Lie Algebraic Methods in Beam Optics

C. Thomas Mottershead
ctmotters@gmail.com
(805) 264-0853
Aug 2021

Introduction

The purpose of this document is to provide a concise heuristic reference collection of formulas and data on Lie algebraic design techniques, and their application to various projects. Beginning with the familiar equations of classical mechanics and electromagnetism, we sketch without derivation or proof the ingredients the Lie algebraic formulation of modern beam dynamics, in particular the connection between magnetic multipoles in the optical system and map non-linearities. This summary is mostly taken from the papers of Dragt and coworkers, which contain the original derivations and proofs. Adapted from a talk given at the Monterey Star Wars conference around 1990.

Classical Hamiltonian Mechanics

Consider a beamline consisting entirely of static magnetic elements and drifts, without RF accelerating fields. The static magnetic field must satisfy Maxwell's equations: $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{B} = 0$. The divergence condition is satified identically by setting $\mathbf{B} = \nabla \times \mathbf{A}$. The system is fully defined by specifying the vector potential $\mathbf{A}(\mathbf{x})$ everywhere. The Hamiltonian for a charged particle under the influence of the magnetic forces implied by $\mathbf{A}$ is

$$H(\mathbf{p}, \mathbf{x}) = c\sqrt{(\mathbf{p} - e\mathbf{A}(\mathbf{x}))^2 + (mc)^2}, \tag{1}$$

where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{p} = (p_1, p_2, p_3)$ are the position and cannonical momentum coordinates of the particle. Hamiltons equations of motion for the particle are

$$\dot{x}_k = +\frac{\partial H}{\partial p_k} = -\{H, x_k\} \qquad \text{and} \qquad \dot{p}_k = -\frac{\partial H}{\partial x_k} = -\{H, p_k\} \tag{2}$$

where $\{f, g\}$ is the usual Poisson bracket of classical mechanics, defined for any two functions f and g on phase space by

$$\{f, g\} \equiv \sum_{k=1}^{3} \frac{\partial f}{\partial x_k}\frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k}\frac{\partial g}{\partial x_k} \tag{3}$$

Beam Coordinates

The beam particle coordinates commonly used in magnetic optics are dimensionless deviations from a reference trajectory, defined as the path through the system of a selected reference particle with specified initial position and momentum. For straight systems, this is usually just the centerline. The longitudinal position s of the reference particle is taken as the independent variable in place of the time. The deviations in time of flight ($\tau = c\Delta t$) and energy ($p_\tau = -\Delta E/p_0 c$) relative to the reference particle are then taken as new longitudinal coordinates. Dimensionless tranverse momenta, essentially the trajectory slopes, are formed by dividing by the longitudinal design momentum of the reference particle : $P_x = p_x/p_0$, $P_y = p_y/p_0$. We often denote the vector of these new phase space coordinates by $\vec{z} \equiv (x, P_x, y, P_y, \tau, p_\tau)$. When

transformed to these new cannonical beam coordinates, the Hamiltonian turns out to be essentially the old p_z:

$$H = -\frac{A_z}{B\rho} - \sqrt{1 + p_\tau^2 - 2\frac{p_\tau}{\beta_0} - (P_x - \frac{A_x}{B\rho})^2 - (P_y - \frac{A_y}{B\rho})^2} \tag{4}$$

The special role of the z-component of the vector potential should be noted. In regions where the transverse components of the vector potential vanish, the Hamiltonian has the form of a kinetic term plus a potential eA_z.

Lie Algebraic Methods

An algebra is a set of mathematical elements $(A, B, C, \cdots)$ for which an addition rule $C = A + B$, and a composition (multiplication) rule $C = A \circ B$ are defined for combining two of the elements to make a third one. A Lie algebra is one in which the composition rule is antisymmetric and satifies the Jacobi identity:

$$A \circ B + B \circ A = 0 \qquad \text{and} \qquad A \circ (B \circ C) + B \circ (C \circ A) + C \circ (A \circ B) = 0 \tag{5}$$

The relevance of Lie algebraic methods to classical mechanics follows from the fact that the Poisson bracket operation, viewed as a rule for combining two functions on phase space to produce a third one, satisfies these requirements for any functions f, g, h:

$$\{f, g\} + \{g, f\} \equiv 0 \qquad \text{and} \qquad \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} \equiv 0 \tag{6}$$

A Lie operator $:f:$ may be associated with any function on phase space $f(\vec{z})$ by using the Poisson bracket to define its effect on an arbitrary function $g(\vec{z})$ as:

$$:f:g(\vec{z}) \equiv \{f, g\}. \tag{7}$$

The Lie operators so defined by the Poisson bracket operation form a Lie algebra.

