PHY 564
Advanced Accelerator Physics
Lecture 6
Matrices and Matrix function
Silvester formulae

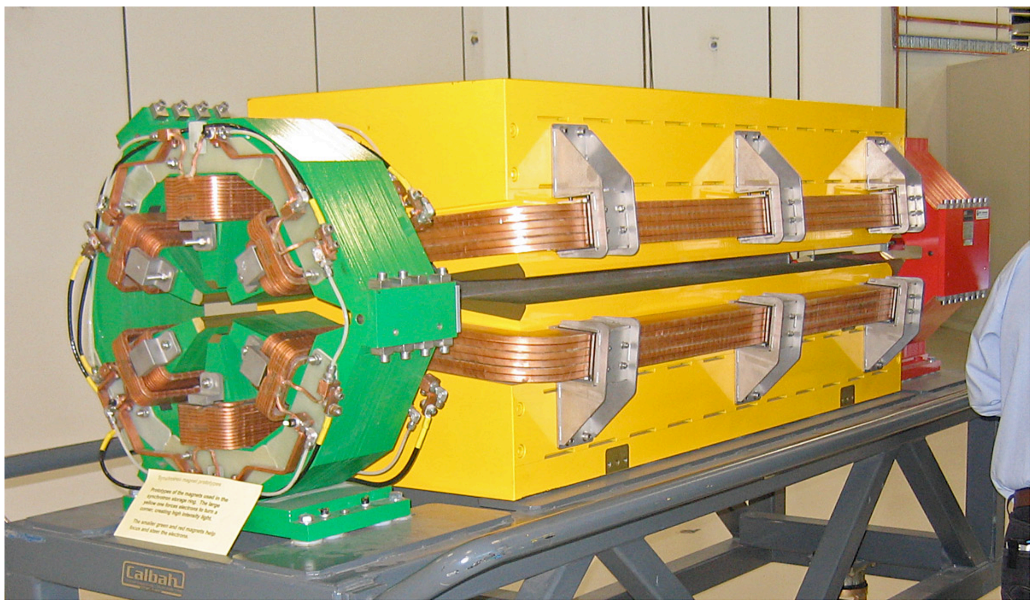
Vladimir N. Litvinenko
Yichao Jing
Gang Wang

CENTER for ACCELERATOR SCIENCE AND EDUCATION
Department of Physics & Astronomy, Stony Brook University
Collider-Accelerator Department, Brookhaven National Laboratory

Matrices and matrix functions. As a practical matter, when somebody wants to build an accelerator, she or he should use some approximations. One of VERY popular design approximation is called "an element (usually a magnet)" with nearly constant parameters. Then our Hamiltonian is s-independent on at part of the trajectory.

$$\begin{gathered}\mathbf{H}=\mathbf{H}_i(s);\;\; \mathbf{H}_i(s)=const;\left\{s_i<s<s_{i+1}\right\};\; \frac{d\mathbf{M}}{ds}=\mathbf{SH}\cdot\mathbf{M};\;\; \mathbf{D}=\mathbf{SH} \\ \mathbf{M}(s_o,s)=\prod_{i=1}\mathbf{M}_i;\quad \mathbf{M}_i(s_i,s)=\exp\left(\mathbf{SH}_i(s-s_i)\right)\end{gathered} \tag{187}$$

e.g. we just need to learn how to calculate $\exp\left(\mathbf{SH}_i(s-s_i)\right)$. Finally, she or he than should try to build such elements. They never ideal but can be relatively close to the ideal boxes…



Typical elements of accelerators are dipoles and quadrupoles (or their combination), sextupoles and octupoles (they a nonlinear), solenoids, wigglers…. Let's start from a linearized Hamiltonian (143) magnetic DC elements – this is typical accelerator beam-line.

$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F\frac{x^2}{2} + Nxy + G\frac{y^2}{2} + L\left(xP_3 - yP_1\right) + \frac{\delta^2}{2p_o}\cdot\frac{m^2c^2}{p_o^2} + g_x x\delta\,; \qquad (188)$$

with

$$\frac{F}{p_o} = \left[\left(\frac{e}{p_oc}B_y\right)^2 - \frac{e}{p_oc}\frac{\partial B_y}{\partial x} + \left(\frac{eB_s}{2p_oc}\right)^2\right]; \frac{G}{p_o} = \left[\frac{e}{p_oc}\frac{\partial B_x}{\partial y} + \left(\frac{eB_s}{2p_oc}\right)^2\right]$$

$$\frac{N}{p_o} = \left[\frac{e}{p_oc}\frac{\partial B_x}{\partial x}\right];\ \ L = \kappa + \frac{e}{2p_oc}B_s; \qquad g_x = -K\frac{c}{\mathrm{v}_o}; \qquad (189)$$

$$\frac{\partial B_y}{\partial x} = \frac{\partial B_x}{\partial y}; \frac{\partial B_x}{\partial x} = -\frac{\partial B_y}{\partial y};$$

If momentum p_o is constant, we can use (134) and rewrite Hamiltonian of the linearized motion as

$$\tilde{h}_n = \frac{\pi_1^2 + \pi_3^2}{2} + f\frac{x^2}{2} + n\cdot xy + g\frac{y^2}{2} + L\left(x\pi_3 - y\pi_1\right) + \frac{\pi_o^2}{2}\cdot\frac{m^2c^2}{p_o^2} + g_x x\pi_o\,; \qquad (188\text{-n})$$

with

$$f = \frac{F}{p_o}; n = \frac{N}{p_o}; g = \frac{G}{p_o};\ ; \qquad (189\text{-n})$$

Focusing/defocusing in transverse direction can come from
(a) a dipole field B_y or in other words, form the curvature of trajectory. Note that it is always focusing.
(b) from quadrupole field $\frac{\partial B_y}{\partial x} = \frac{\partial B_x}{\partial y}$. Note that quadrupole is focusing in one direction and defocusing in the other.
(c) from solenoidal field, B_s. Note that it is always focusing.
The other terms, are responsible for coupling
(a) the transverse motion (x & y): solenoidal field, B_s and torsion κ as well as SQ-quadrupole $\frac{\partial B_x}{\partial x}$.
(b) or transverse and longitudinal motion: $g_x x\delta$ - it is responsible of dependence of the time of flight on transverse coordinate.
Finally, there is $\frac{\delta^2}{2p_o} \cdot \frac{m^2c^2}{p_o^2}$ term which is corresponds to the velocity dependence on the particle energy. It is frequently neglected at very high energies when $m^2c^2 / p_o^2 \approx \gamma^{-2} <<< 1$. But it should be kept for many accelerators, including RHIC.
We should not forget one of the most common element in any accelerator lattice – an empty space, call drift.
In standard accelerator physics book you will find solution (matrices) for various elements of the lattice: drift, bending magnet (with or with field gradient), quadrupole. Then, piecewise, you can see introduction of solenoids, SQ-quadrupoles…. Instead of solving dozen of second, fourth and sixth order differential equations… we will use matrix function approach to find all solutions at once.

