8th February 1999 SRS/APN/99/130

Lie Operator Methods for Single Particle Dynamics

A Wolski
CLRC Daresbury Laboratory

Abstract

Lie operators provide a mathematical tool for studying the dynamics of a wide variety of systems. They are particularly appropriate to the dynamics of particles in accelerators, as they can be used to make approximations allowing real insight into the effects of nonlinear elements. The single-turn map, developed from Lie operators, is a powerful tool for studying the dynamics of particles in circular accelerators. In this note, we review Hamiltonian mechanics, show how Lie operators appear naturally from the theory, and explore some of their important properties. We go on to derive some of the results of linear accelerator physics using Lie operators and mapping methods, and show how the theory extends to cover nonlinear systems. The methods can be used to study the motion of a particle in the full three degrees of freedom, although in this introductory discussion, we limit ourselves to motion in one dimension, transverse to the design orbit. In this context, the procedure of normalisation and the concept of a nonlinear rotation suggest ways to build simple models of complex lattices, allowing a more analytic approach to the study of such systems. We compare the predictions of a simple mapping model, with the results of a more rigorous simulation of a lattice using a conventional tracking code. Initial investigations suggest that Lie operators and mapping methods could provide useful tools for the study of nonlinear behaviour. This could be important for the minimisation of, or allowance made for, nonlinear effects arising, for example, from chromaticity correction using sextupoles.

Contents

1. Introduction

Traditional approaches to particle dynamics in synchrotron accelerators begin by using Newtonian mechanics to write down the equation of motion of particles in various magnetic elements. The equation of motion (Hill's equation) is then integrated to find the paths of particles through the lattice. There are a number of difficulties associated with the Newtonian method. In particular, finding the paths of particles through sextupoles and higher order elements involves solving non-linear differential equations, which gives little insight into the dynamics. It is similarly difficult to incorporate coupling between degrees of freedom in the motion, and transverse and longitudinal motions are treated separately. Hamiltonian mechanics provides a more appropriate formalism for analysing the dynamics in many cases where the Newtonian equations are difficult to solve. In this section we sketch an overview of Hamiltonian mechanics, and define Lie operators which are the mathematical tools we use in the rest of this note for mapping particles from one position in a magnetic lattice to another. We

cannot here give more than a very brief summary of the results of Hamiltonian mechanics relevant to the development of accelerator physics using Lie operators. For further details, the reader is referred to standard texts on classical mechanics, such as Goldstein [1]. It will become apparent that the Lie operator techniques afford much greater insight into the motion of particles in an accelerator, particularly when the motion is nonlinear. Furthermore, the longitudinal motion and effects such as coupling are intrinsic to the theory, rather than requiring separate treatment.

1.1 Overview of Hamiltonian Mechanics.

In Hamiltonian mechanics, the state of a particle is given by its co-ordinates x_i, and momenta p_i. There exists a formal definition of the momenta [1], which are properly called the canonical momenta, or the momenta canonical to x_i. For the present work, the relationship $p_i = m\dot{x}_i$ holds (though not generally true). The *Hamiltonian* $H(x_i, p_i)$ is a function of the co-ordinates and momenta such that:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \tag{1}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}. \tag{2}$$

Equations (1) and (2) are known as Hamilton's equations. They can be extended to cover the case that the Hamiltonian is also an explicit function of time. It can easily be seen that familiar Newtonian mechanics is recovered from Hamilton's equations in the case:

$$H = T + V \tag{3}$$

where $T = \sum_i \frac{p_i^2}{2m}$ is the kinetic energy of the particle, and $V = V(x_i)$ is the potential energy. Therefore, in the case that the Hamiltonian is time-independent and there is no velocity dependent potential, evaluating the Hamiltonian gives a constant value equal to the total energy of the particle.

From Hamilton's equations, we can derive a general expression for the time dependence of any function $f(x_i, p_i)$ of the co-ordinates and momenta, thus:

$$\begin{aligned} \frac{df}{dt} &= \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\frac{dx_i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}\right) \\ &= \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial x_i}\right) \end{aligned} \tag{4}$$

where N is the total number of degrees of freedom of the system. The form of the last line of equation (4) leads us to define the *Poisson bracket* $[f, g]$ of two functions $f(x_i, p_i)$ and $g(x_i, p_i)$:

$$[f, g] = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x_i}\right). \tag{5}$$

In terms of the Poisson bracket, equation (4) for the time evolution of any function can be rewritten as:

$$\frac{df}{dt} = -[H, f]. \tag{6}$$

Note that for any functions f and g, and real or complex number a:

$$[af, g] = [f, ag] = a[f, g]$$
$$[f, g] = -[g, f]$$
$$[f, f] = 0$$

from which it follows using equation (6):

$$\frac{dH}{dt} = -[H, H] = 0. \tag{7}$$

Equation (7) shows that the Hamiltonian is a constant of the motion, as might be expected from conservation of energy. We can construct an operator $:f:$ from any function f, defined by:

$$:f:g = [f, g]. \tag{8}$$

An operator of the form $:f:$ is called a *Lie operator*. Repeated action of $:f:$ is written as $:f:^2$, $:f:^3$ etc., where

$$:f:^2 g =: f:(:f:g) = [f, [f, g]]$$
$$:f:^3 g =: f:(:f:(:f:g)) = [f, [f, [f, g]]]$$

and so on. In this way, we can express an exponential of a Lie operator $:f:$ by means of a power series expansion:

$$\exp(:f:) = \sum_{n=0}^{\infty} \frac{1}{n!} :f:^n \tag{9}$$

An operator of the form $\exp(:f:)$ is called a *Lie transform* [2]. Using (8), we can rewrite (6) as:

$$\frac{df}{dt} = -:H:f \tag{10}$$

which suggests the solution:

$$f(t) = e^{-t:H:} f(0) \tag{11}$$

Note that we have not derived equation (11) rigorously, but the result is apparent from the solution of the analogous first-order ordinary differential equation to (10). Equation (11) is an important result, and forms the basis for much of the work that follows. It can be seen already that the Hamiltonian approach has potential advantages, particularly for nonlinear systems. Using Newtonian methods, the equation of motion is expressed as a differential equation that must be integrated to find the values of variables at particular times. In contrast, since $:H:$ is a differential operator, finding the value of a function f at some time t involves a process of differentiation. Of course, we still have a sum over an infinite number of terms to consider, but in a number of important cases this can be resolved without much trouble.

