PHY 564
Advanced Accelerator Physics
Lecture 11

Vladimir N. Litvinenko
Yichao Jing
Gang Wang

Department of Physics & Astronomy, Stony Brook University
Collider-Accelerator Department, Brookhaven National Laboratory

Lecture 15. Matrix functions and projection operators - continued

In last class we had shown that for if 2nx2n matrix **D** has 2n unequal eigen values $\lambda_k \neq \lambda_i$,

$$\mathbf{D}Y_k = \lambda_k Y_k;\ \det\left[\mathbf{D} - \lambda_k \mathbf{I}\right] = 0 \tag{1}$$

it can be brought to the diagonal form of

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

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

$$f[\mathbf{D}s] = \sum_{k=1}^{2n} f(\lambda_k s) \prod_{j\neq k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j}$$
$$\exp[\mathbf{D}s] = \sum_{k=1}^{2n} e^{\lambda_k s} \prod_{j\neq k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \tag{3}$$

In practice, there are always cases when eigen values have multiplicity, and denominators in (3) 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 (3) again, which just has fewer elements (m instead of 2n):

$$\exp[\mathbf{D}s] = \sum_{k=1}^{m} e^{\lambda_k s} \prod_{\lambda_j \neq \lambda_k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \tag{4}$$

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. $\lambda_k \neq \lambda_j;\ \forall\ \ k \neq j$:

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

$$\mathbf{M} = \mathbf{UGU}^{-1};\ \ \mathbf{G} = \sum_{\oplus \mathbf{k=1,m}} \mathbf{G_k} = \mathbf{G_1} \oplus \mathbf{....} \oplus \mathbf{G_m};\ \ \sum size[\mathbf{G_k}] = M \tag{5}$$

where ⊕ 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

$$\mathbf{G_k} = \sum_{\oplus j=1,p_k} \mathbf{G_k^{\mathit{j}}} = \mathbf{G_1^1} \oplus \mathbf{....} \oplus \mathbf{G_m^{\mathit{p_k}}}; \quad \sum size[\mathbf{G_k^{\mathit{j}}}] = l_k \tag{6}$$

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

$$n_\mathbf{k} = size[\mathbf{G_k^{\mathit{p_k}}}] \leq l_k \tag{7}$$

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

$$\mathbf{G_k^n} = \begin{bmatrix} \lambda_\mathbf{k} & 1 & 0 & 0 \\ 0 & \lambda_\mathbf{k} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_\mathbf{k} \end{bmatrix} \tag{8}$$

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

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

$$\mathbf{M} = \mathbf{UGU^{-1}}; \quad \mathbf{G} = \begin{bmatrix} \lambda_1 & 0 & \\ 0 & ... & \\ & & \lambda_M \end{bmatrix}; \;\; \mathbf{U} = [\mathbf{Y}_1, \mathbf{Y}_2, ...\mathbf{Y}_M]; \;\; \mathbf{M} \cdot \mathbf{Y_k} = \lambda_k \mathbf{Y_k}; \;\; k = 1,...M \qquad (9)$$

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

$$f(\mathbf{M}) = \sum_{i=1}^{\infty} f_i \mathbf{M}^i = \sum_{i=1}^{\infty} f_i \left(\mathbf{UGU^{-1}}\right)^i \equiv \left(\sum_{i=1}^{\infty} f_i \mathbf{U}(\mathbf{G})^i \mathbf{U^{-1}}\right) = \mathbf{U}\left(\sum_{i=1}^{\infty} f \;\; (\mathbf{G})^i\right)\mathbf{U^{-1}} = \mathbf{U} f(\mathbf{G}) \mathbf{U^{-1}} \qquad (10)$$

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

$$f(\mathbf{G}) = \sum_{i=1}^{\infty} f_i \mathbf{G}^i = \sum_{i=1}^{\infty} f_i \begin{bmatrix} \lambda_1 & 0 & \\ 0 & \dots & \\ & & \lambda_M \end{bmatrix}^i = \begin{bmatrix} \sum_{i=1}^{\infty} f_i \lambda_1^{\,i} & 0 & \\ 0 & \dots & \\ & & \sum_{i=1}^{\infty} f_i \lambda_M^{\,i} \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & \dots & \\ & & f(\lambda_M) \end{bmatrix} \quad (11)$$

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

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 $\mathbf{U}^{-1}$:

$$f(\mathbf{M}) = [\mathbf{Y}_1 \dots \mathbf{Y}_\mathbf{k} \dots \mathbf{Y}_M] \cdot \begin{vmatrix} f(\lambda_1) & 0 & \\ 0 & \dots & \\ & & f(\lambda_M) \end{vmatrix} \mathbf{U}^{-1} = \sum_{k=1}^{M} f(\lambda_k)[0 \dots \mathbf{Y}_\mathbf{k} \dots 0]\mathbf{U}^{-1} \quad (12)$$

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

$$\mathbf{P}_k^i = \frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \tag{13}$$

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:

$$\begin{aligned}
\mathbf{P}_k^i \mathbf{Y}_k &= \frac{\mathbf{M}\cdot\mathbf{Y}_k - \lambda_k \mathbf{I}\cdot\mathbf{Y}_k}{\lambda_i - \lambda_k} = \frac{\lambda_k - \lambda_k}{\lambda_i - \lambda_k}\mathbf{Y}_k \equiv 0; \\
\mathbf{P}_k^i \mathbf{Y}_i &= \frac{\mathbf{M}\cdot\mathbf{Y}_i - \lambda_k \mathbf{I}\cdot\mathbf{Y}_i}{\lambda_i - \lambda_k} = \frac{\lambda_i - \lambda_k}{\lambda_i - \lambda_k}\mathbf{Y}_i \equiv \mathbf{Y}_i; \\
\mathbf{P}_k^i \mathbf{Y}_j &= \frac{\mathbf{M}\cdot\mathbf{Y}_j - \lambda_k \mathbf{I}\cdot\mathbf{Y}_j}{\lambda_i - \lambda_k} = \frac{\lambda_j - \lambda_k}{\lambda_i - \lambda_k}\mathbf{Y}_j
\end{aligned} \tag{14}$$

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

$$\mathbf{P}^i_{unit} = \prod_{k\neq i} \mathbf{P}^i_k = \prod_{k\neq i}\left(\frac{\mathbf{M}-\lambda_k \mathbf{I}}{\lambda_i-\lambda_k}\right);$$
$$\mathbf{P}^i_{unit}\mathbf{Y}_j = \delta^i_j \mathbf{Y}_j = \left\{\begin{matrix} \mathbf{Y}_i, & j=i \\ \mathbf{O}, & j\neq i \end{matrix}\right\} \qquad (15)$$