Calculating matrices. Next, we focus on the question of how matrices are calculated. We already discussed general idea than they can be integrates piece-wise wherein the coefficients in the Hamiltonian expansion do not change significantly. In practice, accelerators are build from elements, which, to a certain extent, offers such conditions.

Since method of calculating 6x6 or 4x4 (or even some 2x2) matrices is very similar to that for 2nx2n, where n is arbitrary integer. Hence, initially we will explore a general way of calculating matrices, and then consider few examples. When the matrices **D** are piece-wise constant and the **D** from different elements do not commute, we can write

$$\mathbf{M}(s_o|s) = \prod_i \mathbf{M}(s_{i-1}|s_i); \mathbf{M}(s_{i-1}|s) = \prod_{elements} \exp[\mathbf{D_i}(s - s_{i-1})] \tag{193}$$

The definition of the matrix exponent is very simple

$$\exp[\mathbf{A}] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}; \quad \exp[\mathbf{D} \cdot s] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{D}^k s^k}{k!} \tag{194}$$

According to the general theorem of Hamilton-Kelly, the matrix is a root of its characteristic equation:

$$d(\lambda) = \det[\mathbf{D} - \lambda I]; \quad d(\lambda_k) = 0 \tag{195}$$

$$d(\mathbf{D}) \equiv 0 \tag{196}$$

i.e., a root of a polynomial of order ≤ 2n. There is a theorem in theory of polynomials (rather easy to prove) that any polynomial ***$p_1(x)$*** of power n can be expressed via any polynomial ***$p_2(x)$*** of power m<n as

$$p_1(x) = p_2(x) \cdot d(x) + r(x)$$

where r(x) is a polynomial of power less than m. Accordingly, series (194) can be always truncated to

$$\exp[\mathbf{D}] = I + \sum_{k=1}^{2n-1} c_k \mathbf{D}^k, \tag{197}$$

with the remaining daunting task of finding coefficients c_k!

There are two ways of doing this; one is a general, and the other is case specific, but an easy one. Starting from a specific case when the matrix **D** is nilpotent (m<2n+1), i.e.,

$$\mathbf{D}^m = 0.$$

In this case, $\mathrm{D}^{m+j} = 0$ the truncation is trivial:

$$\exp[D] = I + \sum_{k=1}^{m-1} \frac{D^k}{k!}. \tag{198}$$

We lucky to have such a beautiful case in hand – a drift, where all fields are zero and K=0 and κ=0:

$$\tilde{h} = \frac{\pi_1^{\,2} + \pi_3^{\,2}}{2} + \frac{\pi_\delta^{\,2}}{2} \cdot \frac{m^2c^2}{p_o^{\,2}}; \quad \mathbf{D} = \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{bmatrix}; D = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; D_2 = \begin{bmatrix} 0 & \frac{m^2c^2}{p_o^{\,2}} \\ 0 & 0 \end{bmatrix}; \tag{199}$$

where it is easy to check: $\mathrm{D}^2 = 0$. Hence, the 6x6 matrix of drift with length l will be

$$\mathbf{M}_{drift} = \exp[\mathbf{D} \cdot l] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{D}^k l^k}{k!} = \mathbf{I} + \mathbf{D} \cdot l = \begin{bmatrix} M_t & 0 & 0 \\ 0 & M_t & 0 \\ 0 & 0 & M_\tau \end{bmatrix}; M_t = \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix}; M_\tau = \begin{bmatrix} 1 & l/(\beta_o\gamma_o)^2 \\ 0 & 1 \end{bmatrix}; \text{(200}$$

The general evaluation of the matrix exponent in (193) is straightforward using the eigen values of the D-matrix:

$$\det[\mathbf{D}-\lambda\cdot\mathbf{I}]=\det[\mathbf{SH}-\lambda\cdot\mathbf{I}]=0 \tag{201}$$

When the eigen values are all different (2n numerically different eigen values, $\lambda_i=\lambda_i \Rightarrow i=j$, no degeneration, i.e., D can be diagonalized),

$$\mathbf{D}=\mathbf{U}\Lambda\mathbf{U}^{-1};\ \Lambda=\begin{pmatrix}\lambda_1 & 0 & & 0\\ 0 & \lambda_2 & & 0\\ & & \ldots & 0\\ 0 & 0 & 0 & \lambda_{2n}\end{pmatrix}; \tag{202}$$

we can use Sylvester's formula that is correct for any analytical f(D), http://en.wikipedia.org/wiki/Sylvester's_formula for evaluating (193):

$$\exp[\mathbf{D}s]=\sum_{k=1}^{2n}e^{\lambda_k s}\prod_{j\neq k}\frac{\mathbf{D}-\lambda_j\mathbf{I}}{\lambda_k-\lambda_j} \tag{203}$$

Let's prove this very useful formula. First, let consider a polynomial function

$$f_N(x) = \sum_{k=0}^{N} a_k x^k \tag{204}$$

and apply it to (202)

$$f_N(D) = \sum_{k=0}^{N} a_k D^k = \sum_{k=0}^{N} a_k \left(\mathbf{U}\Lambda\mathbf{U}^{-1}\right)^k = \mathbf{U}\left\{\sum_{k=0}^{N} a_k \Lambda^k\right\}\mathbf{U}^{-1} = \mathbf{U}\cdot f_N(\Lambda)\cdot\mathbf{U}^{-1}$$

$$f_N(\Lambda) \equiv \begin{bmatrix} \dots & 0 & 0 \\ 0 & f_N(\lambda_i) & 0 \\ 0 & 0 & \dots \end{bmatrix} \tag{205}$$

e.g. function of diagonalizable matrix is a similarity transformation of the diagonal matrix with function of it eigen values. Goin to infinite series, we get

$$\exp(D) = \sum_{k=0}^{\infty} \frac{D^n}{k!} = \mathbf{U}\sum_{k=0}^{\infty} a_k (\Lambda)^k \mathbf{U}^{-1} = \mathbf{U}\exp(\Lambda)\mathbf{U}^{-1}$$

$$\exp(\Lambda) \equiv \begin{bmatrix} \dots & 0 & 0 \\ 0 & e^{\lambda_i} & 0 \\ 0 & 0 & \dots \end{bmatrix} \tag{206}$$