Lie Transforms

Hamiltons equation of motion in beam coordinates may be written

$$\frac{dz_a}{ds} = -\{H, z_a\} \equiv - :H: z_a, \tag{8}$$

where the independent variable s is the arclength specifying the location of the reference particle on the reference trajectory, the index a runs from 1 to 6, and $:H:$ denotes the Lie operator associated with the Hamiltonian. If the particles are independent (space charge interactions neglected), the final phase space coordinate of a particle $\vec{w} = \vec{F}(\vec{z})$ depends only on the initial coordinates $\vec{z}$. This vector function is the transfer map of the beamline. It may also be thought of as an operator $\mathcal{M}(s)$ that converts initial phase space coordinates (at $s = 0$) into final ones: $\vec{z}(s) = \mathcal{M}(s)\vec{z}(0)$. Substitution in the equations of motion for $\vec{z}(s)$ imply corresponding formal equations of motion for this operator:

$$\frac{d\mathcal{M}(s)}{ds} = - :H: \mathcal{M}(s), \tag{9}$$

with the initial conditions $\mathcal{M}(0) = \mathcal{I}$, the identity. For a beamline element of length L in which the Hamiltonian H is constant, this equation may be formally integrated

$$\int_z^w \frac{dz}{z} = ln(\frac{w}{z}) \;=\; -\int_0^L ds\, :H: \;=\; -\,L\, :H: . \tag{10}$$

to give the transfer map operator as the exponential of the Hamiltonian Lie operator:

$$\mathcal{M} = e^{-L:H:} \tag{11}$$

The exponential of a Lie operator is called a "Lie Transformation", which may be defined (for arbitrary $f(\vec{z})$) by the usual power series for an exponential:

$$e^{:f:} \equiv \sum_{n=0}^{\infty} \frac{:f:^n}{n!} = 1 + :f: + \frac{1}{2}:f:^2 + \frac{1}{6}:f:^3 + \cdots \tag{12}$$

The application of this Lie transform to a general function $g(\vec{z})$ thus takes the form

$$e^{:f:}g = g + \{f, g\} + \frac{1}{2}\{f, \{f, g\}\} + \frac{1}{6}\{f, \{f, \{f, g\}\}\} + \cdots \tag{13}$$

Both $f(\vec{z})$ and $g(\vec{z})$ are evaluated at the same point of phase space. When the g function is one of the coordinates ($g(\vec{z}) = z_i$), the Poisson brackets simplify to the congugate derivatives

$$\{f, x_i\} = -\frac{\partial f}{\partial p_i} \qquad \text{and} \qquad \{f, p_i\} = \frac{\partial f}{\partial x_i} \tag{14}$$

The whole Lie transform then reduces to:

$$e^{:f:}x_i = x_i - \frac{\partial f}{\partial p_i} - \frac{1}{2}\{f, \frac{\partial f}{\partial p_i}\} + \cdots \qquad \text{and} \qquad e^{:f:}p_i = p_i - \frac{\partial f}{\partial x_i} - \frac{1}{2}\{f, \frac{\partial f}{\partial x_i}\} + \cdots \tag{15}$$

Lie Transform Maps

The Hamiltonian Lie map $\mathcal{M}$ has the symplectic symmetry required by classical mechanics. By the factorization theorem, any such map may be written as a product of Lie transformations of specified order: When the $f_m(\vec{z})$ are homogeneous polynomials of order m,

$$\mathcal{M} = e^{:f_4:}e^{:f_3:}e^{:f_2:} \tag{16}$$

is always a symplectic map. The number of monomials of order m in k variables is

$$N(m, k) = \frac{(m + k - 1)!}{m!(k - 1)!}. \tag{17}$$

For k=6, this says f_2 has 21 independent coefficients, f_3 has 56, and f_4 has 126. MARYLIE 3.0 uses these Lie polynomials up to f_4 to compute 3^{rd} order maps. MARYLIE-IMPACT computes to 5^{th} order using polynomials up to f_6. The term $:f_m:^n z$ in the power series (Eq.12) is of order $n(m-2)+1$.