Observation that

$$\mathbf{P}^k_{unit}\mathbf{U} = \mathbf{P}^k_{unit}\left[\mathbf{Y}_1 \ldots \mathbf{Y_k} \ldots \mathbf{Y}_M\right] = [0 \ldots. \mathbf{Y_k} \ldots 0] \qquad (16)$$

allows us to rewrite eq. (12) in the form which is easy to use:

$$f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)[0 \ldots. \mathbf{Y_k} \ldots 0]\mathbf{U^{-1}} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}^k_{unit}\mathbf{U}\cdot\mathbf{U^{-1}} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}^k_{unit}; \qquad (17)$$

which with (15) give final form of Sylvester formula (for non-degenerated matrixes):

$$f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)\prod_{i\neq k}\left(\frac{\mathbf{M}-\lambda_i \mathbf{I}}{\lambda_k-\lambda_i}\right); \qquad (18)$$

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:

$$g(\lambda) = \det[\mathbf{M} - \lambda\mathbf{I}];\;\; g(\mathbf{M}) \equiv 0; \tag{19}$$

which is polynomial of power M. It means that any polynomial of higher order of matrix **M** can reduced to M-1 order. Equation (18) 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:

$$f(\mathbf{M}) = \sum_{k=1}^{m} f(\lambda_k) \prod_{\lambda_i \neq \lambda_k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right); \tag{18-red}$$

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:

$$f(\mathbf{G}) = \sum_{i=0}^{\infty} f_i \mathbf{G}^i = \sum_{i=0}^{\infty} f_i \begin{bmatrix} \mathbf{G_1^1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{...} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{...} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G_m^{p_m}} \end{bmatrix}^i = \begin{bmatrix} \sum_{i=0}^{\infty} f_i \left(\mathbf{G_1^1}\right)^i & 0 & \\ 0 & ... & \\ & & \sum_{i=0}^{\infty} f_i \left(\mathbf{G_m^{p_m}}\right)^i \end{bmatrix}$$

$$= \begin{bmatrix} f(\mathbf{G_1^1}) & 0 & \\ 0 & ... & \\ & & f(\mathbf{G_m^{p_m}}) \end{bmatrix} == \sum_{\oplus k=1,m,\quad j=1,p_k} f\left(\mathbf{G_k^j}\right) = f\left(\mathbf{G_1^1}\right) \oplus \mathbf{....} \oplus f\left(\mathbf{G_m^{p_m}}\right); \qquad (20)$$

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

$$\mathbf{G} = \begin{bmatrix} \lambda & 1 & ... & 0 \\ 0 & \lambda & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \\ 0 & 0 & ... & \lambda \end{bmatrix}; \quad f(\mathbf{G}) = \begin{bmatrix} f(\lambda) & f'(\lambda)/1! & ...f^{(k)}(\lambda)/k! & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & ... & f^{(n-2)}(\lambda)/(n-2)! \\ ... & .... & ... & ... \\ 0 & 0 & ... & f'(\lambda)/1! \\ 0 & 0 & ... & f(\lambda) \end{bmatrix} \qquad (21)$$

The prove of Eq. 21 is your home-work for today. 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 $\mathbf{G_k^n}$:

$$\mathcal{Y}_k^n \in \left\{\mathbf{Y_k^{n,1}},\ldots,\mathbf{Y_k^{n,q}}\right\}; \quad q = size\left(\mathbf{G_k^n}\right) \tag{22}$$

$$\mathbf{M}\cdot\mathbf{Y_k^{n,1}} = \lambda_k \mathbf{Y_k^{n,1}}; \quad \mathbf{M}\cdot\mathbf{Y_k^{n,\mathit{l}}} = \lambda_k \mathbf{Y_k^{n,\mathit{l}}} + \mathbf{Y_k^{n,\mathit{l}-1}}; \;\; 1 < l \le q \tag{23}$$

It is obvious from equation (21) that projection operator (15) 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. (6) and (21):

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

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\cdots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{A_k^i}_{\lambda_k}\quad\ldots 0\ldots\quad\underbrace{0}_{\lambda_m}\right] \tag{24}$$

$$A_k^i = \left[\underbrace{B_1^{i\ \ k}\quad\ldots.\quad B_{p_k}^{i\ \ k}}_{\lambda_k}\right]; B_n^{i\ \ k} = \left[\underbrace{0\ldots..0}_{i\ \ collumns}\quad Y_k^{n,1}\quad Y_k^{n,q_n-1}\right] \tag{25}$$

i.e.

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\cdots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{\left[\underbrace{\ldots}_{1}\quad\cdots\quad\underbrace{\underbrace{0\ldots..0}_{i\ \ collumns}\quad Y_k^{n,1}\quad Y_k^{n,q_n-1}}_{n-th}\quad\underbrace{\ldots}_{p_k}\right]}_{\lambda_k}\quad\ldots 0\ldots\quad\underbrace{0}_{\lambda_m}\right] \tag{26}$$

From (23) we get:

$$
\begin{aligned}
&[\mathbf{M}-\lambda_k\mathbf{I}]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{k},q}=0;\quad [\mathbf{M}-\lambda_k\mathbf{I}]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},k}=\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},k-1};\quad 1<k\le q\\
&U_1^{\mathbf{n}\ \ \mathbf{k}}=\left[\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q}\right];\\
&[\mathbf{M}-\lambda_k\mathbf{I}]\cdot U_1^{\mathbf{n}\ \ \mathbf{k}}=U_2^{\mathbf{n}\ \ \mathbf{k}}=\left[0,\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-1}\right]\\
&\ldots\ldots\\
&[\mathbf{M}-\lambda_k\mathbf{I}]^j\cdot U_1^{\mathbf{n}\ \ \mathbf{k}}=U_j^{\mathbf{n}\ \ \mathbf{k}}=\Bigg[\underbrace{0..0}_{j\ \ zeros},\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}\ldots\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-j}\Bigg]\\
&\ldots\ldots\\
&[\mathbf{M}-\lambda_k\mathbf{I}]^q\cdot U_1^{\mathbf{n}\ \ \mathbf{k}}=0
\end{aligned}
\tag{27}
$$

$$
\mathbf{U}f(\mathbf{G})=\sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}[\mathbf{M}-\lambda_k\mathbf{I}]^i\left[\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\ldots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{\left[U^{k\ \ 1}\quad\ldots\quad\underbrace{U^{k\ \ n}}_{n-th}\qquad U^{k\ \ p_k}\right]}_{\lambda_k}\quad\ldots0\ldots\quad\underbrace{0}_{\lambda_m}\right]
\tag{28}
$$

i.e. we collected all eigen vectors belonging to the eigen value λ_k. Now we need a projection non-distorting operator on the sub-space of λ_k. First, let's find zero operator for sunspace of λ_i:

$$O_i = [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} \Rightarrow [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} U_1^{r\ \ i} = [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} \left[\mathbf{Y}_\mathbf{k}^{r,1} \ldots \mathbf{Y}_\mathbf{k}^{r,l} \ldots \mathbf{Y}_\mathbf{k}^{r,q}\right] = 0;$$

$$T_k = \prod_{i \neq k} \frac{O_i}{(\lambda_k - \lambda_i)^{n_i}} = \prod_{i \neq k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right)^{n_i} \tag{29}$$

