Numerical Methods: Treatment of linear eigenvalue problems with the help of the Hamilton-Cayley equation.

Karl Hessenberg

Darmstadt, July 27, 1940

This is only a partial translation of Hessenberg's report.

Notation for matrices and vectors

Capital fraktur letters: square matrices, $\mathfrak{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$, $\mathfrak{J}$ identity matrix, $|\mathfrak{A}|$ determinant of $\mathfrak{A}$,

small fraktur letters: vectors $\mathfrak{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Introduction

The homogeneous system of linear equations

$$(1) \quad \lambda x_\mu = \sum_{\nu=1}^{n} a_{\mu,\nu} x_\nu, \ (\mu = 1, \ldots, n),$$

is represented in matrix notation by the equation

$$(1') \quad (\lambda \mathfrak{J} - \mathfrak{A})\mathfrak{x} = 0,$$

which has, as is well known, for those λ-values, a solution different from 0 when the determinant

$$(2) \quad |\lambda \mathfrak{J} - \mathfrak{A}| \equiv \lambda^n + k_1 \lambda^{n-1} + \cdots + k_{n-1}\lambda + k_n \equiv \mathfrak{p}(\lambda)$$

vanishes. The numbers λ_i which satisfy the equation

$$(3) \quad |\lambda \mathfrak{J} - \mathfrak{A}| = 0$$

are called eigenvalues and the corresponding vectors $\mathfrak{x}_i$, eigensolutions.

Numerous different methods can be used to calculate the eigenvalues and eigensolutions. The computation of the determinant $|\lambda\mathfrak{I} - \mathfrak{A}|$ often requires a quite extensive and cumbersome work, especially for multi-row systems of equations.

Approximation methods, especially when only single eigenvalues and eigensolutions are needed, sometimes lead faster to the goal; However, they are applicable with advantage only under special conditions.

In the first part of this paper, a particularly simple method is presented which is based on an application of the Hamilton-Cayley equation and is already known in its basic idea. The method is not only suitable for the determination of the eigenvalues, but also allows the calculation of the corresponding eigensolutions with relatively little work. However, the method has the disadvantage that in certain cases only individual eigenvalues can be found with a satisfactory accuracy, while the others can only be found very imprecisely or not at all. A further method is therefore developed, the implementation of which also requires only a moderate amount of work, but which can be used in any case and which avoids any unnecessary loss of accuracy.

1 A. First method

1.1 A 1. Calculation of the eigenvalues

The Hamilton-Cayley equation states that between the first n powers of any n-row matrix $\mathfrak{A}$ the following relation exists

$$(4) \quad \mathfrak{p}(\mathfrak{A}) \equiv \mathfrak{A}^n + k_1\mathfrak{A}^{n-1} + \cdots + k_{n-1}\mathfrak{A} + k_n\mathfrak{I} = 0,$$

where the coefficients $k_1, \ldots, k_n$ are the same as those of the polynomial (3).

By multiplication of any vector $\mathfrak{z}_0$, with the matrix polynomial $\mathfrak{p}$ one obtains according to equation (4)

$$(5) \quad \mathfrak{p}(\mathfrak{A})\mathfrak{z}_0 = \mathfrak{A}^n\mathfrak{z}_0 + k_1\mathfrak{A}^{n-1}\mathfrak{z}_0 + \cdots + k_{n-1}\mathfrak{A}\mathfrak{z}_0 + k_n\mathfrak{z}_0 = 0.$$

The vectors appearing here

$$(6) \quad \mathfrak{z}_1 = \mathfrak{A}\mathfrak{z}_0,\ \mathfrak{z}_2 = \mathfrak{A}\mathfrak{z}_1 = \mathfrak{A}^2\mathfrak{z}_0, \ldots, \mathfrak{z}_n = \mathfrak{A}\mathfrak{z}_{n-1} = \mathfrak{A}^n\mathfrak{z}_0,$$

can be calculated by repeatedly multiplying the vectors by the matrix $\mathfrak{A}$ with relatively little work. Thus, according to (5) and (6), we obtain the vector equation

$$(7) \quad k_n\mathfrak{z}_0 + k_{n-1}\mathfrak{z}_1 + \cdots + k_1\mathfrak{z}_{n-1} + \mathfrak{z}_n = 0.$$

This vector equation, in which all vectors $\mathfrak{z}_0, \ldots, \mathfrak{z}_n$ are known, contains n linear equations for the sought n coefficients.

