi_j/\Phi_{j-1})\,(B)$, we may, as before, use the relationships (4.14) and (4.19), replacing them in (2.30) to define

$$H_j = (\Phi_j/\Phi_{j-1})\,(B) = \Lambda_j\,(\Xi_j/\Xi_{j-1})(J\,B), \tag{4.25}$$

where each matrix Λ_j is either a diagonal matrix of signs if n_j is even, or a block-diagonal matrix with $\frac{r_j}{2}$ diagonal blocks $\Sigma_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ if n_j is odd. As in § 4.3.2 above, we incorporate this information into Theorem 3.2 to obtain the following straightforward consequence of it:

Theorem 4.8. *Let $\lambda = \pm 1$ be an eigenvalue of a symplectic matrix S with Jordan structure* (2.24). *Let $\widehat{S}(\varepsilon) = (\,I + \varepsilon B + O(\varepsilon^2)\,)\,S$ be an arbitrary symplectic structured perturbation of S as in* (1.3). *Let $j \in \{1, \ldots, q\}$ be given and assume that $\Xi_{j-1}(J\,B)$ is nonsingular if $j > 1$, where $\Xi_{j-1}(\cdot)$ is as in Definition* 4.5. *Then there are $r_j n_j$ eigenvalues of the perturbed matrix $\widehat{S}(\varepsilon)$ admitting first order expansions*

$$\widehat{\lambda}_{j,k,l}(\varepsilon) = \lambda\left(1 + \xi_{j,k}^{\frac{1}{n_j}}\,\varepsilon^{\frac{1}{n_j}}\right) + o\left(\varepsilon^{\frac{1}{n_j}}\right), \qquad k = 0,\,1,\,\cdots,\,n_j - 1, \tag{4.26}$$

where $\xi_{j,k}$ is any of the eigenvalues of the matrix

$$H_j = \Lambda_j\,(\Xi_j/\Xi_{j-1})(J\,B)$$

for $j > 1$, or of the matrix $H_1 = \Lambda_1\,\Xi_1(J\,B)$ if $j = 1$. The matrix Λ_j is either a diagonal of signs if n_j is even, or a block-diagonal matrix with $\frac{r_j}{2}$ diagonal blocks $\Sigma_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ if n_j is odd. The values $\widehat{\lambda}_{j,k,l}(\varepsilon)$ for $l = 1, \ldots, n_j$ are defined by taking the n_j-th roots of $\xi_{j,k}$ (which are pairwise distinct unless $\xi_{j,k} = 0$).

In order to better describe these expansions, we distinguish two cases, depending on the parity of n_j :

4.3.3.1 n_j is even

In this case, $H_j = \mathrm{Diag}\{\varsigma_j^1, \ldots, \varsigma_j^{r_j}\}(\Xi_j/\Xi_{j-1})(J\,B)$, but now the Schur complement $(\Xi_j/\Xi_{j-1})(J\,B)$ is not in general Hermitian: it only is if both $S, \widehat{S}$ are real, but this is just a particular case of the one treated in § 4.3.2. Hence, for complex perturbations, the perturbed eigenvalues may move away from λ in any direction, although we observe that the expansions are still consistent with the symplectic spectral symmetry, since n_j is even, and hence the n_j-th roots of ξ in (4.26) can be split into opposite pairs. If both S, $\widehat{S}$ are real, the perturbed eigenvalues can still move away from λ in any direction, since H_j may have nonreal eigenvalues. In the special case when H_j has some real eigenvalue $\xi_{j,k}$, the behavior of the perturbed eigenvalues is the one described in Table 4.3.2, i.e., the one corresponding to the asymptotic expansions (4.26), which are basically the same as the ones in (4.24) (up to a factor i).

4.3.3.2 n_j is odd

In this case, the left-right eigenvector relationships (4.20) imply that the matrix H_j in the statement of Theorem 4.8 is

$$H_j = D_j\,(\Xi_j/\Xi_{j-1})(JB)\,, \qquad j > 1, \tag{4.27}$$

where $D_j = \overbrace{\Sigma_2 \oplus \ldots \oplus \Sigma_2}^{\frac{r_j}{2}\ \text{times}}$ (for $j = 1$, this formula reduces to $H_1 = D_1\,\Xi_1(JB)$). One can easily check that $D_j^{-1} = D_j^T = -D_j$, and that H_j is D_j-Hamiltonian. Therefore, the spectrum of H_j is symmetric with respect to the origin, and its nonzero eigenvalues can be grouped into pairs $(\xi, -\xi)$, each giving rise to two perturbed eigenvalues with expansions

$$\begin{aligned}\widehat{\lambda}_k(\varepsilon) &= \lambda\left(1 + \xi^{\frac{1}{n_j}}\varepsilon^{\frac{1}{n_j}}\right) + o\left(\varepsilon^{\frac{1}{n_j}}\right), \\ \widehat{\mu}_k(\varepsilon) &= \lambda\left(1 - \xi^{\frac{1}{n_j}}\varepsilon^{\frac{1}{n_j}}\right) + o\left(\varepsilon^{\frac{1}{n_j}}\right),\end{aligned} \tag{4.28}$$

which reflects the symplectic symmetry of the spectrum of the perturbed matrix $\widehat{S}(\varepsilon)$.

We will now show that there are situations when the Schur complement *must* have at least one zero eigenvalue depending on the case whether ρ is odd or even. These situations fall, of course, out of the scope of Theorem 4.8, which only gives relevant information if $\xi_{j,k} \neq 0$.

- **Case 1: ρ is odd**

  We have seen in Remark 4.3 that if r =rank(B) is given by (4.3), then the Schur complement $H_j = (\Phi_j/\Phi_{j-1})\,(B)$ may have at most ρ nonzero eigenvalues. Thus, Theorem 4.8 accounts for at most ρn_j of the perturbed eigenvalues, which will come from ρ destroyed $n_j \times n_j$ Jordan blocks of S.

Since the nonzero eigenvalues of H_j show up in pairs $(\xi, -\xi)$ and the trace of H_j is zero, we conclude that in the case that ρ is odd, at least one eigenvalue of H_j *must* be zero. This zero eigenvalue corresponds to n_j eigenvalues of $\widehat{S}$ unaccounted for by Theorem 4.8. The $(\rho-1)n_j$ perturbed eigenvalues corresponding to the $\rho-1$ generically nonzero eigenvalues of H_j are still given by Theorem 4.8.

This very particular situation, i.e., *when both n_j and ρ are odd*, is what we shall call the *nongeneric case*. Clearly, it requires a separate analysis on its own to describe what happens to the 'atypical' n_j perturbed eigenvalues corresponding to that zero eigenvalue of H_j. This can be done by combining results in [38, 3] with Newton diagram techniques. Since the proof has to be built from scratch, and requires techniques quite different to the ones employed here, this analysis shall be deferred to Chapter 5.

- **Case 2: ρ is even**

  This situation is easier to analyze: the asymptotic expansions for all perturbed eigenvalues with nonzero ξ are those described in (4.28) (in this case, as in any other, there may be zero eigenvalues in the Schur complement, corresponding to nongeneric perturbations. The difference with **Case 1** above is that there the zero eigenvalue *must* be present).

  Notice that both asymptotic expansions in (4.24) and (4.28) are essentially the same, up to a factor i. The main difference is that now we have a more detailed formula for H_j. It is straightforward to check that also in this case the spectrum of H_j may contain arbitrary complex eigenvalues, so the perturbed eigenvalues can move away from λ in any direction. Finally, if $S, \widehat{S} \in \mathbb{R}^{2n\times 2n}$ and if ξ is neither real nor purely imaginary, then the number of Jordan blocks of size n_j destroyed by the perturbation is at least 4, because the quadruple $\{\xi, -\xi, \overline{\xi}, -\overline{\xi}\}$ is in the spectrum of H_j.

4.4 Rank-one structure-preserving perturbations

As announced above, in this section we will specialize the asymptotic formulas we have obtained in Section 4.3 for the special case when the perturbation $\widehat{S}(\varepsilon) - S$ has rank one. As before, we focus only on the two most interesting cases:

1. **The case** $S, \widehat{S} \in \mathbb{R}^{2n\times 2n}$ **with** $|\lambda| = 1$, $\operatorname{rank}(\widehat{S}(\varepsilon) - S) = 1$

   The fact that $\operatorname{rank}(\widehat{S}(\varepsilon) - S) = 1$, together with the J-Hamiltonian character of B, implies that BJ is symmetric with rank one. It follows that either $B = uu^T J$ for some $u \in \mathbb{C}^{2n}$ in the complex case, or $B = \pm uu^T J$ for some $u \in \mathbb{R}^{2n}$ in the real case.

Let $\xi = \text{trace}(\Xi_1(JB))$ be the only nonzero eigenvalue of $H_1 = \Upsilon_1\Xi_1(JB) = \Upsilon_1 X_1^* J B X_1$, where Υ_1 is a $n_1 \times n_1$ diagonal matrix with the signs of the Jordan blocks for λ of size n_1 on its main diagonal. Hence,

$$\xi = \text{trace}(\Xi_1(JB)) = \pm\text{trace}\left(\Upsilon_1 \underbrace{(X_1^* J u)}_{a} \underbrace{(u^T J X_1)}_{-a^*}\right) = \mp\sum_{j=1}^{r_1} \varsigma_j |a_j|^2,$$

where the ς_j are the signs on the main diagonal of Υ_1 and the vector a with entries a_i is $a = X_1^* J u$. The asymptotic expansions for the n_1 perturbed eigenvalues created by the destruction of one largest Jordan block are

$$\widehat{\lambda}_k(\varepsilon) \;=\; \lambda\left(1 + i\, \xi^{\frac{1}{n_1}}\, \varepsilon^{\frac{1}{n_1}}\right) + o\left(\varepsilon^{\frac{1}{n_1}}\right)$$

where

$$\xi = \mp\sum_{j=1}^{r_1} \varsigma_j |a_j|^2 \in \mathbb{R}.$$

Thus, the behavior of the perturbed eigenvalues depends primarily on the sign of ξ and on the remainder modulus 4 of n_1, as already discussed above in Table 4.3.2.

2. **The case** $\lambda \in \{-1, 1\}$, n_1 **even,** $\text{rank}\,(\widehat{S}(\varepsilon) - S) = 1$**.** We distinguish between the cases of real and complex matrices:

   (a) **The real case:** S, $\widehat{S} \in \mathbb{R}^{2n\times 2n}$

   This is just a particular case of Case 1. above, so what we obtain is basically the same, but for $\lambda = \pm 1$ and a real eigenvector matrix X_1. Recall that $B = \pm uu^T J$ and

   $$H_1 = \Upsilon_1\Xi_1(JB) = \pm\Upsilon_1 X_1^T J\, uu^T J\, X_1 = \mp\Upsilon_1\, a\, a^T,$$

   where, $a = X_1^T J\, u \in \mathbb{R}^{r_1}$. Hence,

   $$\xi = \text{trace}\,(\Phi_1(B)) = \mp\sum_{j=1}^{r_1} \varsigma_1^j\, (a_j)^2 \in \mathbb{R}\,.$$

   As before, the fact that n_1 is even reduces the analysis of whether the perturbed eigenvalues stay on the unit circle to a discussion depending on the sign of ξ and the divisibility of n_1 by 4, as discussed in § 4.3.3.1. The four possible situations are summarized in Table 4.3.2, by just taking $\xi_{j,k} = i^{n_1}\,\xi$.

We stress that, although the asymptotic expansions (4.26) give only information about the tangents to the 'escape directions' of the perturbed eigenvalues, and therefore are, in principle, inconclusive as to whether the eigenvalues stay or not on the unit circle, they are actually forced to stay in this case by the symmetry constraints imposed by symplecticity, at least until they meet another eigenvalue with opposite sign in the sign characteristic (see Section 4.4.1 for more details about this).

(b) **The complex case:** $S,\ \widehat{S} \in \mathbb{C}^{2n\times 2n}$

In this case there are no signs, since there is no sign characteristic. Hence, $\Upsilon_1 = I_{r_1}$ and

$$\Phi_1(B) = X_1^T\, J\, uu^T\, J\, X_1 \in \mathbb{C}^{r_1\times r_1},$$

so

$$\xi = -\mathrm{trace}(a\, a^T) = -\sum_{i=1}^{r_1} a_i^2 \in \mathbb{C}.$$

In the case that, by chance, ξ happens to be real, then (again depending on the parity of n and the sign of ξ) there may be two perturbed eigenvalues whose escape direction is tangential to the unit circle and that move away in opposite directions. However, in contrast to case (a) these eigenvalues need not stay on the unit circle, because in this case no structural constraints force them to.

3. **The case** $\lambda \in \{-1, 1\}$, n_1 **odd,** $\mathrm{rank}\,(\widehat{S}(\varepsilon) - S) = 1$**.** The case of odd-sized largest Jordan block and $\mathrm{rank}\,(B) = 1$ constitutes, as already shown, a highly nongeneric one, since one can prove that all the eigenvalues of $\Phi_1(B)$ are zero. Thus, the first order perturbation theory we are using provides no information whatsoever on the leading terms, and a completely different analysis has to be performed. This will be done in Chapter 5 below.

4.4.1 Observations on coalescing eigenvalues

The results obtained in Section 4.3 can be used to explain the behavior of eigenvalues of real symplectic matrices under parameter-dependent rank-one perturbations with special focus on eigenvalues on the unit circle: let $J \in \mathbb{R}^{2n\times 2n}$ be skew-symmetric and invertible, let $S \in \mathbb{R}^{2n\times 2n}$ be J-symplectic and consider a structure-preserving rank-one perturbation of the form

$$S(\tau) = (I + \tau uu^T J)S$$

for some $u \in \mathbb{R}^{2n\times 2n}$. Then for each eigenvalue λ_0 of $S = S(0)$, there is an eigenvalue $\lambda(\tau)$ of $S(\tau)$ such that $\lambda(\tau)$ is continuous as a function of τ and such that $\lambda(0) = \lambda_0$.

In the following, we will investigate what happens if two simple eigenvalues meet at 1 for some value τ_0 to form a Jordan block of size 2×2. At first, this may sound like a highly nongeneric situation and one may think that it will almost never be observed in practice. Indeed, it was observed in [46] that the eigenvalue curves $\lambda(\tau)$ generically do not intersect, if S is a general complex matrix and, furthermore, even if the curves intersect, why should this be at the special point 1? In our particular situation, however, the matrix S is both real and symplectic. Thus, if $S(\tau)$ has a simple eigenvalue $\lambda(\tau)$ on the unit circle for some τ, then for symmetry reasons it has to stay on the unit circle unlees it meets another eigenvalue on the unit circle, so an intersection of eigenvalue curves for eigenvalues on the unit circle can be expected and this is indeed what can be observed in numerical experiments. Since furthermore $S(\tau)$ is real and thus all its eigenvalues are symmetric with respect to the real line, the values ± 1 are natural values where eigenvalues moving along the unit circle can (and do) meet.

Letting τ_0 denote the moment when two eigenvalues meet, and setting $\widetilde{S} = S(\tau_0)$, we can use the results from Table 4.3.2 to understand what happens at and around the moment τ_0 when the two eigenvalues meet and form a Jordan block associated with the eigenvalue one. For $\varepsilon \geq 0$ let us define

$$\begin{aligned} S_f(\varepsilon) &:= (I + \varepsilon uu^T J)\widetilde{S} = \big(I + (\tau_0 + \varepsilon)uu^T J\big)S, \quad \text{and} \\ S_p(\varepsilon) &:= \big(I + \varepsilon uu^T(-J)\big)\widetilde{S} = \big(I + (\tau_0 - \varepsilon)uu^T J\big)S. \end{aligned}$$

Thus, the asymptotic expansion for $S_f(\varepsilon)$ will tell us what happens "in the future" when the 2×2 Jordan block at $\lambda_0 = 1$ will split again, while the asymptotic expansion for $S_p(\varepsilon)$ gives information about "the past", i.e., for the time before the two eigenvalues have met and formed the Jordan block. Observe that if ξ is the unique nonzero eigenvalue of $\Phi_1(uu^T J)$, then $-\xi$ is the unique nonzero eigenvalue of $\Phi_1\big(uu^T(-J)\big)$. Thus, according to Table 4.3.2, we have the following two situations:

1. If $\xi > 0$, then there exists a neighborhood U of τ_0 such that for all $\tau \in U$ we have:
   (a) for increasing τ, $\tau < \tau_0$ the matrix $S(\tau)$ has two conjugate complex eigenvalues on the unit circle close to $\lambda_0 = 1$ that both move towards $\lambda_0 = 1$;
   (b) for increasing τ, $\tau > \tau_0$ the matrix $S(\tau)$ has two real eigenvalues close to $\lambda_0 = 1$ that are reciprocals of each other and that both move away from $\lambda_0 = 1$.
2. If $\xi < 0$, then there exists a neighborhood U of τ_0 such that for all $\tau \in U$ we have:
   (a) for increasing τ, $\tau < \tau_0$ the matrix $S(\tau)$ has two real eigenvalues close to $\lambda_0 = 1$ that are reciprocals to each other and that both move towards $\lambda_0 = 1$;
   (b) for increasing τ, $\tau > \tau_0$ the matrix $S(\tau)$ has two conjugate complex eigenvalues on the unit circle close to $\lambda_0 = 1$ that both move away from $\lambda_0 = 1$.

4.5 Numerical examples

In this final section of this chapter we illustrate with numerical examples the different expansions derived in Section 4.3 above. As was already noted in Remark 4.4, given any structure-preserving perturbation $\widehat{S}$ of the form (1.3), there are infinitely many other symplectic perturbations with the same first-order coefficient matrix B. Some of the eigenvalues of these matrices (those covered by Theorem 3.2 and its offspring) will have the same leading terms as the ones from $\widehat{S}$, but other eigenvalues may behave differently, either due to the $O(\varepsilon^2)$ coefficient matrix, or even because the (common) Schur complements H_j have some zero eigenvalue.