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 crate as follows:

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

$$RU_1 = U_1 + \alpha U_2 \qquad\qquad U_1 = U_1$$

.....

$$RU_{q-1} = U_{q-1} + \alpha U_q \qquad\qquad U_{q-1} = T^{q-2} U_1$$

$$RU_q = U_q \qquad\qquad U_q = T^{q-1} U_1$$

$$Q = \alpha T$$

$$U_q = RU_q = RT^{q-1}U_1$$

$$U_{q-1} = R(I+Q)U_{q-1} = RQT^{q-2}U_1$$

$$U_{q-1} = RQU_{q-1} = RQT^{q-2}U_1$$

.....

$$U_1 = R\left(\sum_j^{q-1} Q^j\right)U_1$$

so, we get it:

$$P_k^i = \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i}\left(\mathbf{I} + \sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k}\right)^j\right) \tag{30}$$

The final stroke is:

$$P_k = \prod_{i \neq k} \left(P_k^i\right)^{n_i} = \prod_{i \neq k} \left\{ \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k - 1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \right\}^{n_i} \tag{31}$$

and

$$f(\mathbf{M}) = \sum_{k=1}^{m} \left[\prod_{i \neq k} \left\{ \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k - 1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \right\}^{n_i} \sum_{i=1}^{n_k - 1} \frac{f^{(i)}(\lambda_k)}{i!} \left[\mathbf{M} - \lambda_k \mathbf{I}\right]^i \right] \tag{32}$$

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:

$$f(\mathbf{M} \cdot s) = \sum_{k=1}^{m} \left[\prod_{i \neq k} \left\{ \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k - 1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \right\}^{n_i} \sum_{i=1}^{n_k - 1} \frac{f^{(i)}(\lambda_k)}{i!} \left[\mathbf{M} - \lambda_k \mathbf{I}\right]^i s^i \right] \tag{33}$$

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):

$$\exp[\mathbf{D}s] = \sum_{k=1}^{m} \left[e^{\lambda_k s} \prod_{i \neq k} \left\{ \frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \sum_{j=0}^{n_k - 1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\}^{n_i} \sum_{p=0}^{n_k - 1} \frac{s^p}{p!} (\mathbf{D} - \lambda_k \mathbf{I})^p \right] \tag{34}$$

where $n_k < 2n$ is height of the eigen value λ_k. It is also shown there that n_k can be replaced in (34) by any number nn > n_k – it will add only term, which are zeros, but can make (34) 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. (34) is a bit uglier than (3), but still can be used with some elegance.

In our (**HAMLTONIAN**) case we again have a shortcut to solutions…. Is not this a wonderful repeating pattern of freebees... Eigen values split into pairs with the opposite sign because it is a Hamiltonian system:

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

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

$$\det[\mathbf{D} - \lambda I] = \prod (\lambda_i - \lambda)(-\lambda_i - \lambda) = \prod \left(\lambda^2 - \lambda_i^{\ 2}\right) = 0 \,. \tag{36}$$

For accelerator elements it is of paramount importance, 1D case is reduces to trivial (38), 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! I always pick quadratic or cubic equation instead of an arbitrary 4th ot 6th order equation – the later also does not have analytical expressions for solutions. Power to **HAMLTONIAN!**

It is also allow us to simplify (3) into

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

$$\exp[\mathbf{D}s] = \sum_{k=1}^{n} \left(\frac{e^{\lambda_k s} + e^{-\lambda_k s}}{2} \mathbf{I} + \frac{e^{\lambda_k s} - e^{-\lambda_k s}}{2\lambda_k} \mathbf{D} \right) \prod_{j \neq k} \left(\frac{\mathbf{D}^2 - \lambda_j{}^2 \mathbf{I}}{\lambda_k{}^2 - \lambda_j{}^2} \right) \qquad (37)$$

where index k goes only through n pairs of $\{\lambda_k, -\lambda_k\}$. While (37) 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

$$\lambda_1 = -\lambda_2 = \lambda; \quad \lambda^2 = -\det[\mathbf{D}] \tag{38}$$

Thus, it is non-degenerated case only when $\det[D] \neq 0$. (34) give us a simple two-piece expression :

$$\exp[\mathbf{D}s] = \sum_{k=1}^{2} e^{\lambda_k s} \prod_{\lambda_j \neq \lambda_k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} = e^{\lambda s} \frac{\mathbf{D} - (-\lambda)\mathbf{I}}{\lambda - (-\lambda)} + e^{-\lambda s} \frac{\mathbf{D} - \lambda \mathbf{I}}{(-\lambda) - \lambda} = e^{\lambda s} \frac{\mathbf{D} + \mathbf{I}}{2\lambda} - e^{-\lambda s} \frac{\mathbf{D} - \lambda \mathbf{I}}{2\lambda} \tag{39}$$

while (37) bring it home right away:

$$\begin{aligned}
&\exp[\mathbf{D}s] = \mathbf{I} \cdot \frac{e^{\lambda s} + e^{-\lambda s}}{2} + \mathbf{D} \frac{e^{\lambda s} - e^{-\lambda s}}{2\lambda}; \\
&\exp[\mathbf{D}s] = \mathbf{I} \cdot \cosh|\lambda| s + \frac{\mathbf{D} \sinh|\lambda| s}{|\lambda|}; \quad \det[\mathbf{D}] < 0; \quad |\lambda| = \sqrt{-\det[\mathbf{D}]} \\
&\exp[\mathbf{D}s] = \mathbf{I} \cdot \cos|\lambda| s + \frac{\mathbf{D} \sin|\lambda| s}{|\lambda|}; \quad \det[\mathbf{D}] > 0; \quad |\lambda| = \sqrt{\det[\mathbf{D}]}
\end{aligned} \tag{40}$$

The case $\det[\mathbf{D}]=0$ means in this case that D is nilpotent: eqs (37) look like follows

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

hence

$$\exp[\mathbf{D}s]=\mathbf{I}+\mathbf{D}s;\quad \det[\mathbf{D}]=0; \tag{41}$$

Naturally, (42) is result of full-blown degenerated case – eq. (34), but it also can be obtained as a limit case of (40) when $|\lambda|\to 0$.

The value of this approach to matrix calculation is that we do not need to memorize all the different ways of deriving the matrices of various elements in accelerators, and ways of solving a myriad of systems of 2, 4, 6… linear differential equations. Just a smart "coach potato principle" allover again….

The elements of 6x6, 4x4, or 2x2 accelerator matrixes (often called R or T) are numerated as R_{ij}, where ***i*** is the line number and ***j*** is the column number. For example, R_{56} will signify an increment in τ (-arrival time by c) caused by the particle's energy change, δ. Let's look at most trivial case of decoupled transverse motion.