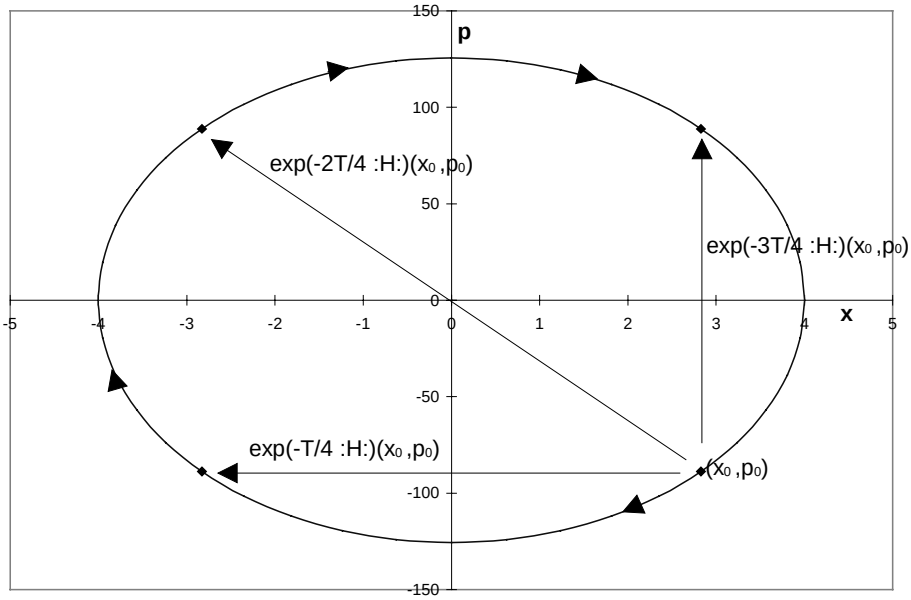

Figure 1

Phase space evolution of a simple harmonic oscillator in one degree of freedom. The oscillator has amplitude 4 and period $T = 0.2$. The initial state of the system is shown (x_0, p_0), and the effects of mapping from this points using a Lie transform with the Hamiltonian $H = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 x^2$.

To illustrate these results, and the use of equation (11) in particular, we consider the simple case of a simple harmonic oscillator in one dimension. Using equation (3), we write for this system:

$$H = \tfrac{1}{2} p^2 + \tfrac{1}{2} \omega^2 x^2 .$$

Note that we work in units in which the mass of the particle is normalised to one. Using equation (11) with $f(x, p) = x$, we have:

$$\begin{aligned} x(t) &= \exp(-t{:}H{:})x(0) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (-t{:}H{:})^n x(0) \end{aligned}$$

Applying the operator $:H:$ to x, we find:

$$\begin{aligned} :H:x &= [H, x] = -p \\ :H:^2 x &=: H:(-p) = -\omega^2 x \end{aligned}$$

etc.

And thus,

$$\sum_{n=0}^{\infty} \frac{1}{n!} (-t{:}H{:})^n x(0) = x(0)\left(1 - \tfrac{1}{2}\omega^2 t^2 + \tfrac{1}{4!}\omega^4 t^4 - \ldots\right) + \frac{p(0)}{\omega}\left(\omega t - \tfrac{1}{3!}\omega^3 t^3 + \tfrac{1}{5!}\omega^5 t^5 - \ldots\right) \qquad (12)$$

Note that despite the repeated operation of $:H:$ in the summation on the left-hand side, this expression is linear in x and p; equation (12) is an example of a *solvable map*, which can be expressed as a finite series in the state variables. The series in ωt are recognised as power series expansions of trigonometric functions, thus:

$$x(t) = x(0)\cos(\omega t) + \frac{p(0)}{\omega}\sin(\omega t). \tag{13}$$

Similarly, we find by applying $\exp(-t{:}H{:})$ to the function $f(x,p) = p$, we find:

$$p(t) = p(0)\cos(\omega t) - \omega x(0)\sin(\omega t). \tag{14}$$

Equations (13) and (14) are of course obtained in Newtonian mechanics by integrating the equation of motion:

$$\frac{d^2x}{dt^2} = -\omega^2 x. \tag{15}$$

Since (15) is linear, there appears little advantage in using the Hamiltonian formalism; the advantages of the approach will become apparent, however, when we come to consider nonlinear systems.

1.2 Canonical Transformations

It is possible to make a transformation of variables from x_i and p_i to new variables X_i and P_i, expressible as functions of the old variables:

$$\begin{aligned} X_i &= X_i\left(x_i, p_i\right) \\ P_i &= P_i\left(x_i, p_i\right). \end{aligned} \tag{16}$$

In general, though, there is no guarantee that the new variables will satisfy a form of Hamilton's equations, i.e. that the momenta P_i will be canonically conjugate to the co-ordinates X_i. But supposing that the new variables do satisfy Hamilton's equations for some Hamiltonian $\tilde{H}$ obtained from a transformation of the Hamiltonian H, then the transformation (16) is called a *canonical transformation*, and we can write:

$$\begin{aligned} \frac{dX_i}{dt} &= \frac{\partial \tilde{H}}{\partial P_i} \\ \frac{dP_i}{dt} &= -\frac{\partial \tilde{H}}{\partial X_i} \end{aligned} \tag{17}$$

A canonical transformation preserves the features of the original system, for example there must be the same number of constants of motion, and these must be related in a way determined by the canonical transformation.

We can represent the evolution of any system in a chosen degree of freedom by a line on a chart of p *vs* x. Such a chart is called a *phase space diagram*; an example for a simple harmonic oscillator is shown in Figure 1. We can see from equation (11) that the action of the operator $\exp(-t{:}H{:})$ on the phase space co-ordinates $\left(x,p\right)$ is to map the system from its state at some initial time to its state at a time t later.

1.3 Symplectic Maps

Suppose we have two vectors $\vec{e}_1 = \left(dx^1, dp^1\right)$ and $\vec{e}_2 = \left(dx^2, dp^2\right)$ in phase space. $\vec{e}_1$ and $\vec{e}_2$ define two sides of a parallelogram, which encloses the area

$$
\begin{aligned}
A &= \vec{e}_1 \times \vec{e}_2 \\
&= dx^1 dp^2 - dp^1 dx^2 \\
&= \begin{pmatrix} dx^1 & dp^1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} dx^2 \\ dp^2 \end{pmatrix}
\end{aligned} \tag{18}
$$

Now, if M is a linear mapping, then its action on any phase space vector $\vec{e}_i$ can be represented by a matrix:

$$
\vec{e}_i \xrightarrow{M} \vec{e}_i' = \mathbf{M} \begin{pmatrix} dx^i \\ dp^i \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} dx^i \\ dp^i \end{pmatrix} \tag{19}
$$

From (18) and (19), we see that under the action of the map represented by the matrix $\mathbf{M}$, the area A becomes:

$$
A \xrightarrow{M} A' = \begin{pmatrix} dx^1 & dp^1 \end{pmatrix} \mathbf{M}^T \mathbf{J} \mathbf{M} \begin{pmatrix} dx^2 \\ dp^2 \end{pmatrix} \tag{20}
$$