Now we start using our refresher on linear algebra. Each eigen value of diagonalizable matrix corresponds to an eigen vector

$$D \cdot Y_i = \lambda_i Y_i \,. \tag{207}$$

(existence comes from statement that $(D - \lambda_i I) Y_i = 0$ has non-trivial solution if $\det(D - \lambda_i I) = 0$). The set of eigen vectors is a full set of vectors, e.g. any arbitrary vector can be expanded as

$$X = \sum_i \alpha_i Y_i \,. \tag{208}$$

This eigen vectors are columns of the matrix used for similarity transform to its diagonal form:

$$\mathbf{U} = \left[Y_1, Y_2, ..., Y_{2n} \right] \tag{209}$$

which is trivial to prove using (208(and (209) and comparing it with (202)

$$\begin{aligned} &\mathbf{DU} = \mathbf{U}\Lambda; \quad \to \mathbf{D} = \mathbf{U}\Lambda\mathbf{U}^{-1} \\ &\mathbf{U}\Lambda \equiv \left[\lambda_1 Y_1, \lambda_2 Y_2 ..., \lambda_{2n} Y_{2n} \right] \end{aligned} \tag{210}$$

Now, let's build a unit projection operator on Y_k :

$$P_k = \prod_{i \neq k} \frac{M - \lambda_i I}{\lambda_k - \lambda_i} \tag{211}$$

It is easy to show that

$$P_k Y_k = Y_k; \;\; P_k Y_{i \neq k} = 0; \tag{212}$$

First, each of the elements of the product (211) is unit on Y_k

$$\frac{M - \lambda_i I}{\lambda_k - \lambda_i} \cdot Y_k = \frac{\lambda_k - \lambda_i}{\lambda_k - \lambda_i} Y_k = Y_k ; i \neq k \tag{213}$$

while there is a zero-operator for all other eigen vectors:

$$\frac{M - \lambda_i I}{\lambda_k - \lambda_i} \cdot Y_i = \frac{\lambda_i - \lambda_i}{\lambda_k - \lambda_i} \cdot Y_i = 0 \tag{214}$$

Now we write

$$P_k \mathbf{U} = \left[\ldots 0, Y_k, 0 \ldots \right] \tag{215}$$

and

$$\begin{gathered} f(D) = \mathbf{U} \cdot f(\Lambda) \cdot \mathbf{U}^{-1} \\ \mathbf{U} \cdot f(\Lambda) = \sum_{k=1}^{2n} f(\lambda_k) [\ldots 0, Y_k, 0 \ldots] = \sum_{k=1}^{2n} f(\lambda_k) \cdot P_k \cdot \mathbf{U} \end{gathered} \tag{216}$$

and finally

$$\begin{gathered} = \mathbf{U} \cdot f(\Lambda) \cdot \mathbf{U}^{-1} \\ f(D) = \mathbf{U} \cdot f(\Lambda) \cdot \mathbf{U}^{-1} == \sum_{k=1}^{2n} f(\lambda_k) \cdot P_k \cdot \mathbf{U} \cdot \mathbf{U}^{-1} = \sum_{k=1}^{2n} f(\lambda_k) \cdot P_k \end{gathered} \tag{217}$$

e.g.

$$f[\mathbf{D}] = \sum_{k=1}^{2n} f(\lambda_k) \prod_{j \neq k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \tag{218}$$

equivalent to

$$f[\mathbf{D}s] = \sum_{k=1}^{2n} f(\lambda_k s) \prod_{j \neq k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \tag{219}$$

we got famous Sylvester formula.

We will use most of the time $f:\exp$ and Sylvester formula in form of (203). Naturally, (219) is comprised of power of matrix $\mathbf{D}$ up to 2n-1 – perfectly with agreement that $\mathbf{D}$ is a root its characteristic equation (196).

Since $\mathbf{D}$ is real matrix, any of its complex eigen values paired with their complex conjugates:

$$\mathbf{D}Y_k = \lambda_k Y_k \iff \mathbf{D}Y_k^* = \lambda_k^* Y_k^* \tag{220}$$

meanwhile real eigen values not always related. One more important ratio for accelerators: trace of $\mathbf{D}$ is equal zero, e.g. sum of it eigen values is also equal zere:

$$Trace[\mathbf{D}] = Trace\left[\mathbf{U}\Lambda\mathbf{U}^{-1}\right] = Trace\left[\mathbf{U}^{-1}\mathbf{U}\Lambda\right] = Trace[\Lambda] = \sum_{k=1}^{2n}\lambda_k \tag{221}$$

It is especially useful for n=1 – you will see it in your home work.

Another easy case is when D can be diagonalized, even though the number of different eigen values is m < 2n (there is degeneration, i.e. some eigen values have multiplicity >1). We can use again simple Sylvester's formula (202) again, which just has fewer elements (m instead of 2n):

$$\exp[\mathbf{D}s] = \sum_{k=1}^{m} e^{\lambda_k s} \prod_{\lambda_j \neq \lambda_k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \tag{225}$$

Furthermore, in most general case when matrix D cannot be diagonalized (i.e. there is degeneracy, some of eigen values have multiplicity, and D can be only reduced to a Jordan form) we can still write a specific from (generalization of Sylvester's formula):

$$\exp[\mathbf{D}s] = \sum_{k=1}^{m} \left[e^{\lambda_k s} \prod_{i \neq k} \left\{ \frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \sum_{j=0}^{n_k - 1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\}^{n_i} \sum_{p=0}^{n_k - 1} \frac{s^p}{p!} (\mathbf{D} - \lambda_k \mathbf{I})^p \right] \qquad (226)$$

where $n_k < 2n$ is so called height of the eigen value λ_k. Details of the definitions and as well the proof of Sevester's formulae are given in Appendix E.

It is also shown there that n_k can be replaced in (226) by any number $nn > n_k$ – it will add only term, which are zeros, but can make (226) look more uniform. One of the logical choices will be $nn = \max\{n_k\}$. The other natural choice will be $nn = 2n+1-m$, especially if computer does it for you. Eq. (226) is a bit uglier than (202), but still can be used with some elegance.

Appendix E. We had shown that for if 2nx2n matrix $\mathbf{D}$ has 2n unequal eigen values $\lambda_k \neq \lambda_i$,

$$\mathbf{D}Y_k = \lambda_k Y_k;\ \det\left[\mathbf{D}-\lambda_k \mathbf{I}\right]=0 \tag{E1}$$

it can be brought to the diagonal form of

$$\mathbf{D}=\mathbf{U}\Lambda\mathbf{U}^{-1};\Lambda=\begin{bmatrix} \lambda_1 & 0 & 0 \\ ... & ... & ... \\ 0 & \lambda_k & 0 \\ .... & ... & ... \\ 0 & 0 & \lambda_{2n} \end{bmatrix};\mathbf{U}=\left[Y_1,...Y_k.,Y_{2n}\right] \tag{E2}$$

The we proved that a straight-forward Sylveter formula for an arbitrary (to be exact, analytical) functions:

$$f[\mathbf{D}s]=\sum_{k=1}^{2n} f(\lambda_k s)\prod_{j\neq k}\frac{\mathbf{D}-\lambda_j\mathbf{I}}{\lambda_k-\lambda_j}$$
$$\exp[\mathbf{D}s]=\sum_{k=1}^{2n} e^{\lambda_k s}\prod_{j\neq k}\frac{\mathbf{D}-\lambda_j\mathbf{I}}{\lambda_k-\lambda_j} \tag{E3}$$