Most accelerators have a flat orbit (κ=0), avoid longitudinal fields ($B_s=0$), and do not have the SQ-quadrupole (N=0). Let us examine a magnetic element (no RF field) and a field in vacuum, where

$$\vec{\nabla}\times\vec{B}=0 \Rightarrow \frac{\partial B_y}{\partial x}=\frac{\partial B_x}{\partial y}.$$

This renders the one-liner Hamiltonian: (the momenta are normalized)

$$\tilde{h} = \left(\frac{\pi_3^{\ 2}}{2} + K_1\frac{y^2}{2}\right) + \frac{\pi_1^{\ 2}}{2p_o} + \left[K^2 - K_1\right]\frac{x^2}{2} + \frac{\pi_\delta^{\ 2}}{2}\cdot\frac{m^2c^2}{p_o^{\ 2}} - K\frac{c}{\mathrm{v}_o}x\pi_\delta;\ K_1 = \frac{e}{p_oc}\frac{\partial B_y}{\partial x}; \quad (42)$$

with a clearly separated vertical (***y***) part of motion. In the presence of the curvature ***K***, i.e., a non-zero dipole field in the reference orbit, both the longitudinal and horizontal (***x***) degrees of freedom remain coupled. In a quadrupole ***K***=0, the Hamiltonian is completely decoupled into three degrees of freedom:

$$\tilde{h} = \left(\frac{\pi_3^{\ 2}}{2} + K_1\frac{y^2}{2}\right) + \left(\frac{\pi_1^{\ 2}}{2} - K_1\frac{x^2}{2}\right) + \frac{\pi_\delta^{\ 2}}{2}\cdot\frac{m^2c^2}{p_o^{\ 2}};\ K_1 = \frac{e}{p_oc}\frac{\partial B_y}{\partial x}; \quad (43)$$

The matrix in the longitudinal direction is the same as that for a drift (29), while the x and y matrices are given by (39). Depending on the sign of the gradient $\partial B_y/\partial x$, the quadrupole focuses in x and defocuses in y, or vice versa:

$$D_x = \begin{bmatrix} 0 & 1 \\ K_1 & 0 \end{bmatrix};\quad D_y = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix};\quad \phi = s\sqrt{K_1}$$
$$M_F = \begin{bmatrix} \cos\phi & \sin\phi/\sqrt{K_1} \\ -\sqrt{K_1}\sin\phi & \cos\phi \end{bmatrix};\quad M_D = \begin{bmatrix} \cosh\phi & \sinh\phi/\sqrt{K_1} \\ \sqrt{K_1}\sinh\phi & \cosh\phi \end{bmatrix} \quad (44)$$

It is worth noting that there is no difference if we use momentum and coordinates, not x, x'.

$$\tilde{h} = \left(\frac{P_3^{\ 2}}{2p_o} + p_oK_1\frac{y^2}{2}\right) + \left(\frac{P_1^2}{2p_o} - p_oK_1\frac{x^2}{2}\right) + \frac{\delta^2}{2p_o}\cdot\frac{m^2c^2}{p_o^{\ 2}};\ K_1 = \frac{e}{p_oc}\frac{\partial B_y}{\partial x}; \quad (45)$$

$$D_x = \begin{bmatrix} 0 & 1/p_o \\ p_oK_1 & 0 \end{bmatrix};\quad D_y = \begin{bmatrix} 0 & 1/p_o \\ -p_oK_1 & 0 \end{bmatrix};\ \phi = s\sqrt{K_1}$$
$$M_F = \begin{bmatrix} \cos\phi & \sin\phi/p_o\sqrt{K_1} \\ -p_o\sqrt{K_1}\sin\phi & \cos\phi \end{bmatrix};\quad M_D = \begin{bmatrix} \cosh\phi & \sinh\phi/p_o\sqrt{K_1} \\ p_o\sqrt{K_1}\sinh\phi & \cosh\phi \end{bmatrix} \quad (46)$$

As we can see, this is not a more complicated that using x,x', but definitely correct for any accelerator.

Matrix of general DC accelerator element (including twisted quads or helical wiggler) can be found using our recipe. With all diversity of possible elements on accelerators, DC (or almost DC) magnets play the most prominent role. In this case energy of the particle stays constant and we can use reduced variables. Furthermore, large number of terms is the Hamiltonian simply disappear and from the previous lecture we have:

$$\tilde{h}_n = \frac{\pi_1^{\,2} + \pi_3^{\,2}}{2} + f\frac{x^2}{2} + n\cdot xy + g\frac{y^2}{2} + L(x\pi_3 - y\pi_1) + \frac{\pi_o^{\,2}}{2}\cdot\frac{m^2c^2}{p_o^{\,2}} + g_x x\pi_o + g_y y\pi_o; \qquad \text{(L2-46-n)}$$

Even though it is tempting to remove electric field, it does not either helps or hurts in general case of an element. Hence, we will keep DC transverse electric fields. We also assume that these fields are in vacuum and $\dfrac{\partial B_y}{\partial x} = \dfrac{\partial B_x}{\partial y}, \ \dfrac{\partial E_x}{\partial x} + KE_x + \dfrac{\partial E_y}{\partial y} = 0$:

$$f = K^2 - \frac{e}{p_o c}\frac{\partial B_y}{\partial x} - \frac{e}{p_o \mathrm{v}_o}\frac{\partial E_y}{\partial y} + \left(\frac{eB_s}{2p_o c}\right)^2 + \left(\frac{meE_x}{p_o^{\,2}}\right)^2;$$

$$g = \frac{e}{p_o c}\frac{\partial B_y}{\partial x} + \frac{e}{p_o \mathrm{v}_o}\frac{\partial E_y}{\partial y} + \left(\frac{eB_s}{2p_o c}\right)^2 + \left(\frac{meE_z}{p_o^{\,2}}\right)^2; \qquad ;\ (47)$$

$$2n = \left[\frac{e}{p_o c}\frac{\partial B_x}{\partial x} - \frac{e}{p_o c}\frac{\partial B_y}{\partial y}\right] - K\cdot\frac{e}{p_o c}B_x - \frac{e}{p_o \mathrm{v}_o}\left(\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x}\right) - 2K\frac{eE_y}{p_o \mathrm{v}_o} + \left(\frac{meE_z}{p_o^{\,2}}\right)\left(\frac{meE_x}{p_o^{\,2}}\right)$$

$$L = \kappa + \frac{e}{2p_o c}B_s; \qquad g_x = \frac{(mc)^2\cdot eE_x}{p_o^3} - K\frac{c}{\mathrm{v}_o}; \ g_y = \frac{(mc)^2\cdot eE_y}{p_o^3};$$

In the absence of longitudinal electric field, the momentum P_2 is constant as well $\pi_o = const$, δ=const. The fact that particle's energy does not changes in such element is rather obvious (It is completely correct for magnetic elements. Presence of electric field makes it less obvious, but it comes from the fact that Hamiltonian does not depend on time!): $\pi_o{}' = -\frac{\partial h}{\partial \tau} = 0$.