In order to distinctly isolate in the numerical examples the behavior of the eigenvalues covered by our analysis, we will associate to each perturbed matrix $\widehat{S}(\varepsilon)$ another symplectic matrix $\widetilde{S}(\varepsilon)$, also of the form (1.3) and with the same leading matrix B, but $\widetilde{S}(\varepsilon)$ will only display those perturbed eigenvalues covered by the theory in Section 4.3. The reason for this is that $\widetilde{S}$ will be constructed in such a way that $\text{rank}\,(B) = \text{rank}\,(\widetilde{S} - S)$, an identity which is not true, in general, for every structure-preserving perturbation of the form (1.3). To be more precise, $\widetilde{S}$ will be just the product of exactly $\text{rank}\,(B)$ rank-one multiplicative J-Hamiltonian perturbations. This will ensure that each successive rank-one multiplicative perturbation destroys just one single Jordan block, and the corresponding perturbed eigenvalues of $\widetilde{S}$ are described by the expansions in §4.3.

The details on the construction and properties of $\widetilde{S}$ are given in the next subsection.

4.5.1 An auxiliary perturbation matrix

Given any structure-preserving perturbation $\widehat{S}(\varepsilon)$ as defined in (1.3), we shall construct a structure-preserving perturbation $\widetilde{S}$, also of the form (1.3) with the same first-order term B, but which can be written as a sequence

$$\widetilde{S} = \widetilde{S}(\varepsilon) = \left(\prod_{k=1}^{r} \left(I + \varepsilon s_k u_k \, u_k^T \, J \right) \right) S \tag{4.29}$$

of $r = \text{rank}\,(B)$ symplectic multiplicative perturbations, each of whose factors is a rank-one J-Hamiltonian perturbation of identity (the u_k are vectors, the s_k are signs, see Corollary 4.10 below). As a consequence of this, the ranks of $\widetilde{S}(\varepsilon) - S$ and B will coincide, and all perturbed eigenvalues of $\widetilde{S}(\varepsilon)$ will be generically covered by the asymptotic expansions derived in Section 4.3. Those eigenvalues, of course, will have the same first-order behavior as the corresponding eigenvalues of $\widehat{S}(\varepsilon)$, as explained in Remark 4.4.

We begin by observing that B is J-Hamiltonian if and only if the matrix $H = -B\,J$ is complex symmetric or, equivalently, $B = H\,J$, for some complex symmetric H. First, we describe the set of all $2n \times 2n$ complex symmetric matrices of given rank r:

Lemma 4.9. *Let* $r \in \{1\,, \ldots\,, 2n\}$. *Then*

$$\left\{ H \in \mathbb{C}^{2n\times 2n} \;:\; H^T = H, \;\; \mathit{rank}(H) = r \right\} = \left\{ \sum_{k=1}^{r} s_k \, u_k \, u_k^T \;:\; \{u_1\,, \ldots\,, u_r\} \in \mathcal{V}_r^{2n}, \;\; s_k = \pm 1 \right\}, \tag{4.30}$$

where $\mathcal{V}_r^{2n}$ *denotes the family of all orthogonal subsets of* r *nonzero vectors in* $\mathbb{C}^{2n}$*. If* $H \in \mathbb{R}^{2n\times 2n}$*, then* $\{u_1\,,\,\dots\,,\,u_r\}$ *can also be chosen to be real.*

Proof: Takagi's factorization (see, e.g., Corollary 4.4.4 in [22]) allows us to factor any complex symmetric matrix $H \in \mathbb{C}^{2n\times 2n}$ as

$$H = U\,D\,U^T\,,$$

where $U \in \mathbb{C}^{2n\times 2n}$ is unitary and D is diagonal with real non-negative diagonal entries. If we denote by $d_k\,,\ k = 1\,,\dots,\,r$, the nonzero entries of D and by $\widetilde{u}_k$ the k-th column of U, then

$$H = UDU^T = \sum_{k=1}^{r} d_k\,\widetilde{u}_k\,\widetilde{u}_k^T = \sum_{k=1}^{r} s_k\,u_k\,u_k^T,$$

where

$$u_k = i^{\frac{1-s_k}{2}}\sqrt{d_k}\,\widetilde{u}_k.$$

If the matrix H is complex the signs $s_k = \pm 1$ can be chosen arbitrarily. If H is real, however, then H is both real and symmetric and, therefore, orthogonally diagonalizable. Hence, $H = W\Lambda W^T$ for some orthogonal matrix W and some real diagonal matrix Λ of rank r. Of course, now some of the eigenvalues λ_k may be negative, but, if we denote by w_k the k-th column of W, by s_k the sign of the corresponding eigenvalue λ_k, and define

$$u_k = \sqrt{|\lambda_k|}\,w_k,$$

we obtain again $H = \sum_{k=1}^{r} s_k\,u_k\,u_k^T$, but now with real vectors u_k. □

As a straightforward consequence of Lemma 4.9, we obtain the following first-order description of the symplectic matrices $\widehat{I}(\varepsilon)$ in formula (4.1):

Corollary 4.10. *Let* $r \in \{1,\,\dots,\,2n\}$ *and let* $\widehat{I}(\varepsilon)$ *be as in formula (4.1) with* $\mathrm{rank}(B) = r$*. Then there exists an orthogonal set of* r *nonzero vectors* $u_1,\dots,u_r \in \mathbb{C}^{2n}$ *and signs* $s_k \in \{-1\,,\,1\},\ k = 1,\dots,r$*, such that*

$$\widetilde{I}(\varepsilon) := \prod_{k=1}^{r}\left(I + \varepsilon s_k u_k\,u_k^T\,J\right) = I_{2n} + \varepsilon B + O(\varepsilon^2).$$

If $\widehat{I}(\varepsilon)$ *is real, then the vectors* $\{u_1\,,\,\dots\,,\,u_r\}$ *can be chosen to be real.*

Hence, if we define $\widetilde{S}(\varepsilon)$ as in (4.29) above for the vectors u_k and the signs s_k in Corollary 4.10, then at least r eigenvalues of both $\widehat{S}(\varepsilon)$ in (4.2) and $\widetilde{S}(\varepsilon)$ in (4.29) will have asymptotic expansions with the same leading term. In other words, Corollary 4.10 allows us to analyze any small symplectic perturbation $\widehat{S}$ of a symplectic matrix S as the effect of the consecutive application of r rank one multiplicative perturbations (at least when analyzing the most likely behavior, described in Section 4.3). We stress that we are not describing here the set of all possible structure-preserving perturbations, but the set of all possible *leading terms* of structure-preserving perturbations. Since generic behavior of the perturbed eigenvalues depends on the *first order* perturbation matrix B in (1.3), this will be enough for our purposes.

4.5.2 Numerical examples

For each situation described in Section 4.3 we shall specify an unperturbed symplectic matrix with the appropriate spectral structure, and then use MATLAB R2016B to first randomly generate thousands of pairs of structure-preserving perturbations $\widehat{S}(\varepsilon)$ and $\widetilde{S}(\varepsilon)$, and then to compute (and plot) their eigenvalues. To be more precise, once the appropriate unperturbed matrix S is chosen,

1. we specify a (low) rank r, and generate r random linearly independent [2] vectors $u_1, \ldots, u_r$, and a random list of signs $\{s_1, \ldots, s_r\}$ to construct a perturbation

$$\widetilde{S}(\varepsilon) = \left(\prod_{k=1}^{r} \left(I + \varepsilon s_k u_k \, u_k^T \, J \right) \right) S = \left(I + \varepsilon \, B + \cdots \right) S$$

   as in (4.29).

2. Once $\widetilde{S}(\varepsilon)$ is constructed, the $2n \times 2n$ matrix B above is partitioned into four $n \times n$ blocks, which are then modified according to Lemma 4.1, i.e., if

$$B = \left[\begin{array}{cc} B_{11} & B_{12} \\ B_{21} & -B_{11}^T \end{array} \right], \qquad B_{12}^T = B_{12}, \quad B_{21}^T = B_{21},$$

   random $n \times n$ matrices C_{11}, C_{12}, C_{21} are generated, with C_{12} and C_{21} symmetric, in order to construct a perturbation

$$\widehat{S}(\varepsilon) = \left[\begin{array}{cc} G & G\,E \\ C\,G & G^{-T} + C\,G\,E \end{array} \right] S,$$

   where $G = I_n + \varepsilon B_{11} + \varepsilon^2 C_{11}$, $\quad E = \varepsilon B_{12} + \varepsilon^2 C_{12}$, $\quad C = \varepsilon B_{21} + \varepsilon^2 C_{21}$. This perturbation will have the same first-order matrix B, but the difference $\widehat{S}(\varepsilon) - S$ will have typically full rank.

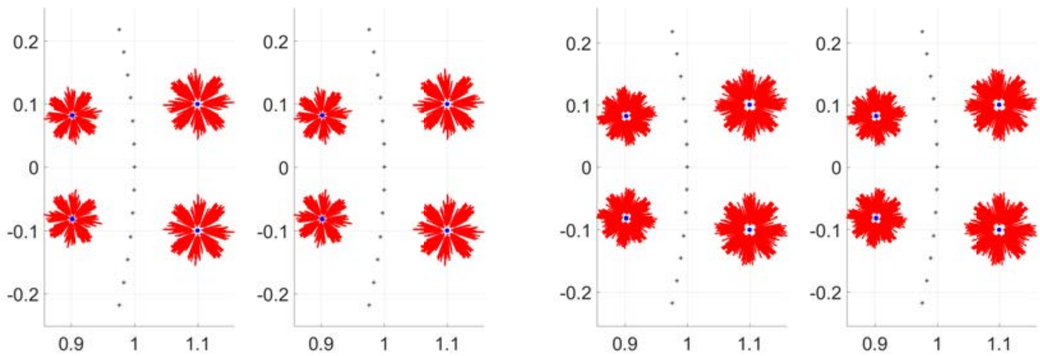
3. Finally, hundreds of equally spaced values for ε are sampled from a certain interval $[\varepsilon_{\min}, \varepsilon_{\max}]$, and the eigenvalues of the corresponding evaluations for each perturbation $\widetilde{S}(\varepsilon)$ and $\widehat{S}(\varepsilon)$ are computed and plotted via MATLAB. Those of $\widetilde{S}(\varepsilon)$ are placed on the left half of the figure, those of $\widehat{S}(\varepsilon)$ on the right half. Perturbed eigenvalues are plotted in red, unperturbed ones in blue.

[2] In order to reduce the workload of the generating process we relax the orthogonality condition on the u_j to just linear independence. This is enough to guarantee both rank and symplecticity

Example 4.11. *Let*

$$S = \left[\begin{array}{c|c} \Lambda & 0 \\ \hline 0 & \Lambda^{-T} \end{array}\right], \qquad \textit{with} \quad \Lambda = I_4 \otimes \left[\begin{array}{cc} a & b \\ -b & a \end{array}\right] + J_8(0),$$

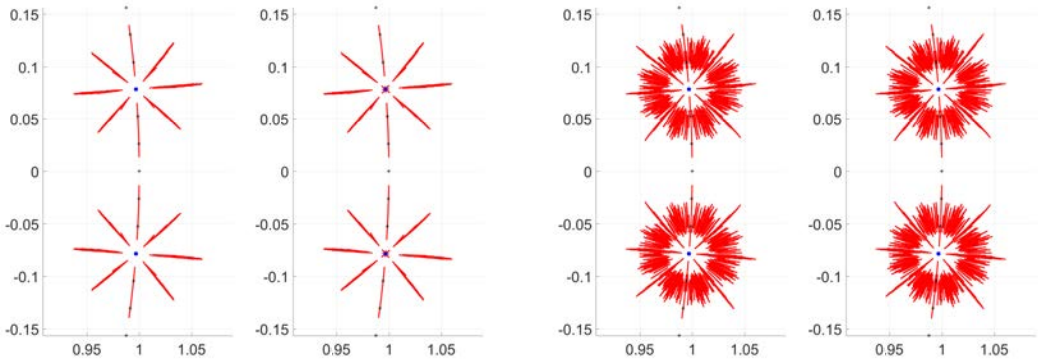
where $a^2 + b^2 \neq 1$*. The matrix* S *is symplectic with four Jordan blocks of size 4, each of them associated with one of the four distinct eigenvalues* $\{\,\lambda\,,\,\lambda^{-1}\,,\,\overline{\lambda}\,,\,\overline{\lambda}^{-1}\,\}$ *of* S*. Figure 4.5.1 displays the eigenvalues of symplectic perturbations to* S *when* $a = 1.1$ *and* $b = 0.1$*. The plot in Figure 4.5.1a corresponds to perturbations with rank* $r = 1$*, while the one in Figure 4.5.1b corresponds to rank* $r = 2$*. In either case, 100 different random structured perturbations* $\widetilde{S}(\varepsilon)$ *have been created by randomly generating appropriate sets of vectors as described in Step 1 above. Then, each of those 100 perturbations is modified to create the corresponding* $\widehat{S}(\varepsilon)$*, as described in Step 2. Finally, for each of pair of perturbations* $\widetilde{S}(\varepsilon)$*,* $\widehat{S}(\varepsilon)$*, the interval* $[\,0\,,\,10^{-5}\,]$ *for* ε *is uniformly sampled with step* 10^{-7}*, and* MATLAB *computes and plots the eigenvalues of the evaluations of both* $\widetilde{S}(\varepsilon)$ *and* $\widehat{S}(\varepsilon)$ *at the values of* ε *given by the samples.*



(a) rank $(B) = 1$ (b) rank $(B) = 2$

Figure 4.5.1: Low rank perturbations of a symplectic matrix S with four e-vals. out of the unit circle

Example 4.12. The case $S, \widehat{S} \in \mathbb{R}^{2n\times 2n}$, $|\lambda| = 1$. *Let* S *be the symplectic matrix in Example 4.11 but with* $a^2 + b^2 = 1$ *and* $b \neq 0$*.* S *is symplectic with four Jordan blocks of size 4, two of them associated with each of the two distinct eigenvalues* $\{\lambda\,,\,\overline{\lambda}\}$*, where* $\lambda = a + i\,b$ *lies on the unit circle. Figure 4.5.2 displays the eigenvalues of symplectic perturbations to* S *when* $a = \cos(\pi/40)$ *and* $b = \sin(\pi/40)$*. The plot in Figure 4.5.2a corresponds to perturbations with rank* $r = 1$*, while the one in Figure 4.5.2b corresponds to rank* $r = 2$*. In either case, 100 different random structured perturbations* $\widetilde{S}(\varepsilon)$ *have* *been created by randomly generating appropriate sets of vectors as described in Step 1 above. Then, each of those 100 perturbations is modified to create the corresponding* $\widehat{S}(\varepsilon)$*, as described in Step 2. Finally, for each of pair of perturbations* $\widetilde{S}(\varepsilon)$, $\widehat{S}(\varepsilon)$*, the interval* $[\,10^{-6}\,,\,10^{-5}\,]$ *for* ε *is uniformly sampled with step* 10^{-7}.



(a) $\operatorname{rank}(B) = 1$ (b) $\operatorname{rank}(B) = 2$

Figure 4.5.2: Low rank perturbations of a symplectic matrix S with two e-vals. on the unit circle

First, we observe that the difference in behavior for the two ranks is consistent with the expansions in Section 4.3.2: take first Figure 4.5.2a, where $\operatorname{rank}(B) = 1$. Since we are in the real case, the number $\xi = \xi_{1,k}$ in the asymptotic expansions (4.24) is a single real number. The 8-legged star visible at each unperturbed eigenvalue on Figure 4.5.2a corresponds to the superposition of two rotated versions of the four-legged star depicted by the fourth roots of unity, one for each of the two possible signs for ξ: if $\xi > 0$, then two of the four perturbed eigenvalues stay on the unit circle, while the other two move away from it orthogonally, as predicted in Table 4.3.2. When $\xi < 0$, the fourth roots of ξ give rise to the four remaining escape directions in the figure. As to the plot on Figure 4.5.2b, where $\operatorname{rank}(B) = 2$, both pairs of Jordan blocks of S associated with λ may have opposite signs in the sign characteristic, so the matrix $\Phi_1(B)$ may have nonreal eigenvalues. Thus, its fourth roots may be anywhere and, as shown in the plot, the spectra of both $\widetilde{S}$ and $\widehat{S}$ can move away from λ in any direction.

Notice that in Figure 4.5.2b, the rank of B is enough for $\widetilde{S}$ to break the two Jordan blocks associated with each eigenvalue. Hence, the asymptotic expansions (4.24) explain the behavior of all perturbed eigenvalues, and no significant differences between the spectra of $\widetilde{S}$ and $\widehat{S}$ can be appreciated with the naked eye. In Figure 4.5.2a, however, one can spot a very small additional 8-legged star in the graph on the right, corresponding to perturbations of type $\widehat{S}$, which is not present in the one at the left for the auxiliary perturbations $\widetilde{S}$. The reason for this is that $\operatorname{rank}(B) = 1$ is not enough to destroy the whole

Jordan structure associated with each eigenvalue. The perturbations $\widetilde{S}$ destroy only one Jordan block, while $\widehat{S}$ is in general of full rank and is thus able to undo the complete λ-Jordan structure of S. However, the behavior of those perturbed eigenvalues which escape fastest away from the unperturbed ones is still the same for both perturbations.