Where $\mathbf{M}^T$ represents the transpose of the matrix $\mathbf{M}$, and $\mathbf{J}$ is the antisymmetric matrix:

$$
\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{21}
$$

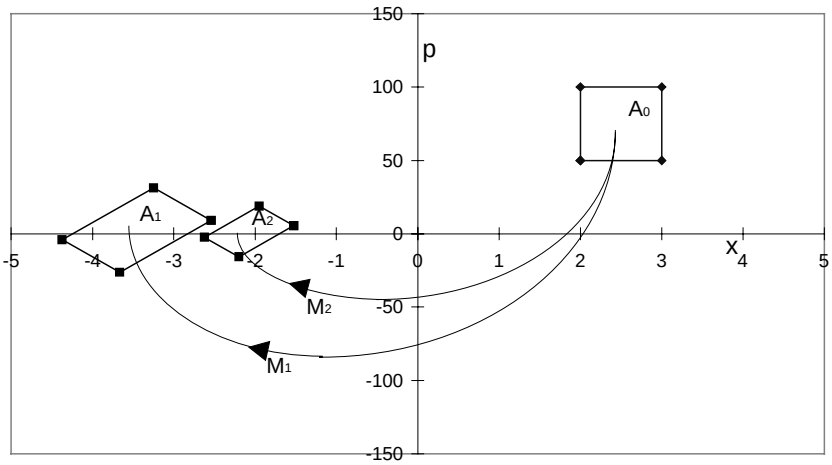

Figure 2

Symplectic and non-symplectic mappings in phase space. Points within A_0 are mapped to within A_1 by the symplectic map $\mathbf{M}_1$; because the area enclosed by the boundary stays the same, the density of points in phase space remains the same. Points within A_0 are mapped to within A_2 by the map $\mathbf{M}_2$; because the map is non-symplectic, the area changes. In this case, the density of points within phase space increases. Liouville's theorem is not obeyed, because the system is not Hamiltonian.

From (20) we see that if

$$
\mathbf{M}^T \mathbf{J} \mathbf{M} = \mathbf{J} \tag{22}
$$

then

$$A' = A,$$

i.e. the phase space area defined by the vectors $\vec{e}_1$ and $\vec{e}_2$ is preserved. A matrix $\mathbf{M}$ that satisfies (22) is called a *symplectic* matrix[1]; any mapping (linear or nonlinear) that preserves the area of elements of phase space in this way is called a *symplectic mapping*; see Figure 2.

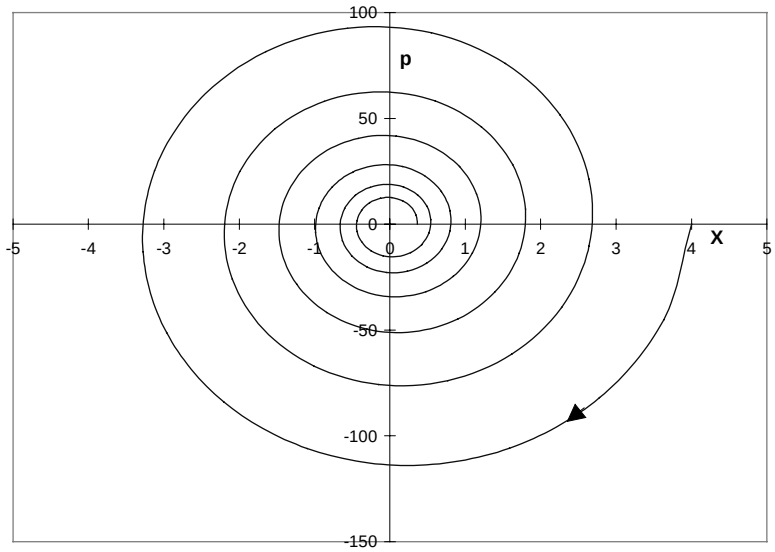

Figure 3

Phase space plot for a damped harmonic oscillator. The system is non-symplectic, because the density of states in phase space increases with time, as all initial states head towards the origin. Six periods of oscillation are shown.

Hamiltonian systems are intrinsically symplectic; this is expressed in Liouville's theorem[2] (see, for example [1]), which says that for any Hamiltonian system, the density of states in phase space remains constant as the states are evolved in time. In contrast, Liouville's theorem is not generally valid for non-Hamiltonian systems, in which a set of states may be seen converging to or diverging from a point in phase space. For example, a line in phase space representing the motion of a damped harmonic oscillator will inevitably end up at the origin, no matter where the starting point. Similarly, a line representing a harmonic oscillator driven at resonance will spiral away from the origin without limit, from any starting point. In accelerator physics, electron motion is generally non-Hamiltonian (non-symplectic) because of the effects of synchrotron radiation; hopefully, the rate of energy loss is small, being of order 1% per hundred turns of a circular accelerator. In linear accelerators, energy losses from synchrotron radiation can generally be ignored. However, any theory must be able to take into account damping and driving effects. The Hamiltonian approach leads to a mapping theory based on the use of Lie operators, which always generate symplectic maps. By contrast, the Newtonian approach makes it necessary to use approximations which render the motion of a particle apparently *non*-symplectic, even *before* effects such as damping and driving are included. The author contends that it is preferable to start with a physical symplectic theory to which non-symplectic effects can be added, rather than add such effects to a non-physical description of particle motion. In the words of Willeke [2], "a non-symplectic model might cause an artificial instability which in practice may be very difficult to distinguish from genuine weak instability."

[1] It can easily be shown that for a 2×2 matrix $\mathbf{M}$ to satisfy (22), it must have unit determinant.
[2] Liouville's theorem is itself a direct consequence of the Hamiltonian equations of motion.

Lie transforms must be symplectic mappings, because they give solutions to Hamilton's equations. Briefly, this is because any analytic function H of the phase space co-ordinates can serve as the Hamiltonian, and the evolution of states in phase space will always be symplectic. Thus, any operator of the form $\exp(-t{:}H{:})$, which gives the evolution of the system from any initial state, will be a symplectic mapping.

1.4 Transfer Maps and Koopman Maps

In equation (19) we represented the action of a map by a matrix acting on variables in phase space. We refer to such a map as a *transfer map*, and it is familiar from traditional treatments of particle dynamics. As we have already indicated, the motion of a Hamiltonian system through phase space can be determined by a mapping of the form (11). Note, however, that Lie operators act on functions of the phase space variables, not on the variables themselves [3][4]. This can easily lead to confusion, as we are frequently interested in the case, for example, where a Lie operator acts on $f(x,p) = x$. An operator acting on a function is a different kind of mapping from a transfer map; to distinguish the two, we call the effect of an operator on a function a *Koopman map*.