Equations of motion become specific:

$$\mathbf{X}^T = [x, \pi_1, y, \pi_3, \tau, \pi_o] = [X^T, \tau, \pi_o]; \; X^T = [x, \pi_1, y, \pi_3], \tag{48}$$

$$\frac{d\mathbf{X}}{ds} = \mathbf{D}(s)\cdot\mathbf{X}; \; \mathbf{D} = \mathbf{S}\cdot\mathbf{H}(s) = \begin{bmatrix} 0 & 1 & -L & 0 & 0 & 0 \\ -f & 0 & -n & -L & 0 & g_x \\ L & 0 & 0 & 1 & 0 & 0 \\ -n & L & -g & 0 & 0 & g_y \\ g_x & 0 & g_y & 0 & 0 & m^2c^2/p_o^2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \tag{49}$$

and can be rewritten in a slightly different (just deceivingly looking better) way:

$$\frac{dX}{ds} = D\cdot X + \pi_o \cdot C;$$

$$\frac{d\tau}{ds} = g_x x + g_y y + \pi_o \cdot m^2c^2/p_o^2; D = \begin{bmatrix} 0 & 1 & -L & 0 \\ -f & 0 & -n & -L \\ L & 0 & 0 & 1 \\ -n & L & -g & 0 \end{bmatrix}; \; C = \begin{bmatrix} 0 \\ g_x \\ 0 \\ g_y \end{bmatrix}. \tag{50}$$

Hence, solution for transverse motion (4-vector) in such an element can be written as combination general solution of homogeneous equation plus specific solution of inhomogeneous one:

$$X = M(s)\cdot X_o + \pi_o \cdot R(s);\;\; M = e^{D(s-s_o)};\;\; R' = D\cdot R + C;\;\; R(s_o) = 0. \qquad (51)$$

It worth noting that C=0 as soon as there is no field on the orbit – E=0, B=0. In this case R=0.

Before finding 4x4 matrixes M and vector R, let's see what we will know about the 6x6 matrix after that. First, the obvious:

$$\mathbf{M}_{6x6} = \begin{bmatrix} & \mathbf{M}_{4x4} & & & 0 & R \\ R_{51} & R_{52} & R_{53} & R_{54} & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (52)$$

with a natural question of what are non-trivial R_{5k} elements? Usually these elements, with exception of R_{56} are not even mentioned in most of textbooks.

Fortunately for us, Mr. Hamiltonian gives us a hand in the form of symplecticity of transport matrixes. Using (18) and (18-1) we can find that:

$$\mathbf{M}_{6x6}{}^{T}\mathbf{S}\mathbf{M}_{6x6} == \begin{bmatrix} \mathbf{M}^{T}{}_{4x4} & L^{T} & 0 \\ 0 & 1 & 0 \\ R^{T} & R_{56} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{S}_{4x4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{4x4} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{4x4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{M}^{T}{}_{4x4}\mathbf{S}_{4x4} & 0 & L^{T} \\ 0 & 0 & 1 \\ R^{T}\mathbf{S}_{4x4} & -1 & R_{56} \end{bmatrix} \begin{bmatrix} \mathbf{M}_{4x4} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}^{T}{}_{4x4}\mathbf{S}_{4x4}\mathbf{M}_{4x4} & 0 & 0 \\ 0 & 0 & 1 \\ R^{T}\mathbf{S}_{4x4}\mathbf{M}_{4x4} - L & -1 & R^{T}\mathbf{S}_{4x4}R \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{4x4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

where we used $L = [R_{51}, R_{52}, R_{53}, R_{54}]$. We should note what $\mathbf{X^{T}SX=0}$ for any vector, $\mathbf{M}^{T}{}_{4x4}\mathbf{S}_{4x4}\mathbf{M}_{4x4} = \mathbf{S}_{4x4}$ and only non-trivial condition from the equation above is:

$$R^{T}\mathbf{S}_{4x4}\mathbf{M}_{4x4} - L = 0$$

which gives us very valuable dependence of the arrival time on the transverse motions:

$$L = R^{T}\mathbf{S}_{4x4}\mathbf{M}_{4x4}; \quad or \quad L^{T} = -\mathbf{M}^{T}{}_{4x4}\mathbf{S}_{4x4}R. \qquad (53)$$

Element R_{56} is decoupled form the symplectic condition in this case and should be determined by direct integration - no magic here:

$$\tau(s) = \tau(s_o) + \pi_o \cdot \left\{ m^2c^2/p_o(s-s_o) + {}^2 \int_{s_o}^{s} \left(g_x R(s)_{16} + g_y R_{36}(s) \right) ds \right\}$$
$$R_{56} = m^2c^2/p_o(s-s_o) + {}^2 \int_{s_o}^{s} \left(g_x R(s)_{16} + g_y R_{36}(s) \right) ds \qquad (54)$$

Let's find the solutions for 4x4 matrixes of arbitrary element. First, let solve characteristic equation for D:

$$\det[D - \lambda I] = \lambda^4 + \lambda^2\left(f + g + 2L^2\right) + fg + L^4 - L^2(f+g) - n^2 = 0 \qquad (55)$$

with easy roots:

$$\lambda^2 = a \pm b;\ \ a = -\frac{f+g+2L^2}{2};\ \ b^2 = \frac{(f-g)^2}{4} + 2L^2(f+g) + n^2 \qquad (56)$$

$$\lambda^2 = a \pm b; a = -\frac{f+g+2L^2}{2}; b^2 = \frac{(f-g)^2}{4} + 2L^2(f+g) + n^2$$

Before starting classification of the cases, let's note that

$$f+g = K^2 + 2\left(\frac{eB_s}{2p_o c}\right)^2 + \left(\frac{meE_x}{p_o{}^2}\right)^2 + \left(\frac{meE_z}{p_o{}^2}\right)^2 \geq 0$$

i.e. $a \leq 0;\ b^2 \geq 0;\ \mathrm{Im}(b) = 0$. The solutions can be classified as following: remember that the full set of eigen values is $\lambda_1, -\lambda_1, \lambda_2, -\lambda_2$:

I. $\lambda_1 = \lambda_2 = 0;\quad a = 0;\ b = 0;$

II. $\lambda_1 = \lambda_2 = i\omega;\quad a = -\omega^2;\ b = 0;$

III. $\lambda_1 = 0;\quad \lambda_2 = i\omega;\quad a + b = 0;\quad 2b = \omega^2$

IV. $\lambda_1 = i\omega_1;\quad \lambda_2 = i\omega_2;\quad \omega_1{}^2 = -a - b;\ \omega_2{}^2 = -a + b;\ |a| > b$

V. $\lambda_1 = i\omega_1;\quad \lambda_2 = \omega_2;\ \omega_1{}^2 = -a - b;\ \omega_2{}^2 = b - a;\ b > |a|$

Before going to case-by-case calculations, lets use Sylvester's formulae and try to find solution of inhomogeneous equation:

$$\frac{d\mathbf{R}}{ds} = \mathbf{D} \cdot \mathbf{R} + \mathbf{C}; \;\; \mathbf{R}(0) = 0 \, . \tag{57}$$