Remark 4.13. *The discussion above also explains the different behavior of the two matrices S_1 and S_2 in (1.2) observed in Figure 1. A calculation of the symplectic canonical forms of S_1 and S_2, respectively, reveals that the two Jordan blocks of size $n_1 = 1$ associated with the eigenvalue $\lambda = i$ have opposite signs ± 1 in the sign characteristic for S_1, but equal signs for S_2. An analogous observation holds for the eigenvalue $\lambda = -i$. Thus, if we consider random symplectic perturbations of S_1, then the matrix Diag$\{+1, -1\}H_1$ will generically be nonsymmetric and may thus have complex eigenvalues, so the perturbed eigenvalues may escape into any direction. For S_2, in contrast, we are exactly in the situation highlighted above, where all signs in the sign characteristic corresponding to Jordan blocks of size n_1 are equal, resulting in the fact that the eigenvalues $\xi_{k,l}$ are real. Since $n_1 = 1$ is odd, the perturbed eigenvalues coming from each of the two Jordan blocks associated with $\lambda = i$ (or $\lambda = -i$, respectively) have to stay on the unit circle.*

Example 4.14. The case $\lambda \in \{-1, 1\}$, n_j even. *Let*

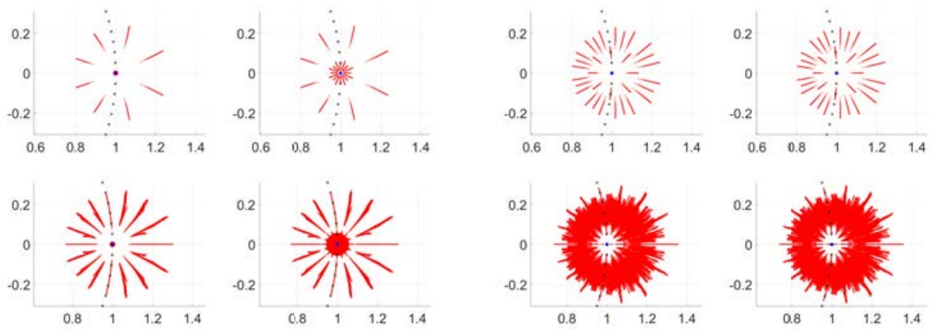
$$S = \left[\begin{array}{c|c} \Lambda & 0 \\ \hline 0 & \Lambda^{-T} \end{array}\right], \qquad \textit{for} \qquad \Lambda = \mathcal{J}_8(1) \oplus \mathcal{J}_8(1),$$

where $\mathcal{J}_8(1)$ stands for a Jordan block of size 8 associated with $\lambda = 1$. The matrix S is symplectic with four Jordan blocks of size 8. As before, Figure 4.5.3 displays the eigenvalues of 100 *random structured symplectic perturbations of S of the form* (1.3), *with* $\operatorname{rank}(B) = 1$ *in Figure 4.5.3a and with* $\operatorname{rank}(B) = 4$ *in Figure 4.5.3b. Within each of the (a),(b) versions, the perturbations of type $\widetilde{S}$ are displayed on the left, and those of type $\widehat{S}$ on the right. On the top of each figure we single out one amongst the* 100 *randomly generated perturbations, at the bottom we show the superposition of the 99 remaining ones. The values for ε have been sampled from the interval $[\,0\,,\,10^{-5}\,]$ at uniform steps of length 10^{-7}.*

The discussion here is more or less the same as in Example 4.12. In fact, for even-sized blocks associated with ± 1, the behavior of perturbed eigenvalues is similar to that of those coming from eigenvalues elsewhere on the unit circle, since the asymptotic expansions (4.26), though more detailed, are basically the same as those in (4.24).

The main differences between the auxiliary perturbations $\widetilde{S}$ and the standard ones $\widehat{S}$ can be seen, as in Example 4.12 above, only when the rank of B is not enough by itself to destroy the whole Jordan structure associated with λ: in Figure 4.5.3a, where $\operatorname{rank}(B) = 1$, the perturbations $\widetilde{S}(\varepsilon)$ can only break one single 8×8 Jordan block associated with $\lambda = \pm 1$, while in general, every perturbation $\widehat{S}(\varepsilon)$ destroys the whole λ-Jordan structure. Hence, there is again an additional cluster of eigenvalues for $\widehat{S}(\varepsilon)$ very close to $\lambda = 1$. Since these eigenvalues come from the $O(\varepsilon^2)$ terms in $\widehat{S}(\varepsilon)$, their leading exponent is larger than $1/8$ and, therefore, they move away from λ much slower than the ones coming from the first order term B. These coincide, to first order, with the eigenvalues of $\widetilde{S}$.

In Figure 4.5.3b, where $\operatorname{rank}(B) = 4$, the first order terms have enough rank to break all Jordan blocks corresponding to $\lambda = 1$. Thus, all perturbed eigenvalues are described by Theorem 4.8, and no differences are visible between the left and right plots.



(a) $r = 1$ (b) $r = 4$

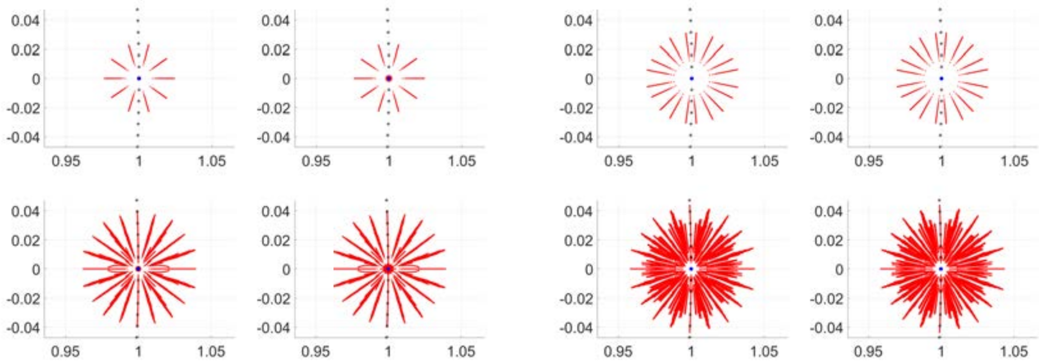
Figure 4.5.3:

Example 4.15. . **The case** $\lambda \in \{-1, 1\}$, n_j **odd**. *Let* S *be a* 20×20 *symplectic matrix with the same block structure as the one in Example 4.14, but now taking*

$$\Lambda = \mathcal{J}_5(1) \oplus \mathcal{J}_5(1)\,.$$

Then S *has a single eigenvalue* $\lambda = 1$ *with four Jordan blocks of size 5. Notice that, since all blocks are of the same size,* $\operatorname{rank}(B) = \rho$*, where* ρ *is given by* (4.3)*. Hence,* ρ *is even whenever* $\operatorname{rank}(B)$ *is either* 2 *or* 4 *(any rank larger than or equal to* 4 *destroys all four Jordan blocks). Figure 4.5.4 displays the eigenvalues of* 100 *random symplectic perturbations of* S *with the same format as in Figure 4.5.3 (i.e., one perturbation singled out at the top, the remaining 99 at the bottom, the auxiliary perturbations* $\widetilde{S}$ *on the left, and* $\widehat{S}$ *on the right). Now the values for* ε *are taken from partitioning the interval* $[\,0\,,\,10^{-7}\,]$ *with step* 10^{-9}.

We only discuss the case when ρ is even (the case when ρ is odd will be dealt with in Chapter 5): Figure 4.5.4a corresponds to $\operatorname{rank}(B) = 2$ and Figure 4.5.4b to $\operatorname{rank}(B) = 4$. As in the two previous examples, no significant differences are visible in Figure 4.5.4b, since the perturbation $\widetilde{S}$ has enough rank to break all Jordan blocks. In



(a) $r = 2$ (b) $r = 4$

Figure 4.5.4:

Figure 4.5.4a, on the other hand, we see on top the superposition of two 5-legged stars, one corresponding to the fifth roots of ξ, the other to the fifth roots of $-\xi$, as explained in (4.28). As in the previous examples, a tiny additional cluster of eigenvalues is visible on the right of Figure 4.5.4a for the perturbation $\widehat{S}$, due to the additional rank provided by the $O(\varepsilon^2)$ terms.

Chapter 5

Structured perturbation of eigenvalues of symplectic matrices: The nongeneric case

In this Chapter we analyze in detail the special situation identified in §4.3.3.2, where the usual unstructured perturbation behavior is incompatible with the spectral constraints imposed by symplecticity: as already explained in the Abstract, what we have called the nongeneric case corresponds to the one in which, in the absence of structure, the rank of the perturbation would break an odd number of odd-sized Jordan blocks corresponding either to the eigenvalue 1 or -1. Since this is not allowed by symplecticity, one among that odd number of Jordan blocks does not break, but increases its size by one becoming an even-sized block.

In our analysis above, in Chapter 4, that corresponds to a situation when the matrix H_j has at least one zero eigenvalue. Thus, the corresponding perturbed eigenvalue is not covered by any of our previous results, and a new, separate analysis has to be done to obtain the appropriate asymptotic expansion. This is precisely what we do in this Chapter.

5.1 The nongeneric case: $\lambda \in \{ -1 , 1\}$, n_j odd, ρ odd.

As mentioned above, the case when $\lambda \in \{-1, 1\}, n_j$ is odd and r =rank$(B) = \sum_{k=1}^{j-1} r_k+$ ρ, $1 \leq \rho \leq r_j$ with odd ρ is not fully covered by Theorem 3.2, since under these conditions the Schur complement $(\Phi_j/\Phi_{j-1})\,(B)$ has rank ρ, but at most $\rho - 1$ nonzero eigenvalues. We stress that Theorem 3.2 still applies to the perturbed eigenvalues corresponding to the nonzero eigenvalues of the Schur complement. However , to obtain the asymptotic expansions associated with the zero eigenvalues of $(\Phi_j/\Phi_{j-1})\,(B)$, we will have to start from scratch: as can be seen in §3.1.2, one possible way of deriving these asymptotic expansions is by making use of the Newton Diagram (see §2.2.2.1 for more details).

5.1.1 The main result

Going back to the case $\lambda \in \{-1\,,\,1\}$, n_j odd, ρ odd under scrutiny in this section, since we are in the nongeneric case, the relevant information will not be available just from using eigenvectors only. As we shall see, we will need to make use of the second vectors in the Jordan chains as well: let $X_j^{(2)}$ (resp., $Y_j^{(2)}$) be the $2n \times r_j$ matrix (resp., the $r_j \times 2n$ matrix) whose columns (resp., rows) are the second vectors in the right (resp., left) Jordan chains of S associated with λ for all Jordan blocks of size n_j, placed in the same order as they appear in the Jordan structure. Now, define $r_j \times r_j$ matrices $\Phi_j^{(R)}(B)$ and $\Phi_j^{(L)}(B)$ as

$$\Phi_j^{(R)}(B) = W_j\, B\, X_j^{(2)}, \qquad \Phi_j^{(L)}(B) = Y_j^{(2)}\, B\, Z_j, \tag{5.1}$$

where W_j, Z_j are as defined in §2.2.3. Now, for every index $k \in \{1, \ldots, r_j\}$, we define the matrix $\Phi_j^{(k,L)}(B)$ (resp., $\Phi_j^{(k,R)}(B)$) as the result of replacing the k-th row (resp., the k-th column) of $\Phi_j(B)$ by the k-th row of $\Phi_j^{(L)}(B)$ (resp., the k-th column of $\Phi_j^{(R)}(B)$).

In terms of these matrices, we define the real number

$$\begin{aligned}\widehat{\alpha} \;=\; & \sum_{k=1}^{r_j} \left(\pi_\rho \left(\left(\Phi_j^{(k,L)} / \Phi_{j-1} \right)(B) \right) + \pi_\rho \left(\left(\Phi_j^{(k,R)} / \Phi_{j-1} \right)(B) \right) \right) + \\ & \qquad + \delta_{n_j - n_{j+1},1} \sum_{\omega} \pi_\rho \left(\left(\Phi_{j+1}[\omega] / \Phi_{j-1} \right)(B) \right),\end{aligned} \tag{5.2}$$

where $\pi_\rho(\cdot)$ stands for the product of the ρ nonzero eigenvalues of its matrix argument[1], each matrix quotient $(M_1/M_2)(B)$ stands for the Schur complement of $M_2(B)$ in $M_1(B)$, and $\Phi_{j+1}[\omega]$ stands for the principal submatrix of $\Phi_{j+1}(B)$ corresponding to the indices in ω, where ω is any index subset of $\{1, \ldots, f_{j+1}\}$ (with f_{j+1} is as defined in (2.28)) with cardinality r such that

(i) its first f_{j-1} indices are $\{1,\, \ldots,\, f_{j-1}\}$,

(ii) its next $\rho - 1$ indices are taken from the subset $\{f_{j-1} + 1,\, \ldots,\, f_j\}$, and

(iii) its last index is taken from the subset $\{f_j + 1,\, \ldots\, f_{j+1}\}$.

---

[1]Since either $\Phi_{j+1}[\omega]$, $\Phi_j^{(k,L)}(B)$ or $\Phi_j^{(k,R)}(B)$ have rank at most $\operatorname{rank} B = r$ and $\Phi_{j-1}(B)$ is supposed to be nonsingular, the Schur complement of $\Phi_{j-1}(B)$ in either of $\Phi_{j+1}[\omega]$, $\Phi_j^{(k,L)}(B)$ or $\Phi_j^{(k,R)}(B)$ has generically $r - f_{j-1} = \rho$ nonzero eigenvalues, where f_{j-1} is as defined in (2.28).

Notice that each matrix of $\Phi_{j+1}[\omega]$, $\Phi_j^{(k,L)}(B)$ or $\Phi_j^{(k,R)}(B)$ can be written as a product of three matrices with B in the middle, so the rank of either of the three matrices is at most the rank of B.

With this terminology, we are now in the position of stating the main result in this section:

Theorem 5.1. *Let S be a $2n \times 2n$ symplectic matrix with Jordan structure (2.24) associated with an eigenvalue $\lambda \in \{-1, 1\}$. Let $\widetilde{S}$ be an arbitrary structured multiplicative perturbation of S of the form* (4.29) *with first order perturbation matrix B, let r =rank(B) and let ρ be as defined in* (4.3)*. If r is less than the geometric multiplicity of λ, and both n_j and ρ are odd, then there are generically $n_j(\rho-1)$ perturbed eigenvalues with asymptotic expansions (4.26), and $n_j - 1$ perturbed eigenvalues with asymptotic expansions*

$$\widehat{\lambda}_k = \lambda\left(1 + \xi^{\frac{1}{n_j-1}}\varepsilon^{\frac{1}{n_j-1}}\right) + o\left(\varepsilon^{\frac{1}{n_j-1}}\right),$$

where $k = 1, \ldots, n_j - 1$ runs over all distinct complex (n_j-1)-th roots of

$$\xi = \frac{\widehat{\alpha}}{\prod_{i=1}^{\rho-1}\xi_i}, \tag{5.3}$$

where $\widehat{\alpha}$ is given by (5.2)*, and the ξ_i are the nonzero eigenvalues of the Schur complement $(\Phi_j/\Phi_{j-1})(B)$.*

Since proving Theorem 5.1 is somewhat involved, we shall first outline the general setup for its proof, decoupling the influence of eigenvalues other than λ. Then, we will work out in §5.1.2 the details of an illustrative low-dimensional example, where the main ideas of the proof will be presented. The proof itself in full generality is the content of §5.1.3.

Assuming, as before, that the unperturbed symplectic matrix S has Jordan form (2.24), the characteristic polynomial of the perturbed matrix $\widetilde{S}(\varepsilon)$ can be written as

$$\det\left(zI - \widetilde{S}(\varepsilon)\right) = \det\left(zI - (I + \varepsilon B + O(\varepsilon^2))\,S\right) =$$
$$= \det\left(zI - \left(I + \varepsilon\widetilde{B} + O(\varepsilon^2)\right)\left[\begin{array}{c|c} \mathcal{J} & \\ \hline & \widehat{\mathcal{J}} \end{array}\right]\right),$$

where $\mathcal{J}$ contains the Jordan blocks associated with λ, $\widehat{\mathcal{J}}$ contains the Jordan blocks corresponding to eigenvalues other than λ, and

$$\widetilde{B} = \left[\begin{array}{c} Q \\ \hline \widehat{Q} \end{array}\right] B \left[\begin{array}{c|c} P & \widehat{P} \end{array}\right] = \left[\begin{array}{c|c} QBP & QB\widehat{P} \\ \hline \widehat{Q}BP & \widehat{Q}B\widehat{P} \end{array}\right] = \left[\begin{array}{c|c} \widetilde{B}_{11} & \widetilde{B}_{12} \\ \hline \widetilde{B}_{21} & \widetilde{B}_{22} \end{array}\right].$$

We shall see later that the terms of order ε^2 play no role whatsoever in the leading terms of the asymptotic expansion, so we may disregard the $O(\varepsilon^2)$ terms above. If we partition the matrix zI, and make use of Schur complements, we may factorize

$$P(z\,,\,\varepsilon) = \det\left(\, zI - (I + \varepsilon B)\, S\,\right) = \widehat{p}(z\,,\,\varepsilon)\, p(z\,,\,\varepsilon)$$

for

$$\begin{aligned} \widehat{p}(z\,,\,\varepsilon) &= \det\,(M)\,, \\ p(z\,,\,\varepsilon) &= \det\Big(zI - \big(I + \varepsilon \widetilde{B}_{11}\big)\mathcal{J} - \varepsilon^2\left(\Psi(z\,,\,\varepsilon)\right)\Big)\,, \end{aligned}$$

where

$$\begin{aligned} M &= zI - \big(I + \varepsilon \widetilde{B}_{22}\big)\widehat{\mathcal{J}}, \\ \Psi(z\,,\,\varepsilon) &= \widetilde{B}_{12}\widehat{\mathcal{J}}\; M^{-1}\, \widetilde{B}_{21}\mathcal{J}. \end{aligned}$$

For small values of ε, the eigenvalues of $\big(I + \varepsilon \widetilde{B}_{22}\big)\widehat{\mathcal{J}}$ are close to the ones of $\widehat{\mathcal{J}}$. Hence, if $\lambda(\varepsilon)$ is an eigenvalue of the perturbed matrix $\widetilde{S}(\varepsilon)$ close to λ, it cannot be a root of $\widehat{p}(z\,,\,\varepsilon)$, so it must be a root of $p(z\,,\,\varepsilon)$. Thus, we expand

$$\begin{aligned} p(z\,,\,\varepsilon) &= \det\Big(zI - \Big(I + \varepsilon \widetilde{B}_{11}\Big)\,\mathcal{J} + O\left(\varepsilon^2\right)\Big) = \\ &= \det\left((z-\lambda)I - H(\varepsilon)\right), \end{aligned}$$

where

$$H(\varepsilon) = \mathcal{J}_0 + \varepsilon\left(\lambda \widetilde{B}_{11} + \widetilde{B}_{11}\,\mathcal{J}_0\right) + O\left(\varepsilon^2\right)$$

for $\mathcal{J}_0 = \mathcal{J} - \lambda\, I$. Again, since the terms of order ε^2 will play no role whatsoever in the leading terms of the asymptotic expansion, we replace the matrix $H(\varepsilon)$ by its leading matrix

$$\widehat{H}(\varepsilon) = \mathcal{J}_0 + \varepsilon\left(\lambda \widetilde{B}_{11} + \widetilde{B}_{11}\,\mathcal{J}_0\right). \tag{5.4}$$

Hence, all we need to identify the leading terms in the asymptotic expansions of the perturbed eigenvalues is to find the Newton Diagram for this leading matrix $\widehat{H}(\varepsilon)$.