In practice, there are always cases when eigen values have multiplicity, and denominators in (E3) turn into zeros, e.g. we have a degeneration of this simple form. Another easy case is when D can be diagonalized, even though the number of different eigen values is m < 2n (there is degeneration, i.e. some eigen values have multiplicity >1). We can use again simple Sylvester's formula (E3) again, which just has fewer elements (m instead of 2n):

$$\exp[\mathbf{D}s]=\sum_{k=1}^{m} e^{\lambda_k s}\prod_{\lambda_j\neq\lambda_k}\frac{\mathbf{D}-\lambda_j\mathbf{I}}{\lambda_k-\lambda_j} \tag{E4}$$

But the full consideration requires a bit more work – here we are walking through a general case. An arbitrary matrix **M** can be reduced to an unique matrix, which in general case has a Jordan form: for a matrix with arbitrary height of eigen values the set of eigen values $\{\lambda_1,...,\lambda_m\}$ contains only unique eigen values, i.e. $\lambda_k \neq \lambda_j;\ \forall\ k \neq j$:

$$size[\mathbf{M}] = M;\ \{\lambda_1,.....,\lambda_m\};\ m \leq M;\ \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{U}\mathbf{G}\mathbf{U}^{-1};\ \ \mathbf{G} = \sum_{\oplus \mathbf{k}=\mathbf{1},\mathbf{m}} \mathbf{G}_\mathbf{k} = \mathbf{G}_\mathbf{1} \oplus .... \oplus \mathbf{G}_\mathbf{m};\ \ \sum size[\mathbf{G}_\mathbf{k}] = M \qquad \text{(E5)}$$

where $\oplus$ means direct sum of block-diagonal square matrixes $\mathbf{G_k}$ which correspond to the eigen vector sub-space adjacent to the eigen value λ_k. Size of $\mathbf{G_k}$, which we call l_k, is equal to the multiplicity of the root λ_k of the characteristic equation

$$\det[\lambda \mathbf{I} - \mathbf{M}] = \prod_{k=1,m} (\lambda - \lambda_k)^{l_k}.$$

In general case, $\mathbf{G_k}$ is also a block diagonal matrix comprised of orthogonal sub-spaces belonging to the same eigen value

$$\mathbf{G}_\mathbf{k} = \sum_{\oplus\, j=1,p_k} \mathbf{G}_\mathbf{k}^j = \mathbf{G}_\mathbf{1}^1 \oplus .... \oplus \mathbf{G}_\mathbf{m}^{p_k};\ \ \sum size[\mathbf{G}_\mathbf{k}^j] = l_k \qquad \text{(E6)}$$

where we assume that we sorted the matrixes by increasing size: $size[\mathbf{G}_\mathbf{k}^{j+1}] \geq size[\mathbf{G}_\mathbf{k}^j]$, i.e. the

$$n_\mathbf{k} = size[\mathbf{G}_\mathbf{k}^{p_k}] \leq l_k \qquad \text{(E7)}$$

is the maximum size of the Jordan matrix belonging to the eigen value λ_k. General form of the Jordan matrix is:

$$\mathbf{G}_\mathbf{k}^\mathbf{n} = \begin{bmatrix} \lambda_\mathbf{k} & 1 & 0 & 0 \\ 0 & \lambda_\mathbf{k} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_\mathbf{k} \end{bmatrix} \qquad \text{(E8)}$$

This is obviously includes non-degenerate case when matrix **M** has M independent eigen values and all is just perfectly simple: matrix is reducible to a diagonal one

$$size[\mathbf{M}] = M; \ \{\lambda_1,.....,\lambda_M\}; \ \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{U}\mathbf{G}\mathbf{U}^{-1}; \quad \mathbf{G} = \begin{bmatrix} \lambda_1 & 0 & \\ 0 & ... & \\ & & \lambda_M \end{bmatrix}; \quad \mathbf{U} = [\mathbf{Y}_1, \mathbf{Y}_2, ...\mathbf{Y}_M]; \ \mathbf{M} \cdot \mathbf{Y_k} = \lambda_k \mathbf{Y_k}; \ k = 1,...M \qquad \text{(E9)}$$

An arbitrary analytical matrix function of **M** can be expended into Taylor series and reduced to the function of its Jordan matrix **G** :

$$f(\mathbf{M}) = \sum_{i=1}^{\infty} f_i \mathbf{M}^i = \sum_{i=1}^{\infty} f_i (\mathbf{U}\mathbf{G}\mathbf{U}^{-1})^i \equiv \left(\sum_{i=1}^{\infty} f_i \mathbf{U}(\mathbf{G})^i \mathbf{U}^{-1} \right) = \mathbf{U} \left(\sum_{i=1}^{\infty} f \ \ (\mathbf{G})^i \right) \mathbf{U}^{-1} = \mathbf{U} f(\mathbf{G}) \mathbf{U}^{-1} \quad \text{(E10)}$$

Before embracing complicated things, let's again look at the trivial case, when Jordan matrix is diagonal:

$$f(\mathbf{G}) = \sum_{i=1}^{\infty} f_i \mathbf{G}^i = \sum_{i=1}^{\infty} f_i \begin{bmatrix} \lambda_1 & 0 & \\ 0 & ... & \\ & & \lambda_M \end{bmatrix}^i = \begin{bmatrix} \sum_{i=1}^{\infty} f_i \lambda_1^{\ i} & 0 & \\ 0 & ... & \\ & & \sum_{i=1}^{\infty} f_i \lambda_M^{\ i} \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & ... & \\ & & f(\lambda_M) \end{bmatrix} \text{(E11)}$$

$$f(\mathbf{M}) = \mathbf{U} \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & ... & \\ & & f(\lambda_M) \end{bmatrix} \mathbf{U}^{-1}$$

The last expression can be rewritten as a sum of a product of matrix U containing only specific eigen vector (other columns are zero!) with matrix $\mathbf{U}^{-1}$:

$$f(\mathbf{M}) = [\mathbf{Y}_1 \ldots \mathbf{Y}_\mathbf{k} \ldots \mathbf{Y}_M] \cdot \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & \ldots & \\ & & f(\lambda_M) \end{bmatrix} \mathbf{U}^{-1} = \sum_{k=1}^{M} f(\lambda_k)[0 \ldots . \mathbf{Y}_\mathbf{k} \ldots 0]\mathbf{U}^{-1} \quad \text{(E12)}$$