When matrix detD≠0, (57) can be inversed using a $\mathbf{R} = A + e^{Ds} \cdot B$ as a guess and the boundary condition $\mathbf{R}(0) = 0$:

$$\mathbf{R} = \left(\mathbf{M_{4x4}(s)} - \mathbf{I}\right) \cdot \mathbf{D}^{-1} \cdot \mathbf{C} \tag{58}$$

is the easiest solution. Prove is just straight forward:

$$\mathbf{R}' = \mathbf{D} \cdot \mathbf{M_{4x4}}^{-1} \cdot \mathbf{C};$$

$$\mathbf{D} \cdot \left(\mathbf{M} - \mathbf{I}\right) \cdot \mathbf{D}^{-1} \cdot \mathbf{C} + \mathbf{C} = \mathbf{D} \cdot \mathbf{M_{4x4}}^{-1} \cdot \mathbf{C} \quad \#$$

In all cases we can use method of variable constants to find it:

$$\frac{dR}{ds} = R' = \mathbf{D} \cdot R + \mathbf{C}; \quad \mathbf{M}' = \mathbf{DM};$$

$$R = \mathbf{M}(s) A(s) \Rightarrow \mathbf{M}' A + \mathbf{M} A' = \mathbf{DM} A + C; \quad R(0) = 0 \Rightarrow A_o = 0 \tag{59}$$

$$A' = \mathbf{M}^{-1}(s) C \Rightarrow A = \int_0^s \mathbf{M}^{-1}(z) C dz = \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot C; \quad R = e^{\mathbf{D}s} \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot C$$

It is important to remember that $\mathbf{M}^{-1}(\mathbf{s})$ is just the $\mathbf{M(-s)} = \mathbf{e}^{-\mathbf{Ds}}$. Hence in all our formulae for matrixes from previous lectures we need to replace s by –s to get $\mathbf{M}^{-1}(\mathbf{s})$**.** Other vice, we have to use general formula (33) for the homogeneous solution and use method of variable constants (see Appendix F) to find it:

$$R(s) = \sum_{k=1}^{m}\left\{\prod_{i\neq k}\left[\frac{\mathbf{D}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\right]\sum_{j=0}^{n_k-1}\left(\frac{\mathbf{D}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right\}\sum_{n=0}^{n_k-1}(\mathbf{D}-\lambda_k\mathbf{I})^n\frac{s^n}{n!}\cdot\sum_{p=0}^{n_k-1}(-1)^{p+1}(\mathbf{D}-\lambda_k\mathbf{I})^p\cdot\mathbf{C}\cdot\left[\sum_{q=0}^{p1}\frac{s^{p-q}}{(p-q)!\,\lambda_k{}^{q+1}}-\frac{e^{\lambda_k}}{\lambda_k{}^{p+1}}\right] \tag{60}$$

In all specific cases I, II, III, IV and V, integrating (L-53) directly is usually easier that using general form of (60).

Case I.

$$\frac{f+g+2L^2}{2}=0;\ \frac{(f-g)^2}{4}+2L^2(f+g)+n^2=0;$$

$$f+g=pos^2\geq 0\Rightarrow (f-g)^2=0;\ L^2(f+g)=0;\ n^2=0$$

$$f+g+2L^2=pos^2+2L^2=0\Rightarrow L=0;\ f+g=0\Rightarrow$$

$$f-g=0\Rightarrow f=g=L=n=0!!!$$

means that there is nothing in the Hamiltonian but p^2– is this the drift section matrix of which we already know. Hence, there is not curvature as well and R=0.

$$\mathbf{M_{4x4}}=\begin{bmatrix}1 & s & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & s\\ 0 & 0 & 0 & 0\end{bmatrix};\ \mathbf{R}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}. \qquad \text{(I-1)}$$

The only not trivial (ha-ha – it is also as trivial as it can be) is R_{56}:

$$R_{56}=\frac{m^2c^2}{p_o^{\ 2}}s \qquad \text{(I-2)}$$

we already had seen it when studied nilpotent case…

Case II: $b = \dfrac{(f-g)^2}{4} + 2L^2(f+g) + n^2 = 0;$

$$f = g;\ n = 0 \ \ and \ \ L^2(f+g) = L^2\left(K^2 + \Omega^2 + El^2\right) = 0; \Omega = eB_s / p_o c; E_\perp = 0.$$

i.e. there are two cases: L=0 or $f + g = 0$.

If both are equal zero, i.e. $f + g = 0;\ \ L = 0$, this is equivalent to the case I above.

Case II a: $f + g = 0$, K≠0, Bs=0 -> L=κ. Thus, this is just a drift (straight section) with rotation, whose matrix is trivial: Drift + rotation. There is not transverse force – hence R=0.

$$\mathbf{M_{4x4}} = \begin{bmatrix} M_d \cdot \cos\kappa s & -M_d \cdot \sin\kappa s \\ M_d \cdot \sin\kappa s & M_d \cdot \cos\kappa s \end{bmatrix};\ M_d = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix};\ \ \mathbf{R} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \qquad \text{(IIa-1)}$$

$R_{56}$is as for a drift:

$$R_{56} = \frac{m^2c^2}{p_o^2} s \qquad \text{(IIa-2)}$$

Case II b: L=0; $f = g = \left(K^2 + \Omega^2\right)/2; \kappa = -\Omega$; i.e. the motion is uncoupled:

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -f & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -f & 0 \end{bmatrix}; \; C = \begin{bmatrix} 0 \\ g_x \\ 0 \\ g_y \end{bmatrix}.$$

$$\mathbf{M_{4x4}} = \begin{bmatrix} M & 0 \\ 0 & M \end{bmatrix}; M = \begin{bmatrix} \cos\omega s & \sin\omega s/\omega \\ -\omega\sin\omega s & \cos\omega s \end{bmatrix} \qquad \text{(IIb-1)}$$