Before we deal with the general case in §5.1.3, we shall briefly explore a specific example, which will help us to highlight the main ideas behind the proof of Theorem 5.1.

5.1.2 An illustrative example

Suppose that S is a 12-by-12 symplectic matrix, with a single eigenvalue $\lambda = 1$ and a Jordan structure comprising four blocks of size three. Then, our leading matrix $\widehat{H}(\varepsilon)$, defined in (5.4) above, can be partitioned as

$$\widehat{H}(\varepsilon) = \left[\begin{array}{ccc|ccc|ccc|ccc} * & +1 & * & * & * & * & * & * & * & * & * & * \\ \clubsuit & * & +1 & \clubsuit & * & * & \clubsuit & * & * & \clubsuit & * & * \\ \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * \\ \hline * & * & * & * & +1 & * & * & * & * & * & * & * \\ \clubsuit & * & * & \clubsuit & * & +1 & \clubsuit & * & * & \clubsuit & * & * \\ \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * \\ \hline * & * & * & * & * & * & * & +1 & * & * & * & * \\ \clubsuit & * & * & \clubsuit & * & * & \clubsuit & * & +1 & \clubsuit & * & * \\ \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * \\ \hline * & * & * & * & * & * & * & * & * & * & +1 & * \\ \clubsuit & * & * & \clubsuit & * & * & \clubsuit & * & * & \clubsuit & * & +1 \\ \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * & \blacklozenge & \spadesuit & * \end{array}\right],$$

where the '$+1$' entries indicate entries of the form $1 + O(\varepsilon)$. We choose to highlight the entries marked with $\clubsuit$, $\spadesuit$, $\blacklozenge$ because, as we shall see, the entries in those positions will be the ones generically involved in the formulas for the leading coefficients. Notice that the $\blacklozenge$ entries belong to $\Phi_1(B)$, and the $\clubsuit$ belong to the matrix $\Phi_1^{(L)}(B)$ defined in (5.1). As to the $\spadesuit$s, each one is the sum of an entry of $\Phi_1^{(R)}(B)$ plus the $\blacklozenge$ in its same block (the latter comes from the product $\widetilde{B}_{11}\mathcal{J}_0$ in (5.4)[2]).

Notice that this situation is the same as in Example 4.15 above. However, in this case we will explore the behavior of perturbations of rank $r = 3$, instead of ranks 2 or 4. If the perturbation matrix B has rank 3, we know from [38] that $\widetilde{S}(\varepsilon)$ will have, generically, one Jordan block of size four for the eigenvalue 1, and 8 perturbed eigenvalues moving away from $\lambda = 1$. Recall that formula (4.27) shows that $\Phi_1(B) = D_1 H_1$ for D_1, H_1 as in §4.3.3.2. Hence, $\Phi_1(B)$ has trace zero and its nonzero eigenvalues are coupled in opposite pairs. Since $\operatorname{rank}\Phi_1(B) \le r = 3$, we conclude that $\Phi_1(B)$ has generically two nonzero eigenvalues $\pm\xi$, which, according to Theorem 3.2, give rise to six perturbed eigenvalues with asymptotic expansions

$$\widehat{\lambda}_k(\varepsilon) = 1 \pm \xi^{\frac{1}{3}}\varepsilon^{\frac{1}{3}} + O(\varepsilon^{\frac{2}{3}}), \qquad k = 1, 2, 3, \tag{5.5}$$

where the different values for k correspond to the three distinct complex cubic roots of ξ.

In this situation, however, we are mostly interested in the asymptotic expansions of the remaining two perturbed eigenvalues. One can check that the characteristic polynomial

[2] As shall be seen later, this last $\blacklozenge$ plays no role whatsoever in the formulas, so one can think of each $\spadesuit$ as just being the corresponding entry of $\Phi_1^{(R)}(B)$.

of $\widetilde{S}(\varepsilon)$ for this example can be factorized as

$$p(z\,,\,\varepsilon) = z^4\,\left[z^8 + \alpha_1(\varepsilon)z^7 + \ldots + \alpha_6(\varepsilon)z^2 + \alpha_7(\varepsilon)z + \alpha_8(\varepsilon)\right], \tag{5.6}$$

where

$$\alpha_k(\varepsilon) = (-1)^k \mathbb{E}_k\left[\widehat{H}(\varepsilon)\right] = \widehat{\alpha}_k \varepsilon^{a_k} + o(\varepsilon^{a_k})\ ,$$

and for each k the expression $\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right]$ stands for the sum of all k-by-k principal minors of $\widehat{H}(\varepsilon)$. Furthermore, each leading exponent a_k of $\alpha_k(\varepsilon)$ turns out to be the minimum number of Jordan blocks associated with λ whose sizes sum up to at least k. The reason for this is that the lowest possible exponent in ε for the term in z^{12-k} is attained at those $k \times k$ principal minors in the sum $\mathbb{E}_k\left[\widehat{H}(\varepsilon)\right]$ containing as many '$+1$' entries as possible. But a '$+1$' entry is in a principal submatrix of $\widehat{H}(\varepsilon)$ only if two consecutive rows in the same Jordan block are included in that submatrix. In fact, the best way to include as many '$+1$'s as possible is to include in the principal submatrix full Jordan blocks of maximal dimensions, since each full Jordan block of size, say, p contributes to the principal minor with just a $O(\varepsilon)$ factor (multiplied by the $p-1$ entries '$+1$' in the block, which do not contribute to the exponent). The larger the Jordan blocks, the fewer blocks are needed to construct the $k \times k$ principal minor, hence the lowest exponent[3] in ε.

Thus, the Newton diagram associated to $p(z\,,\,\varepsilon)$ is as follows.

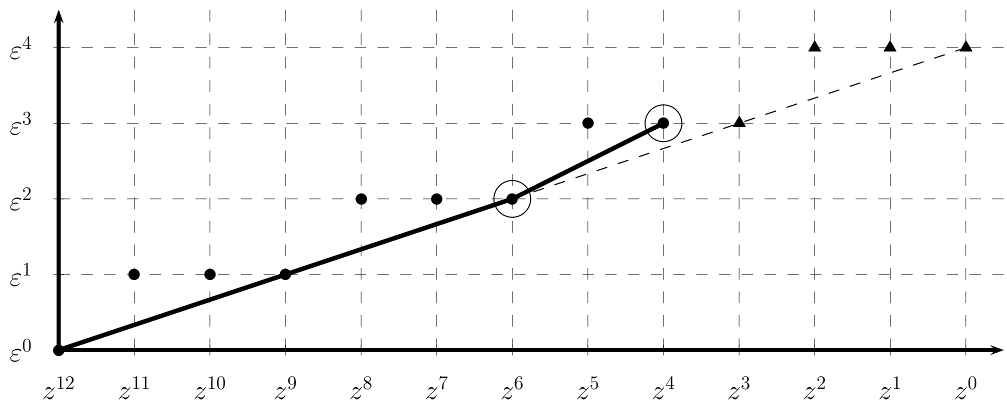


Figure 5.1.1: *Newton diagram for* $p(z\,,\,\varepsilon) = 0$.

The first segment joining the origin and the point of coordinates $(6, 2)$ corresponds to the six $O(\varepsilon^{1/3})$ roots (5.5), but here we are interested mostly in what happens beyond $(6, 2)$: according to our description above of the a_k, the leading exponent of the term $\alpha_7(\varepsilon)$ in $z^5 = z^{12-7}$ is $a_7 = 3$, since we need to sample from at least three 3×3 Jordan

[3] At this point it becomes clear that including any $O(\varepsilon^2)$ term in the submatrix instead of an $O(\varepsilon)$ term would only cause an unnecessary increase in the order in ε.

blocks in order to get 7 rows and columns. Similarly, for the term in $z^4 = z^{12-8}$ we get $a_8 = 3$ as well, since we still need to choose rows and columns from three Jordan blocks. These are the two dots at height 3 plotted in the Figure, and no other dots appear. Hence, provided the coefficient $\widehat{\alpha}_8$ does not vanish, the Newton Diagram for rank $r = 3$, which is obtained as a lower boundary of a convex hull, contains the segment of slope $1/2$ joining the two circled points $(6, 2)$ and $(8, 3)$ (the dashed line represents what is called in [40] the *Newton envelope* corresponding to this Jordan form, i.e., the lowest possible segments which would occur for unstructured perturbations *of any rank*).

Furthermore, the leading coefficients provided by the Newton Diagram for the two asymptotic expansions are the roots of the polynomial equation

$$\widehat{\alpha}_6\,\gamma^2 + \widehat{\alpha}_8 = 0, \tag{5.7}$$

or, equivalently, the two complex square roots of $-\dfrac{\widehat{\alpha}_8}{\widehat{\alpha}_6}$, provided that both $\widehat{\alpha}_6$ and $\widehat{\alpha}_8$ are different from zero (in fact, since both are real, the leading coefficients will be either both real, or both purely imaginary).

Now, notice that since the point $(6, 2)$ is in the Newton Diagram for any generic additive unstructured perturbation to S (i.e., it lies on the Newton envelope), one can use the theory in §2.2.4 to show that $\widehat{\alpha}_6$ is just the sum of all 2×2 principal minors of $\Phi_1(B)$, or, equivalently, the product $-\xi^2$ of its two nonzero eigenvalues.

To find the coefficient $\widehat{\alpha}_8$ (which corresponds to the point $(8, 3)$, which no longer lies on the Newton envelope), we need to characterize all 8-by-8 principal minors of $\widehat{H}(\varepsilon)$ containing exactly five "$+1$" entries; notice that this is just the maximum number of such entries one can get into a 8×8 principal minor of $\widehat{H}(\varepsilon)$. One can easily check that to get those five "$+1$" entries one needs to include in the principal submatrix two full Jordan blocks plus two consecutive rows from a third Jordan block; one such possibility is, for example, choosing the first 8 rows of $\widehat{H}(\varepsilon)$. We show in the left side of Figure 5.1.2 below the result of expanding the minor along the corresponding '+1' entries: notice that not only the ♦ entries are involved in $\widehat{\alpha}_8$, but also the ♣ ones. On the right half of Figure 5.1.2 we show the principal minor corresponding to choosing the first six rows of $\widehat{H}(\varepsilon)$ plus its 8th and 9th rows: notice that in this case it is the ♠ entries, instead of the ♣ ones, which are involved. Also, since we are including five '+1' terms, and each such term changes the sign of the minor when developing it along the '+1' entries, the resulting minor is *minus* the determinant of the 3×3 matrix formed by the unshaded entries.

One can check that all possible choices of two full Jordan blocks plus two consecutive rows from a third one give rise to all possible determinants of 3×3 submatrices which result from replacing one single row (resp., one single column) in a 3×3 principal submatrix of $\Phi_1(B)$ by the corresponding row (resp., column) of the matrix $\Phi_1^{(L)}(B)$

$$
\left[\begin{array}{ccc|ccc|cc}
\ast & +1 & \ast & \ast & \ast & \ast & \ast & \ast \\
\clubsuit & \ast & +1 & \clubsuit & \ast & \ast & \clubsuit & \ast \\
\blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit \\
\hline
\ast & \ast & \ast & \ast & +1 & \ast & \ast & \ast \\
\clubsuit & \ast & \ast & \clubsuit & \ast & +1 & \clubsuit & \ast \\
\blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit \\
\hline
\ast & \ast & \ast & \ast & \ast & \ast & \ast & +1 \\
\clubsuit & \ast & \ast & \clubsuit & \ast & \ast & \clubsuit & \ast
\end{array}\right]
\qquad
\left[\begin{array}{ccc|ccc|cc}
\ast & +1 & \ast & \ast & \ast & \ast & \ast & \ast \\
\clubsuit & \ast & +1 & \clubsuit & \ast & \ast & \ast & \ast \\
\blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit & \ast & \spadesuit & \ast \\
\hline
\ast & \ast & \ast & \ast & +1 & \ast & \ast & \ast \\
\clubsuit & \ast & \ast & \clubsuit & \ast & +1 & \ast & \ast \\
\blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit & \ast & \spadesuit & \ast \\
\hline
\clubsuit & \ast & \ast & \clubsuit & \ast & \ast & \ast & +1 \\
\blacklozenge & \spadesuit & \ast & \blacklozenge & \spadesuit & \ast & \spadesuit & \ast
\end{array}\right]
$$

Figure 5.1.2: *Two different* 8×8 *principal minors of* $\widehat{H}(\varepsilon)$*: on the left, the one including the rows and columns* $\{1, 2, 3, 4, 5, 6, 7, 8\}$*. On the right, the one including the rows and columns* $\{1, 2, 3, 4, 5, 6, 8, 9\}$.

(resp., $\Phi_1^{(R)}(B)$), i.e.,

$$
\underbrace{\begin{bmatrix} \blacklozenge & \blacklozenge & \blacklozenge \\ \blacklozenge & \blacklozenge & \blacklozenge \\ \blacklozenge & \blacklozenge & \blacklozenge \end{bmatrix}}_{\textit{principal submatrix of } \Phi_1(B)} \mapsto \underbrace{\left\{ \begin{bmatrix} \blacklozenge & \blacklozenge & \blacklozenge \\ \blacklozenge & \blacklozenge & \blacklozenge \\ \clubsuit & \clubsuit & \clubsuit \end{bmatrix}, \begin{bmatrix} \blacklozenge & \blacklozenge & \spadesuit \\ \blacklozenge & \blacklozenge & \spadesuit \\ \blacklozenge & \blacklozenge & \spadesuit \end{bmatrix}, \begin{bmatrix} \blacklozenge & \blacklozenge & \blacklozenge \\ \clubsuit & \clubsuit & \clubsuit \\ \blacklozenge & \blacklozenge & \blacklozenge \end{bmatrix}, \dots \right\}}_{\textit{leading coefficients}} \tag{5.8}
$$

Hence, we may describe $\widehat{\alpha}_8$, as the sum of all determinants of this form, i.e.,

$$
\widehat{\alpha}_8 = -\sum_{\sigma} \sum_{k=1}^{3} \left(\overbrace{\det\left(\Phi_1^{(\sigma(k),L)}(B)[\sigma]\right)}^{\text{leading coeff. with } \clubsuit} + \overbrace{\det\left(\Phi_1^{(\sigma(k),R)}(B)[\sigma]\right) + \det(\Phi_1(B)[\sigma])}^{\text{leading coeff. with } \spadesuit} \right),
$$

where σ runs over all ordered index subsets of $\{1, \dots, 4\}$ with cardinality 3, and the matrix $\Phi_1^{(\sigma(k),R)}(B)[\sigma]$ (resp. $\Phi_1^{(\sigma(k),L)}(B)[\sigma]$) is defined as right after (5.1). The matrices on the right in (5.8), for instance, correspond respectively to $\Phi_1^{(3,L)}(B)[\sigma]$, $\Phi_1^{(3,R)}(B)[\sigma]$, and $\Phi_1^{(2,L)}(B)[\sigma]$. Also, in this case there are no additional terms in $\widehat{\alpha}_8$, since there are no 2×2 Jordan blocks.

Notice, however, that

$$
\sum_{\sigma} \sum_{k=1}^{3} \det(\Phi_1(B)[\sigma]) = 0,
$$

since, up to a sign, this is just the leading coefficient of the term of order 3 in the characteristic polynomial of $\Phi_1(B)$. Alternatively, it is also equal to the sum of all possible products of three among the eigenvalues of $\Phi_1(B)$. Recall that, as already mentioned, $\Phi_1(B)$ has generically only two nonzero eigenvalues, so the double sum above is zero.