This counting rule specifies the number of terms needed to reach a given order in the series expansion of the exponentials. The first two terms $e^{:f_m:} \simeq 1 + :f_m:$ are sufficient to 3^{rd} order for f_4, and to 5^{th} order for $m > 4$. For $m = 3$, $:f_3:^n z$ is of order $n+1$. The 3^{rd} order map uses only the first 3 terms in the expansion because n=4 is a fifth order term:

$$e^{:f_3:} \simeq 1 + :f_3: + \frac{1}{2}:f_3:^2 \tag{18}$$

For $m = 2$ this counting rule says $:f_2:^n z$ is of order 1 for all powers n. The whole infinite series then sums to the linear Lie Transform $\mathcal{R} \equiv e^{:f_2:}$. MARYLIE stores and uses the matrix representation of $\mathcal{R}$ in place of f_2.

Multipole Dependencies of the Lie Coefficients

To see the contribution of a given multipole to the overall map, we note that the Lie polynomials that represent a given element come from the series expansion of its Hamiltonian, which depends in turn on its vector potential $\mathbf{A}(\vec{z})$. The form of this vector potential follows directly from the transverse multipole solutions of Maxwell's equations. $\nabla \times \mathbf{B} = 0$ in cylindrical coordinates reads:

$$\frac{\partial B_z}{\partial \theta} = r\frac{\partial B_\theta}{\partial z}, \qquad \frac{\partial B_z}{\partial r} = \frac{\partial B_r}{\partial z}, \qquad \frac{\partial B_r}{\partial \theta} = \frac{\partial (rB_\theta)}{\partial r}. \tag{19}$$

Therefore B_z must be a constant (namely zero) in the main field where $\mathbf{B}$ is independent of z. Actual determination of the vector potential $\mathbf{A}$ requires a choice of gauge, because the gradient of an arbitrary scalar can be added to $\mathbf{A}$ without affecting $\mathbf{B}$. Without loss of generality, we may choose a gauge where $A_\theta = 0$. Then $\mathbf{B} = \nabla \times \mathbf{A}$ reads:

$$B_r = \frac{1}{r}\frac{\partial A_z}{\partial \theta}, \qquad B_\theta = \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}, \qquad B_z = -\frac{1}{r}\frac{\partial A_r}{\partial \theta}. \tag{20}$$

Thus $B_z = 0$ is consistent with $A_r = 0$, leaving only $A_z \neq 0$, with the simple solution for the m^{th} multipole

$$A_z = a_m(x + iy)^m = a_m r^m[cos(m\theta) + isin(m\theta)]. \tag{21}$$

The real part of this potential corresponds to a normal multipole, the imaginary part to a skew multipole. This simplicity is lost in the fringe field region where variation of $\mathbf{B}$ with z requires $B_z \neq 0$, hence $A_r \neq 0$.

The series expansion of the vector potential for an m^{th} order multipole thus begins with terms of order m in the transverse coordinates. Quadrupoles contribute to terms of order $m = 2$ and higher, sextupoles to $m = 3$ and higher, octupoles to $m = 4$ and higher, etc. Consequently, the Lie polynomial of order m that describes the map of an m^{th} order multipole is linear in its strength (field gradient).