To clarify this distinction, consider the general linear function $f(x,p) = ax + bp$, where a and b are constants, and the general homogeneous second order Lie operator $:A:$, where $A = -\frac{1}{2}\left(\gamma x^2 + 2\alpha xp + \beta p^2\right)$. We can see at once that

$$:A: f(x,p) = (\alpha a - \gamma b)x + (\beta a - \alpha b)p\,. \tag{23}$$

The form of equation (23) suggests that the Koopman map can be thought of as acting on the coefficients of the phase space variables in f rather than on the variables themselves. This has important consequences. In the present (linear) case, if we wished to find a matrix representation of $:A:$, we would write:

$$:A: f(x,p) = \tilde{f}(x,p) = \tilde{a}x + \tilde{b}p$$

where

$$\begin{pmatrix}\tilde{a}\\ \tilde{b}\end{pmatrix} = \begin{pmatrix}\alpha & -\gamma\\ \beta & -\alpha\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \mathbf{A}\begin{pmatrix}a\\ b\end{pmatrix}.$$

Here, $\mathbf{A}$ is the required matrix representation of $:A:$. Rearranging (23) gives

$$:A: f(x,p) = a(\alpha x + \beta p) + b(-\gamma x - \alpha p) = f(\bar{x},\bar{p})$$

where

$$\begin{pmatrix}\bar{x}\\ \bar{p}\end{pmatrix} = \begin{pmatrix}\alpha & \beta\\ -\gamma & -\alpha\end{pmatrix}\begin{pmatrix}x\\ p\end{pmatrix} = \mathbf{M}\begin{pmatrix}x\\ p\end{pmatrix}.$$

Note that $\mathbf{M}$ is the (matrix representation) of the transfer map corresponding to the Koopman map with matrix representation $\mathbf{A}$, and that we have:

$$\mathbf{A} = \mathbf{M}^T. \tag{24}$$

If we apply a sequence of transfer maps $\ldots\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1$, because of the transpose relationship in (24), the corresponding Koopman map has matrix representation $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\ldots$, i.e. the maps are applied in the reverse order [4]. This result holds true even in the nonlinear case, and is an intrinsic property of the Lie operator.

1.5 Action of Koopman Maps on Lie Operators

Suppose we have a Koopman map $\mathcal{K}$ which acts on a function $f(x,p)$:

$$\tilde{f}(x,p) = \mathcal{K}f(x,p).$$

If we consider the effect of $\mathcal{K}$ on $f_1(x,p) = x$ and $f_2(x,p) = p$, then from equations (17) we can say that $\mathcal{K}$ is a canonical transformation if we can find a transformation $\tilde{H}$ of the Hamiltonian H, such that

$$\frac{d\tilde{f}_i}{dt} = -:\tilde{H}:\tilde{f}_i = -:\tilde{H}:(\mathcal{K}f_i). \tag{25}$$

If $\mathcal{K}$ is independent of time, we can write

$$\frac{d\tilde{f}_i}{dt} = \frac{d(\mathcal{K}f_i)}{dt} = \mathcal{K}\frac{df_i}{dt} = -\mathcal{K}(:H:f_i). \tag{26}$$

Combining equations (25) and (26) gives the result:

$$:\tilde{H}: = \mathcal{K}:H:\mathcal{K}^{-1}. \tag{27}$$

1.6 Invariance of the Poisson Bracket

Another important result is the invariance of the Poisson bracket under symplectic mappings, which we express as

$$e^{:k:}[g,h] = [e^{:k:}g, e^{:k:}h] \tag{28}$$

for any functions k, g and h. For a proof of equation (28), the reader is referred to [3]. As an illustration of the use of this result, we look again at equation (27), and consider the case that $\mathcal{K}$ can be written as a Lie transform, $\mathcal{K} = e^{:k:}$. If we apply $:\tilde{H}:$ to an arbitrary function f, we get

$$\begin{aligned} :\tilde{H}:f &= e^{:k:}:H:e^{-:k:}f \\ &= e^{:k:}\left[H, e^{-:k:}f\right] \\ &= \left[e^{:k:}H, f\right] \end{aligned} \tag{29}$$

where we have used the invariance of the Poisson bracket (28) in the last line of (29). It then follows immediately that

$$:\tilde{H}: = e^{:k:}:H:e^{-:k:} =: e^{:k:}H:$$

Again using the invariance of the Poisson bracket, we find the general result for any differentiable functions f and g:

$$e^{:f:}\exp(:g:)e^{-:f:} = \exp(:e^{:f:}g:). \tag{30}$$

2. Linear Beam Dynamics

Having argued for a Hamiltonian theory of beam dynamics and established some of the basic properties of Lie operators, we now show how the theory can be applied to a linear magnetic lattice, before moving on to consider the more general nonlinear case. Much of work of this section will be based on the idea of the 'single-turn map'. In simple terms, this is a rule for determining the position of a

particle in phase space after one complete pass round a circular lattice, from any starting position. We shall consider in detail only the motion in one (transverse) degree of freedom, but it should become clear how the theory can be extended to cover the full three degrees of freedom available to real particles. We shall further neglect radiation damping or the energy gain from the RF system, effectively assuming that the motion of the particle is symplectic.

2.1 Constants of the Single-Turn Map

Consider the effect of a Lie operator formed from a general second order homogeneous polynomial, acting on linear functions of the phase space co-ordinates:

$$\begin{aligned} F &= \tfrac{1}{2}\left(\gamma x^2 + 2\alpha xp + \beta p^2\right) \\ :F:x &= -\alpha x - \beta p \\ :F:p &= \gamma x + \alpha p \end{aligned} \tag{31}$$

We can construct a symplectic map from $:F:$ by using it as the exponent in a Lie transform (multiplied, if we wish, by an arbitrary constant); since F is second order and homogeneous, the effect of the Lie transform will always be a linear mapping. Thus, a symplectic linear map on a function $f(x,p)$ can be written:

$$\mathcal{M}f\left(x,p\right) = \exp(-\mu:F:)\, f\,(x,p) \tag{32}$$

Since, by assumption, the motion of a particle through our lattice is linear and symplectic, equation (32) gives the form of a general single-turn map. Since

$$:F:F = [F,F] = 0$$

we have at once

$$\mathcal{M}F(x,p) = \exp(-\mu:F:)F(x,p) = F(x,p)$$

and hence $\gamma x^2 + 2\alpha xp + \beta p^2 = 2F$ is a constant of the single turn map. In other words, after each pass around the machine, the particle's phase space co-ordinates must satisfy

$$\gamma x^2 + 2\alpha xp + \beta p^2 = \varepsilon \text{ (constant)}, \tag{33}$$

which is the equation of an ellipse, shown in Figure 4. The left hand side of equation (33) corresponds to the Courant-Snyder invariant, familiar from traditional accelerator theory. Note that the motion of the particle in phase space as measured at a single point in the ring is not continuous; rather the particle moves round the ellipse in discrete steps. The parameters α, β and γ are functions of position around the ring.