Example: Let

$$\mathfrak{A} = \begin{pmatrix} 10 & 8 & 2 \\ 5 & 6 & 3 \\ 1 & 2 & 4 \end{pmatrix}.$$

By repeatedly multiplying an arbitrarily chosen vector vector, e.g. a unit vector $\mathfrak{z}_0$ by the matrix $\mathfrak{A}$ we get

$$\mathfrak{z}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ \mathfrak{z}_1 = \begin{pmatrix} 10 \\ 5 \\ 5 \end{pmatrix},\ \mathfrak{z}_2 = \begin{pmatrix} 142 \\ 83 \\ 24 \end{pmatrix},\mathfrak{z}_3 = \begin{pmatrix} 2132 \\ 1280 \\ 404 \end{pmatrix}.$$

For the sought coefficients k_1, k_2, k_3 of equation (4) we have, according to (7), the system linear equations

$$\begin{aligned} k_3 + 10\,k_2 + 142\,k_1 + 2132 &= 0 \\ 5\,k_2 + 83\,k_1 + 1280 &= 0 \\ k_2 + 24\,k_1 + 404 &= 0 \end{aligned}$$

whose solution is $k_1 = -20$, $k_2 = 76$ and $k_3 = -52$. It yields the equation

$$\lambda^3 - 20\lambda^2 + 76\lambda - 52 = 0,$$

whose roots are

$$\lambda_1 = 15.23575,\ \lambda_2 = 3.88595,\ \lambda_3 = 0.87830.$$

[Note from the translator: the eigenvalues are $\lambda_1 = 15.23574506553870$, $\lambda_2 = 3.885957537026334$, $\lambda_3 = 0.8782973974349532$.]

1.2 A 2. Calculation of the eigensolutions

If the n eigenvalues λ_i are all different from each other, then to each eigenvalue λ_i corresponds exactly one eigensolution $\mathfrak{x}_i$, determined except for an arbitrarily selectable proportionality factor; The eigensolutions corresponding to the n different eigenvalues $\mathfrak{x}_1, \ldots, \mathfrak{x}_n$ are linearly independent.

If there were a linear relation between them

$$(8) \quad \sum \alpha_i \mathfrak{x}_i = \alpha_1 \mathfrak{x}_1 + \alpha_2 \mathfrak{x}_2 + \cdots + \alpha_n \mathfrak{x}_n = 0,$$

in which at least one of the α_i may be different from zero, then we would also have

$$(8') \quad \mathfrak{A} \sum \alpha_i \mathfrak{x}_i = \lambda_1\, \alpha_1 \mathfrak{x}_1 + \lambda_2\, \alpha_2 \mathfrak{x}_2 + \cdots + \lambda_n\, \alpha_n \mathfrak{x}_n = 0$$

and in general

$$(8'') \quad \mathfrak{A}^p \sum \alpha_i \mathfrak{x}_i = \lambda_1^p\, \alpha_1 \mathfrak{x}_1 + \lambda_2^p\, \alpha_2 \mathfrak{x}_2 + \cdots + \lambda_n^p\, \alpha_n \mathfrak{x}_n = 0$$

The equations (8) to (8") for $p = 0, \ldots, n-1$ form an n-row homogeneous system of equations for the vectors $\alpha_i \mathfrak{x}_i$. Since the determinant of this system of equations, the Vandermonde determinant

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{vmatrix} = \prod_{1 \leq k < i \leq n} (\lambda_i - \lambda_k)$$

vanishes only if, contrary to the presupposition, at least two eigenvalues are equal to each other, so all the $\alpha_i \mathfrak{x}_i = 0$, i.e. a linear relationship (8) between the eigensolutions $\mathfrak{x}_i$ $(i = 1, \ldots, n)$ cannot exist.

[...]

2 B. Second method

2.1 B 1. Preliminary note

The procedure considered so far offers advantages of the common elementary methods: a particularly simple calculation procedure, relatively little work and the possibility to determine also the eigensolutions with moderate computational work.

However, these advantages are counterbalanced by the disadvantage that one of the searched eigenvalues of the matrix is often only obtained with a satisfying accuracy and the others, however, inaccurately or not at all. The reason for such a failure can be both in the unfavorable choice of the initial vector $\mathfrak{z}_0$ and in the properties of the matrix. In the first case, if necessary, the calculation can be repeated with a different initial vector. In the second case, which will occur frequently, especially with multi-row systems of equations, the smaller or little different eigenvalues can often only be obtained by considering a very high number of decimal places or by a fundamental change of the calculation procedure.

A procedure modified in this sense, which avoids any unnecessary loss of accuracy is discussed below. In this method, the formation of small differences from large numbers is avoided by eliminating certain components from each vector calculated by multiplication with the matrix $\mathfrak{A}$ and by ensuring a sufficient linear independence of the vectors used for further calculation from the beginning.