Single Element Maps and Concatenation

The map of a whole beamline is built up by concatenating the maps of the individual elements. The concatenation of maps is defined by substituting the first mapping function inside the second one: If $\vec{z}_1 = \vec{F}(\vec{z}_0)$, and $\vec{z}_2 = \vec{G}(\vec{z}_1)$, then

$$\vec{H}(\vec{z}_0) \equiv \vec{G}(\vec{F}(\vec{z}_0)) \tag{22}$$

defines the operator product $H = GF$. This definition implies rules for concatenating the Lie polynomials that represent the maps:

$$h_3(\vec{z}) = g_3(\vec{z}) + f_3(G\vec{z}) \qquad \text{and} \qquad h_4(\vec{z}) = g_4(\vec{z}) + f_4(G\vec{z}) + \frac{1}{2}\{f_3(G\vec{z}), g_3(\vec{z})\} \tag{23}$$

These formulas imply that the concatenation process spreads the effects of a given order upward, but not downward. That is, the higher order parts of the map of a full beamline collect contributions from the lower order parts of the component maps, but the lower order parts of the full map are independent of the higher order parts of the component maps. Thus h_3, which here represents accumulated second order effects, also contains implicit first order effects due to the $G\vec{z}$ in the argument of f_3. Similarly h_4, representing accumulated third order effects, contains contributions from first and second order (f_3 and g_3) terms in the component maps.

Integrating the Hamiltonian

Substitution of the explicit form $\mathcal{M} = e^{:f_4:} e^{:f_3:} e^{:f_2:}$ into the equation of motion of the map reduces it to a set of coupled non-linear differential equations for the Lie polynomials. The Hamiltonian itself, which depends on the vector potential, must also be Taylor expanded into a series of homogeneous polynomials:

$$H(\vec{z}, s) = \sum_{n=1}^{\infty} H_n(\vec{z}, s). \tag{24}$$

Finally, the first order behavior of the system is taken out by transforming to the "interaction representation" in which only deviations from first order motion are considered. The Interaction Representation of the n^{th} order part of the Hamiltonian is:

$$H_n^I(z, s) \equiv H_n(R(s)\, z, s), \tag{25}$$

where $R(s)$ is the linear (matrix) part of the map $\mathcal{M}(s)$, whose motion is determined by $H_2(s)$ alone. After all of this, the resultant coupled non-linear differential equations for the Lie polynomials f_3 and f_4 are:

$$\frac{\partial f_3(z, s)}{\partial s} = -H_3^I(z, s) \qquad \text{and} \qquad \frac{\partial f_4(z, s)}{\partial s} = -H_4^I(z, s) + \frac{1}{2}\left[f_3(z, s), \frac{\partial f_3(z, s)}{\partial s}\right] \tag{26}$$

These equations are further separated power by power into coupled non-linear differential equations for the 209 polynomial coefficients (for m = 1, 2, 3, and 4) that represent the transfer map to third order. The integrals are done analytically for the library of hard edge elements. MARYLIE-IMPACT integrates 923 equations to generate 5^{th} order maps.

Taylor Maps

For calculation of final coordinates, the components of the transfer map $\vec{F}$ are expanded as a Taylor series in the components of the initial coordinates $\vec{z}$ (the indices a, b, c, d run from 1 to 6):

$$w_a = F_a(\vec{z}) = \sum_b R_{ab} z_b + \sum_{b,c} T_{abc} z_b z_c + \sum_{b,c,d} U_{abcd} z_b z_c z_d + \cdots \tag{27}$$

The fact that $\vec{z} = 0$ maps into $\vec{w} = 0$ is a consequence of taking the origin of coordinates as the position of a reference particle that flys an actual physical trajectory through the system. The R-matrix of linear optics is the first sum in this expansion. The second sum presents the second order corrections as a quadratic form in the six variables, for each of the six final components. Each quadratic form has 21 terms, so there are 126 second order Taylor map coefficients T_{abc}, which are linear combinations of the 56 Lie coefficients in f_3. The third sum expresses third order optics as 6 homogeneous cubics in 6 variables. The same counting rule shows $6 \times 56 = 336$ cubic coefficients U_{abcd}. All of these depend on the smaller number of Lie polynomial coefficients that represent the same map. By Eq 17, they are quadratic in the f_3 coefficients,and linear in the f_4.