We conclude that

$$\widehat{\alpha}_8 = -\sum_{\sigma} \sum_{k=1}^{3} \left(\det \left(\Phi_1^{(\sigma(k),R)}(B)[\sigma] \right) + \det \left(\Phi_1^{(\sigma(k),L)}(B)[\sigma] \right) \right)$$

is one possible way of describing $\widehat{\alpha}_8$. But we can write it in a more intrinsic way if we interchange the two summations above, i.e.,

$$\widehat{\alpha}_8 = -\sum_{k=1}^{4} \sum_{\sigma} \left(\det \left(\Phi_1^{(k,R)}(B)[\sigma] \right) + \det \left(\Phi_1^{(k,L)}(B)[\sigma] \right) \right) ,$$

where σ runs over all index sets with cardinality 3 of $\{ 1 \, , \, \ldots \, , \, 4 \}$ and both $\Phi_1^{(k,L)}(B)$ and $\Phi_1^{(k,R)}(B))$ are as defined right after (5.1). Notice that now each inner sum

$$\sum_{\sigma} \det \left(\Phi_1^{(k,*)}(B)[\sigma] \right) ,$$

where $*$ stands either for L or R, is the sum of all possible products of three eigenvalues of the matrix $\Phi_1^{(k,*)}(B)$. But that matrix has rank at most three, since $\operatorname{rank} B = 3$, so we can write

$$\sum_{\sigma} \det \left(\Phi_1^{(k,*)}(B)[\sigma] \right) = \pi_3 \left(\Phi_1^{(k,*)}(B)[\sigma] \right) ,$$

where $\pi_3(\cdot)$ stands for the product of the three only nonzero eigenvalues of its matrix argument. Therefore,

$$\widehat{\alpha}_8 = -\sum_{k=1}^{4} \left(\pi_3 \left(\Phi_1^{(k,R)}(B) \right) + \pi_3 \left(\Phi_1^{(k,L)}(B) \right) \right) . \tag{5.9}$$

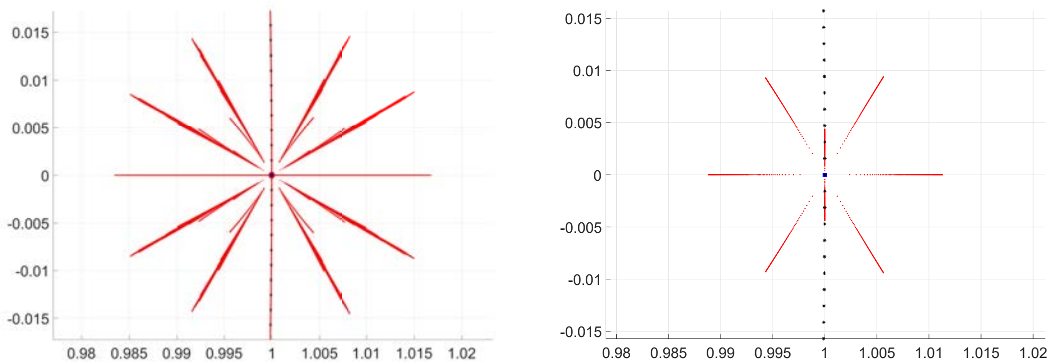
Summarizing, in our example we have that, generically, under a rank three perturbation (4.29), there will be six perturbed eigenvalues with $O(\varepsilon^{1/3})$ asymptotic expansions (5.5), and another two with $O(\varepsilon^{1/2})$ asymptotic expansions

$$\widehat{\lambda}_k = 1 + \left(\frac{\widehat{\alpha}_8}{\xi^2} \right)^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} + o\left(\varepsilon^{\frac{1}{2}} \right) , \tag{5.10}$$

where $\widehat{\alpha}_8$ is given by (5.9) and $(-\xi \, , \, \xi)$ is the pair of nonzero eigenvalues of $\Phi_1(B)$.

The behavior of all these eigenvalues is illustrated in Figure 5.1.3 below on a specific example: we consider the same unperturbed matrix S as in Example 4.15, but now for perturbations of rank three. Figure 5.1.3a on the left displays the spectra of 100 randomly generated structure-preserving perturbations (4.29) of S for ε ranging from 0 to 10^{-5} in increasing steps of 10^{-7}, as before. The rather special configuration of Figure 5.1.3a is explained by the confluence of two special features of the example we have chosen, namely S being real and the rank of B being 3: on one hand, since S is real, then Φ_1

is also real, and real spectral symmetry forces the two only nonzero eigenvalues $-\xi$, ξ of Φ_1 to be either both real or both purely imaginary[4]. As a consequence, the quotient $\widehat{\alpha}_8/\xi^2$ in (5.10) is always real, since both numerator and denominator are real, and the two $O(\varepsilon^{1/2})$-perturbed eigenvalues can only escape from $\lambda = 1$ either both horizontally or both vertically, in opposite directions. As to the six $O(\varepsilon^{1/3})$-perturbed eigenvalues, they escape along a six-legged star anchored either on the real or on the imaginary axis (such six-legged star is actually the union of two three-legged stars, each one being the reflection of the other with respect to the imaginary or the real axis, respectively). To clarify, Figure 5.1.3b shows the eigenvalues of *one single* perturbation to S (i.e., *for one specific random choice* of the vectors u_1, u_2, u_3 in (4.29)). For that particular choice, ξ turns out to be real (thus the six-legged star anchored to the real axis, corresponding to the six $O(\varepsilon^{1/3})$ perturbed eigenvalues), while the quotient $\widehat{\alpha}_8/\xi^2$ turns out to be negative, thus the vertical escape of the two $O(\varepsilon^{1/2})$ perturbed eigenvalues.



(a) (b)

Figure 5.1.3:

5.1.3 Proof of Theorem 5.1

We now analyze the general situation in the statement of Theorem 5.1, in which $\text{rank}(B) = r$ and all Jordan blocks of size larger than n_j, together with $\rho - 1$ Jordan blocks of size n_j, are completely destroyed. This amounts to $r - 1$ blocks: according to [38], only $n_j - 1$ eigenvalues from the last block of size n_j will be perturbed away from λ, and the remaining eigenvalue λ generically persists and joins another Jordan block of size n_j in order to build a new λ-Jordan block of size $n_j + 1$. Again, we are mostly interested in the asymptotic expansions of those $n_j - 1$ eigenvalues, since the ones for the perturbed eigenvalues coming from the $r - 1$ broken blocks are already covered by the theory in Section 4.3.

[4]This is no longer true for perturbations of rank $r \geq 4$, which would, in principle, allow for all four eigenvalues ξ, $\bar{\xi}$, $-\xi$, $-\bar{\xi}$ to be in the spectrum of Φ_1. The same would happen if S were not real.

To simplify the proof we assume that $\lambda = 1$ (if $\lambda = -1$, one just considers the matrix $-S$, which is also symplectic). We know from (4.3.3.2) that the $r_j \times r_j$ Schur complement $(\Phi_j/\Phi_{j-1})(B) = D_j H_j$ has trace zero and that its nonzero eigenvalues can be grouped into opposite pairs $\{-\xi, \xi\}$. Hence, $(\Phi_j/\Phi_{j-1})(B)$, which is known to have rank less than or equal to ρ, will generically have $\rho - 1$ nonzero eigenvalues, plus a zero eigenvalue. The $\rho - 1$ nonzero eigenvalues of the Schur complement give rise to $n_j(\rho - 1)$ perturbed eigenvalues with asymptotic expansions (4.26), since in that case Theorem 3.2 applies.

We now concentrate on the $n_j - 1$ eigenvalues coming from the zero eigenvalue of the Schur complement: for the same reasons as in the example in §5.1.2, the Newton Diagram in this general situation will be as in Figure 5.1.4 below: from the origin up to the leftmost of the two circled dots, the Newton Diagram coincides with the Newton envelope, i.e., it consists of $j - 1$ consecutive segments of slopes $1/n_k$ and horizontal projections of respective lengths $r_k n_k$, $k = 1, \ldots, j - 1$, followed by another segment of slope $1/n_j$ and horizontal projection of length $(\rho - 1)n_j$. Following the parallel with §5.1.2, the factorization (5.6) becomes now

$$p(z, \varepsilon) = z^{2n-K(r)} \left(z^{K(r)} + \alpha_1(\varepsilon) z^{K(r)-1} + \ldots + \alpha_{K(r)-1}(\varepsilon) z + \alpha_{K(r)}(\varepsilon)\right),$$

where

$$K(r) = \sum_{k=1}^{j-1} r_k n_k + \rho n_j - 1.$$

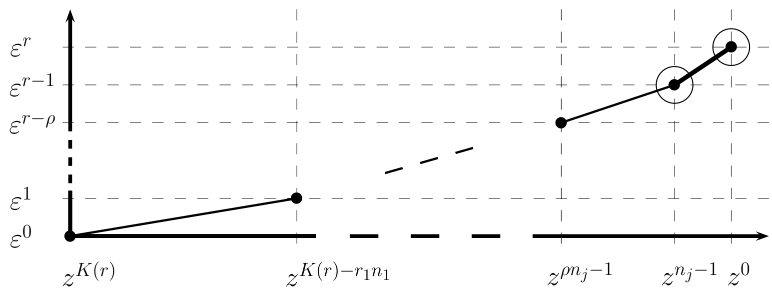


Figure 5.1.4: *Newton diagram* .

With this notation, the leftmost circled point in Figure 5.1.3 is $(K(r) - n_j + 1, r - 1)$ and its presence on the Newton Diagram is equivalent to our assumption that $(\Phi_j/\Phi_{j-1})(B)$ has $\rho - 1$ nonzero eigenvalues. In fact, since that circled point is on the Newton envelope, and following the ideas in §2.2.4, we know that $\widehat{\alpha}_{K(r)-2}$ is the sum of all principal minors of $\Phi_j(B)$ of size $r - 1$ containing their first f_{j-1} rows, with f_{j-1} as in (2.28). This corresponds to the "♦" entries in $\widehat{H}(\varepsilon)$ associated with Jordan blocks of size larger than

n_j , i.e.

$$\widehat{\alpha}_{K(r)-2} = (-1)^{K(r)-2+\overbrace{K(r)-(n_j-1)-(r-1)}^{\text{expanding the principal minors of } \widehat{H}(\varepsilon) \text{ by the "+1"}}} \sum_{\sigma} \det\left(\Phi_j(B)[\sigma]\right)$$

where $\det\left(\Phi_j(B)[\sigma]\right)$ stands for the principal minor of $\Phi_j(B)$ corresponding to rows and columns in the index set σ, and σ runs over all index subsets of $\{1, \ldots, f_j\}$ with cardinality $r-1$ and f_{j-1} first indices $\{1, \ldots, f_{j-1}\}$.

In fact, the (unsigned) sum above is the product of all principal minors of size $\rho-1$ of the Schur complement $(\Phi_j/\Phi_{j-1})(B)$, multiplied by $\det \Phi_{j-1}(B)$. Therefore, it is equal to the product of the $\rho-1$ distinct eigenvalues of $(\Phi_j/\Phi_{j-1})(B)$ by the determinant of $\Phi_{j-1}(B)$, i.e.,

$$\widehat{\alpha}_{K(r)-2} = (-1)^{r+1} \det\, \Phi_{j-1}(B) \prod_{k=1}^{\rho-1} \xi_k, \tag{5.11}$$

where $\xi_1, \ldots, \xi_{\rho-1}$ are the $\rho-1$ nonzero eigenvalues of the Schur complement $(\Phi_j/\Phi_{j-1})(B)$.

Now we show that the rightmost circled point $(K(r), r)$ in Figure 5.1.4 is the last point on the Newton Diagram by showing that the associated coefficient $\widehat{\alpha}$ is generically nonzero: for that we need to identify all principal minors of $\widehat{H}(\varepsilon)$ of size $K(r)$ which are $O(\varepsilon^r)$. We know the index set corresponding to any such minor must contain at least the f_{j-1} first indices (corresponding to all Jordan blocks of size larger than n_j) plus the indices corresponding to the rows and columns of $\rho-1$ among the r_j possible full Jordan blocks of size n_j. Hence, only n_j-1 free indices are left for us to choose. At this point, we distinguish two cases:

- $n_{j+1} < n_j - 1$: in this case, principal minors of order ε^r and size $K(r)$ can only be completed by choosing n_j-1 consecutive rows (i.e., all rows except the first one, or all rows except the last one) from one of the $r_j-\rho$ remaining Jordan block of size n_j (i.e., one whose rows and columns were not yet chosen).
- $n_{j+1} = n_j - 1$: this case is special, since now we may additionally obtain an exponent r for ε by choosing in the last stage the rows and columns of a full Jordan block among the r_{j+1} blocks of size $n_{j+1} = n_j - 1$.

This shows that the coefficient $\widehat{\alpha}_{K(r)}$ can be written as

$$\begin{aligned}\widehat{\alpha}_{K(r)} \;&=\; (-1)^r \Bigg[\sum_{\sigma}\sum_{k=1}^{\rho}\Big(\det\, \Phi_j^{(\sigma(k),R)}(B)[\sigma] \;+\; \det\, \Phi_j^{(\sigma(k),L)}(B)[\sigma] + \det(\Phi_j(B)[\sigma])\Big) + \\ &\qquad + \delta_{n_j-n_{j+1},1} \sum_{\omega} \det\, \Phi_{j+1}(B)[\omega]\Bigg],\end{aligned} \tag{5.12}$$

where σ runs over all ordered index subsets of $\{1,\,\dots,\,f_j\}$ with cardinality r and their first f_{j-1} indices fixed to be $\{1,\,\dots,\,f_{j-1}\}$. The matrices $\Phi_j^{(\sigma(k),R)}(B)[\sigma]$ (resp. $\Phi_j^{(\sigma(k),L)}(B)[\sigma]$) are as defined above, right after (5.1). The δ is a Kronecker delta, i.e., whenever $n_{j+1} = n_j - 1$ (and only in that case), $\widehat{\alpha}$ includes an additional sum of principal minors $\det\,\Phi_{j+1}(B)[\omega]$ of $\Phi_{j+1}(B)$, each one corresponding to the rows and columns indexed by ω, where ω stands for any subset of indices of $\{1,\dots,f_{j+1}\}$ with cardinality r such that

(i) its first f_{j-1} indices are $\{1,\,\dots,\,f_{j-1}\}$,

(ii) its next $\rho-1$ indices are taken from the subset $\{f_{j-1}+1,\,\dots,\,f_j\}$, and

(iii) its last index is taken from among $\{f_j+1,\,\dots,\,f_{j+1}\}$.

First, we notice that, just as in §5.1.2,

$$\sum_{\sigma} \det(\Phi_j(B)[\sigma]) = 0. \tag{5.13}$$

In order to prove it, recall that $\Phi_{j-1}(B)$ is nonsingular, so we may write

$$\sum_{\sigma} \det(\Phi_j(B)[\sigma]) = \det(\Phi_{j-1}(B)) \sum_{\sigma} \det\left(\left(\Phi_j[\sigma]/\Phi_{j-1}\right)(B)\right) \tag{5.14}$$

and the sum in the right hand side is just the sum of all possible products of ρ eigenvalues of the Schur complement $(\Phi_j/\Phi_{j-1})(B)$. But we know that the Schur complement has rank at most $\rho-1$, which leads to (5.13).