2.2 B 2. Derivation of the method

Let us take for example again $\mathfrak{z}_0 = (\,1 \quad 0 \quad \cdots \quad 0\,)^T$, and

$$\mathfrak{z}_1 = \mathfrak{A}\mathfrak{z}_0 = (\,a_{11} \quad a_{21} \quad \cdots \quad a_{n1}\,)^T \,.$$

Already in this vector the first component is eliminated:

$$\mathfrak{z}_1' = \mathfrak{z}_1 - \alpha_{10}\,\mathfrak{z}_0, \quad \alpha_{10} = a_{11}.$$

One then calculates the vector $\mathfrak{z}_2' = \mathfrak{A}\mathfrak{z}_1'$ and eliminates from this, first with $\mathfrak{z}_0$, then with $\mathfrak{z}_1'$ the second component

$$\mathfrak{z}_2' - \alpha_{20}\mathfrak{z}_0 - \alpha_{21}\mathfrak{z}_1' = \mathfrak{z}_2''.$$

Likewise, from $\mathfrak{z}_3'' = \mathfrak{A}\mathfrak{z}_2''$, the first three components are eliminated and, in general, from the vector

$$(52) \quad \mathfrak{z}_\nu^{(\nu-1)} = \mathfrak{A}\mathfrak{z}_{\nu-1}^{(\nu-1)}, \ (\nu = 1, 2, \ldots, n)$$

the first ν components are eliminated:

$$(53) \quad \mathfrak{z}_\nu^{(\nu-1)} - \alpha_{\nu 0}\mathfrak{z}_0 - \alpha_{\nu 1}\mathfrak{z}_1' - \cdots - \alpha_{\nu,\nu-1}\mathfrak{z}_{\nu-1}^{(\nu-1)} = \mathfrak{z}_\nu^{(\nu)}.$$

Elimination of all n components from the vector $\mathfrak{z}_n^{(n-1)} = \mathfrak{A}\mathfrak{z}_{n-1}^{(n-1)}$, thus results in a linear relation

$$(54) \quad \mathfrak{z}_n^{(n-1)} - \alpha_{n0}\mathfrak{z}_0 - \alpha_{n1}\mathfrak{z}_1' - \cdots - \alpha_{n,n-1}\mathfrak{z}_{n-1}^{(n-1)} = \mathfrak{z}_n^{(n)} = 0$$

between $\mathfrak{z}_n^{(n-1)}$ and the vectors $\mathfrak{z}_\nu^{(\nu-1)}$ $(\nu = 0, 1, \ldots, n-1)$.

The coefficients $\alpha_{\nu,\mu}$ $(\nu = 1, 2, \ldots, n;\ \mu = 0, 1, \ldots, \nu - 1)$ can be used in a simple way to determine the sought coefficients $k_1, \ldots, k_n$, in equation (7), as will be shown below. However, since the vectors $\mathfrak{z}_2, \ldots, \mathfrak{z}_n$ contained in equation (7) are not known, the relationship between them and the vectors $\mathfrak{z}_\nu^{(\nu-1)}$ and $\mathfrak{z}_\nu^{(\nu)}$, respectively, calculated according to equations (52) and (53) must first be established.

The n vectors $\mathfrak{z}_\nu^{(\nu)}$ $(\nu = 0, 1, \ldots, n-1)$ can be combined in a matrix

$$(55) \quad \mathfrak{Z}' = \begin{pmatrix} \mathfrak{z}_0 & \mathfrak{z}_1' & \mathfrak{z}_2'' & \cdots & \mathfrak{z}_{n-1}^{(n-1)} \end{pmatrix}.$$

Between the vector $\mathfrak{z}_\nu^{(\nu-1)}$ and the columns of the matrix $\mathfrak{Z}'$ it exists according to equation (52) and (53) the relation

$$(56) \quad \mathfrak{z}_\nu^{(\nu-1)} = A\mathfrak{z}_{\nu-1}^{(\nu-1)} = \alpha_{\nu 0}\mathfrak{z}_0 + \alpha_{\nu 1}\mathfrak{z}_1' + \cdots + \alpha_{\nu,\nu-1}\mathfrak{z}_{\nu-1}^{(\nu-1)} + \mathfrak{z}_\nu^{(\nu)} = \mathfrak{Z}' \begin{pmatrix} \alpha_{\nu 0} \\ \alpha_{\nu 1} \\ \vdots \\ \alpha_{\nu,\nu-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

