Discrete Mathematics and Theoretical Computer Science **1**, 1997, 149–157

A Lie connection between Hamiltonian and Lagrangian optics

Alex J. Dragt

Physics Department, University of Maryland, College Park, MD 20742, USA

It is shown that there is a non-Hamiltonian vector field that provides a Lie algebraic connection between Hamiltonian and Lagrangian optics. With the aid of this connection, geometrical optics can be formulated in such a way that all aberrations are attributed to ray transformations occurring only at lens surfaces. That is, in this formulation there are no aberrations arising from simple transit in a uniform medium. The price to be paid for this formulation is that the Lie algebra of Hamiltonian vector fields must be enlarged to include certain non-Hamiltonian vector fields. It is shown that three such vector fields are required at the level of third-order aberrations, and sufficient machinery is developed to generalize these results to higher order.

Keywords: Lie algebra, Hamiltonian and Lagrangian optics

1 Introduction

In Hamiltonian optics rays are described using the Hamiltonian

$$H = -(n^2 - \boldsymbol{p}^2)^{1/2} \tag{1}$$

The use of a Hamiltonian formulation is advantageous because Hamiltonian flows produce symplectic maps, and there is a well developed calculus, using both characteristic functions and Lie algebraic methods, for handling symplectic maps in an efficient and economical way [1, 2].

However, the use of a Hamiltonian approach has the consequence, perhaps at first surprising, that the map describing simple transit in a uniform medium (*free flight* in optical parlance, and a *drift* in accelerator parlance), is nonlinear. Therefore, in a Hamiltonian approach to optics, aberrations (nonlinearities) arise not only from transfer maps associated with lens interfaces, but also from simple transit within and between lenses. This circumstance is perhaps of less consequence in graded index optics, where one expects to have aberration effects associated with transit (at least within lenses), but it might be viewed as a drawback for the use of Hamiltonian methods for optics involving only uniform media.

Let $\mathcal{D}$ be the map for transit (drift) over a distance ℓ in a uniform medium having refractive index n. Then, in Lie algebraic notation and using canonical coordinates, this map may be written in the form

$$\mathcal{D} = \exp :\ell(n^2 - \boldsymbol{p}^2)^{1/2}: \tag{2}$$

for the action of this map on $\boldsymbol{q}$ and $\boldsymbol{p}$ we find the result

$$\mathcal{D}\boldsymbol{q} = \boldsymbol{q} + \ell\boldsymbol{p}/(n^2 - \boldsymbol{p}^2)^{1/2} \tag{3}$$

$$\mathcal{D}\boldsymbol{p} = \boldsymbol{p} \tag{4}$$

We observe, as advertised, that relation (3) is nonlinear.

Instead of using the canonical coordinates $\boldsymbol{q}, \boldsymbol{p}$ to specify a point in phase (ray) space, as is done in the Hamiltonian formulation, one might instead use the Lagrangian coordinates $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$. From (1) we have the result

$$\dot{\boldsymbol{q}} = \partial H/\partial \boldsymbol{p} = \boldsymbol{p}/(n^2 - \boldsymbol{p}^2)^{1/2} \tag{5}$$

Suppose we let the map $\mathcal{D}$ act on the pair $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$. From (3)–(5) we find the result

$$\mathcal{D}\boldsymbol{q} = \boldsymbol{q} + \ell\dot{\boldsymbol{q}} \tag{6}$$

$$\mathcal{D}\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}} \tag{7}$$

We see that in Langrangian coordinates, unlike the case of canonical Hamiltonian coordinates, the map for simple transit is completely linear. Consequently, in a Lagrangian formulation of optics and in the case of uniform media, all aberrations arise only from transfer maps associated with lens interfaces. The purpose of this note is to study the relation between Hamiltonian and Lagrangian coordinates and dynamics in more detail.