Furthermore, in order to rewrite $\widehat{\alpha}_{K(r)}$, we first observe that we may swap the two summations in (5.12): each of the first two sums in (5.12) can be written as

$$\sum_{k=1}^{r_j} \sum_{\widetilde{\sigma}} \det\left(\Phi_j^{(k,*)}(B)[\widetilde{\sigma}]\right), \tag{5.15}$$

where $*$ is either L or R, and $\widetilde{\sigma}$ runs over all ordered index subsets of $\{1,\,\dots,\,f_j\}$ with cardinality r and $f_{j-1}+1$ indices fixed as $\{1,\,\dots,\,f_{j-1}\} \cup \{f_{j-1}+k\}$. Now, for each index k, consider all subsets $\widehat{\sigma}$ of $\{1,\,\dots,\,f_j\}$ with cardinality r such that $f_{j-1}+k$ is not in $\widehat{\sigma}$. Then,

$$\sum_{k=1}^{r_j} \sum_{\sigma} \det\left(\Phi_j^{(k,*)}(B)[\sigma]\right) = \sum_{k=1}^{r_j} \sum_{\widetilde{\sigma}} \det\left(\Phi_j^{(k,*)}(B)[\widetilde{\sigma}]\right) + \sum_{k=1}^{r_j} \sum_{\widehat{\sigma}} \det\left(\Phi_j^{(k,*)}(B)[\widehat{\sigma}]\right), \tag{5.16}$$

where σ runs over all subsets of $\{1,\dots,f_j\}$ with cardinality r and first f_{j-1} indices $\{1,\dots,f_{j-1}\}$. We will show that the second sum on the right hand side is zero: first, since $k \notin \widehat{\sigma}$, we have $\Phi_j^{(k,*)}(B)[\widehat{\sigma}] = \Phi_j(B)[\widehat{\sigma}]$, so

$$\sum_{k=1}^{r_j} \sum_{\widehat{\sigma}} \det\left(\Phi_j^{(k,*)}(B)[\widehat{\sigma}]\right) = \sum_{k=1}^{r_j} \sum_{\widehat{\sigma}} \det\left(\Phi_j(B)[\widehat{\sigma}]\right).$$

Now, for each fixed subset in σ, there are $r_j - \rho$ indices it does not contain, so

$$\sum_{k=1}^{r_j} \sum_{\widehat{\sigma}} \det \left(\Phi_j^{(k,*)}(B)[\widehat{\sigma}] \right) = (r_j - \rho) \sum_{\sigma} \det \left(\Phi_j(B)[\sigma] \right),$$

where σ is as above. But the sum on the right is zero, as shown in (5.13). This proves that the sum (5.15) can be alternatively written as

$$\sum_{k=1}^{r_j} \sum_{\sigma} \det \left(\Phi_j^{(k,*)}(B)[\sigma] \right).$$

Each of the submatrices $\Phi_j^{(k,*)}(B)[\sigma]$ contains $\Phi_{j-1}(B)$ as its upper left block, so, using the same argument in (5.14) above, we obtain that

$$\sum_{\sigma} \det \left(\Phi_j^{(k,*)}(B)[\sigma] \right) = \det(\Phi_{j-1}(B)) \, \pi_\rho \left(\left(\Phi_j^{(k,*)} / \Phi_{j-1} \right) (B) \right).$$

Finally, in order to simplify the last sum in (5.12), we take Schur complements of $\Phi_{j-1}(B)$ in the submatrices $\Phi_{j+1}(B)[\omega]$, which leads to

$$\sum_{\omega} \det(\Phi_{j+1}(B)) = \det(\Phi_{j-1}(B)) \sum_{k=1}^{r_{j+1}} \sum_{\omega_k} \det \left(\left(\Phi_{j+1}[\omega_k] / \Phi_{j-1} \right) (B) \right) \tag{5.17}$$

where, for every fixed k, the subset ω_k contains the index $r_j + k$ plus $\rho - 1$ indices taken from $\{1, \ldots, r_j\}$. As before, if we denote by $\widehat{\omega}_k$ all index subsets of $\{1, \ldots, r_j + r_{j+1}\}$ with cardinality ρ which do not contain $r_j + k$, then one can prove that

$$\sum_{k=1}^{r_{j+1}} \sum_{\widehat{\omega}} \det \left(\left(\Phi_{j+1}[\omega_k] / \Phi_{j-1} \right) (B) \right) = 0,$$

just as we showed that the second sum in the right hand side of (5.16) vanishes. But notice that for each k the reunion of $\widehat{\omega}_k$ and ω_k is the total set of index subsets of $\{1, \ldots, r_{j-1} + r_j\}$ with cardinality ρ. Consequently, we may replace the sum over ω_k in (5.17) with the sum over ω, which allows us to write the sum of determinants over ω as a product π_ρ of the ρ nonzero eigenvalues of the submatrix corresponding to ω_k for every k. This leads finally to

$$\begin{aligned} \widehat{\alpha}_{K(r)} = (-1)^r \det(\Phi_{j-1}(B)) \Bigg[\sum_{k=1}^{r_j} \left(\pi_\rho \left(\left(\Phi_j^{(k,L)} / \Phi_{j-1} \right) (B) \right) + \pi_\rho \left(\left(\Phi_j^{(k,R)} / \Phi_{j-1} \right) (B) \right) \right) + \\ + \delta_{n_j - n_{j+1}, 1} \sum_{k=1}^{r_{j+1}} \pi_\rho \left(\left(\Phi_{j+1}[\omega_k] / \Phi_{j-1} \right) (B) \right) \Bigg]. \end{aligned}$$

(The sign $(-1)^r$ comes from keeping track of the successive sign changes due to expanding the principal minors along the positions of the '+1' entries). Hence, the Newton Diagram provides the leading coefficients in the asymptotic expansions of the $n_j - 1$ eigenvalues of $\widehat{S}(\varepsilon)$ as the roots of the polynomial equation

$$\widehat{\alpha}_{K(r)} + \widehat{\alpha}_{K(r)-2}\,\gamma^{n_j-1} = 0.$$

Using the expression above for $\widehat{\alpha}_{K(r)}$ and (5.11) for $\widehat{\alpha}_{K(r)-2}$ leads to the formula (5.3) for the leading coefficients. □

We conclude this section by noting that in the real case, even though the spectrum of $(\Phi_j/\Phi_{j-1})(B)$ may contain nonreal eigenvalues, it is still closed under complex conjugation; recall that in this case $(\Phi_j/\Phi_{j-1})(B)$ is real, so the product $\prod_{i=1}^{\rho-1}\xi_i$ is also real, and $\widehat{\alpha}_{K(r)}$ is also real since it is the sum of principal minors of real matrices. Hence, $\xi \in \mathbb{R}$.
Thus, if $\xi > 0$, two perturbed eigenvalues stay on the real line, and if $\xi < 0$, then two eigenvalues stay on the unit circle, and the remaining $n_j - 3$ will move away, half outwards and half inwards the unit circle. These two behaviors overlap whenever the remainder of n_j modulus four is 1.

5.2 Rank-one structure preserving perturbations: the case when $\mathbf{n_1}$ is odd

In this section we address the special case when the perturbation $\widehat{S}(\varepsilon) - S$ has rank one and the size n_1 of the largest Jordan block of S is odd. This corresponds to the last case analyzed in §4.4.
In this case it is known [38] that generically only $n_1 - 1$ new eigenvalues are created, since one copy of λ remains unchanged and, in fact, a Jordan block for λ of size $n_1{+}1$ is created out of that unchanged eigenvalue and the 'partner' Jordan block of size n_1, which must exist due to S being symplectic. This is consistent with the fact that all eigenvalues of the matrix $\Phi_1(B)$ are zero in this case (see below). Hence, this case is not covered by our results in § 4.3 above, which are valid under the assumption that the rank-one matrix $\Phi_1(B)$ has one nonzero eigenvalue ξ. This means that the horizontal projection of the Newton diagram has length $n_1 - 1$, instead of n_1. This is in fact a particular case of the one studied in § 5.1, but easier to solve. From there we have that the $n_1 - 1$ perturbed eigenvalues have asymptotic expansions

$$\widehat{\lambda}_k(\varepsilon) = \lambda\left(1 + \xi^{\frac{1}{n_1-1}}\varepsilon^{\frac{1}{n_1-1}}\right) + o\left(\varepsilon^{\frac{1}{n_1-1}}\right),$$

where the different values for k correspond to the distinct complex $(n_1 - 1)$-th roots of the number ξ in (5.3).
Now, what we need is to simplify the formula for ξ by applying the particularities of this

case. Then

$$\xi = -\widehat{\alpha}_{K(1)}\,,$$

where, $K(1) = n_1 - 1$, and

$$\widehat{\alpha}_{n_1-1} = -\left(\mathrm{trace}\left(\Phi_1^{(R)}(B)\right) + \mathrm{trace}\left(\Phi_1^{(L)}(B)\right) + \delta_{n_1-n_2,1}\mathrm{trace}\left(\Phi_2(B)\right)\right)\,.$$

It is important to remark that in this particular case $\Phi_1^{(R)}(B)$ and $\Phi_1^{(R)}(B)$ are square real matrices and recall that

$$\Phi_1^{(R)}(B) = Y_1\, B\, X_1^{(2)} \quad , \; \Phi_1^{(L)}(B) = Y_1^{(2)}\, B\, X_1\,. \tag{5.18}$$

Finally, we have that

$$\xi = \sum_{k=1,\,3,\,\cdots,\,\frac{r_1-2}{2}} \left(y_1^k\, B\, x_{(1)}^k + y_{(1)}^k\, B\, x_1^k + y_1^{k+1}\, B\, x_{(1)}^{k+1} + y_{(1)}^{k+1}\, B\, x_1^{k+1}\right) + \delta_{n_1-n_2,1} \sum_{k=1}^{r_2} y_2^k\, B\, x_2^k$$

where $\delta_{n_1-n_2,1}$ is a Kronecker delta (i.e., it is zero unless $n_2 - n_1 = 1$) and, for each $k = 1,\, 3,\, 5,\, \ldots,\, \frac{r_1-2}{2}$, we denote by y_1^k (resp., x_1^k) the left (resp., right) eigenvector, and by $y_{(1)}^k$(resp. $x_{(1)}^k$) the second left (resp., right) generalized eigenvector in the same Jordan chain as y_1^k (resp., as x_1^k). Finally, y_2^k (resp., x_2^k) are the left (resp., right) eigenvectors associated with the Jordan blocks of the second largest size n_2.

We now make use of the relations found in **Case 1.b** of § 4.2.2.1 between left and right generalized eigenvectors in order to write the coefficient ξ in a more compact form: a first observation is that first row of matrix D_{n_1} in formula (4.18) is $e_{n_1}^T$ and its second row is $he_{n_1} - e_{n_1-1}$, where h is a real number; recall that D_{n_1} is defined in (4.8). The relations (4.18) for $j = 1$,

$$(P_1^k)^T\, J = -D_{n_1}\, Q_1^{k+1}, \qquad \left(P_1^{k+1}\right)^T\, J = D_{n_1}\, Q_1^k \tag{5.19}$$

involve both the right vectors x_1^k and $x_{(1)}^k$, which are just the first and second column of P_1^k, as well as the left vectors $y_{(1)}^k$ and y_1^k who are just, in this order, the last two rows of Q_1^k. Taking this into account, the formulas above lead to

$$\begin{aligned}
J\, x_1^k &= \left(y_1^{k+1}\right)^T \\
J\, x_{(1)}^k &= -\left(y_{(1)}^{k+1}\right)^T + h\left(y_1^{k+1}\right)^T \\
J\, x_1^{k+1} &= -\left(y_1^k\right)^T \\
J\, x_{(1)}^{k+1} &= \left(y_{(1)}^k\right)^T - h\left(y_1^k\right)^T
\end{aligned}$$

If we introduce these relationships into the formula for ξ above, and we replace B by uu^TJ, we obtain

$$\xi = 2 \sum_{k=1,3,\cdots,\frac{r_1-2}{2}} \left(y_1^{k+1}\, u\, u^T \left(y_{(1)}^k\right)^T - y_1^k\, u\, u^T \left(y_{(1)}^{k+1}\right)^T \right) + \delta_{(n_1-n_2),1} \sum_{k=1}^{r_2} \varsigma_2^k\, y_2^k\, u\, u^T \left(y_2^k\right)^T$$

This formula can be simplified if we denote

$$a = Y_1 u \in \mathbb{C}^{r_1}, \quad a^{(1)} = Y_{(1)} u \in \mathbb{C}^{r_1}, \quad b = Y_2 u \in \mathbb{C}^{r_2},$$

where

$$Y_{(1)} = \begin{bmatrix} y_{(1)}^1 \\ \vdots \\ y_{(1)}^{r_1} \end{bmatrix}.$$

With this notation, the leading coefficient ξ can be written as

$$\xi = -\mathrm{trace}\left(a^{(1)} a^T D + \delta_{n_1-n_2,1} b\, b^T \Upsilon_2\right)$$

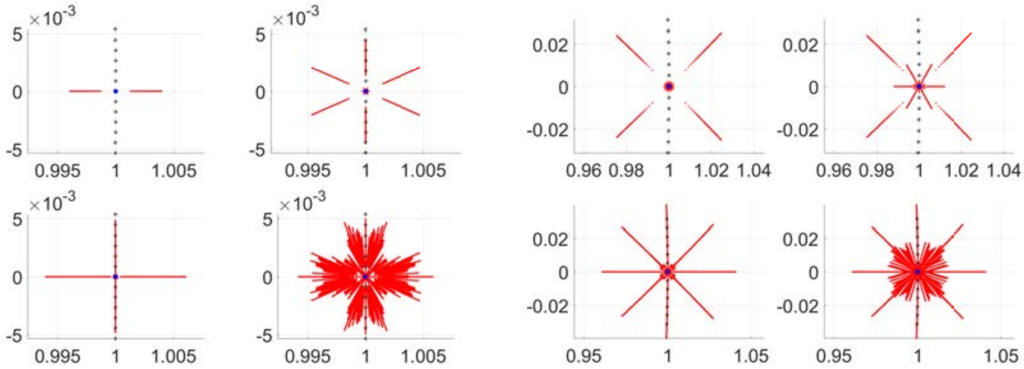
where $D = \Sigma_2 \oplus \cdots \oplus \Sigma_2$, the 2×2 matrix Σ_2 is as defined in (2.2) and $\Upsilon_2 = \mathrm{Diag}\,\{\varsigma_2^1, \cdots, \varsigma_2^{r_2}\}$, where $\varsigma_2^1, \ldots, \varsigma_2^{r_2}$ are the signs of the blocks of size $n_1 - 1$ (if there are any) associated with λ in the Jordan form of S.

Example 5.2. . *Let S be a $4n_1 \times 4n_1$ symplectic matrix with n_1 equal either to 3 or to 5, and with the same block structure as the one in Example 4.15, but now taking*

$$\Lambda = \mathcal{J}_{n_1}(1) \oplus \mathcal{J}_{n_1}(1)\,.$$

Then S has a single eigenvalue $\lambda = 1$ with four Jordan blocks of size n_1. Figure 5.2.1 displays the eigenvalues of 100 random symplectic perturbations of S with the same format as in Figure 4.5.4 (i.e., one perturbation singled out at the top, the remaining 99 at the bottom, the auxiliary perturbations $\widetilde{S}$ on the left, and $\widehat{S}$ on the right). Figure (a) corresponds to choosing $n_1 = 3$ and Figure (b) to choosing $n_1 = 5$. Now the values for ε are taken from partitioning the interval $[\,0\,,\,10^{-5}\,]$ with step 10^{-7} in the case of $n_1 = 3$, and the interval $[\,0\,,\,10^{-6}\,]$ with step 10^{-8} in the case of $n_1 = 5$.

This example clearly shows a big difference between the nongeneric case analyzed in this chapter and the ones addressed in Chapter 4: in the latter, the asymptotic expansions given in §4.3 have non-zero leading coefficients for as many of the largest Jordan blocks as the rank of B, so at least those perturbed eigenvalues of the matrix $\widehat{S}(\varepsilon)$ must behave locally in much the same way as the corresponding eigenvalues of $\widetilde{S}(\varepsilon)$ (in other words, as described in the appropriate Theorem). In the nongeneric case, however, the statements in Theorem 5.1 are valid only for perturbations of the form $\widetilde{S}$, and we do not know, in principle, if they will also hold for some eigenvalues of the more general perturbations



(a) $n_1 = 3$ (b) $n_1 = 5$

Figure 5.2.1:

$\widehat{S}$. It looks like it in Figure 5.2.1 b), where the four perturbed eigenvalues of $\widehat{S}$ split in the same directions as those for $\widetilde{S}$. In Figure 5.2.1 a), however, it becomes clear that completely different behaviors for all eigenvalues of $\widehat{S}$ and $\widetilde{S}$ are possible: the figures on the left correspond to the behavior predicted by Theorem 5.1 for $\widetilde{S}$: the rank one perturbation breaks just one 3×3 Jordan block, but one of those three eigenvalues is used to create a new 4×4 Jordan block. Hence, only two $O(\varepsilon^{1/2})$ are created, which split either horizontally or vertically. The figures on the right, however, show that, since $\widehat{S} - S$ has rank typically larger than one, more than one Jordan block is broken. The top figure on the right, for instance, suggests two triplets of $O(\varepsilon^{1/3})$ perturbed eigenvalues, anchored to both vertical semiaxes. In other words, in this case not only the first order perturbation coefficient B is involved in the leading term, but also the coefficients with higher order come into play.

We end with a final comment: preliminary analyses using Newton diagram techniques strongly support the idea that $n_1 = 3$ is an anomaly among odd values of n_1, i.e., that it is likely that one can prove for n_1 odd, $n_1 \geq 5$, that the leading terms for the eigenvalues of $\widehat{S}(\varepsilon)$ and $\widetilde{S}(\varepsilon)$ coincide, just as in the generic cases. However, no formal proof is available at the moment.

Chapter 6

Conclusions, publications and open problems

6.1 Conclusions and original contributions

Below, we provide a summary of the most relevant original results included in this work

Chapter 3: Asymptotic expansions in the parameter ε for the eigenvalues of a multiplicatively perturbed matrix $\widehat{A}(\varepsilon) = (I + \varepsilon B)A(I + \varepsilon C)$ have been obtained for arbitrary matrices B and C, and for any Jordan form of A. The case in which the unperturbed eigenvalue is not zero is easiest to analyze (see Theorem 3.2), because Lidskii's classical formulas for additive perturbations are directly applicable, and readily provide the appropriate asymptotic expansions. Analyzing the other case, the one associated with singular Jordan blocks, was much more complicated, and required completely different techniques, since the leading coefficients that classical additive theory would provide are all zero, so the corresponding asymptotic expansions are useless. However, making use of the Newton diagram technique, first order asymptotic expansions for the perturbed eigenvalues have also been obtained for zero unperturbed eigenvalues (see Theorem 3.3). Additionally, first order asymptotic expansions have also been obtained for singular values by using the Jordan-Wielandt form of A (see Theorem 3.10).

Chapter 4: The results in this chapter are based on three main ingredients:

(i) first, we show that any small J-symplectic perturbation of a J-symplectic matrix can be written multiplicatively, with a J-Hamiltonian first-order coefficient matrix. Thus, the multiplicative theory in Chapter 3 can be applied in our structured context;

(ii) next, we reveal the special connection induced by symplectic structure between left and right eigenvector in the Jordan canonical form of symplectic matrices;

(iii) finally, we incorporate this special connection between left and right eigenvectors into the asymptotic multiplicative expansions, specializing them to the case of symplectic structure preservation. This provides us with explicit first order expansions for eigenvalues of structure-preserving perturbations for all kinds of unperturbed eigenvalues (unimodular or not), revealing for each situation the special perturbation behavior induced by symplectic structure.

Unsurprisingly, the most interesting results turn out to be either the exceptional eigenvalues ± 1 (whose reciprocal coincides with itself – see Theorem 4.8), or eigenvalues on the unit circle of *real* symplectic matrices (see Theorem 4.6). In both cases, the behavior of perturbed eigenvalues heavily depends on their so-called sign characteristics. This is made especially clear in §4.4, where perturbations of rank 1 are considered. For any other eigenvalues of symplectic matrices, the behavior under structure-preserving perturbation is not much different from the behavior under arbitrary ones, in the sense that perturbed eigenvalues can move in any direction. There is one special case, however, already identified as atypical in [38] when analyzing the possible outcoming Jordan forms of structure-preserving symplectic perturbations, for which some of the asymptotic expansions cannot be obtained from the approach followed in this chapter.