The second order Taylor map may also be thought of as expressing the linear $\vec{z}$ dependence of the R-matrix:

$$w_a = \sum_b \left[R_{ab} + \vec{z} \cdot \vec{T}_{ab}\right] z_b \qquad \text{with} \qquad \vec{z} \cdot \vec{T}_{ab} \equiv \sum_c T_{abc} z_c. \tag{28}$$

This says the 2^{nd} order corrections to R_{ab} vanish in the phase space hyperplane orthogonal to $\vec{T}_{ab}$, offering the possiblity of avoiding them by imposing correlations that confine the initial particle set to this plane. But there are 36 such vectors and hyperplanes. The only point in *all* of them is generally the origin, so it is usually necessary to just find and eliminate the worst T_{abc}.

Second Order Maps

The factorization theorem implies that the map may be applied in successive orders: First apply the linear map $\mathcal{R}$ to convert the initial coordinates into ideal (linear) final coordinates, then apply the high-order corrections. This means the second order Lie polynomials are expressed in linear final coordinates. Write these intermediate linear coordinates as $\vec{v} \equiv (q_1, p_1, q_2, p_2, q_3, p_3)$. The components of $\vec{v}$ are defined by the linear R-matrix:

$$v_a \equiv \sum_b R_{ab} z_b \tag{29}$$

Application of the Lie transformation $e^{:f_3:}$ to this intermediate coordinate is

$$e^{:f_3:} v_a = v_a + \{f_3, v_a\} + \frac{1}{2}\{f_3, \{f_3, v_a\}\} + \cdots \tag{30}$$

The first term of a comparison with the Taylor map (Eq.27) verifies that v_a is given by the linear map. The order counting rule says the third term involving double Poisson brackets is the contribution of f_3 to the 3^{rd} order components of the map. The single Poisson bracket is the only quadratic form, so must be the second order map:

$$\{f_3(\vec{v}), v_a\} = \sum_{b,c} T_{abc} z_b z_c \tag{31}$$

Index a marks the final coordinate w_a, as well as v_a. Odd $a = 2n-1, (n = 1, 3)$ are position coordinates $(v_a = q_n)$, while even $a = 2n$ are momenta $(v_a = p_n)$. The Lie polynomial $f_3(\vec{v})$ is a homogeneous cubic in 6 variables, with 56 possible terms (many absent due to design symmetries). Since the Poisson bracket with a coordinate is the conjugate derivative, we write the p_n dependence explicitly in the form

$$f_3(\vec{v}) = A_n(\vec{v}) + B_n(\vec{v}) p_n + C_n(\vec{v}) {p_n}^2 + D_n {p_n}^3, \tag{32}$$

where $A_n(\vec{v})$ is also a cubic in all variables other than p_n, $B_n(\vec{v})$ is quadratic in these variables, $C_n(\vec{v})$ is linear, and D_n is a constant. The subscript n indicates a different arrangement of monomials for each plane. We then have for the coordinate terms:

$$\sum_{b,c} T_{abc} z_b z_c = \{f_3(\vec{v}), q_n\} = -\frac{\partial f_3}{\partial p_n} = -[B_n(\vec{v}) + 2C_n(\vec{v}) p_n + 3D_n {p_n}^2] \tag{33}$$

Note that the cubic term $A_n(\vec{v})$ has vanished due to lack of p_n dependence. Only quadratic terms are left. A corresponding process is needed for the even $(a = 2n)$ momentum terms. This time we arrange the cubic Lie generator in powers of q_n:

$$f_3(\vec{v}) = E_n(\vec{v}) + F_n(\vec{v}) q_n + G_n(\vec{v}) {q_n}^2 + H_n {q_n}^3, \tag{34}$$

and derive

$$\sum_{b,c} T_{abc} z_b z_c = \{f_3(\vec{v}), p_n\} = +\frac{\partial f_3}{\partial p_n} = F_n(\vec{v}) + 2G_n(\vec{v}) q_n + 3H_n {q_n}^2. \tag{35}$$

A term by term comparison of powers of z_a in these quadratic forms then relates the second order Taylor map coeficients T_{abc} to the Lie coeficients. The monomial counting rule (Eq. 18) says the quadratics $B_n(\vec{v})$ and $F_n(\vec{v})$ contain 15 Lie coefficients each, the vectors $C_n(\vec{v})$ and $G_n(\vec{v})$ contain 5, while the constants D_n and H_n are a single coefficient.