Here we may have non-zero R: yes, it may be! It is simple integrals to be taken care of:

$$C_{x,y} = -g_{x,y}\begin{bmatrix}0\\1\end{bmatrix};\; M^{-1}(z) = \begin{bmatrix}\cos\omega z & -\sin\omega z/\omega\\ \omega\sin\omega z & \cos\omega z\end{bmatrix} C_{x,y} = g_{x,y}\begin{bmatrix}\sin\omega z/\omega\\ -\cos\omega s\end{bmatrix};$$

$$\int_0^s \mathbf{M}^{-1}(z)C_{x,y}dz = g_{x,y}\begin{bmatrix}\int_0^s \sin(\omega z)dz/\omega\\ -\int_0^s \cos(\omega z)dz\end{bmatrix} = \frac{g_{x,y}}{\omega}\begin{bmatrix}(1-\cos\omega s)/\omega\\ -\sin(\omega s)\end{bmatrix}$$

$$C_{x,y} = -g_{x,y}\begin{bmatrix}0\\1\end{bmatrix}; \int_0^s \mathbf{M}^{-1}(z)dz\cdot C_{x,y} = -g_{x,y}\begin{bmatrix}0\\1\end{bmatrix};$$

$$M(s)\int_0^s \mathbf{M}^{-1}(z)C_{x,y}dz = \frac{g_{x,y}}{\omega}\begin{bmatrix}\cos\omega s & \sin\omega s/\omega\\ -\omega\sin\omega s & \cos\omega s\end{bmatrix}\cdot\begin{bmatrix}(1-\cos\omega s)/\omega\\ -\sin(\omega z)\end{bmatrix} = \frac{g_{x,y}}{\omega^2}\begin{bmatrix}\cos\omega s - 1\\ -\omega\sin\omega s\end{bmatrix}$$

$$R_{56} = s\cdot m^2c^2/p_o + \int_0^s \left(g_x R(z)_{16} + g_y R_{36}(z)\right)dz =$$

$$\int_0^s \left(g_x R(z)_{16} + g_y R_{36}(z)\right)dz = \frac{g_x{}^2 + g_y{}^2}{\omega^2}\int_0^s (\cos\omega z - 1)dz = \frac{g_x{}^2 + g_y{}^2}{\omega^2}\left(\frac{\sin\omega s}{\omega} - s\right)$$

with the result:

$$R = \begin{bmatrix}\frac{g_x}{\omega^2}(\cos\omega s - 1)\\ -\frac{g_x}{\omega}\sin\omega s\\ \frac{g_y}{\omega^2}(\cos\omega s - 1)\\ -\frac{g_x}{\omega}\sin\omega s\end{bmatrix};\quad R_{56} = \frac{m^2c^2}{p_o{}^2}s + \frac{g_x{}^2 + g_y{}^2}{\omega^2}\left(\frac{\sin\omega s}{\omega} - s\right) \quad \text{(IIb-2)}$$

Case III: $a+b=0$; $\det D=0$; $\omega^2=2b$; $\lambda_{1,2}=\pm i\omega; \lambda_3=0$; $m=3$.

We have to use degenerated case formula, but the maximum height of the eigen vector is 2 and only for 3-rd eigen value. Since it is not scary at all: $n_1=1; n_2=1; n_3=2$

Because of the Hamilton-Kelly theorem, $\mathbf{D}^2(\mathbf{D}^2+\omega^2 I)=0$. Let's do it

$$\exp[\mathbf{D}s]=\sum_{k=1}^{3}\left[e^{\lambda_k s}\prod_{i\neq k}\left\{\frac{\mathbf{D}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\sum_{j=0}^{n_k-1}\left(\frac{\mathbf{D}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right\}^{n_i}\sum_{p=0}^{n_k-1}\frac{s^p}{p!}(\mathbf{D}-\lambda_k\mathbf{I})^p\right]=$$

$$\lambda_1\lambda_2=\omega^2;\ i\omega$$

$$k=3;\ \left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)^2(\mathbf{I}+s\mathbf{D});\ \left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)^2=\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)+\frac{\mathbf{D}^2}{\omega^2}\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)\downarrow_0=\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)$$

$$k=3;\ \left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)(\mathbf{I}+s\mathbf{D})$$

$$k=1+2;\ e^{i\omega s}\frac{\mathbf{D}+i\omega\mathbf{I}}{-2i\omega}\frac{\mathbf{D}^2}{\omega^2}+c.c.=-\frac{\mathbf{D}^2}{\omega^2}\left(\mathbf{I}\cos\omega s+\frac{\mathbf{D}}{\omega}\right)$$

$$M=\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)(\mathbf{I}+s\mathbf{D})-\frac{\mathbf{D}^2}{\omega^2}\left(\mathbf{I}\cos\omega s+\frac{\mathbf{D}}{\omega}\sin\omega s\right)$$

$$\mathbf{M}_{\mathbf{4x4}}=\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)(\mathbf{I}+s\mathbf{D})-\frac{\mathbf{D}^2}{\omega^2}\left(\mathbf{I}\cos\omega s+\frac{\mathbf{D}}{\omega}\sin\omega s\right)\qquad\text{(III-1)}$$

Similarly

$$R=\left\{\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)\left(\mathbf{I}s+\mathbf{D}\frac{s^2}{2}\right)+\frac{\mathbf{D}^2}{\omega^4}\left(\mathbf{D}(\cos\omega s-1)-\mathbf{I}\omega\sin\omega s\right)\right\}C \qquad \text{(III-2)}$$

Next is just

$$\int_o^s C^T\left\{\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)\left(\mathbf{I}z+\mathbf{D}\frac{z^2}{2}\right)+\frac{\mathbf{D}^2}{\omega^4}\left(\mathbf{D}(\cos\omega z-1)-\mathbf{I}\omega\sin\omega z\right)\right\}Cdz=$$

$$C^T\left\{\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)\left(\mathbf{I}\frac{s^2}{2}+\mathbf{D}\frac{s^3}{6}\right)+\frac{\mathbf{D}^2}{\omega^4}\left(\mathbf{D}\left(\frac{\sin\omega s}{\omega}-s\right)\mathbf{I}(\cos\omega z-1)\right)\right\}C$$

with result of:

$$R_{56}=m^2c^2/p_o s+C^T\left\{\left(\mathbf{I}+\frac{\mathbf{D}^2}{\omega^2}\right)\left(\mathbf{I}\frac{s^2}{2}+\mathbf{D}\frac{s^3}{6}\right)+\frac{\mathbf{D}^2}{\omega^4}\left(\mathbf{D}\left(\frac{\sin\omega s}{\omega}-s\right)\mathbf{I}(\cos\omega z-1)\right)\right\}C \qquad \text{(III-3)}$$

Case IV: all roots are different, no degeneration. Use formula (36)

$$\exp[\mathbf{D}s] = \sum_{k=1}^{2}\left(\frac{e^{\lambda_k s}+e^{-\lambda_k s}}{2}\mathbf{I} + \frac{e^{\lambda_k s}-e^{-\lambda_k s}}{2\lambda_k}\mathbf{D}\right)\prod\left(\frac{\mathbf{D}^2-\lambda_j{}^2\mathbf{I}}{\lambda_k{}^2-\lambda_j{}^2}\right)$$

with only one term in the product:

$$\mathbf{M_{4x4}} = \frac{1}{\omega_1{}^2-\omega_2{}^2}\left\{\left(\mathbf{I}\cos\omega_1 s + \mathbf{D}\frac{\sin\omega_1 s}{\omega_1}\right)\left(\mathbf{D}^2+\omega_2{}^2\mathbf{I}\right) - \left(\mathbf{I}\cos\omega_2 s + \mathbf{D}\frac{\sin\omega_2 s}{\omega_2}\right)\left(\mathbf{D}^2+\omega_1{}^2\mathbf{I}\right)\right\} \qquad \text{(IV-1)}$$