Chapter 5: This special kind of perturbations, which correspond to a situation in which the spectral structural constraints induced by symplecticity go against the 'natural' behavior of eigenvalues under non-structured perturbations, has been analyzed by reconstructing the associated Newton diagram from scratch. A new analysis has been performed from which both leading exponents and explicit formulas for the leading coefficients have been derived through the Newton diagram technique. These formulas not only identify the leading exponents, but also show that, in this nongeneric situation, the directions along which perturbed eigenvalues move away depend not only on normalized left and right eigenvectors associated with the Jordan blocks of size n_k, but also on first generalized Jordan vectors for these blocks, as well as on left and right eigenvectors associated with Jordan blocks of size n_{k-1} (see Theorem 5.1).

6.2 Future Work

We describe here some possible topics for future research which might improve, extend or clarify some of the results presented in this dissertation:

- The numerical examples in §4.5 show that our theory (as any other first-order one) only covers those perturbed eigenvalues with asymptotic expansions whose leading terms depend only on the first order coefficient B of the perturbation (in other words, it only applies to perturbations of the form (4.29), see the discussion in

§4.5.1). It is a natural question to ask how representative this kind of perturbations is among all small symplectic perturbations. To be more specific, given a J-symplectic matrix S and a small J-symplectic perturbation $\widehat{S}$ of S such that $\operatorname{rank}(\widehat{S} - S) = r$, is it possible to find vectors $u_1, \ldots, u_r$ such that

$$\widehat{S} = \prod_{k=1}^{r} \left(I + u_k u_k^T J\right) \cdot S\,?$$

Answering this question would help us assess the actual scope of the first order perturbation theory developed here among all structure-preserving J-symplectic perturbations of given rank.

- The discussion of Example 5.2 makes it clear that in the nongeneric case the behavior described by Theorem 5.1 may not apply to *any* of the eigenvalues of an arbitrary J-symplectic perturbation $\widehat{S}$, at least when the size of the largest Jordan block is $n_1 = 3$. As already mentioned, there are reasons to believe that this is an anomaly exclusive to that particular value of n_1. Hence, we consider the possibility of showing that for n_1 odd, $n_1 \geq 5$, the leading terms for the eigenvalues of perturbation of type $\widehat{S}(\varepsilon)$ and of type $\widetilde{S}(\varepsilon)$ coincide, just as in the generic cases analyzed in Chapter 4.

- The theory we have developed for J-symplectic structure-preserving perturbations can be translated, on a case-by-case basis, to the closely related context of J-*Hamiltonian* structure-preserving perturbations. The main tool to perform this translation is the so called *Cayley transform*, a transformation

$$\mathcal{C} : A \mapsto (A + I_n)(A - I_n)^{-1}.$$

on sets of real or complex $n \times n$ matrices A, which is known to map the set of symplectic matrices not having 1 in their spectrum to the set of Hamiltonian matrices; conversely, the set of Hamiltonian matrices not having 1 in their spectrum is mapped into the set of symplectic matrices not having 1 in their spectrum (in the latter case also -1 will be excluded from the spectrum of A because of the Hamiltonian spectral symmetry; this will have the effect that zero is not in the spectrum of $\mathcal{C}(A)$). Using the Cayley transform to go back and forth between symplectic and Hamiltonian matrices, and vice versa, we expect to get a mirror image for Hamiltonian matrices of the symplectic theory we have derived here, with the unit circle replaced by the maginary axis, and the exceptional points ± 1 replaced by the origin. This is work in progress, and a corresponding publication should be submitted soon.

Appendices

Appendix A

Proof of (3.21)

Given $l = f_{j-1} + \rho$ for some index $j \in \{1, \ldots, q\}$ with $n_j \geq 2$ and $1 \leq \rho \leq r_j$, our goal is to find the leading coefficient $\det(\widetilde{M_1}[\beta \,|\, \gamma]) \det(\widetilde{M_2}[\theta \,|\, \beta])$ of the product $\det(M_1[\beta \,|\, \gamma]) \det(M_2[\theta \,|\, \beta])$, where $\gamma = \{1, \ldots, k(l) + l\} \setminus \Omega$ with $k(l)$ given by (3.8), and β and θ are index sets satisfying conditions 2 – 4 in (3.19).

The choice of β gives rise to the lists $\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4$ of indices labelling which diagonal blocks of $\widetilde{M_1}[\beta \,|\, \gamma]$, $\widetilde{M_2}[\theta \,|\, \beta]$ are in each of the four cases described in Table 3.1.6. Also, recall that the only entries in these matrices which are not 0 or 1 are those placed on the rows and columns not containing a $O(1)$ entry, and they are just entries of $\Phi_j(B)$ and $\Phi_j(C)$, respectively.

Now, we need to expand both determinants along the rows and columns where the 1 entries lie in order to simplify the formula. This will lead to a product of a sign and two minors of matrices $\Phi_j(B)$ and $\Phi_j(C)$, and the sign will be given by the position of those 1s in the matrices $\widetilde{M_1}[\beta \,|\, \gamma]$ and $\widetilde{M_2}[\theta \,|\, \beta]$ (see Table 3.1.6). Such positions, however, can widely vary, so a direct analysis becomes impractical. To avoid this, we rearrange both matrices in such a way that all the 1 entries lie either on the main diagonal or on the first superdiagonal. This will largely simplify the analysis of the sign: first consider $\widetilde{M_1}[\beta \,|\, \gamma]$ with the block partition induced by (2.26), and permute its rows and columns to obtain a new matrix

$$\widehat{M_1} = P_1 \widetilde{M_1}[\beta \,|\, \gamma] P_2$$

for appropriate permutation matrices P_1, P_2 in such a way that the diagonal blocks of $\widehat{M_1}$ are those of $\widetilde{M_1}[\beta \,|\, \gamma]$ in the following order

- first, all the diagonal blocks in Case 2 (see Table 3.1.6),
- next, all the diagonal blocks in Case 3,
- then, alternate pairs of one block in Case 1 and one block in Case 4,

all this without changing the relative order among the original blocks in the same Case.

In Example 3.6 above, for instance, where $4, 7, 8 \in \beta$, we would have

$$\widetilde{M}_1[\beta \,|\, \gamma] = \left[\begin{array}{cc|c|c|c} 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \clubsuit_{31} & 0 & 0 & 0 & 0 \end{array}\right] , \qquad \widehat{M}_1 = \left[\begin{array}{cc|c|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \clubsuit_{31} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

for permutation matrices

$$P_1 = \left[\begin{array}{c|c|c} 0 & I_2 & 0 \\ \hline 0 & 0 & I_2 \\ \hline I_1 & 0 & 0 \end{array}\right] , \qquad P_2 = \left[\begin{array}{c|c|c} 0 & 0 & I_2 \\ \hline I_2 & 0 & 0 \\ \hline 0 & I_1 & 0 \end{array}\right] . \tag{A.1}$$

In general, according to Table 3.1.6, the diagonal blocks in Case 2 are all identity matrices, and those in Case 3 are also square with the 1s placed on the superdiagonal. Although the blocks in Cases 1 and 4 are not square, one can easily check that by pairing them we obtain larger square matrices, of dimension $(n_i + n_j - 2) \times (n_i + n_j - 2)$, with all its 1 entries on the main diagonal. Hence, the total amount of 1s on the superdiagonal is $\widehat{M}_1$ is $\sum_{i \in v_3}(n_i - 2)$, and all the remaining 1 entries of $\widehat{M}_1$ lie on the main diagonal

To rearrange $\widetilde{M}_2[\theta \,|\, \beta]$ we use the same permutations, but transposed, i.e.,

$$\widehat{M}_2 = P_2^T \, \widetilde{M}_2[\theta \,|\, \beta] \, P_1^T.$$

This produces the exact same order of Cases in the diagonal blocks as above, since the Case for each diagonal block is fixed by the choice of β, and while β selects rows in $\widetilde{M}_1[\beta \,|\, \gamma]$, it selects columns in $\widetilde{M}_2[\theta \,|\, \beta]$, thus the transposes. In Example 5, we get

$$\widetilde{M}_2[\theta \,|\, \beta] = \left[\begin{array}{c|cc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & \spadesuit_{12} & 0 & \spadesuit_{13} & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & \spadesuit_{22} & 0 & \spadesuit_{23} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] , \qquad \widehat{M}_2 = \left[\begin{array}{cc|cc|c} 0 & 1 & 0 & 0 & 0 \\ \spadesuit_{22} & 0 & \spadesuit_{23} & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \spadesuit_{12} & 0 & \spadesuit_{13} & 0 & 0 \end{array}\right]$$

for the same permutations P_1, P_2 in (A.1). Now, according to the last column in Table 3.1.6, the upper left blocks in Case 2 are square with the 1s on the superdiagonal, the next diagonal blocks in Case 3 are identity matrices, and each pair of diagonal blocks in Cases 1 and 4 forms a square $(n_i+n_j-2)\times(n_i+n_j-2)$ block with $\sum_{i\in v_1}(n_i-1)+\sum_{j\in v_4}(n_j-2)$ entries equal to one on the superdiagonal of $\widehat{M}_2$.

Hence, after reordering, both $\widehat{M}_1$ and $\widehat{M}_2$ have all their 1 entries either on the main diagonal or on the superdiagonal, and the number of superdigonal 1 entries in both matrices together is

$$\sum_{i \in v_2}(n_i - 2) + \sum_{i \in v_3}(n_i - 2) + \sum_{i \in v_1}(n_i - 1) + \sum_{i \in v_4}(n_i - 2) = k(l) - l + \operatorname{card}(v_1).$$

As a consequence of this,

$$\begin{aligned} & \det(\widetilde{M}_1[\beta\,|\,\gamma])\det(\widetilde{M}_2[\theta\,|\,\beta]) = \det(\widehat{M}_1)\det(\widehat{M}_2)\det(P_1)^2\det(P_2)^2 = \\ = \ & (-1)^{k(l)-l+\mathrm{card}(\upsilon_1)}\det\left(\Phi_j(C)[(\upsilon_3,\upsilon_1)\,|\,(\upsilon_3,\upsilon_4)]\right)\det\left(\Phi_j(B)[(\upsilon_2,\upsilon_4)\,|\,(\upsilon_2,\upsilon_1)]\right) \end{aligned} \tag{A.2}$$

Now, since the entries of $\Phi_j(C)[(\upsilon_3,\upsilon_1)\,|\,(\upsilon_3,\upsilon_4)],\ \Phi_j(B)[(\upsilon_2,\upsilon_4)\,|\,(\upsilon_2,\upsilon_1)]$ are not in their natural order on $\Phi_j(B)$ and $\Phi_j(C)$, we need to restore the order changed by reordering the diagonal blocks in $\widetilde{M}_1[\beta\,|\,\gamma]$ and $\widetilde{M}_2[\theta\,|\,\beta]$: let $\vartheta = \upsilon_3+\upsilon_1$ be the increasingly ordered tuple containing the indices in both υ_3 and υ_1, and, similarly, $\zeta = \upsilon_3+\upsilon_4$. Then, using the notation introduced in §3.1.6 for signs of permutations,

$$\begin{aligned} \det\left(\Phi_j(C)[(\upsilon_3,\upsilon_1)\,|\,(\upsilon_3,\upsilon_4)]\right) &= \mathrm{sgn}(\upsilon_3,\upsilon_1)\,\mathrm{sgn}(\upsilon_3,\upsilon_4)\det\left(\Phi_j(C)[\vartheta\,|\,\zeta]\right) \\ \det\left(\Phi_j(B)[(\upsilon_2,\upsilon_4)\,|\,(\upsilon_2,\upsilon_1)]\right) &= \mathrm{sgn}(\upsilon_2,\upsilon_4)\,\mathrm{sgn}(\upsilon_2,\upsilon_1)\det\left(\Phi_j(B)[\vartheta^c\,|\,\zeta^c]\right) \end{aligned} \tag{A.3}$$

where, as before, υ^c and ζ^c denote the complementary in $\{1,\dots,l\}$.

Let us now prove that

$$\mathrm{sgn}\,(\upsilon_3,\upsilon_1)\,\mathrm{sgn}\,(\upsilon_2,\upsilon_4)\,\mathrm{sgn}\,(\upsilon_3,\upsilon_4)\,\mathrm{sgn}\,(\upsilon_2,\upsilon_1) = (-1)^{\mathrm{card}(\upsilon_1)}\mathrm{sgn}\,(\vartheta\,,\ \vartheta^c)\,\mathrm{sgn}\,(\zeta\,,\ \zeta^c) \tag{A.4}$$

In order to do that, consider the auxiliary l-tuples

$$\begin{aligned} \upsilon_R &= (\upsilon_3,\upsilon_1,\upsilon_2,\upsilon_4)\,, \\ \upsilon_C &= (\upsilon_3,\upsilon_4,\upsilon_2,\upsilon_1)\,. \end{aligned}$$

First we check how many transpositions are needed to transform υ_R into υ_C. Below we detail, step by step, the required transformations:

$$\begin{aligned} (\upsilon_3,\upsilon_1,\upsilon_2,\upsilon_4) &\longmapsto (\upsilon_3,\upsilon_2,\upsilon_1,\upsilon_4) &&: \ \mathrm{card}(\upsilon_1)\,\mathrm{card}(\upsilon_2)\text{ transpositions} \\ (\upsilon_3,\upsilon_2,\upsilon_1,\upsilon_4) &\longmapsto (\upsilon_3,\upsilon_2,\upsilon_4,\upsilon_1) &&: \ \mathrm{card}(\upsilon_1)\,\mathrm{card}(\upsilon_4)\text{ transpositions} \\ (\upsilon_3,\upsilon_2,\upsilon_4,\upsilon_1) &\longmapsto (\upsilon_3,\upsilon_4,\upsilon_2,\upsilon_1) &&: \ \mathrm{card}(\upsilon_2)\,\mathrm{card}(\upsilon_4)\text{ transpositions} \end{aligned}$$

Hence, the total number of transpositions needed to transform υ_R into υ_C is

$$\mathrm{card}(\upsilon_1)\,\mathrm{card}(\upsilon_2) + \mathrm{card}(\upsilon_1)\,\mathrm{card}(\upsilon_4) + \mathrm{card}(\upsilon_2)\,\mathrm{card}(\upsilon_4)$$

But we know that $\mathrm{card}(\upsilon_4) = \mathrm{card}(\upsilon_1)$, so

$$\begin{aligned} \mathrm{sgn}\,(\upsilon_R) &= (-1)^{\mathrm{card}(\upsilon_1)^2+2\mathrm{card}(\upsilon_1)\mathrm{card}(\upsilon_2)}\mathrm{sgn}\,(\upsilon_C) \\ &= (-1)^{\mathrm{card}(\upsilon_1)}\mathrm{sgn}\,(\upsilon_C) \end{aligned} \tag{A.5}$$

Finally, take for instance υ_R: if we rearrange increasingly both its first half (υ_3,υ_1) and its second half (υ_2,υ_4), we obtain (ϑ,ϑ^c). The same goes for υ_C and (ζ,ζ^c). Therefore,

$$\begin{aligned} \mathrm{sgn}(\upsilon_R) = \mathrm{sgn}\,(\upsilon_3,\upsilon_1,\upsilon_2,\upsilon_4) &= \mathrm{sgn}\,(\upsilon_3,\upsilon_1)\,\mathrm{sgn}\,(\upsilon_2,\upsilon_4)\,\mathrm{sgn}\,(\vartheta\,,\ \vartheta^c)\,, \\ \mathrm{sgn}(\upsilon_C) = \mathrm{sgn}\,(\upsilon_3,\upsilon_4,\upsilon_2,\upsilon_1) &= \mathrm{sgn}\,(\upsilon_3,\upsilon_4)\,\mathrm{sgn}\,(\upsilon_2,\upsilon_1)\,\mathrm{sgn}\,(\zeta\,,\ \zeta^c)\,. \end{aligned}$$

Combining this with (A.5) proves (A.4). Finally, if we substitute (A.3) into (A.2) and make use of (A.4), we obtain (3.21), as claimed.