2 Transformation between Hamiltonian and Lagrangian Coordinates

To describe the relation between Hamiltonian and Lagrangian optics, introduce the transformation $\mathcal{T}$ with the property

$$\mathcal{T}\boldsymbol{q} = \boldsymbol{q} \tag{8}$$

$$\mathcal{T}\boldsymbol{p} = \dot{\boldsymbol{q}} = \boldsymbol{p}/(n^2 - \boldsymbol{p}^2)^{1/2} \tag{9}$$

Thus, $\mathcal{T}$ maps the Hamiltonian variables $\boldsymbol{q}, \boldsymbol{p}$ into the Langrangian variables $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$. Note from (8) and (9) that $\mathcal{T}$ is not symplectic, since it does not preserve the fundamental Poisson brackets between $\boldsymbol{q}$ and $\boldsymbol{p}$.

Let us employ $\mathcal{T}$ to rewrite relations (3), (4), (6) and (7). From (6), (7), (8) and (9), we deduce the relations

$$\begin{aligned} \mathcal{D}\mathcal{T}\boldsymbol{q} &= \mathcal{D}\boldsymbol{q} = \boldsymbol{q} + \ell\dot{\boldsymbol{q}} \\ &= \mathcal{T}\boldsymbol{q} + \ell\mathcal{T}\boldsymbol{p} = \mathcal{T}(\boldsymbol{q} + \ell\boldsymbol{p}) \\ &= \mathcal{T}\exp[(-\ell/2) :\boldsymbol{p}^2:]\boldsymbol{q} \end{aligned} \tag{10}$$

$$\begin{aligned} \mathcal{D}\mathcal{T}\boldsymbol{p} &= \mathcal{D}\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}} = \mathcal{T}\boldsymbol{p} \\ &= \mathcal{T}\exp[(-\ell/2) :\boldsymbol{p}^2:]\boldsymbol{p} \end{aligned} \tag{11}$$

Upon comparing (2), (10) and (11), we obtain the operator result

$$\mathcal{T}^{-1}\mathcal{D}\mathcal{T} = \exp[(-\ell/2) : \boldsymbol{p}^2:] \tag{12}$$

or, more explicitly,

$$\mathcal{T}^{-1}\exp[: \ell(n^2 - \boldsymbol{p}^2)^{1/2} :]\mathcal{T} = \exp[(-\ell/2) :\boldsymbol{p}^2:] \tag{13}$$

3 Explicit Form for Transformation

It would be useful to have $\mathcal{T}$ itself in explicit operator form. Although we know that $\mathcal{T}$ is not symplectic, it may still be possible to write $\mathcal{T}$ in Lie form with the aid of some non-Hamiltonian vector field. Consider the vector field V defined by the equation

$$V = \boldsymbol{p}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) \tag{14}$$

Evidently, V has the properties

$$V\boldsymbol{q} = 0 \tag{15}$$

$$V\boldsymbol{p} = \boldsymbol{p}^2\boldsymbol{p} \tag{16}$$

Let us use V to generate an autonomous flow parameterized by an independent variable that we will call t. Doing so gives the differential equations

$$\boldsymbol{q}' = V\boldsymbol{q} = 0 \tag{17}$$

$$\boldsymbol{p}' = V\boldsymbol{p} = \boldsymbol{p}^2\boldsymbol{p} \tag{18}$$

Here we have used a prime to denote the differentiation (d/dt).

Let us integrate (17) and (18). In doing so, we will find that we are well on our way to discovering an explicit representation for $\mathcal{T}$. From (18) we find the result

$$\boldsymbol{p}\cdot\boldsymbol{p}' = (d/dt)(1/2)(\boldsymbol{p}^2) = (\boldsymbol{p}^2)^2 \tag{19}$$