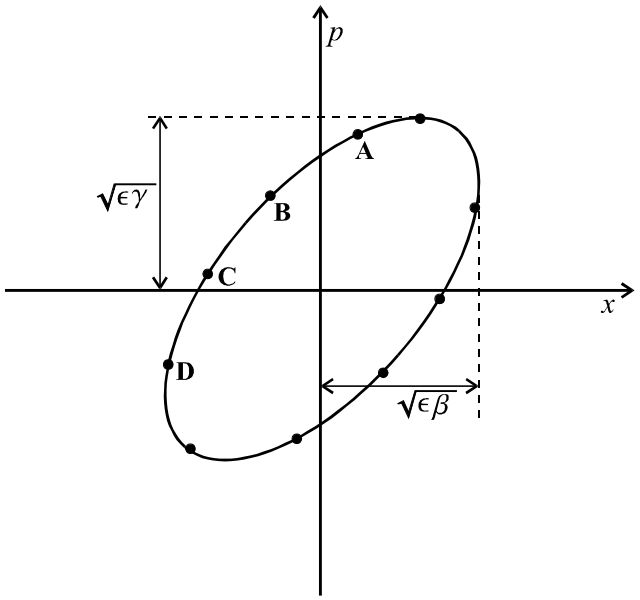

Figure 4

Phase space ellipse mapped out by a particle on successive turns around an accelerator. The initial position of the particle in phase space is indicated by point **A**, the position after one turn is point **B** and so on. The area of the enclosed ellipse is $\pi\varepsilon_x$, where ε_x is the emittance.

2.2 Single-Turn Transfer Matrix

To work out the transfer matrix, we need to apply (32) to the functions $f(x,p)=x$, and $f(x,p)=p$. We find, by repeated application of $:F:$:

$$\begin{aligned}\exp(-\mu:F:)x &= x\left(\cos(\mu\Delta)+\frac{\alpha}{\Delta}\sin(\mu\Delta)\right)+p\frac{\beta}{\Delta}\sin(\mu\Delta)\\ \exp(-\mu:F:)p &= -\frac{\gamma}{\Delta}x\sin(\mu\Delta)+p\left(\cos(\mu\Delta)-\frac{\alpha}{\Delta}\sin(\mu\Delta)\right)\end{aligned} \tag{34}$$

where

$$\Delta=\sqrt{\beta\gamma-\alpha^2}\ .$$

From equations (34), it is clear that the transfer matrix corresponding to the single turn map is

$$\mathbf{M}=\begin{pmatrix}\cos(\mu\Delta)+\frac{\alpha}{\Delta}\sin(\mu\Delta) & \frac{\beta}{\Delta}\sin(\mu\Delta)\\ -\frac{\gamma}{\Delta}\sin(\mu\Delta) & \cos(\mu\Delta)-\frac{\alpha}{\Delta}\sin(\mu\Delta)\end{pmatrix}. \tag{35}$$

It is easy to verify that the matrix in (35) always has unit determinant, as is necessary for symplecticity. However, we notice that there are apparently too many independent variables: the most general symplectic 2×2 matrix can be written in terms of three independent variables (four matrix elements, with one condition imposed by symplecticity), and (35) has four independent variables, α,β,γ and μ. We are therefore free to impose one relationship between these variables; for convenience, we choose

$$\Delta=\sqrt{\beta\gamma-\alpha^2}=1, \tag{36}$$

and the transfer matrix then becomes

$$\mathbf{M} = \begin{pmatrix} \cos(\mu) + \alpha\sin(\mu) & \beta\sin(\mu) \\ -\gamma\sin(\mu) & \cos(\mu) - \alpha\sin(\mu) \end{pmatrix}. \tag{37}$$

Note that with the choice for Δ given by (36), the transfer matrix (37) has a simple interpretation: namely, that on each pass round the machine, a particle moves around an ellipse in phase space given by (33), with phase advance μ per revolution. Note also that

$$\mathrm{Tr}(\mathbf{M}) \le \mathbf{2}$$

with equality holding for a phase advance equal to integer multiple of 2π.

One alternative choice for Δ is

$$\Delta = \sqrt{\beta\gamma - \alpha^2} = i\,.$$

With this choice, the transfer matrix becomes

$$\mathbf{M} = \begin{pmatrix} \cosh(\mu) + \alpha\sinh(\mu) & \beta\sinh(\mu) \\ -\gamma\sinh(\mu) & \cosh(\mu) - \alpha\sinh(\mu) \end{pmatrix}.$$

In this case, however, the motion of a particle is no longer stable, but the phase space coordinates increase hyperbolically with the number of turns. The trace of the transfer matrix now satisfies for $\mu > 0$,

$$\mathrm{Tr}(\mathbf{M}) > \mathbf{2}\,.$$

2.3 Evolution of the Lattice Parameters

The parameters α, β and γ, which we refer to as the lattice parameters, are clearly dependent on the structure of the magnetic lattice in the accelerator. To find how they vary through the lattice, we can apply the ideas of Hamiltonian mechanics, and will need to make some assumptions about the Hamiltonian in the lattice. First, because we are interested for the moment in motion in only one degree of freedom, we assume that a particle travels at constant speed through the lattice, and thus we can replace the time variable with the distance s measured from a chosen starting point in the lattice. Equation (11) then becomes,

$$f(s+ds) = e^{-ds:H:} f(s)$$

which gives the change in a function with distance through the lattice. We represent the mapping from s to $s+ds$ by $\mathcal{M}_{s\to s+ds}$, and the single-turn map at position s by $\mathcal{M}_s$. We can deduce the single-turn map $\mathcal{M}_{s+ds}$ at $s+ds$, since this must be obtained by mapping from $s+ds$ back to s, applying the single turn map at s, then mapping from s to $s+ds$. Thus,

$$\begin{aligned} \mathcal{M}_{s+ds} &= \mathcal{M}_{s+ds\to s}\mathcal{M}_s\mathcal{M}_{s\to s+ds} \\ &= e^{ds:H:}\exp(-\mu:\varepsilon_s:)e^{-ds:H:} \end{aligned}$$

where ε_s is the Courant-Snyder invariant at s. Using equation (30), this becomes

$$\mathcal{M}_{s+ds} = \exp(-\mu:e^{ds:H:}\varepsilon_s:)\,. \tag{38}$$

But we can also write

$$\mathcal{M}_{s+ds} = \exp(-\mu:\varepsilon_{s+ds}:)\,. \tag{39}$$

Comparing equations (38) and (39), we see

$$\begin{aligned}\varepsilon_{s+ds} &= e^{ds:H:}\varepsilon_s \\ &= \varepsilon_s + ds[H,\varepsilon_s] + O(ds^2)\end{aligned}$$

and thus

$$\frac{d\varepsilon_s}{ds} = \left[H,\varepsilon_s\right]. \tag{40}$$

From the way in which it was derived, equation (40) gives the variation in the Courant-Snyder invariant resulting from variation in the lattice functions, *not* from the dynamics of the particle; it should not be confused with the Hamiltonian equations of motion (6), which *do* describe the dynamics of the particle. We are working under the assumption that the lattice is linear; therefore the Hamiltonian at any point can be written as a homogeneous second order polynomial:

$$H = \tfrac{1}{2}H_{20}x^2 + H_{11}xp + \tfrac{1}{2}H_{02}p^2$$

where H_{20}, H_{11} and H_{02} are independent of position. From (40), and writing

$$\varepsilon_s = \gamma(s)x^2 + 2\alpha(s)xp + \beta(s)p^2$$

we then have

$$\begin{aligned}\frac{d\gamma}{ds} &= 2H_{20}\alpha - 2H_{11}\gamma \\ \frac{d\alpha}{ds} &= H_{20}\beta - H_{02}\gamma \\ \frac{d\beta}{ds} &= 2H_{11}\beta - 2H_{02}\alpha\end{aligned} \tag{41}$$