For R we invoke a simplest formula:

$$\mathbf{R} = (\mathbf{M_{4x4}(s)} - \mathbf{I})\mathbf{D}^{-1}\cdot\mathbf{C} \qquad \text{(IV-2)}$$

For R56 it is tedious but easy:

$$R_{56} = m^2c^2/p_o s + C^T \mathsf{M}\mathbf{D}^{-1}C;$$

$$\mathsf{M} = \frac{1}{\omega_1{}^2-\omega_2{}^2}\left\{\begin{array}{l}\left(\mathbf{I}\dfrac{\sin\omega_1 s}{\omega_1}+\mathbf{D}\dfrac{1-\cos\omega_1 s}{\omega_1{}^2}\right)\left(\mathbf{D}^2+\omega_2{}^2\mathbf{I}\right) - \\ \left(\mathbf{I}\dfrac{\sin\omega_2 s}{\omega_2}+\mathbf{D}\dfrac{1-\cos\omega_2 s}{\omega_2{}^2}\right)\left(\mathbf{D}^2+\omega_1{}^2\mathbf{I}\right) - \mathbf{I}\cdot s\end{array}\right\} \qquad \text{(IV-3)}$$

Case V: all roots are different, no degeneration. Use formula (36) again

$$\mathbf{M_{4x4}} = \frac{1}{\omega_1^{\;2} + \omega_2^{\;2}} \left\{ \left(\mathbf{I}\cos\omega_1 s + \mathbf{D}\frac{\sin\omega_1 s}{\omega_1} \right)\left(\mathbf{D}^2 - \omega_2^{\;2}\mathbf{I}\right) - \left(\mathbf{I}\cosh\omega_2 s + \mathbf{D}\frac{\sinh\omega_2 s}{\omega_2} \right)\left(\mathbf{D}^2 + \omega_1^{\;2}\mathbf{I}\right) \right\} \quad \text{(V-1)}$$

$$\mathbf{R} = \left(\mathbf{M_{4x4}(s)} - \mathbf{I}\right)\mathbf{D^{-1}} \cdot \mathbf{C} \quad \text{(V-2)}$$

$$R_{56} = m^2c^2/p_o s + C^T \mathsf{M}\mathbf{D}^{-1}C;$$

$$\mathsf{M} = \frac{1}{\omega_1^{\;2} + \omega_2^{\;2}} \left\{ \begin{aligned} &\left(\mathbf{I}\frac{\sin\omega_1 s}{\omega_1} + \mathbf{D}\frac{1-\cos\omega_1 s}{\omega_1^{\;2}} \right)\left(\mathbf{D}^2 - \omega_2^{\;2}\mathbf{I}\right) - \\ &\left(\mathbf{I}\frac{\sinh\omega_2 s}{\omega_2} + \mathbf{D}\frac{\cosh\omega_2 s - 1}{\omega_2^{\;2}} \right)\left(\mathbf{D}^2 + \omega_1^{\;2}\mathbf{I}\right) - \mathbf{I}\cdot s \end{aligned} \right\} \quad \text{(V-3)}$$

Before going into the discussion of the parameterization of the motion, we need to finish discussion of few remaining topics for 6x6 matrix of an accelerator. First is multiplication of the 6x6 matrixes for purely magnetic elements:

$$\mathbf{M}_{k(6x6)} = \begin{bmatrix} \mathbf{M}_{k(4x4)} & 0 & R_k \\ L_k & 1 & R_{56_k} \\ 0 & 0 & 1 \end{bmatrix};$$

$$\mathbf{M}_{2(6x6)}\mathbf{M}_{1(6x6)} = \begin{bmatrix} \mathbf{M}_{(4x4)} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_2\mathbf{M}_1 & 0 & R_2 + \mathbf{M}_2 R_1 \\ L_2 + L_1\mathbf{M}_2 & 1 & R_{56_1} + R_{56_2} + L_2 R_1 \\ 0 & 0 & 1 \end{bmatrix} \quad (61)$$

i.e. having transformation rules for mixed size objects: a 4x4 matrix M, 4-elemetn column R, 4 element line L, and a number R_{56}. As you remember, L is dependent (L4-7) and expressed as **L= R$^{\mathbf{T}}$SM**. Thus:

$$\mathbf{M}_{(4x4)} = \mathbf{M}_2\mathbf{M}_1;\ R = \mathbf{M}_2 R_1 + R_2;\ L = L_2\mathbf{M}_1 + L_1;\ R_{56} = R_{56_1} + R_{56_2} + L_2 R_1 \quad (62)$$

One thing is left without discussion so far – the energy change. Thus, we should look into a particle passing through an RF cavity, which has alternating longitudinal field. Again, for simplicity we will assume that equilibrium particle does not gain energy, i.e. p_o stays constant and we can continue using reduced variables. We will also assume that the is no transverse field, neither AC or DC. In this case the Hamiltonian reduces to a simple, fully decoupled:

$$\tilde{h} = \frac{\pi_1^2 + \pi_3^2}{2} + \frac{\pi_o^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + u \frac{\tau^2}{2};; \qquad \text{(L2-46)}$$

$$\frac{dX}{ds} = \mathbf{D} \cdot X;\ \mathbf{D} = \begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_l \end{bmatrix};\ D_x = D_y = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; D_l = \begin{bmatrix} 0 & \frac{m^2 c^2}{p_o^2} \\ -u & 0 \end{bmatrix};$$

$$\mathbf{M} = \begin{bmatrix} M_x & 0 & 0 \\ 0 & M_y & 0 \\ 0 & 0 & M_l \end{bmatrix}; M_x = M_y = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}; \qquad \omega = \sqrt{|\det D_l|} = \frac{mc}{p_o}\sqrt{|u|} \qquad (63)$$

$$M_l = \begin{bmatrix} \cos\omega s & \frac{m^2c^2}{p_o^2}\sin\omega s/\omega \\ -u\sin\omega s/\omega & \cos\omega s \end{bmatrix}; u > 0;\ M_l = \begin{bmatrix} \cosh\omega s & \frac{m^2c^2}{p_o^2}\sinh\omega s/\omega \\ -u\sinh\omega s/\omega & \cosh\omega s \end{bmatrix}; u < 0;$$

In majority of the cases $\omega s \ll 1$ ($mc/p_o \sim 1/\gamma$) and RF cavity can be represented as a thin lens located in its center:

$$\mathbf{M} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_l \end{bmatrix};\quad M_l = \begin{bmatrix} 1 & 0 \\ -q & 1 \end{bmatrix};\quad q = u \cdot l_{RF} = -\frac{e}{p_o c}\frac{\partial V_{rf}}{\partial t} \qquad (64)$$