Let ρ denote the quantity $\boldsymbol{p}^2$,

$$\rho = \boldsymbol{p}^2 \tag{20}$$

With this notation, (19) can be written in the forms

$$\rho' = 2\rho^2 \tag{21}$$

or

$$d\rho/\rho^2 = 2dt \tag{22}$$

Equation (22), with the initial condition $\rho(0) = \rho^0$, can be integrated by quadrature to give the result

$$\rho(t) = \rho^0(1 - 2t\rho^0)^{-1} \tag{23}$$

Next substitute (23) into (18). Then, for the components p_j of $\boldsymbol{p}$ we find the differential equations

$$p_j' = \rho^0(1 - 2t\rho^0)^{-1}p_j \tag{24}$$

These equations can be rewritten in the form

$$dp_j/p_j = \rho^0(1 - 2t\rho^0)^{-1}dt \tag{25}$$

and thus can also be integrated by quadrature. The results, with the initial conditions $p_j(0) = p_j^0$, are the relations

$$\log(p_j/p_j^0) = -(1/2)\log(1 - 2t\rho^0) \tag{26}$$

or

$$p_j = p_j^0/(1-2t\rho^0)^{1/2} \tag{27}$$

Finally, upon combining (20) and (27), we find the result

$$\boldsymbol{p}(t) = \boldsymbol{p}^0/[1-2t(\boldsymbol{p}^0)^2]^{1/2} \tag{28}$$

In terms of vector fields, (28) is equivalent to the result

$$\exp(tV)\boldsymbol{p} = \boldsymbol{p}/[1-2t\boldsymbol{p}^2]^{1/2} \tag{29}$$

which has as a special case the result

$$\exp[(1/2)V]\boldsymbol{p} = \boldsymbol{p}/[1-\boldsymbol{p}^2]^{1/2} \tag{30}$$

Note that the left-hand side of (30) resembles the left-hand side of (9). In addition, we immediately have from (15) the result

$$\exp[(1/2)V]\boldsymbol{q} = \boldsymbol{q} \tag{31}$$

Next, let W be the vector field defined by the equation

$$W = \boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) \tag{32}$$

For this field we have the result

$$W\boldsymbol{q} = 0 \tag{33}$$

$$W\boldsymbol{p} = \boldsymbol{p} \tag{34}$$

The vector field W can also be used to generate a flow. For this flow we get the result

$$\exp[-(\log n)W]\boldsymbol{q} = \boldsymbol{q} \tag{35}$$

$$\exp[-(\log n)W]\boldsymbol{p} = \exp(-\log n)\boldsymbol{p} = (1/n)\boldsymbol{p} \tag{36}$$

Now consider the joint effect of $\exp[-(\log n)W]$ and $\exp[(1/2)V]$.
From (31) and (35) we find the relation

$$\exp[-(\log n)W]\exp[(1/2)V]\boldsymbol{q} = \boldsymbol{q} \tag{37}$$

From (30) and (36) we find

$$\begin{aligned} & \exp\,[-(\log n)W]\exp[(1/2)V]\boldsymbol{p} \\ = \; & \exp\,[-(\log n)W]\{\boldsymbol{p}/[1-\boldsymbol{p}^2]^{1/2}\} \\ = \; & \boldsymbol{p}/[n^2-\boldsymbol{p}^2]^{1/2} \end{aligned} \tag{38}$$

Upon comparing (8) and (9) with (37) and (38), we conclude that $\mathcal{T}$ is given by the operator relation

$$\mathcal{T} = \exp[-(\log n)W]\exp[(1/2)V] \tag{39}$$

4 Variation

Consider the map $\mathcal{U}$ defined by writing

$$\begin{aligned}\mathcal{U} &= \mathcal{T}\exp[(\log n)W] \\ &= \exp[-(\log n)W]\exp[(1/2)V]\exp[(\log n)W] \end{aligned} \tag{40}$$

From (35) and (36) we immediately have the relations

$$\exp[(\log n)W]\boldsymbol{q} = \boldsymbol{q} \tag{41}$$

$$\exp[(\log n)W]\boldsymbol{p} = n\boldsymbol{p} \tag{42}$$

Combining (8) with (41) gives the result

$$\mathcal{U}\boldsymbol{q} = \boldsymbol{q} \tag{43}$$

Combining (9) with (42) gives the result

$$\mathcal{U}\boldsymbol{p} = \mathcal{T}n\boldsymbol{p} = n\mathcal{T}\boldsymbol{p} = n\dot{\boldsymbol{q}} \tag{44}$$

Note that $\mathcal{U}$, like $\mathcal{T}$, is also not symplectic.