Still both eigen vector and $\mathbf{U}^{-1}$ in is very complicated (and generally unknown) functions of **M**…. Hmmmmm! We only need to find a matrix operator, which makes projection onto individual eigen vector. Because all eigen values are different, we have a very clever and simple way of designing projection operators. Operator

$$\mathbf{P}_k^i = \frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \quad \text{(E13)}$$

has two important properties: it is unit operator for $\mathbf{Y_i}$, it is zero operator for $\mathbf{Y_k}$ and multiply the rest of them by a constant:

$$\mathbf{P}_k^i \mathbf{Y}_k = \frac{\mathbf{M} \cdot \mathbf{Y}_k - \lambda_k \mathbf{I} \cdot \mathbf{Y}_k}{\lambda_i - \lambda_k} = \frac{\lambda_k - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_k \equiv 0;$$

$$\mathbf{P}_k^i \mathbf{Y}_i = \frac{\mathbf{M} \cdot \mathbf{Y}_i - \lambda_k \mathbf{I} \cdot \mathbf{Y}_i}{\lambda_i - \lambda_k} = \frac{\lambda_i - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_i \equiv \mathbf{Y}_i; \quad \text{(E14)}$$

$$\mathbf{P}_k^i \mathbf{Y}_j = \frac{\mathbf{M} \cdot \mathbf{Y}_j - \lambda_k \mathbf{I} \cdot \mathbf{Y}_j}{\lambda_i - \lambda_k} = \frac{\lambda_j - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_j$$

I.e. it project U into a subspace orthogonal to $\mathbf{Y_k}$. We should note the most important quality of this operator: it comprises of known matrixes: **M** and unit one. Also, zero operators for two eigen vectors commute with each other – being combination of **M** and **I** makes it obvious. Constructing unit projection operator $\mathbf{Y_i}$ which is also zero for remaining eigen vectors is straight forward from here: it is a product of all M-1 projection operators

$$\mathbf{P}^i_{unit} = \prod_{k\neq i} \mathbf{P}^i_k = \prod_{k\neq i}\left(\frac{\mathbf{M}-\lambda_k \mathbf{I}}{\lambda_i - \lambda_k}\right); \qquad \text{(E15)}$$

$$\mathbf{P}^i_{unit}\mathbf{Y}_j = \delta^i_j \mathbf{Y}_j = \begin{cases} \mathbf{Y}_i, & j = i \\ \mathbf{O}, & j \neq i \end{cases}$$

Observation that

$$\mathbf{P}^k_{unit}\mathbf{U} = \mathbf{P}^k_{unit}\left[\mathbf{Y}_1 \ldots \mathbf{Y_k} \ldots \mathbf{Y}_M\right] = [0 \ldots . \mathbf{Y_k} \ldots 0] \qquad \text{(E16)}$$

allows us to rewrite eq. (E12) in the form which is easy to use:

$$f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)[0 \ldots . \mathbf{Y_k} \ldots 0]\mathbf{U}^{-1} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}^k_{unit}\mathbf{U}\cdot\mathbf{U}^{-1} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}^k_{unit}; \qquad \text{(E17)}$$

which with (E15) give final form of Sylvester formula (for non-degenerated matrixes):

$$f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)\prod_{i\neq k}\left(\frac{\mathbf{M}-\lambda_i \mathbf{I}}{\lambda_k - \lambda_i}\right); \qquad \text{(E18)}$$

One can see that this is a polynomial of power M-1 of matrix M, as we expected from the theorem of Jordan and Kelly that matrix is a root of its characteristic equation:

$$g(\lambda) = \det[\mathbf{M} - \lambda\mathbf{I}];\ \ g(\mathbf{M}) \equiv 0; \qquad \text{(E19)}$$

which is polynomial of power M. It means that any polynomial of higher order of matrix **M** can reduced to M-1 order. Equation (E18) gives specific answer how it can be done for the arbitrary series.

If matrix **M** is reducible to diagonal form, where some eigen values have multiplicity, we need to sum only by independent eigen values:

$$f(\mathbf{M}) = \sum_{k=1}^{m} f(\lambda_k) \prod_{\lambda_i \neq \lambda_k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right); \qquad \text{(E18-red)}$$

and it has maximum power of **M** of m-1. Prove it trivial using the above.

Let's return to most general case of Jordan blocks, i.e. a degenerated case when eigen values have non-unit multiplicity. For a general form of the Jordan matrix we can only say that it is direct sum of the function of the Jordan blocks:

$$f(\mathbf{G})=\sum_{i=0}^{\infty} f_i \mathbf{G}^i=\sum_{i=0}^{\infty} f_i \begin{bmatrix} \mathbf{G}_1^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_{\mathbf{m}}^{\mathbf{p_m}} \end{bmatrix}^i = \begin{bmatrix} \sum_{i=0}^{\infty} f_i \left(\mathbf{G}_1^1\right)^i & 0 & \\ 0 & \dots & \\ & & \sum_{i=0}^{\infty} f_i \left(\mathbf{G}_{\mathbf{m}}^{\mathbf{p_m}}\right)^i \end{bmatrix}$$

$$= \begin{bmatrix} f(\mathbf{G}_1^1) & 0 & \\ 0 & \dots & \\ & & f(\mathbf{G}_{\mathbf{m}}^{\mathbf{p_m}}) \end{bmatrix} = \sum_{\oplus k=1,m,\; j=1,p_k} f\left(\mathbf{G}_{\mathbf{k}}^{j}\right) = f\left(\mathbf{G}_1^1\right) \oplus \dots \oplus f\left(\mathbf{G}_{\mathbf{m}}^{p_m}\right); \qquad \text{(E20)}$$

Function of a Jordan block of size n contains not only the function of corresponding eigen value λ, but also its derivatives to $(n\text{-}1)^{th}$ order:

$$\mathbf{G}=\begin{bmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & \lambda \end{bmatrix};\; f(\mathbf{G})=\begin{bmatrix} f(\lambda) & f'(\lambda)/1! & \dots f^{(k)}(\lambda)/k! & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & \dots & f^{(n-2)}(\lambda)/(n-2)! \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f'(\lambda)/1! \\ 0 & 0 & \dots & f(\lambda) \end{bmatrix} \qquad \text{(E21)}$$

The prove of Eq. 21 is your take-home task – use polynomial as a function.