One can now combine the vectors $\mathfrak{z}_\nu^{(\nu-1)}$ $(\nu = 1, 2, \ldots, n)$ also in a matrix, namely according to equation (55) and (56)

$$(57) \quad \begin{pmatrix} \mathfrak{z}_1 & \mathfrak{z}_2' & \mathfrak{z}_3'' & \cdots & \mathfrak{z}_n^{(n-1)} \end{pmatrix} = \mathfrak{A}\mathfrak{Z}' = \mathfrak{Z}'\mathfrak{N},$$

where the matrix $\mathfrak{N}$ is a notation for

$$(58) \quad \mathfrak{N} = \begin{pmatrix} \alpha_{10} & \alpha_{20} & \cdots & \alpha_{n-1,0} & \alpha_{n,0} \\ 1 & \alpha_{21} & \cdots & \alpha_{n-1,1} & \alpha_{n,1} \\ 0 & 1 & \cdots & \alpha_{n-1,2} & \alpha_{n,2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \alpha_{n,n-1} \end{pmatrix}.$$

As it will be shown later, it is always possible to achieve that the determinant of the matrix $\mathfrak{Z}'$ is different from zero, if necessary by slightly changing the calculation procedure. Under this condition, equation (57) can also be written in the following way by multiplying both sides by $(\mathfrak{Z}')^{-1}$,

$$(59) \quad (\mathfrak{Z}')^{-1}\mathfrak{A}\mathfrak{Z}' = \mathfrak{N}, \quad (\mathfrak{Z}')\mathfrak{N}(\mathfrak{Z}')^{-1} = \mathfrak{A}.$$

This follows also for every power $\mathfrak{A}^\nu$ $(\nu = 1, 2, \ldots)$ and also for each polynomial $f(\mathfrak{A})$ of the matrix $\mathfrak{A}$,

$$(60) \quad (\mathfrak{Z}')^{-1}\mathfrak{A}^\nu\mathfrak{Z}' = [(\mathfrak{Z}')^{-1}\mathfrak{A}\mathfrak{Z}']^\nu = \mathfrak{N}^\nu,$$

$$(61) \quad (\mathfrak{Z}')^{-1}f(\mathfrak{A})\mathfrak{Z}' = f[(\mathfrak{Z}')^{-1}\mathfrak{A}\mathfrak{Z}'] = f(\mathfrak{N}).$$

If one sets here for $f(\mathfrak{A})$ the polynomial of Hamilton-Cayley's equation (4)

$$\mathfrak{p}(\mathfrak{A}) = \mathfrak{A}^n + k_1\mathfrak{A}^{n-1} + \cdots + k_{n-1}\mathfrak{A} + k_n\mathfrak{I},$$

it follows from $\mathfrak{p}(\mathfrak{A}) = 0$ and equation (61) that

$$(62) \quad \mathfrak{p}(\mathfrak{N}) = \mathfrak{N}^n + k_1\mathfrak{N}^{n-1} + \cdots + k_{n-1}\mathfrak{N} + k_n\mathfrak{I} = 0.$$

Thus, Hamilton-Cayley's equation remains satisfied if the matrix $\mathfrak{N}$ defined by equation (58) is substituted for the matrix $\mathfrak{A}$.

For the sake of completeness let us show that the matrices $\mathfrak{A}$ and $\mathfrak{N}$ have exactly the same eigenvalues, so that each $p_{(i)}$-fold eigenvalue $\lambda_{(i)}$, of the matrix $\mathfrak{A}$ is also exactly the $p_{(i)}$-fold eigenvalue $\lambda_{(i)}$, of the matrix $\mathfrak{N}$.

As shown in section A 4, for the matrix $\mathfrak{A}$ corresponding to its m single or multiple eigenvalues $\lambda_{(i)}$, there are m unit submatrices $\mathfrak{Y}_{(i)}$ which satisfy the conditions (33) and (35), namely corresponding to each $p_{(i)}$-fold eigenvalue $\lambda_{(i)}$, a unit submatrix $\mathfrak{Y}_{(i)}$, of rank $p_{(i)}$. By means of equations (33') and (25) (35') it was proved that this decomposition is unique.