Suppose we use $\mathcal{U}$ to rewrite the relations (3), (4), (6) and (7). From (6), (7), (43) and (44), we find the results

$$\begin{aligned}\mathcal{D}\mathcal{U}\boldsymbol{q} &= \mathcal{D}\boldsymbol{q} = \boldsymbol{q} + \ell\dot{\boldsymbol{q}} \\ &= \mathcal{U}\boldsymbol{q} + (\ell/n)\mathcal{U}\boldsymbol{p} = \mathcal{U}[\boldsymbol{q} + (\ell/n)\boldsymbol{p}] \\ &= \mathcal{U}\exp\{[-\ell/(2n)]\ {:}\boldsymbol{p}^2{:}\}\boldsymbol{q} \end{aligned} \tag{45}$$

$$\begin{aligned}\mathcal{D}\mathcal{U}\boldsymbol{p} &= \mathcal{D}n\dot{\boldsymbol{q}} = n\mathcal{D}\dot{\boldsymbol{q}} \\ &= n\dot{\boldsymbol{q}} = \mathcal{U}\boldsymbol{p} \\ &= \mathcal{U}\exp\{-[\ell/(2n)]\ {:}\boldsymbol{p}^2{:}\}\boldsymbol{p} \end{aligned} \tag{46}$$

Consequently, we have the operator relation

$$\mathcal{U}^{-1}\mathcal{D}\mathcal{U} = \exp\{-[\ell/(2n)]\ {:}\boldsymbol{p}^2{:}\} \tag{47}$$

or, more explicitly,

$$\mathcal{U}^{-1}\exp[{:}\ \ell(n^2 - \boldsymbol{p}^2)^{1/2}\ {:}]\mathcal{U} = \exp\ {:}-[\ell/(2n)]\boldsymbol{p}^2{:} \tag{48}$$

The relations (47) and (48) are particularly appealing because they relate the exact map for transit to the map for transit in the Gaussian approximation.

According to (40), $\mathcal{U}$ can also be written in exponential form. Indeed, we see from (40) that $\mathcal{U}$ is written in the form of a similarity transformation of the map $\exp[(1/2)V]$, and therefore it should be particularly easy to combine all exponents into a single exponent. This is indeed the case. One procedure for doing so is to use the commutation rule

$$\{W, V\} = WV - VW = 2V \tag{49}$$

which is easily checked from the definitions (14) and (32). A more direct approach, based on (15) and (29), is to observe the relations

$$\exp\{[1/(2n^2)]V\}\boldsymbol{q} = \boldsymbol{q} \tag{50}$$

$$\begin{aligned}\exp\{[1/(2n^2)]V\}\boldsymbol{p} &= \boldsymbol{p}/[1-n^{-2}\boldsymbol{p}^2]^{1/2}\\ &= n\boldsymbol{p}/[n^2-\boldsymbol{p}^2]^{1/2} = n\dot{\boldsymbol{q}}\end{aligned} \tag{51}$$

Upon comparing (43) with (50) and (44) with (51), we find the immediate result

$$\mathcal{U}(n) = \exp\{[1/(2n^2)]V\} \tag{52}$$

Here, for future use, we have explicitly noted that $\mathcal{U}$ depends on n.

5 Applications

Consider an optical system composed of a collection of drifts and lenses. Also suppose, for simplicity, that the system begins and ends with drifts. Then the map for the complete system is a product of maps of the form

$$\mathcal{M} = \mathcal{D}_1\mathcal{S}_1\mathcal{D}_2\mathcal{S}_2\mathcal{D}_3\mathcal{S}_3\cdots\mathcal{D}_{m-1}\mathcal{S}_{m-1}\mathcal{D}_m \tag{53}$$

Here, as before, the $\mathcal{D}_j$ denote drift maps (both between and within lenses), and the $\mathcal{S}_j$ denote *surface* transfer maps describing the effects of lens interfaces. For the drift maps we may use the relations (12) and (48) to write the representations

$$\mathcal{D} = \mathcal{T}\exp[(-\ell/2)\ {:}\boldsymbol{p}^2{:}]\mathcal{T}^{-1} \tag{54}$$

$$\mathcal{D} = \mathcal{U}\exp\{-[\ell/(2n)]\ {:}\boldsymbol{p}^2{:}\}\mathcal{U}^{-1} \tag{55}$$

Note that the maps on the right-hand sides of (12) and (48) are both linear.