Equations (41) form a set of coupled linear differential equations for the lattice functions α, β and γ. In the special case that

$$\begin{aligned}H_{11} &= 0 \\ H_{02} &= 1 \\ H_{20} &= k\end{aligned} \tag{42}$$

we find

$$\begin{aligned}\frac{d\gamma}{ds} &= 2k\alpha \\ \frac{d\alpha}{ds} &= k\beta - \gamma \\ \frac{d\beta}{ds} &= -2\alpha\end{aligned} \tag{43}$$

With the help of the relationship (36), we can eliminate the functions α and γ, to give

$$\beta\frac{d^2\beta}{ds^2} - \frac{1}{2}\left(\frac{d\beta}{ds}\right)^2 + 2k\beta^2 = 2 \tag{44}$$

which we recognise as the envelope equation giving the variation in the beta function through a focusing quadrupole of strength k. The purpose of a linear lattice code is to find periodic solutions to

(41), piecewise through different magnet types, once the Hamiltonian for each type has been determined.

We have so far managed to obtain many of the fundamental results of linear beam dynamics with only the most general assumptions, using Hamiltonian mechanics and the concept of the single-turn map. The reader is invited to compare the derivation of results such as (33) and (44) given here with the derivation based on Newtonian mechanics and the solution of Hill's equation given in many introductory texts (see, for example, [5]). The author feels that once the investment has been made in understanding Hamiltonian mechanics and Lie operator methods, the linear dynamics of particles can be understood much more clearly, the important equations can be derived much more elegantly and directly, and the assumptions involved are more apparent at each stage. In fact, we have not so far, apart from implicitly in (42), needed to consider in any detail the magnetic fields through which the particle is passing. It is appropriate to do this before moving on to nonlinear dynamics, where the power of Lie operator methods will really become apparent, and in the next section, we write down the Hamiltonian for a particle moving through various types of magnet.

Finally, it is worth remarking that quantities such as dispersion and chromaticity can of course be analysed with Hamiltonian tools, just as in the conventional theory, but they are beyond the scope of the present work. The interested reader is referred to Forest's book on beam dynamics [3].

3. The Hamiltonian in a Magnetic Field

In Newtonian mechanics, the equation of motion of a particle in a magnetic field is given by Newton's Second Law of Motion,

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}$$

where the force, $\mathbf{F}$, is given by the Lorentz expression

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{45}$$

Where q is the electric charge of the particle, $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field and $\mathbf{v}$ is the velocity of the particle. The electric and magnetic fields are subject to Maxwell's equations. Classically, Newton's laws of motion (or the corresponding relativistic equations), the Lorentz force and Maxwell's equations have the status of natural laws: they are the foundations of a description of physical phenomena, and it is not 'necessary' to ask for any underlying principles. It is sufficient that the description agrees accurately with the results of experiments. Modern theories, in particular quantum field theory, do reveal underlying principles, and thereby describe an even wider range of physical phenomena. But since we are interested in the classical regime, we could be justified in simply writing down an expression for the Hamiltonian in a magnetic field, as long as it turns out to describe accurately the motion of a charged particle in the field when used in Hamilton's equations. Such a Hamiltonian would correspond with the Lorentz force, which we are not required to justify on any theoretical grounds.

Such a situation would probably leave many readers dissatisfied, but at the same time it is beyond the scope of this note to give a rigorous 'derivation' of expressions for the Hamiltonian in an accelerator, which can be very complicated. We therefore satisfy ourselves with indicating general rules for writing down Hamiltonians which are simple enough to use in analytic calculations, but sufficiently realistic to reveal the significant features of particle dynamics in interesting situations. For a more rigorous general treatment, including the full form of the relativistic Hamiltonian which is highly relevant to accelerator physics, the reader is referred to [6].

3.1 Particle in a Quadrupole Field

The general Hamiltonian for a particle in an electromagnetic field is given by

$$H = \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}\right)^2 + q\varphi \tag{46}$$

where m is the mass of the particle, q is the particle's electric charge, $\mathbf{p}$ is the canonical momentum, $\mathbf{A}$ is the magnetic vector potential and φ is the electric scalar potential. It is important to understand that the canonical momentum is not simply the product of the particle's mass and velocity, but is given by

$$\mathbf{p} = m\dot{\mathbf{x}} + q\mathbf{A}\,. \tag{47}$$

This follows from the formal definition of the canonical momentum in terms of the Lagrangian, for which the reader is referred to standard texts on classical mechanics, for example [1]. We are forbidden from substituting (47) into (46) since the Hamiltonian must be expressed in terms of the coordinates and canonical momenta, not in terms of the velocity. It is instructive to show that the Hamiltonian (46) gives the same equation of motion as the Lorentz force used in Newton's Second Law of Motion[3].

In a focusing quadrupole, the magnetic field is given by

$$B_x = \frac{ck}{qE}y \quad B_y = \frac{ck}{qE}x \quad B_s = 0$$

where c is the speed of light, E is the energy of the particle, and k is a constant giving the strength of the quadrupole. Note that we are using a co-ordinate system where x is the distance horizontally from the closed design orbit, y is the distance vertically from this orbit, and s is the longitudinal distance through the lattice. The field $\mathbf{B}$ is related to the vector potential by

$$\mathbf{B} = \nabla \times \mathbf{A}$$

From which it is clear that a possible form[4] for $\mathbf{A}$ is

$$A_x = 0 \quad A_y = 0 \quad A_s = \frac{ck}{2qE}\left(y^2 - x^2\right). \tag{48}$$

Substituting (48) into (46), and with $\varphi = 0$, gives

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_s^2\right) + \frac{ckp_s}{2mE}\left(x^2 - y^2\right) + \frac{c^2k^2}{2mE^2}\left(x^2 - y^2\right)^2.$$

For a weak field, and if the amplitude of betatron oscillations is small, then we have

$$\frac{ck}{E}\left(x^2 - y^2\right) << p_s \approx \frac{E}{c}\,.$$

Since p_s is approximately constant, and we can add or subtract any constant to the Hamiltonian without affecting the dynamics, we can write:

---

[3] Curiously, this is not often done in standard texts; in fact, the author is unaware of any reference which demonstrates explicitly the correspondence between (46) and (45) - even Jackson [6] leaves it as an exercise for the reader. For this reason, we give a possible proof in Appendix A of this note, where the importance of using (47) correctly is apparent.