Because of the substitution $v_a \equiv \sum_b R_{ab} z_b$, the coefficients of these linear combinations of Lie monomials are themselves quadratic forms in the matrix elements R_{ab}.

The most general case usually simplifies greatly in practice, because many terms vanish by design or by symmetry in real beamlines. For telescopes, the momentum kicks due to high-order aberrations are well defined functions of unperturbed (x,y) position in the output aperture. These are very close to the actual final positions. For the $-I$ identity lenses used in proton radiography, the linear "Interaction Representation" coordinates are just the input coordinates with the sign flipped. For the EEX project however, these intermediate coordinates are very different from the input coordinates because of the linear exchange of the transverse (x-plane) and longitudinal (t-plane) phase spaces.

Second Order Lie/Taylor Map Comparisons

The following table lists the Lie coefficients used in each of the six Taylor map quadratic forms, labled T1 through T6. Note the alternating order due to use of the conjugate derivatives in the Poisson brackets. The coefficient $f(n)$ is followed by a compressed list of exponents for the linear coordinates v_a. All 56 map coefficients are used, but only 21 for any given Tn component. The first row lists the six that are only used once each as the constants D and H. Thirty more occur twice each in the next ten rows - once in the C and F vectors, and again in the diagonals of the quadratic forms B and E. Finally, twenty more Lie coefficients occur three times each in the last ten rows specifying the bilinear cross terms in the quadratic forms. In typical practical cases, more than half of these Lie coefficients will be negligible, or exactly zero.

Lie Coefficients for 2nd Order Taylor Map

T2	T1	T4	T3	T6	T5
constants D or H					
f(28):300000	f(49):030000	f(64):003000	f(74):000300	f(80):000030	f(83):000003
5 linear terms in C or F vectors					
f(29):210000	f(34):120000	f(39):102000	f(43):100200	f(46):100020	f(48):100002
f(30):201000	f(50):021000	f(54):012000	f(58):010200	f(61):010020	f(63):010002
f(31):200100	f(51):020100	f(65):002100	f(68):001200	f(71):001020	f(73):001002
f(32):200010	f(52):020010	f(66):002010	f(75):000210	f(77):000120	f(79):000102
f(33):200001	f(53):020001	f(67):002001	f(76):000201	f(81):000021	f(82):000012
5 diagonal terms in B or E quadratic forms					
f(34):120000	f(29):210000	f(30):201000	f(31):200100	f(32):200010	f(33):200001
f(39):102000	f(54):012000	f(50):021000	f(51):020100	f(52):020010	f(53):020001
f(43):100200	f(58):010200	f(68):001200	f(65):002100	f(66):002010	f(67):002001
f(46):100020	f(61):010020	f(71):001020	f(77):000120	f(75):000210	f(76):000201
f(48):100002	f(63):010002	f(73):001002	f(79):000102	f(82):000012	f(81):000021
10 bilinear cross terms in B or E quadratic forms					
f(35):111000	f(35):111000	f(35):111000	f(36):110100	f(37):110010	f(38):110001
f(36):110100	f(36):110100	f(40):101100	f(40):101100	f(41):101010	f(42):101001
f(37):110010	f(37):110010	f(41):101010	f(44):100110	f(44):100110	f(45):100101
f(38):110001	f(38):110001	f(42):101001	f(45):100101	f(47):100011	f(47):100011
f(40):101100	f(55):011100	f(55):011100	f(55):011100	f(56):011010	f(57):011001
f(41):101010	f(56):011010	f(56):011010	f(59):010110	f(59):010110	f(60):010101
f(42):101001	f(57):011001	f(57):011001	f(60):010101	f(62):010011	f(62):010011
f(44):100110	f(59):010110	f(69):001110	f(69):001110	f(69):001110	f(70):001101
f(45):100101	f(60):010101	f(70):001101	f(70):001101	f(72):001011	f(72):001011
f(47):100011	f(62):010011	f(72):001011	f(78):000111	f(78):000111	f(78):000111