We are half-way through. There is sub-space of eigen vectors $\mathcal{Y}_k^n$ which corresponds to the eigen value λ_k and the block $\mathbf{G}_\mathbf{k}^\mathbf{n}$:

$$\mathcal{Y}_k^n \in \left\{\mathbf{Y}_\mathbf{k}^{\mathbf{n,1}},....,\mathbf{Y}_\mathbf{k}^{\mathbf{n,q}}\right\};\; q = size\left(\mathbf{G}_\mathbf{k}^\mathbf{n}\right) \tag{E22}$$

$$\mathbf{M}\cdot\mathbf{Y}_\mathbf{k}^{\mathbf{n},1} = \lambda_k \mathbf{Y}_\mathbf{k}^{\mathbf{n},1};\;\; \mathbf{M}\cdot\mathbf{Y}_\mathbf{k}^{\mathbf{n},l} = \lambda_k \mathbf{Y}_\mathbf{k}^{\mathbf{n},l} + \mathbf{Y}_\mathbf{k}^{\mathbf{n},l-1};\; 1 < l \le q \tag{E23}$$

It is obvious from equation (E21) that projection operator (E15) will not be zero operator for $\mathcal{Y}_k^n$, and it also will not be unit operator for $\mathcal{Y}_i^n$. Now, let's look on how we can project on individual sub-spaces, eigen vectors, including zero-operator for specific sub-spaces. Just step by step (from eq. (E6) and (E21):

$$f(\mathbf{M}) = \mathbf{U}f(\mathbf{G})\mathbf{U}^{-1}$$

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\begin{bmatrix} \underbrace{0}_{\lambda_1} & \underbrace{0}_{\lambda_2} & \cdots & \underbrace{0}_{\lambda_{k-1}} & \underbrace{A_k^i}_{\lambda_k} & ...0... & \underbrace{0}_{\lambda_m} \end{bmatrix} \tag{E24}$$

$$A_k^i = \begin{bmatrix} \underbrace{B_1^{i\,k} \quad .... \quad B_{p_k}^{i\,k}}_{\lambda_k} \end{bmatrix}; B_n^{i\,k} = \begin{bmatrix} \underbrace{0.....0}_{i\ collumns} & Y_k^{n,1} & Y_k^{n,q_n-i} \end{bmatrix} \tag{E25}$$

i.e.

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\begin{bmatrix} \underbrace{0}_{\lambda_1} & \underbrace{0}_{\lambda_2} & \cdots & \underbrace{0}_{\lambda_{k-1}} & \underbrace{\begin{bmatrix} \underbrace{\cdots}_{1} & \cdots & \underbrace{\underbrace{0.....0}_{i\ collumns} \; Y_k^{n,1} \; Y_k^{n,q_n-i}}_{n-th} & \underbrace{\cdots}_{p_k} \end{bmatrix}}_{\lambda_k} & ...0... & \underbrace{0}_{\lambda_m} \end{bmatrix} \tag{E26}$$

From (E23) we get a shifting operator for eigen value λ_k:

$$\begin{aligned}
&\left[\mathbf{M}-\lambda_k\mathbf{I}\right]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{k},q}=0;\ \left[\mathbf{M}-\lambda_k\mathbf{I}\right]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},j}=\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},j-1};\ 1<j\le q\\
&U_0^{\mathbf{n}\,\mathbf{k}}=\left[\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q}\right];\\
&\left[\mathbf{M}-\lambda_k\mathbf{I}\right]\cdot U_0^{\mathbf{n}\,\mathbf{k}}=U_1^{\mathbf{n}\,\mathbf{k}}=\left[0,\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-1}\right]\\
&\ldots\ldots\\
&\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^j\cdot U_0^{\mathbf{n}\,\mathbf{k}}=U_j^{\mathbf{n}\,\mathbf{k}}=\left[\underbrace{0..0}_{j\ zeros},\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-j}\right]\\
&\ldots..\\
&\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^q\cdot U_0^{\mathbf{n}\,\mathbf{k}}=0
\end{aligned}\tag{E27}$$

$$\mathbf{U}f\left(\mathbf{G}\right)=\sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}\left(\lambda_k\right)}{i!}\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^i\left[\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\ldots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{\left[U^{k\,1}\quad\ldots\quad\underbrace{U^{k\,n}}_{n-th}\quad U^{k\,p_k}\right]}_{\lambda_k}\quad\ldots0\ldots\quad\underbrace{0}_{\lambda_m}\right]\tag{E28}$$

i.e. we collected all eigen vectors belonging to the eigen value λ_k. Now we need a non-distorting projection operator on the sub-space of λ_k.

First, let's find zero operator for subspace of λ_i: it is obvious

$$O_i = \left[\mathbf{M} - \lambda_i \mathbf{I}\right]^{n_i} \Rightarrow \left[\mathbf{M} - \lambda_i \mathbf{I}\right]^{n_i} U_o^{r\,i} = \left[\mathbf{M} - \lambda_i \mathbf{I}\right]^{n_i} \left[\mathbf{Y}_{\mathbf{k}}^{r,1} \ldots \mathbf{Y}_{\mathbf{k}}^{r,l} \ldots \mathbf{Y}_{\mathbf{k}}^{r,q}\right] = 0;$$

$$T_k = \prod_{i \neq k} \frac{O_i}{\left(\lambda_k - \lambda_i\right)^{n_i}} = \prod_{i \neq k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i}\right)^{n_i}$$

T_k is projection operator of sub-space of λ_k, but it is not unit one! To correct that we need an operator, which we create as follow using shift operator $T = \mathbf{M} - \lambda_k \mathbf{I};$:

$$R = \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i};\; T = \mathbf{M} - \lambda_k \mathbf{I};\; \alpha = \alpha_{k,i} = 1/(\lambda_k - \lambda_i)$$

$$RU_0 = U_0 + \alpha U_1 \qquad U_0 = U_0$$

.....

$$RU_{q-2} = U_{q-2} + \alpha U_{q-1} \qquad U_{q-2} = T^{q-2} U_0$$

$$RU_{q-1} = U_{q-1} \qquad U_{q-1} = T^{q-1} U_0$$

$$Q = \alpha T$$

$$U_{q-1} = RU_{q-1} = RT^{q-1} U_0$$

$$U_{q-2} = R\left(I + Q\right) U_{q-2} = RQT^{q-2} U_0$$

.....

.....

$$U_0 = R\left(\sum_j^{q-1} Q^j\right) U_0$$

so, we get it:

]

$$P_k^i = \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i}\left(\mathbf{I} + \sum_{j=1}^{n_k - 1}\left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k}\right)^j\right) \qquad \text{(E30)}$$

The final stroke is:

$$P_k = \prod_{i\neq k}\left(P_k^i\right)^{n_i} = \prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i} \qquad \text{(E31)}$$

and

$$f(\mathbf{M}) = \sum_{k=1}^{m}\left[\prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i}\sum_{i=0}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^i\right] \qquad \text{(E32)}$$

This is most general expression for any matrix function with $f^{(m)}(\lambda_k) \equiv \left.\frac{\partial^m f(\lambda)}{\partial\lambda^m}\right|_{\lambda=\lambda_k}$.

Note that we are using s as a variable which generates polynomials:

$$f(\mathbf{M}\cdot s) = \sum_{k=1}^{m}\left[\prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i}\sum_{i=0}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^i s^i\right] \qquad \text{(E33)}$$

with eigen values of $\det(\mathbf{M}-\lambda_i\mathbf{I})=0$ to be found.

Note that we are using s as a variable which generates polynomials:

$$f(\mathbf{M}\cdot s)=\sum_{k=1}^{m}\left[\prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}[\mathbf{M}-\lambda_k\mathbf{I}]^i s^i\right] \qquad \text{(E33)}$$

with eigen values of $\det(\mathbf{M}-\lambda_i\mathbf{I})=0$ to be found.