[4] There are of course different possible forms for the vector potential giving the same magnetic field. Since the field is the derivative of the potential, for example, we can add a constant to $\mathbf{A}$ without affecting $\mathbf{B}$. We avoid issues of gauge invariance, and the reader is referred to standard texts on electromagnetism, for example [6].

$$H \approx \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{k}{2m}\left(x^2 - y^2\right). \tag{49}$$

Similarly, for a defocusing quadrupole,

$$H \approx \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{k}{2m}\left(y^2 - x^2\right). \tag{50}$$

Furthermore, since $A_x = A_y = 0$, the momenta p_x and p_y in (49) and (50) although technically canonical momenta, are effectively just $m\dot{x}$ and $m\dot{y}$ respectively. Finally, assuming that the particle moves at constant speed through the lattice, by applying a suitable scaling to k, we can change variables from time dependence to s dependence. Although it is clear that (49) and (50) are only approximate forms for the Hamiltonian, they are simple enough for many analytic manipulations, but reproduce accurately the features of the dynamics we are interested in. We have thus achieved our objectives for this section, and will, for the rest of this note, use (49) and (50) as quadrupole Hamiltonians without continual reference to the approximations we have made.

3.2 Hamiltonians for Common Magnet Types

For simplicity, we work in units in which the particles in our accelerator have unit mass. Also, the time-like variable is the longitudinal distance s, so that

$$p_x = x' = \frac{dx}{ds}$$

Hamiltonians for various magnet types can be obtained in much the same way as the Hamiltonian (49) for a quadrupole. Making the same approximations, we find:

Table 1: Hamiltonians for various lattice elements

Magnet Type	Hamiltonian
Drift Space	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right)$
Bend	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right)$
Normal quadrupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) + \frac{1}{2}k_1\left(x^2 - y^2\right)$
Skew quadrupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) + k_1 xy$
Normal sextupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) + \frac{1}{6}k_2\left(x^3 - 3xy^2\right)$
Skew sextupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) - \frac{1}{6}k_2\left(y^3 - 3x^2 y\right)$
Normal octupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) + \frac{1}{24}k_3\left(x^4 - 6x^2y^2 + y^4\right)$
Skew octupole	$H = \frac{1}{2}\left(p_x^2 + p_y^2\right) + \frac{1}{6}k_3\left(x^3 y - xy^3\right)$

Note that with our simplifying assumptions, there is no distinction between a drift space and a bend, and we do not distinguish between different types of bending magnets. We also neglect effects such as fringe fields and field errors. More realistic Hamiltonians will incorporate relativistic effects and longitudinal momentum deviation, and are required for reliable tracking codes, such as MAD [7]. However, it will be seen that by using the expressions in Table 1, we can reproduce significant features of the dynamics of particles in accelerators, while not becoming lost in the complexities of the algebra.

4. Manipulating Lie Operators to Find Solvable Approximations

Having obtained the Hamiltonians for common lattice elements, we are in a position to construct the Lie operators that will give us the mapping from the entrance to the exit of a given magnet. For example, the mapping for a focusing sextupole of length l is just

$$\exp(-l{:}H{:}) = \exp\left(-l{:}\tfrac{1}{2}\left(p_x^2 + p_y^2\right) + \tfrac{1}{6}k_2\left(x^3 - 3xy^2\right){:}\right). \tag{51}$$

Unfortunately, the series representation of maps such as (51) generally extend to infinite order. This is inconvenient to say the least, particularly if we are interested in using a computer to model the dynamics of particles in a lattice. In this section, therefore, we consider how we can find solvable approximations to maps such as (51). For simplicity, we shall work in one degree of freedom, though it is clear from (51) that the structure of Lie operators implicitly allows as many degrees of freedom as desired.

4.1 Solvable Maps

We can easily apply (51) to the functions x and p to find the mapping in phase space resulting from a quadrupole:

$$\exp\left(-l{:}\tfrac{1}{2}p^2 + \tfrac{1}{2}k_1x^2{:}\right)x = x\cos\left(\sqrt{k_1}l\right) + \frac{p}{\sqrt{k_1}}\sin\left(\sqrt{k_1}l\right)$$
$$\exp\left(-l{:}\tfrac{1}{2}p^2 + \tfrac{1}{2}k_1x^2{:}\right)p = -x\sqrt{k_1}\sin\left(\sqrt{k_1}l\right) + p\cos\left(\sqrt{k_1}l\right)$$

The quadrupole Hamiltonian generates a *solvable* map, since it results in a finite series when applied to x and p. The effect is a rotation (phase advance) in phase space, through an angle $\sqrt{k_1}l$. The situation for a sextupole, however, is rather different:

$$\begin{aligned}\exp\left(-l{:}\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3{:}\right)x &= x + lp - \tfrac{1}{4}k_2l^2x^2 - \tfrac{1}{6}k_2l^3xp - \tfrac{1}{48}k_2l^4p^2 + O(3)\\ \exp\left(-l{:}\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3{:}\right)p &= p - \tfrac{1}{2}k_2lx^2 - \tfrac{1}{2}k_2l^2xp - \tfrac{1}{6}k_2l^3p^2 + O(3)\end{aligned} \tag{52}$$

where $O(3)$ stands for terms of order 3 and higher in x and p. In fact, the map in (52) is not solvable, because it generates an infinite power series in x and p. If we wish to implement a map such as this, for example, on a computer, we are faced with making an approximation. The most obvious choice is to truncate the series at some chosen order. This produces a *truncated power series approximation* (TPSA). The drawback with the TPSA approach is that the approximated map is no longer symplectic. This is a potentially serious issue, since we are often interested in applying the map many times to represent repeated passes of the particle through the magnet. A non-symplectic map will introduce features into the dynamics that are not physical; often, a stable system will appear to be unstable. There exist techniques for 'symplectifying' a truncated series [2], but these techniques are not guaranteed to succeed, and at best introduce further approximations.