For simplicity, we will make use of (55). Let us write (55) in the form

$$\mathcal{D} = \mathcal{U}\mathcal{D}^g\mathcal{U}^{-1} \tag{56}$$

where $\mathcal{D}^g$ denotes the map for Gaussian drift,

$$\mathcal{D}^g = \exp\{-[\ell/(2n)]\ {:}\boldsymbol{p}^2{:}\} \tag{57}$$

With this representation, the product (53) can be rewritten in the form

$$\mathcal{M} = \mathcal{U}_1\mathcal{D}_1^g\mathcal{U}_1^{-1}\mathcal{S}_1\mathcal{U}_2\mathcal{D}_2^g\mathcal{U}_2^{-1}\mathcal{S}_2\mathcal{U}_3^{-1}\cdots\mathcal{U}_{m-1}^{-1}\mathcal{S}_{m-1}\mathcal{U}_m\mathcal{D}_m^g\mathcal{U}_m^{-1} \tag{58}$$

Here we have also used the notation

$$\mathcal{U}_j = \mathcal{U}(n_j) \tag{59}$$

See (52). Next, define modified surface maps $\tilde{\mathcal{S}}_{j-1}$ by the relations

$$\tilde{\mathcal{S}}_{j-1} = \mathcal{U}_{j-1}^{-1}\mathcal{S}_{j-1}\mathcal{U}_j \tag{60}$$

Note that these modified maps are also entitled to be called surface maps, since they depend only on the shape of the lens surface and the values of the surrounding indices of refraction. With the aid of the definitions (60), we can rewrite (58) to obtain a representation for $\mathcal{M}$ of the form

$$\mathcal{M} = \mathcal{U}_1[\mathcal{D}_1^g \tilde{\mathcal{S}}_1 \mathcal{D}_2^g \tilde{\mathcal{S}}_2 \cdots \mathcal{D}_{m-1}^g \tilde{\mathcal{S}}_{m-1} \mathcal{D}_m^g]\mathcal{U}_m^{-1} \tag{61}$$

In this form we see that $\mathcal{M}$, apart from the initial and final maps $\mathcal{U}_1$ and $\mathcal{U}_m^{-1}$, can be viewed as the result of Gaussian transits (which yield linear maps, and thus are not a source of aberrations), and modified surface maps (which, in this representation, are the source of all aberrations). That is, in this form all aberrations arise from surface effects. Note also that if $n_1 = n_m$, as is often the case, then $\mathcal{U}_1 = \mathcal{U}_m$.

6 Complications

What are the Lie algebraic generators for the modified surface maps $\tilde{\mathcal{S}}_j$? Suppose, for the present purposes, we restrict our discussion to Gaussian and third order. We know that the original surface maps $\mathcal{S}_j$ have (assuming axial symmetry) the Gaussian generator $:\boldsymbol{q}^2:$ and the third-order aberration generators $:(\boldsymbol{q}^2)^2:$, $:\boldsymbol{q}^2\boldsymbol{q}\cdot\boldsymbol{p}:$, and $:\boldsymbol{q}^2\boldsymbol{p}^2:$. [Other Gaussian generators such as $:\boldsymbol{p}^2:$ and $:\boldsymbol{p}\cdot\boldsymbol{q}:$ and other third-order generators such as $:(\boldsymbol{q}\cdot\boldsymbol{p})^2:$, $:\boldsymbol{p}^2\boldsymbol{q}\cdot\boldsymbol{p}:$, and $:(\boldsymbol{p}^2)^2:$ are allowed by axial symmetry and the symplectic condition, but in fact do not occur.] According to (60) and (52), if we wish to use the $\tilde{\mathcal{S}}_j$ we must add V to the list of generators; and we must also add all the generators that arise from commuting V with the generators we already have.