Furthermore, in most general case when matrix D cannot be diagonalized (i.e. there is degeneracy, some of eigen values have multiplicity, and D can be only reduced to a Jordan form) we can still write a specific from (generalization of Sylvester's formula):

$$\exp[\mathbf{D}s]=\sum_{k=1}^{m}\left[e^{\lambda_k s}\prod_{i\neq k}\left\{\frac{\mathbf{D}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\sum_{j=0}^{n_k-1}\left(\frac{\mathbf{D}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right\}^{n_i}\sum_{p=0}^{n_k-1}\frac{s^p}{p!}(\mathbf{D}-\lambda_k\mathbf{I})^p\right] \qquad \text{(E34)}$$

where $n_k < 2n$ is height of the eigen value λ_k. It is also shown there that n_k can be replaced in (E34) by any number $nn > n_k$ – it will add only term, which are zeros, but can make (E34) look more uniform. One of the logical choices will be $nn = \max\{n_k\}$. The other natural choice will be nn =2n+1–m, especially if computer does it for you. Eq. (E34) is a bit uglier than (E3), but still can be used with some elegance.

Eigen values split into pairs with the opposite sign because it is a Hamiltonian system:

$$\begin{aligned}&\det[\mathbf{SH}-\lambda\cdot\mathbf{I}]=\det[\mathbf{SH}-\lambda\cdot\mathbf{I}]^T=\det[-\mathbf{HS}-\lambda\cdot\mathbf{I}]=\\&(-1)^{2n}\det[\mathbf{H}S+\lambda\cdot\mathbf{I}]=\det\left(\mathbf{S}^{-1}[\mathbf{HS}+\lambda\cdot\mathbf{I}]\mathbf{S}\right)=\det[\mathbf{SH}+\lambda\cdot\mathbf{I}]\#\end{aligned}. \quad \text{(L3-35)}$$

First, it makes finding eigen values a easier problem, because characteristic equation is bi-quadratic:

$$\det[\mathbf{D}-\lambda I]=\prod(\lambda_i-\lambda)(-\lambda_i-\lambda)=\prod\left(\lambda^2-\lambda_i^{\,2}\right)=0. \quad \text{(L3-35-1)}$$

For accelerator elements it is of paramount importance, 1D case is reduces to trivial (L3-37), 2D case is reduced to solution of quadratic equation and 3D case (6D phase space) required to solve cubic equation. For analytical work it gives analytical expressions – compare it with attempt to write analytical formula for roots of a generic polynomial of 6-order? It simply does not exist! Thus, we have an extra gift for accelerator physics – the roots can be written and studied! It is also allow us to simplify (202) into

$$\begin{aligned}&\exp[\mathbf{D}s]=\left\{\sum_{k=1}^{n}e^{\lambda_k s}\frac{\mathbf{D}+\lambda_k\mathbf{I}}{2\lambda_k}\prod_{j\neq k}\left(\frac{\mathbf{D}^2-\lambda_j^{\,2}\mathbf{I}}{\lambda_k^{\,2}-\lambda_j^{\,2}}\right)-e^{-\lambda_k s}\frac{\mathbf{D}-\lambda_k\mathbf{I}}{2\lambda_k}\prod_{j\neq k}\left(\frac{\mathbf{D}^2-\lambda_j^{\,2}\mathbf{I}}{\lambda_k^{\,2}-\lambda_j^{\,2}}\right)\right\}\\&\exp[\mathbf{D}s]=\sum_{k=1}^{n}\left(\frac{e^{\lambda_k s}+e^{-\lambda_k s}}{2}\mathbf{I}+\frac{e^{\lambda_k s}-e^{-\lambda_k s}}{2\lambda_k}\mathbf{D}\right)\prod_{j\neq k}\left(\frac{\mathbf{D}^2-\lambda_j^{\,2}\mathbf{I}}{\lambda_k^{\,2}-\lambda_j^{\,2}}\right)\end{aligned} \quad \text{(L3-36)}$$

where index k goes only through n pairs of $\{\lambda_k,-\lambda_k\}$. While (L3-36) does not look simpler, it really makes it easier (4 times less calculations) when we do it by hands… For example we can look at 1D case. First, we can easily see that

$$\lambda_1=-\lambda_2=\lambda;\quad \lambda^2=-\det[\mathbf{D}] \quad \text{(L3-37)}$$

Thus, it is non-degenerated case only when $\det[D] \neq 0$. (202) give us a simple two-piece expression :

$$\exp[\mathbf{D}s] = e^{\lambda s}\frac{\mathbf{D}-\lambda\mathbf{I}}{2\lambda} - e^{-\lambda s}\frac{\mathbf{D}+\lambda \mathrm{I}}{2\lambda} \qquad \text{(L3-38)}$$

while (L3-36) bring it home right away:

$$\exp[\mathbf{D}s] = \mathbf{I}\cdot\frac{e^{\lambda s}+e^{-\lambda s}}{2} + \mathbf{D}\frac{e^{\lambda s}-e^{-\lambda s}}{2\lambda};$$

$$\exp[\mathbf{D}s] = \mathbf{I}\cdot\cosh|\lambda|s + \frac{\mathbf{D}\sinh|\lambda|s}{|\lambda|}; \quad \det[\mathbf{D}]<0; \quad |\lambda| = \sqrt{-\det[\mathbf{D}]} \qquad \text{(L3-39)}$$

$$\exp[\mathbf{D}s] = \mathbf{I}\cdot\cos|\lambda|s + \frac{\mathbf{D}\sin|\lambda|s}{|\lambda|}; \qquad \det[\mathbf{D}]>0; \quad |\lambda| = \sqrt{\det[\mathbf{D}]}$$

The case $\det[\mathbf{D}]=0$ means in this case that D is nilpotent: eqs (195-25) look like follows

$$\det\mathbf{D}=0 \Rightarrow \lambda_1 = -\lambda_2 = 0; \; d(\lambda) = \det[\mathbf{D}-\lambda I] = (\lambda_1-\lambda)(-\lambda_1-\lambda) = \lambda^2 \;\; \Rightarrow \mathbf{D}^2 = 0 \qquad \text{(L3-40)}$$

hence

$$\exp[\mathbf{D}s] = \mathbf{I}+\mathbf{D}s; \quad \det[\mathbf{D}]=0; \qquad \text{(L3-41)}$$

Naturally, (L3-41) is result of full-blown degenerated case – eq. (L3-33), but it also can be obtained as a limit case of (L3-39) when $|\lambda| \to 0$.

What we learned today?