Appendix B

Proof of Lemma 3.9

Let $D = \det(M+N)$. Then

$$D = \sum_{\sigma \in S_l} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{l} \left(m_{i,\sigma(i)} + n_{i,\sigma(i)} \right) \right),$$

where S_l denotes the group of permutations of l elements. The product inside can be expanded as

$$\prod_{i=1}^{l} \left(m_{i,\sigma(i)} + n_{i,\sigma(i)} \right) = \sum_{i=0}^{l} \left(\sum_{\vartheta \in \Lambda_i} \left(\prod_{j=1}^{i} m_{\vartheta(j),\sigma(\vartheta(j))} \prod_{j=1}^{l-i} n_{\vartheta^c(j),\sigma(\vartheta^c(j))} \right) \right),$$

where Λ_i is the family of all increasingly ordered lists of indices, taken from $\{1, \ldots, l\}$ with length i, and $\vartheta^c \in \Lambda_{l-i}$ denotes, as before, the complement of ϑ in $\{1, \ldots, l\}$. Substituting this expression in the determinant formula above we obtain

$$D = \sum_{\sigma \in S_l} \operatorname{sgn}(\sigma) \left(\sum_{i=0}^{l} \left(\sum_{\vartheta \in \Lambda_i} \left(\prod_{j=1}^{i} m_{\vartheta(j),\sigma(\vartheta(j))} \prod_{j=1}^{l-i} n_{\vartheta^c(j),\sigma(\vartheta^c(j))} \right) \right) \right).$$

Now, since the sum over i and the sum over ϑ are finite and independent of σ, we may swap the three of them,

$$D = \sum_{i=0}^{l} \left(\sum_{\vartheta \in \Lambda_i} \left(\sum_{\sigma \in S_l} \left(\operatorname{sgn}(\sigma) \prod_{j=1}^{i} m_{\vartheta(j),\sigma(\vartheta(j))} \prod_{j=1}^{l-i} n_{\vartheta^c(j),\sigma(\vartheta^c(j))} \right) \right) \right).$$

Our next step is to partition S_l in order to split the sum over σ. Given $\vartheta \in S_l$, let ζ be any set in Λ_i and let $S_{\vartheta,\zeta} \subset S_l$ be the family of all permutations transforming ϑ into ζ, i.e.,

$$S_{\vartheta,\zeta} = \{ \sigma \in S_l \mid \ \sigma(a) \in \{ \zeta(j) \mid 1 \leq j \leq i \}, \forall a \in \{ \vartheta(j) \mid 1 \leq j \leq i \} \}.$$

Clearly, the symmetric group S_l is a disjoint union of all $S_{\vartheta,\zeta}$,

$$S_l = \bigcup_{i=0}^{l} \bigcup_{\vartheta,\zeta\in\Lambda_i} S_{\vartheta,\zeta},$$

so we can split the sum over S_l as

$$\sum_{\sigma\in S_l} (\qquad) = \sum_{i=0}^{l} \sum_{\vartheta,\zeta\in\Lambda_i} \sum_{\sigma\in S_{\vartheta,\zeta}} (\qquad)$$

We also split the sign of each $\sigma \in S_l$ into the product of signs of other permutations in order to further simplify the formula: let $\vartheta, \zeta \in \Lambda_i$, and suppose $\sigma \in S_{\vartheta,\zeta}$. We consider the following auxiliary permutations:

$$\begin{aligned}
\widehat{\sigma} &= \begin{pmatrix} 1 & \dots & i & i+1 & \dots & l \\ \sigma(\vartheta(1)) & \dots & \sigma(\vartheta(i)) & \sigma(\vartheta^c(1)) & \dots & \sigma(\vartheta^c(l-i)) \end{pmatrix} \in S_l \\
\sigma_\vartheta &= \begin{pmatrix} 1 & \dots & i & i+1 & \dots & l \\ \vartheta(1) & \dots & \vartheta(i) & \vartheta^c(1) & \dots & \vartheta^c(l-i) \end{pmatrix} \in S_l \\
\sigma_\zeta &= \begin{pmatrix} 1 & \dots & i & i+1 & \dots & l \\ \zeta(1) & \dots & \zeta(i) & \zeta^c(1) & \dots & \zeta^c(l-i) \end{pmatrix} \in S_l \\
\sigma_1 &= \begin{pmatrix} \zeta(1) & \zeta(2) & \dots & \zeta(i) \\ \sigma(\vartheta(1)) & \sigma(\vartheta(2)) & \dots & \sigma(\vartheta(i)) \end{pmatrix} \in S_i \\
\sigma_2 &= \begin{pmatrix} \widetilde{\zeta}(1) & \widetilde{\zeta}(2) & \dots & \widetilde{\zeta}(l-i) \\ \sigma\left(\widetilde{\vartheta}(1)\right) & \sigma(\vartheta^c(2)) & \dots & \sigma(\vartheta^c(l-i)) \end{pmatrix} \in S_{l-i}
\end{aligned}$$

Notice, on one hand, that $\widehat{\sigma}(i) = \sigma(\sigma_\vartheta(i))$, so $\widehat{\sigma}$ is the composition of σ and σ_ϑ, and, consequently,

$$\operatorname{sgn}(\widehat{\sigma}) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\sigma_\vartheta) \tag{B.1}$$

On the other hand, the concatenation $(\sigma_1, \sigma_2) \in S_l$ satisfies $\widehat{\sigma}(i) = (\sigma_1, \sigma_2)(\sigma_\zeta(i))$, so $\widehat{\sigma}$ is the composition of the concatenation (σ_1, σ_2) with σ_ζ. Hence,

$$\operatorname{sgn}(\widehat{\sigma}) = \operatorname{sgn}(\sigma_\zeta) \operatorname{sgn}(\sigma_1) \operatorname{sgn}(\sigma_2) \tag{B.2}$$

Combining equations (B.1) and (B.2) we get

$$\begin{aligned}
\operatorname{sgn}(\sigma) &= \operatorname{sgn}(\sigma_\zeta) \operatorname{sgn}(\sigma_\vartheta) \operatorname{sgn}(\sigma_1) \operatorname{sgn}(\sigma_2) \\
&= \operatorname{sgn}(\zeta, \zeta^c) \operatorname{sgn}(\vartheta, \vartheta^c) \operatorname{sgn}(\sigma_1) \operatorname{sgn}(\sigma_2)
\end{aligned} \tag{B.3}$$

Finally, the sum over $\sigma \in S_{\vartheta,\zeta}$, can be rewritten as a double sum as

$$\sum_{\sigma\in S_l}(\,\ldots\,)=\sum_{i=0}^{l}\sum_{\vartheta,\zeta\in\Lambda_i}\sum_{\sigma_1\in S_i}\sum_{\sigma_2\in S_{l-i}}(\,\ldots\,)$$

Taking into account all of the above, we rewrite the formula for D as

$$D=\sum_{i=0}^{l}\sum_{\vartheta,\,\zeta\in\Lambda_i}\sum_{\sigma_1\in S_i}\sum_{\sigma_2\in S_{l-i}}\left(\operatorname{sgn}(\zeta,\zeta^c)\operatorname{sgn}(\vartheta,\vartheta^c)\operatorname{sgn}(\sigma_1)\ \operatorname{sgn}(\sigma_2)\,P(\vartheta\,,\,\zeta)\right)$$

where

$$P(\vartheta\,,\,\zeta)=\prod_{j=1}^{i}m_{\vartheta(j),\sigma_1(\zeta(i))}\prod_{j=1}^{l-i}n_{\vartheta^c(j),\sigma_2(\zeta^c(j))}.$$

But

$$\begin{aligned}\det\left(M\,[\vartheta\,|\,\zeta]\right) &= \sum_{\sigma_1}\operatorname{sgn}(\sigma_1)\prod_{j=1}^{i}m_{\vartheta(j),\sigma_1(\zeta(i))} \qquad \text{and}\\ \det\left(N\,[\vartheta^c\,|\,\zeta^c]\right) &= \sum_{\sigma_2}\operatorname{sgn}(\sigma_2)\prod_{j=1}^{l-i}n_{\vartheta^c(j),\sigma_2(\zeta^c(j))},\end{aligned}$$

so finally

$$D \;=\; \sum_{i=0}^{l}\left(\sum_{\vartheta,\,\zeta\in\Lambda_i}\operatorname{sgn}(\zeta,\zeta^c)\operatorname{sgn}(\vartheta,\vartheta^c)\det\left(M\,[\vartheta\,|\,\zeta]\right)\det\left(N\,[\vartheta^c\,|\,\zeta^c]\right)\right),$$

which completes the proof .

□

Bibliography

[1] Arvind, B. Dutta, N. Mukunda and R. Simon. The Real Symplectic Groups in Quantum Mechanics and Optics. *Pramana* , 45 (6): 471 – 497, 1995.

[2] J. Barlow and J. Demmel. Computing accurate eigensystems of scaled diagonally dominant matrices. *SIAM Journal on Numerical Analysis*, 27 (3): 762 – 791, 1990.

[3] L. Batzke, C. Mehl, A. C. Ran, and L. Rodman. Generic rank-k perturbations of structured matrices. In *Operator Theory, Function Spaces, and Applications*, 255: 27 – 48. Birkhäuser, 2016.

[4] H. Baumgärtel. *Analytic Perturbation Theory for Matrices and Operators*. Birkhäuser Verlag, 1985.

[5] P. Benner and H. Faßbender. An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra and its Applications*, 263: 75 – 111, 1997.

[6] P. Benner and H. Faßbender. The symplectic eigenvalue problem, the butterfly form, the SR algorithm, and the Lanczos method. *Linear Algebra and its Applications*, 275 - 276: 19 – 47, 1998.

[7] E. Brieskorn and H. Knörrer. *Plane Algebraic Curves*. Springer, 2012.

[8] A. Bunse-Gerstner and V. Mehrmann. A symplectic QR like algorithm for the solution of the real algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 31 (12): 1104 – 1113, 1986.

[9] J. V. Burke and M. L. Overton. Stable perturbations of nonsymmetric matrices. *Linear Algebra and its Applications*, 171: 249 – 273, 1992.

[10] J. V. Burke and M. L. Overton. Differential properties of the spectral abscissa and the spectral radius for analytic matrix-valued mappings. *Nonlinear Analysis*, 23 (4): 467 – 488, 1994.

[11] F. Chatelin. *Eigenvalues of Matrices: Revised Edition*. SIAM, 2012.

[12] F. M. Dopico and C. R. Johnson. Parametrization of the matrix symplectic group and applications. *SIAM Journal on Matrix Analysis and Applications*, 31 (2): 650 – 673, 2009.

[13] S. C. Eisenstat and I. C. Ipsen. Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices. *BIT Numerical Mathematics*, 38 (3): 502 – 509, 1998.

[14] H. Faßbender. *Symplectic Methods for the Symplectic Eigenproblem*. Kluwer Academic/Plenum Publishers, 2000.

[15] K. Friedrichs. On the perturbation of continuous spectra. *Communications on Pure and Applied Mathematics*, 1 (4): 361 – 406, 1948.

[16] S. Godunov and M. Sadkane. Spectral analysis of symplectic matrices with application to the theory of parametric resonance. *SIAM Journal on Matrix Analysis and Applications*, 28 (4): 1083 – 1096, 2006.

[17] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Springer, 2005.

[18] G. H. Golub and C. F. Van Loan. *Matrix Computations*, Volume 3. The Johns Hopkins University Press, 2013.

[19] V. Guillemin and S. Sternberg. *Symplectic Techniques in Physics*. Cambridge University Press, 1984.

[20] E. Hille and R. S. Phillips. *Functional Analysis and Semi-Groups*. American Mathematical Soc, 1957.

[21] M. A. Hörmander, L. A remark on perturbations of compact operators. *Mathematica Scandinavica*, 75 (2): 255 – 262, 1994.

[22] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[23] T. Kato. *Perturbation Theory for Linear Operators*. Springer Verlag, 1980.

[24] D. Kressner. *Numerical Methods for General and Structured Eigenvalue Problems*. Springer, 2005.

[25] H. Langer and B. Najman. Remarks on the perturbation of analytic matrix functions II. *Integral Equations and Operator Theory*, 12 (3): 392 – 407, 1989.

[26] H. Langer and B. Najman. Remarks on the perturbation of analytic matrix functions III. *Integral Equations and Operator Theory*, 15 (5): 796 – 806, 1992.

[27] H. Langer and B. Najman. Leading coefficients of the eigenvalues of perturbed analytic matrix functions. *Integral Equations and Operator Theory*, 16 (4): 600 – 604, 1993.

[28] C.-K. Li and R. Mathias. The Lidskii-Mirsky-Wielandt theorem–additive and multiplicative versions. *Numerische Mathematik*, 81 (3): 377 – 413, 1999.

[29] R.-C. Li. Relative perturbation theory. III. more bounds on eigenvalue variation. *Linear Algebra and its Applications*, 266: 337 – 345, 1997.

[30] R.-C. Li. Relative perturbation theory: I. eigenvalue and singular value variations. *SIAM Journal on Matrix Analysis and Applications*, 19 (4): 956 – 982, 1998.

[31] V. B. Lidskii. Perturbation theory of non-conjugate operators. *USSR Computational Mathematics and Mathematical Physics*, 6 (1): 73 – 85, 1966.

[32] W.-W. Lin. A new method for computing the closed-loop eigenvalues of a discrete-time algebraic Riccati equation. *Linear Algebra and its Applications*, 96: 157 – 180, 1987.

[33] J. H. Maddocks and M. L. Overton. Stability theory for dissipatively perturbed Hamiltonian systems. *Communications on Pure and Applied Mathematics*, 48 (6): 583 – 610, 1995.

[34] A. Malyshev. *Stability radii of symplectic and Hamiltonian matrices*. Rep ort No 160, Department of Informatics, University of Bergen, 5020 Bergen, Norway, November 1998.

[35] C. Mehl. Essential decomposition of polynomially normal matrices in real indefinite inner product spaces. *Electronic Journal of Linear Algebra*, 15: 84 – 106, 2006.

[36] C. Mehl. On classification of normal matrices in indefinite inner product spaces. *Electronic Journal of Linear Algebra*, 15: 50 – 83, 2006.

[37] C. Mehl, V. Mehrmann, A. C. M. Ran, and L. Rodman. Eigenvalue perturbation theory of classes of structured matrices under generic structured rank one perturbations. *Linear Algebra and its Applications*, 435 (3): 687 – 716, 2011.

[38] C. Mehl, V. Mehrmann, A. C. M. Ran, and L. Rodman. Eigenvalue perturbation theory of symplectic, orthogonal, and unitary matrices under generic structured rank one perturbations. *BIT Numerical Mathematics*, 54 (1): 219 – 255, 2014.

[39] V. Mehrmann. A symplectic orthogonal method for single input or single output discrete time optimal quadratic control problems. *SIAM Journal on Matrix Analysis and Applications*, 9 (2): 221 – 247, 1988.

[40] J. Moro, J. V. Burke, and M. L. Overton. On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM Journal on Matrix Analysis and Applications*, 18 (4): 793 – 817, 1997.

[41] J. Moro and F. M. Dopico. First order eigenvalue perturbation theory and the Newton diagram. In *Applied Mathematics and Scientific Computing*, pages 143–175. Z. Drmac et al (eds.), Kluwer Academic Publishers, 2002.

[42] J. Moro and F. M. Dopico. Low rank perturbation of Jordan structure. *SIAM Journal on Matrix Analysis and Applications*, 25 (2): 495 – 506, 2003.

[43] A. M. Ostrowski. A quantitative formulation of Sylvester's law of inertia. *Proceedings of the National Academy of Sciences*, 45 (5): 740 – 744, 1959.

[44] V. Puiseux. Recherches sur les fonctions algébriques. *Journal de Matheématiques Pures et Appliqueées* 1 (15): 365 – 480, 1850.

[45] R. Ralha. Perturbation splitting for more accurate eigenvalues. *SIAM Journal on Matrix Analysis and Applications*, 31 (1): 75 – 91, 2009.

[46] A. C. Ran and M. Wojtylak. Eigenvalues of rank one perturbations of unstructured matrices. *Linear Algebra and its Applications*, 437 (2): 589–600, 2012.

[47] F. Rellich. Störungstheorie der spektralzerlegung. I. mitteilung. Analytische störung der isolierten punkteigenwerte eines beschränkten operators. *Mathematische Annalen*, 113: 600 – 619, 1937.

[48] F. Rellich. Störungstheorie der spektralzerlegung. II. mitteilung. Stetige abhängigkeit der spektralschar von einem parameter. *Mathematische Annalen*, 113: 677 – 685, 1937.

[49] F. Rellich. Störungstheorie der spektralzerlegung. III. mitteilung. Analytische, nicht notwendig beschränkte störung. *Mathematische Annalen*, 116: 555 – 570, 1939.

[50] F. Rellich. Störungstheorie der spektralzerlegung. IV. *Mathematische Annalen*, 117: 356 – 382, 1940/1941.

[51] F. Rellich. Störungstheorie der spektralzerlegung. V. *Mathematische Annalen*, 118: 462 – 484, 1941/1943.

[52] F. Rellich *Perturbation Theory of Eigenvalue Problems*. Notes on mathematics and its applications, Gordon and Breach, 1969.

[53] S. V. Savchenko. On the change in the spectral properties of a matrix under perturbations of sufficiently low rank. *Functional Analysis and Its Applications*, 38 (1): 69 – 71, 2004.

[54] E. Schrödinger Quantisierung als Eigenwertproblem III. Störungstheorie, mit Anwendung auf den Starkeffekt der Balmer-Linien. *Annalen der Physik*, 80 (4): 437 – 490, 1926.

[55] E. Schrödinger *Collected Papers on Wave Mechanics*. Collected papers of scientists, Chelsea Publishing Company, 1928.

[56] A. P. Seyranian and A. A. Mailybaev *Multiparameter Stability Theory with Mechanical Applications*. World Scientific, 2003.

[57] F. Sosa and J. Moro. First order asymptotic expansions for eigenvalues of multiplicatively perturbed matrices. *SIAM Journal on Matrix Analysis and Applications*, 37 (4): 1478 – 1504, 2016.

[58] G. Stewart and J. guang Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

[59] J. W. Strutt. *The Theory of Sound*. Cambridge University Press, 1877.

[60] B. Szőkefalvi-Nagy. Perturbations des transformations autoadjointes dans l'espace de Hilbert. *Commentarii Mathematici Helvetici*, 19 (1): 347 – 366, 1946/47.

[61] M. M. Vainberg and V. A. Trenogin. *Theory of Branching of Solutions of Nonlinear Equations*. Wolters-Noordhoff B.V, 1974.

[62] M. I. Vishik and L. A. Lyusternik. The solution of some perturbation problems for matrices and selfadjoint or non-selfadjoint differential equations I. *Russian Mathematical Surveys*, 15 (3): 1, 1960.

[63] J. H. Wilkinson and J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press Oxford, 1965.