Let L_- denote the Lie operator given by the relation

$$L_- = :\boldsymbol{q}^2/2: = \boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \tag{62}$$

A simple computation gives the results

$$\{L_-, V\} = 2\boldsymbol{q}\cdot\boldsymbol{p}\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) + \boldsymbol{p}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \tag{63}$$

$$\{L_-, \{L_-, V\}\} = 2\boldsymbol{q}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) + 4\boldsymbol{q}\cdot\boldsymbol{p}\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \tag{64}$$

$$\{L_-, \{L_-, \{L_- V\}\}\} = 6\boldsymbol{q}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \tag{65}$$

We note that V and the generators given by (63)–(65) contribute only to third (and higher) order aberrations. Therefore, we need not consider the commutators of them among themselves or with third-order aberration generators of the form $:(\boldsymbol{q}^2)^2:$, $:\boldsymbol{q}^2\boldsymbol{q}\cdot\boldsymbol{p}:$, etc. since the Lie operators thereby produced generate even higher order aberrations. Also, we note the relation

$$:(\boldsymbol{q}^2)^2: = 4\boldsymbol{q}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \tag{66}$$

Therefore, the Lie operator given by (65) is not new. Consequently, we conclude that the $\tilde{\mathcal{S}}_j$ are generated by the Lie operators $:\boldsymbol{q}^2:$, $:(\boldsymbol{q}^2)^2:$, $:\boldsymbol{q}^2\boldsymbol{q}\cdot\boldsymbol{p}:$, $:\boldsymbol{q}^2\boldsymbol{p}^2:$, V, and those given by the right-hand sides of (63) and (64).

Suppose, in the course of calculation, we wish to combine various $\tilde{\mathcal{S}}_j$ with various $\mathcal{D}_k^g$. Then we must also consider the Gaussian generator $:\boldsymbol{p}^2:$ and its commutators with the generators of the previous paragraph. Define two other Gaussian generators L_+ and L_0 by the relations

$$L_+ = -:\boldsymbol{p}^2/2: = \boldsymbol{p}\cdot(\partial/\partial\boldsymbol{q}) \tag{67}$$

$$L_0 = :\boldsymbol{p}\cdot\boldsymbol{q}/2: = (1/2)\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) - (1/2)\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{q}) \tag{68}$$

We find the results

$$\{L_0, V\} = V \tag{69}$$

$$\{L_+, V\} = -\boldsymbol{p}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{q}) \tag{70}$$

Also, we note the relation

$$:(\boldsymbol{p}^2)^2: = -4\boldsymbol{p}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{q}) \tag{71}$$

Therefore, the Lie operators given by (69) and (70) are not new. We conclude that the new operators are V, $\{L_-, V\}$, and $\{L_-, \{L_-, V\}\}$.

These three new operators, along with those already known from the usual Hamiltonian treatment, can be organized in such a way that they have pleasing transformation properties under the action of L_+, L_- and L_0. Define the operator R_1^1 by the relation

$$R_1^1 = V - (1/4)\ :\boldsymbol{p}^2\boldsymbol{p}\cdot\boldsymbol{q}: \tag{72}$$

Then it is easily verified from (63)–(66) and (69)–(72) that R_1^1 has the properties

$$\{L_0, R_1^1\} = R_1^1 \tag{73}$$

$$\{L_+, R_1^1\} = 0 \tag{74}$$

$$\{L_-, \{L_-, \{L_-, R_1^1\}\}\} = 0 \tag{75}$$

Now define operators R_0^1 and R_{-1}^1 by the relations

$$\{L_-, R_1^1\} = (2)^{1/2}R_0^1 \tag{76}$$

$$\{L_-, R_0^1\} = (2)^{1/2}R_{-1}^1 \tag{77}$$

From these definitions we find that the R_m^1 obey the relations

$$\{L_0, R_m^1\} = mR_m^1 \tag{78}$$

$$\{L_-, R_m^1\} = [(1+m)(1-m+1)]^{1/2}R_{m-1}^1 \tag{79}$$

$$\{L_+, R_m^1\} = [(1-m)(1+m+1)]^{1/2}R_{m+1}^1 \tag{80}$$

As the notation is meant to suggest, these relations are analogous to those for a spin 1 object in the Lie algebraic treatment of angular momentum.