According to equations (26) and (30), the matrix $\mathfrak{Y}_{(i)}$ is a polynomial of the matrix $\mathfrak{A}$:

$$\mathfrak{Y}_{(i)} = f(\mathfrak{A}).$$

Let the corresponding polynomial of matrix $\mathfrak{R}$ be

$$\mathfrak{W}_{(i)} = f(\mathfrak{R}).$$

From equation (61) it follows

$$(63) \quad (\mathfrak{Z}')^{-1}\mathfrak{Y}_{(i)}\mathfrak{Z}' = \mathfrak{W}_{(i)}.$$

Likewise,

$$(\mathfrak{Z}')^{-1}(\mathfrak{A} - \lambda_{(i)}\mathfrak{I})^{p_{(i)}}\mathfrak{Z}' = (\mathfrak{Z}')^{-1}(\mathfrak{R} - \lambda_{(i)}\mathfrak{I})^{p_{(i)}}\mathfrak{Z}',$$

and consequently

$$(\mathfrak{R} - \lambda_{(i)}\mathfrak{I})^{p_{(i)}}\mathfrak{W}_{(i)} = (\mathfrak{Z}')^{-1}(\mathfrak{A} - \lambda_{(i)}\mathfrak{I})^{p_{(i)}}\mathfrak{Y}_{(i)}Z'.$$

Thus, according to equation (33)

$$(64) \quad (\mathfrak{R} - \lambda_{(i)}\mathfrak{I})^{p_{(i)}}\mathfrak{W}_{(i)} = 0.$$

On the other hand, it follows from (63) and (35) that

$$(65) \quad \sum_{i=1}^{m}\mathfrak{W}_{(i)} = \sum_{i=1}^{m}(\mathfrak{Z}')^{-1}\mathfrak{Y}_{(i)}\mathfrak{Z}' = (\mathfrak{Z}')^{-1}[\sum_{i=1}^{m}\mathfrak{Y}_{(i)}]\mathfrak{Z}' = (\mathfrak{z}')^{-1}\mathfrak{I}\mathfrak{Z}' = \mathfrak{I}.$$

The matrices $\mathfrak{W}_{(i)}$, $(i = 1, \ldots, m)$ thus satisfy, with equations (64) and (65), the sufficient conditions (33) and (35) as unit submatrices of the matrix $\mathfrak{R}$. Since, according to equation (63), every matrix $\mathfrak{W}_{(i)}$, as well as the corresponding matrix $\mathfrak{Y}_{(i)}$ has the rank $p_{(i)}$, it follows from the remarks of section A 4 that $\lambda_{(i)}$ is an exactly $p_{(i)}$-fold eigenvalue of the matrix $\mathfrak{R}$.

Thus, after having determined the entries $\alpha_{\nu,\mu}$ of the matrix $\mathfrak{R}$ we can use any of the known methods to compute the eigenvalues of the matrix $\mathfrak{R}$. These agree exactly with the eigenvalues of the matrix $\mathfrak{A}$; it is

$$(66) \quad |\lambda\mathfrak{I} - \mathfrak{R}| = |\lambda\mathfrak{I} - \mathfrak{A}| = \mathfrak{p}(\lambda) = \prod_{i=1}^{m}|\lambda - \lambda_{(i)}|^{p_{(i)}}.$$

The computation of the determinant

$$(67) \quad |\lambda\mathfrak{I} - \mathfrak{R}| = \begin{vmatrix} \lambda - \alpha_{10} & -\alpha_{20} & \cdots & -\alpha_{n-1,0} & -\alpha_{n,0} \\ -1 & \lambda - \alpha_{21} & \cdots & -\alpha_{n-1,1} & -\alpha_{n,1} \\ 0 & -1 & \cdots & -\alpha_{n-1,2} & -\alpha_{n,2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & \lambda - \alpha n, n-1 \end{vmatrix}$$

can be carried out according to the known elementary procedures. The calculation is especially simple and clear, if one computes the subdeterminants

$$F_1 = \lambda - \alpha_{10},\ F_2 = \begin{vmatrix} \lambda - \alpha_{10} & -\alpha_{20} \\ -1 & \lambda - \alpha_{21} \end{vmatrix},\ F_3 = \begin{vmatrix} \lambda - \alpha_{10} & -\alpha_{20} & -\alpha_{30} \\ -1 & \lambda - \alpha_{21} & -\alpha_{31} \\ 0 & -1 & \lambda - \alpha_{32} \end{vmatrix},$$

and so on, up to the determinant

$$F_n = |\lambda\mathfrak{I} - \mathfrak{R}| = \mathfrak{p}(\lambda).$$

Between these expressions F_ν, $(\nu = 1, 2, \ldots, n)$ one has, if one develops the individual subdeterminants with the elements of their last column, the following relations: (69)

$$\begin{aligned} F_2 &= (\lambda - \alpha_{21})F_1 - \alpha_{20}, \\ F_3 &= (\lambda - \alpha_{32})F_2 - \alpha_{31}F_1 - \alpha_{30}, \\ F_4 &= (\lambda - \alpha_{43})F_3 - \alpha_{42}F_2 - \alpha_{41}F_1 - \alpha_{40}, \\ \vdots &= \vdots \\ F_n &= (\lambda - \alpha_{n,n-1})F_{n-1} - \alpha_{n,n-2}F_{n-2} - \cdots - \alpha_{n,1}F_1 - \alpha_{n,0}. \end{aligned}$$

This calculation is carried out at the end of this paper using a numerical example.