- Linear ordinarary equations with constant coefificents (D-matrix) have a natural solution as $exp(D \cdot s)$
- We can use functions of matricies and built entire method have analytical expression of matrix function as soon as we know eigen values of matrix D
- Matrix function have a very simple and elegant form – callse Sylveter formula- when eigen values are unique (e.g. in non-degenrating case) and *D* can be diagonalized
- But even in a most general case, we can write analytical expression for matrix function
- In linear Hamiltonian case, eigne values split in pair of $(\lambda,-\lambda)$ and the expression can be even further simplified
- The remaining task for linear matrices if accelerators is to find analytical expression for eigen values – the job for next class

Appendix F: Inhomogeneous solution

Even though calculations are tedious, they are also transparent and straightforward. General expression for the inhomogeneous equation of 2n ordinary linear differential equations is found by a standard trick of variable constants (method developed by Lagrange), i.e. assuming that $R = \mathbf{M}(s)A(s)$:

$$\begin{aligned} &\frac{dR}{ds} = R' = \mathbf{D}\cdot R + \mathbf{C}; \quad \mathbf{M}' = \mathbf{D}\mathbf{M}; \\ &R = \mathbf{M}(s)A(s) \Rightarrow \mathbf{M}'A + \mathbf{M}A' = \mathbf{D}\mathbf{M}A + C \\ &R(0) = 0 \Rightarrow A_o = 0 \\ &A' = \mathbf{M}^{-1}(s)C \Rightarrow A = \int_0^s \mathbf{M}^{-1}(z)C dz = \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot C \end{aligned} \tag{F-1}$$

with well known result of:

$$R = e^{\mathbf{D}s}\left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot C . \tag{F-2}$$

or

$$R = M_{4x4}(s)\left(\int_0^s M^{-1}{}_{4x4}(z) dz \right) \cdot C . \tag{F-3}$$

If you use computer, eq. (F-3) is one to use. For analytical folks, you should go though a tedious job is combining all terms together into final form:

$$R(s) = \sum_{k=1}^{m} \left\{ \prod_{i\neq k} \left[\frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right] \sum_{j=0}^{n_k - 1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\} \sum_{n=0}^{n_k - 1} (\mathbf{D} - \lambda_k \mathbf{I})^n \frac{s^n}{n!} \cdot \sum_{p=0}^{n_k - 1} (-1)^{p+1} (\mathbf{D} - \lambda_k \mathbf{I})^p \cdot \mathbf{C} \cdot \left[\sum_{q=0}^{p1} \frac{s^{p-q}}{(p-q)!\,\lambda_k{}^{q+1}} - \frac{e^{\lambda_k}}{\lambda_k{}^{p+1}} \right] \tag{F-4}$$

Proof of eq. (E-21):

$$\mathbf{G}^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & & 1 \end{bmatrix}; \ \mathbf{G}^1 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & ... & 1 \\ 0 & 0 & & \lambda \end{bmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & ... & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & ... & 1 \\ 0 & 0 & & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^2 & \lambda & 1... & 0 \\ 0 & \lambda^2 & \lambda... & 1 \\ 0 & 0 & ... & \lambda \\ 0 & 0 & 0 & \lambda^2 \end{bmatrix}$$

Induction:

$$\mathbf{G}^n == \begin{vmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-1}/2! \ ... & ... \\ 0 & \lambda^n & n\lambda^{n-1}/1! & ... \\ 0 & 0 & ... & n\lambda^{n-1}/1! \\ 0 & 0 & 0 & \lambda^n \end{vmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & ... & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-1}/2! \ ... & ... \\ 0 & \lambda^n & n\lambda^{n-1}/1! & ... \\ 0 & 0 & ... & n\lambda^{n-1}/1! \\ 0 & 0 & 0 & \lambda^n \end{bmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (n+1)\lambda^n/1! & (n(n-1)+2n)\lambda^{n-1}/2! \ ... & ... \\ 0 & \lambda^{n+1} & (n+1)\lambda^n/1! & ... \\ 0 & 0 & ... & (n+1)\lambda^n/1! \\ 0 & 0 & 0 & \lambda^{n+1} \end{bmatrix}$$

$$\begin{vmatrix} \lambda^n & C_1^n\lambda^{n-1} & C_2^n\lambda^{n-2} & ... & C_k^n\lambda^{n-k} & C_{k+1}^n\lambda^{n-k-1} & .. \\ 0 & \lambda^n & C_1^n\lambda^{n-1} & ... & C_{k-1}^n\lambda^{n+1-k} & C_k^n\lambda^{n-k} & \\ .... & .... & .... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{vmatrix}$$

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & ... & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{vmatrix} \lambda^n & C_1^n\lambda^{n-1} & C_2^n\lambda^{n-2} & ... & C_k^n\lambda^{n-k} & C_{k+1}^n\lambda^{n-k-1} & .. \\ 0 & \lambda^n & C_1^n\lambda^{n-1} & ... & C_{k-1}^n\lambda^{n+1-k} & C_k^n\lambda^{n-k} & \\ .... & .... & .... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{vmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (C_1^n+1)\lambda^n & (C_2^n+C_1^n)\lambda^{n-2} & ... & (C_k^n+C_{k-1}^n)\lambda^{n-k+1} & (C_{k+1}^n+C_k^n)\lambda^{n-k} & .. \\ 0 & \lambda^{n+1} & (C_1^n+1)\lambda^n & ... & .... & (C_k^n+C_{k-1}^n)\lambda^{n-k+1} & \\ .... & .... & .... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^{n+1} \end{bmatrix}$$

polynomial coefficients: $C_k^{n+1} = C_k^n + C_{k-1}^n$; $C_k^n = n!/k!/(n-k)!$ proves the point.

Hence, we can now calculate a polynomial functions or any function expandable into a Taylor series:

$$f(\mathbf{G}) = \sum_{n=0}^{\infty} f_n \mathbf{G}^n = \sum_{n=0}^{\infty} f_n \begin{bmatrix} \lambda^n & C_1^n \lambda^{n-1} \ldots C_k^n \lambda^{n-k} \ldots & \ldots \\ 0 & \ldots & \ldots \\ 0 & \ldots & \lambda^n \end{bmatrix}^i =$$

$$\begin{bmatrix} \sum_{n=0}^{\infty} f_n \lambda^n \ldots & \sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k} & \ldots \\ 0 & \ldots & \ldots \\ 0 & 0 & \sum_{n=0}^{\infty} f_n \lambda^n \end{bmatrix}$$

The final stroke is noting that

$$\sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k} = \sum_{n=0}^{\infty} f_n \cdot \frac{n!\,\lambda^{n-k}}{k!\cdot(n-k)!} = \frac{1}{k!} \sum_{n=0}^{\infty} f_n \cdot \frac{n!\,\lambda^{n-k}}{(n-k)!} = \frac{1}{k!} \sum_{n=0}^{\infty} f_n \cdot \lambda^{n-k} \prod_{j=0}^{k-1} (n-j)$$

$$= \frac{1}{k!} \frac{d^k}{d\lambda^k} \sum_{n=0}^{\infty} f_n \cdot \lambda^n = \frac{1}{k!} \frac{d^k f}{d\lambda^k} \quad \#$$

Good HW exercise.