Indeed, the connection with spin 1 can also be made in another way. Let L be any Lie operator. For any L define an associated *superoperator*, denoted by $\#L\#$, that acts on any other Lie operator K by the rule

$$\#L\#K = \{L, K\} \tag{81}$$

The quantity $\#L\#$ is sometimes called the *adjoint* of L. Products and powers of superoperators are defined by repeated commutation. For example, $(\#L\#)^2$ is defined by

$$(\#L\#)^2K = \#L\#\#L\#K = \#L\#\{L, K\} = \{L, \{L, K\}\} \tag{82}$$

Now consider the 'Casimir' operator $\mathcal{C}$ defined by

$$\mathcal{C} = (\#L_0\#)^2 + (1/2)\#L_+\#\#L_-\# + (1/2)\#L_-\#\#L_+\# \tag{83}$$

Then one finds from (78)–(80) the result

$$\mathcal{C}R^1_m = 2R^1_m = (1)(1+1)R^1_m \tag{84}$$

Note that the factor $(1)(1+1)$ is of the form $j(j+1)$ with $j = 1$.

Finally, for the record, let us list the explicit forms of the R^1_m. We find the results

$$R^1_1 = (3/4)\boldsymbol{p}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) + (1/2)(\boldsymbol{p}\cdot\boldsymbol{q})\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{q}) + (1/4)\boldsymbol{p}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{q}) \tag{85}$$

$$\begin{aligned} R^1_0 = (2)^{-1/2}\boldsymbol{q}\cdot\boldsymbol{p}\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) &+ (2)^{-1/2}\boldsymbol{q}\cdot\boldsymbol{p}\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{q}) + (8)^{-1/2}\boldsymbol{p}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) \\ &+ (8)^{-1/2}\boldsymbol{q}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{q}) \end{aligned} \tag{86}$$

$$R^1_{-1} = (3/4)\boldsymbol{q}^2\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{q}) + (1/2)(\boldsymbol{p}\cdot\boldsymbol{q})\boldsymbol{q}\cdot(\partial/\partial\boldsymbol{p}) + (1/4)\boldsymbol{q}^2\boldsymbol{p}\cdot(\partial/\partial\boldsymbol{p}) \tag{87}$$

We conclude from the work of this section that the Gaussian and third-order treatment of the modified surface maps $\tilde{S}_j$ requires the introduction of the three additional non-Hamiltonian Lie operators R^1_m. This is the price to be paid at third order for a formalism that attributes all aberrations to (modified) surface maps. Evidently, our discussion can be extended to still higher order, and at these higher orders still more additional non-Hamiltonian Lie operators will be required. Of course, when all these operators (both Hamiltonian and non-Hamiltonian) are combined to compute $\mathcal{M}$ as in (61), then all the net coefficients of all the non-Hamiltonian operators must vanish identically because we know that $\mathcal{M}$, since it is a symplectic map, is expressible entirely in terms of Hamiltonian Lie operators.

Acknowledgements

This work was supported in part by Department of Energy grant DEFG0592-ER40748. It was begun at the Netherlands Research Institute for Applications of Computer Algebra, and completed at the Stanford Linear Accelerator Center. The support and hospitality of these two institutions is also gratefully acknowledged.

References

[1] Dragt, A. J. (1982). A Lie Algebraic Theory of Geometrical Optics and Optical Aberrations. *J. Opt. Sci. Am.* **72** 372.

[2] Dragt, A. J., Forest, E. and Wolf, K. (1986). Foundations of a Lie Algebraic Theory of Geometrical Optics. In: Mondragon, J. S., Wolf, K. B., editors, *Lie Methods in Optics*. Springer-Verlag. (See also A. J. Dragt and E. Forest, *Foundations of a Lie Algebraic Theory of Geometrical Optics*, University of Maryland Physics Department, Technical report.)