One could also, in order to determine the eigenvalues of the matrix $\mathfrak{R}$, apply again the procedure described in section A 1, e.g. by calculating the vector series

$$u_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; u_1 = \mathfrak{R}u_0 = \begin{pmatrix} \alpha_{10} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; u_2 = \mathfrak{R}u_1 = \begin{pmatrix} \alpha_{10}^2 + \alpha_{20} \\ \alpha_{10} + \alpha_{21} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots$$

starting from a unit vector u_0. In the matrix $\mathfrak{U} = \begin{pmatrix} u_0 & u_1 & \cdots & u_{n-1} \end{pmatrix}$ all elements below the main diagonal disappear, and all elements of the main diagonal itself are equal to 1. The solution of the linear system of equations corresponding to equation (7) for $k_1, \ldots, k_n$

$$k_n u_0 + k_{n-1}u_1 + \cdots + k_1 u_{n-1} = -u_n,$$

therefore requires only a small amount of work in itself. However, in this calculation computational difficulties can occur again due to the formation of small differences from large numbers. This disadvantage is avoided by directly multiplying out the determinant $|\lambda\mathfrak{I} - \mathfrak{R}|$ according to equation (69).

2.3 B 3. Carrying out the procedure with complete or approximate disappearance of a vector $\mathfrak{z}_\nu^{(\nu)}$

It remains to prove that it is always possible to achieve that the determinant of the matrix $\mathfrak{Z}'$ is different from zero without a fundamental change of the procedure. With the rule given at the beginning of section B 2 for determining the vectors $\mathfrak{z}_\nu^{(\nu)}$, equation (55) for $\mathfrak{Z}'$ results in a matrix in which all elements above the main diagonal disappear. If all the main diagonal elements of the matrix $\mathfrak{Z}'$, that is, the first, not intentionally eliminated components of the vectors $\mathfrak{z}_\nu^{(\nu)}$, are different from zero, then the determinant of the matrix $\mathfrak{Z}'$ is equal to the product of these diagonal elements, also different from zero.

However, if the elimination of the first ν components from the vector $\mathfrak{z}_\nu^{(\nu_1)}$ according to equation (53), the $(\nu+1)$-th component disappears, two cases have to be distinguished:

α) The vector $\mathfrak{z}_\nu^{(\nu)}$ $(\nu < n)$ disappears completely. This case occurs if, for one of the reasons discussed in Sections A 3 and 4, the vectors $\mathfrak{z}_0, \mathfrak{z}_1, \ldots, \mathfrak{z}_\nu$ are linearly dependent. As the vector $\mathfrak{z}_\nu^{(\nu)}$ vanishes, all of the following columns of the matrix $\mathfrak{Z}'$ also vanish, as well as the determinant of this matrix.

However, one can consider the vector $\mathfrak{z}_\nu^{(\nu)}$, as the limit of an expression $\delta\bar{\mathfrak{z}}_\nu^{(\nu)}$, where $\bar{\mathfrak{z}}_\nu^{(\nu)}$ is an arbitrary vector different from zero and δ is a scalar factor vanishing in the limiting case. For $\bar{\mathfrak{z}}_\nu^{(\nu)}$ one chooses best the $(\nu+1)$-th column of the unit matrix, that is, a unit vector, whose $(\nu+1)$-th component is equal to 1 and all remaining components equal to 0.

Instead of the exact notation $\mathfrak{z}_\nu^{(\nu)} = \lim_{\delta\to 0}\bar{\mathfrak{z}}_\nu^{(\nu)}$ one may use in the following $\mathfrak{z}_\nu^{(\nu)} = \delta\bar{\mathfrak{z}}_\nu^{(\nu)}$ where in the limit case $\delta = 0$ is to be set.

For the further calculation instead of the vanishing vector $\mathfrak{z}_\nu^{(\nu)}$, the vector $\bar{\mathfrak{z}}_\nu^{(\nu)}$ different from zero is used. One calculates $\bar{\mathfrak{z}}_\nu^{(\nu+1)} = \mathfrak{A}\bar{\mathfrak{z}}_\nu^{(\nu)}$ according to equation (52) and eliminates its $(\nu+1)$ first components according to equation (53)

$$(72)\quad \bar{\mathfrak{z}}_\nu^{(\nu+1)} - \bar{\alpha}_{\nu+1,0}\mathfrak{z}_0 - \bar{\alpha}_{\nu+1,1}\mathfrak{z}_1' - \cdots - \bar{\alpha}_{\nu+1,\nu}\bar{\mathfrak{z}}_\nu^{(\nu)} = \bar{\mathfrak{z}}_{\nu+1}^{(\nu+1)}.$$

It only has to be noted that in order to get $\mathfrak{z}_{\nu+1}^{(\nu)}$ or $\mathfrak{z}_{\nu+1}^{(\nu+1)}$, one has to multiply the vectors $\bar{\mathfrak{z}}_{\nu+1}^{(\nu)}$ or $\bar{\mathfrak{z}}_{\nu+1}^{(\nu+1)}$ by the factor δ to obtain (73)

$$\begin{aligned}
\mathfrak{z}_{\nu+1}^{(\nu+1)} = \delta\bar{\mathfrak{z}}_{\nu+1}^{(\nu+1)} &= \delta\bar{\mathfrak{z}}_{\nu+1}^{(\nu)} - \sum_{\mu=0}^{\nu-1}\delta\bar{\alpha}_{\nu+1,\mu}\mathfrak{z}_\mu^{(\mu)} - \delta\bar{\alpha}_{\nu+1,\nu}\bar{\mathfrak{z}}_\nu^{(\nu)}, \\
&= \mathfrak{z}_{\nu+1}^{(\nu)} - \sum_{\mu=0}^{\nu-1}(\delta\bar{\alpha}_{\nu+1,\mu})\mathfrak{z}_\mu^{(\mu)} - \bar{\alpha}_{\nu+1,\nu}\mathfrak{z}_\nu^{(\nu)}.
\end{aligned}$$

This gives the coefficients

$$\alpha_{\nu+1,\mu} = \delta\bar{\alpha}_{\nu+1,\mu} \to 0, \qquad \text{for } \mu = 0, 1, \ldots, \nu-1$$

and

$$\alpha_{\nu+1,\nu} = \bar{\alpha}_{\nu+1,\nu}.$$

to be included in matrix $\mathfrak{R}$.

In general, for the vectors $\bar{\mathfrak{z}}_{\nu+\rho}^{(\nu+\rho)}$ and $\mathfrak{z}_{\nu+\rho}^{(\nu+\rho)}$ we get

$$(72') \quad \bar{\mathfrak{z}}_{\nu+\rho}^{(\nu+\rho)} = \mathfrak{A}\bar{\mathfrak{z}}_{\nu+\rho-1}^{(\nu+\rho-1)} - \sum_{\mu=0}^{\nu-1} \bar{\alpha}_{\nu+\rho,\mu}\mathfrak{z}_{\mu}^{(\mu)} - \sum_{\mu=\nu}^{\nu+\rho-1} \bar{\alpha}_{\nu+\rho,\mu}\bar{\mathfrak{z}}_{\mu}^{(\mu)}$$

$$(73') \quad \mathfrak{z}_{\nu+\rho}^{(\nu+\rho)} = \delta\bar{\mathfrak{z}}_{\nu+\rho}^{(\nu+\rho)} = \mathfrak{A}\mathfrak{z}_{\nu+\rho-1}^{(\nu+\rho-1)} - \sum_{\mu=0}^{\nu-1} (\delta\bar{\alpha}_{\nu+\rho,\mu})\mathfrak{z}_{\mu}^{(\mu)} - \sum_{\mu=\nu}^{\nu+\rho-1} \bar{\alpha}_{\nu+\rho,\mu}\mathfrak{z}_{\mu}^{(\mu)},$$

and

$$(74) \quad \alpha_{\nu+\rho,\mu} = \delta\bar{\alpha}_{\nu+\rho,\mu} \to 0 \text{ for } \mu = 0, 1, \ldots, \nu-1$$

as well as

$$(75) \quad \alpha_{\nu+\rho,\mu} = \bar{\alpha}_{\nu+\rho,\mu} \text{ for } \mu = \nu, \nu+1, \ldots, \nu+\rho-1.$$

If one of the vectors $\bar{\mathfrak{z}}_{\nu+\rho}^{(\nu+\rho)}$ also vanishes in the further course of the calculation, a finite vector can also be introduced here again as $\bar{\mathfrak{z}}_{\nu+\rho}^{(\nu+\rho)} = \delta_2\bar{\bar{\mathfrak{z}}}_{\nu+\rho}^{(\nu+\rho)}$ with $\delta_2 \to 0$ and this approach can also be repeated for even further vanishing vectors.

The numerical example used at the end of section A 3 serves as an explanation,

$$\mathfrak{A} = \begin{pmatrix} 6 & 3 & -3 & -1 \\ 3 & 5 & 3 & -6 \\ -3 & 3 & 14 & -9 \\ -1 & -6 & -9 & 21 \end{pmatrix}, \quad \mathfrak{z}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathfrak{z}_1 = \begin{pmatrix} 6 \\ 3 \\ -3 \\ -1 \end{pmatrix}, \quad \mathfrak{z}_1' = \mathfrak{z}_1 - 6\mathfrak{z}_0 = \begin{pmatrix} 0 \\ 3 \\ -3 \\ -1 \end{pmatrix}$$

$$\mathfrak{z}_2' = \begin{pmatrix} 19 \\ 12 \\ -24 \\ -12 \end{pmatrix}, \quad \mathfrak{z}_2'' = \mathfrak{z}_2' - 19\mathfrak{z}_0 - 4\mathfrak{z}_1' = \begin{pmatrix} 0 \\ 0 \\ -12 \\ -8 \end{pmatrix}, \quad \mathfrak{z}_3'' = \begin{pmatrix} 44 \\ 12 \\ -96 \\ -60 \end{pmatrix},$$

$$\mathfrak{z}_3''' = \mathfrak{z}_3'' - 44\mathfrak{z}_0 - 4\mathfrak{z}_1' - 7\mathfrak{z}_2'' = 0.$$

To be able to continue the calculation, $\mathfrak{z}_3''' = \delta\bar{\mathfrak{z}}_3'''$ with

$$\bar{\mathfrak{z}}_3''' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \bar{\mathfrak{z}}_4''' = \begin{pmatrix} -1 \\ -6 \\ -9 \\ 21 \end{pmatrix}, \quad \bar{\mathfrak{z}}_4^{(4)} = \bar{\mathfrak{z}}_4''' + \mathfrak{z}_0 + 2\mathfrak{z}_1' - 1.25\mathfrak{z}_2'' - 29\bar{\mathfrak{z}}_3''' = 0.$$

Thus, by combining the coefficients $\alpha_{\nu,\mu}$ as shown, one obtains

$$\mathfrak{R} = \begin{pmatrix} 6 & 19 & 44 & -\delta \\ 1 & 4 & 4 & -2\delta \\ 0 & 1 & 7 & -1.25\delta \\ 0 & 0 & 1 & 29 \end{pmatrix}$$

in which one has to set $\delta = 0$.

You can easily convince yourself that all eigenvalues of this matrix correspond to those of the matrix $\mathfrak{A}$.

β) In the vector $\mathfrak{z}_\nu^{(\nu)}$, in addition to the first ν components, the $(\nu+1)$-th component also disappears, but at least one, e.g. the $(\nu+\rho)$-th component of the vector is different from zero.

In this case, a fundamental change in the procedure is not necessary. One must eliminate then only in the following vectors $\mathfrak{z}_{\nu+1}^{(\nu)}$ instead of the $(\nu+1)$-th component the $(\nu+\rho)$-th component.

Such a change in the order of the components to be eliminated may be advantageous in other cases as well. In order to avoid even the slightest unnecessary loss of accuracy due to subtraction, it is advisable, when calculating the vector $\mathfrak{z}_{\nu+1}^{(\nu+1)}$, to eliminate the component that occurs in $\mathfrak{z}_\nu^{(\nu)}$ with the largest amount in addition to those components that have already been eliminated. However, a significant influence on the accuracy of the results is only to be expected if the components differ considerably from $\mathfrak{z}_\nu^{(\nu)}$.

[...]

Summary

For the numerical solution of the equation systems occurring in the calculation of free periodic oscillations, two methods are discussed which are based on an application of Hamilton-Cayley's equation.

The first method is characterized by a particularly simple and clear process and, compared with other methods, requires only a small amount of computational work. Also the calculation of the eigensolutions is here done in a simple manner and with small amount of work. A disadvantage of this method is that in certain cases, as a result of the calculation of small differences from large numbers, some of the eigenvalues and eigensolutions are found only inaccurately or not found at all.

Therefore, based on the same basic idea, a second method has been developed which avoids this disadvantage. By suitable arrangement, the calculation can also be made quite simple and clear with this method. Here too, after determining the eigenvalues, the eigensolutions can be determined easily and with little computational work.

The calculation of a numerical example explains the practical application of the two methods.