An alternative approach is to introduce the approximation into the Lie operator. Our aim here is to produce a Lie operator (or sequence of Lie operators) which are approximately equal to the original operator, and generate solvable maps. Since Lie operators are guaranteed to produce symplectic maps (see Section 1.3), we preserve at least this important feature of the original map. For example, if we make the approximation

$$\exp\left(-l{:}\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3{:}\right) \approx \exp\left(-l{:}\tfrac{1}{2}p^2{:}\right)\exp\left(-l{:}\tfrac{1}{6}k_2x^3{:}\right) \tag{53}$$

and examine the map of the operator on the right hand side of (53), we find:

$$\begin{aligned}\exp\left(-l:\tfrac{1}{2}p^2:\right)\exp\left(-l:\tfrac{1}{6}k_2x^3:\right)x &= x+lp\\ \exp\left(-l:\tfrac{1}{2}p^2:\right)\exp\left(-l:\tfrac{1}{6}k_2x^3:\right)p &= p-\tfrac{1}{2}k_2lx^2-k_2l^2xp-\tfrac{1}{2}k_2l^3p^2\end{aligned} \tag{54}$$

The expressions in (54) are exact, and thus the map is solvable. If we look at the map, we see that it corresponds to a drift space of length l, followed by a sextupole kick of integrated strength k_2l. At this point, it is sensible to ask how we can produce solvable maps that are also good approximations to our original map. The answer requires some knowledge of how maps are combined.

4.2 The Baker-Campbell-Hausdorff Formula

Suppose we have two non-commuting operators A and B, i.e.

$$\{A,B\} = AB - BA \neq 0 \tag{55}$$

where we use curly brackets $\{\ \}$ to represent the commutator of two operators. Now consider the product e^Ae^B. As in equation (9), we represent the exponential of an operator by its power series expansion. Thus:

$$\begin{aligned}e^Ae^B &= \left(1+A+\tfrac{1}{2}A^2+...\right)\left(1+B+\tfrac{1}{2}B^2+...\right)\\ &= 1+A+B+\tfrac{1}{2}A^2+AB+\tfrac{1}{2}B^2+...\end{aligned} \tag{56}$$

On the other hand, if we consider $e^{(A+B)}$, we find:

$$\begin{aligned}e^{(A+B)} &= 1+A+B+\tfrac{1}{2}(A+B)^2+...\\ &= 1+A+B+\tfrac{1}{2}A^2+\tfrac{1}{2}AB+\tfrac{1}{2}BA+\tfrac{1}{2}B^2+...\end{aligned} \tag{57}$$

Note that equation (55) prevents us from combining the terms in AB and BA. Comparing (56) and (57), we see that

$$e^Ae^B \approx \exp\left(A+B+\tfrac{1}{2}\{A,B\}+...\right). \tag{58}$$

Equation (58) is true for any sort of operators for which we can define an exponential, and is called the Baker-Campbell-Hausdorff (BCH) formula.

If we are to apply the BCH formula, it will clearly be an advantage to have an expression for the commutator of two Lie operators. A suitable formula is not hard to find. Consider, for any three functions $f(x,p)$, $g(x,p)$ and $h(x,p)$,

$$\begin{aligned}\{:f:,:g:\}h &=:f:(:g:h)-:g:(:f:h)\\ &=\left[f,\left[g,h\right]\right]-\left[g,\left[f,h\right]\right]\\ &=\left[f,\left[g,h\right]\right]+\left[g,\left[h,f\right]\right]\end{aligned} \tag{59}$$

where in the last line we have used the antisymmetry of the Poisson bracket. We now use the Jacobi identity[5]

$$\left[f,\left[g,h\right]\right]+\left[g,\left[h,f\right]\right]+\left[h,\left[f,g\right]\right]=0 \tag{60}$$

to rewrite the last line of (60),

[5] The Jacobi identity is normally written in terms of the commutator, but (60) is easily verified for the Poisson bracket.

$$\begin{aligned}\{:f:,:g:\}h &= -[h,[f,g]] \\ &= [[f,g],h]\end{aligned}$$

We thus arrive at:

$$\{:f:,:g:\}h =: (:f:g):h\,. \tag{61}$$

As an illustration of the use of the BCH formula, we return to the factorisation of the sextupole map. Using the BCH formula, the right hand side of (53) becomes

$$\exp(-l:\tfrac{1}{2}p^2:)\exp(-l:\tfrac{1}{6}k_2x^3:) = \exp(-l:\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3: + \tfrac{1}{24}l^2k_2\{:p^2:,:x^3:\} + \ldots)$$

Using (61), we find

$$\{:p^2:,:x^3:\} = -6:x^2p:$$

and thus

$$\exp(-l:\tfrac{1}{2}p^2:)\exp(-l:\tfrac{1}{6}k_2x^3:) = \exp(-l:\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3: - \tfrac{1}{4}l^2k_2:x^2p + \cdots:)\,. \tag{62}$$

Equation (62) indicates a correction to the factorisation of the sextupole map that is first order in the sextupole strength k_3, and second order in the magnet length l; it might be expected, therefore, that this factorisation is not a good approximation to the full map. To improve the approximation, we turn to a technique called *symmetric factorisation*.

4.3 Symmetric Factorisation and Yoshida Theory

Instead of factorising the map of a long sextupole into a drift space followed by a kick, we can place the kick in the centre of the sextupole:

$$\exp(-l:\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3:) \approx \exp(-\tfrac{1}{4}l:p^2:)\exp(-\tfrac{1}{6}lk_2:x^3:)\exp(-\tfrac{1}{4}l:p^2:) \tag{63}$$

Because of the apparent drift-kick-drift symmetry of the operators on the right hand side, equation (63) is called a symmetric factorisation of the sextupole map. It is straightforward to check that this factorisation is solvable. We also find, using the BCH formula, that

$$\exp(-\tfrac{1}{4}l:p^2:)\exp(-\tfrac{1}{6}lk_2:x^3:)\exp(-\tfrac{1}{4}l:p^2:) \approx \exp(-l:\tfrac{1}{2}p^2 + \tfrac{1}{6}k_2x^3: + \tfrac{1}{4}l^3k_2:xp^2: + \ldots)\,. \tag{64}$$

Note that the correction term in (64) is now third order in the magnet length, although still first order in the sextupole strength; for a short (even if not infinitesimally thin) magnet, symmetric factorisation thus improves the approximation to the full map.

It is possible, by further factorisations, to arrive at a still better approximation. We continue to produce solvable maps, but with each factorisation, we increase the order of the series representation of the map. For example, our original factorisation (53) gives a series that is first order (i.e. linear) for the mapping of x, though second order in the mapping of p. The symmetric factorisation (63) gives:

$$\begin{aligned}\exp(-l:\tfrac{1}{4}p^2:)\exp(-lk_2:\tfrac{1}{6}x^3:)\exp(-l:\tfrac{1}{4}p^2:)x &= x + lp - \tfrac{1}{4}k_2l^2x^2 - \tfrac{1}{4}k_2l^3xp - \tfrac{1}{16}k_2l^4p^2 \\ \exp(-l:\tfrac{1}{4}p^2:)\exp(-lk_2:\tfrac{1}{6}x^3:)\exp(-l:\tfrac{1}{4}p^2:)p &= p - \tfrac{1}{2}k_2lx^2 - \tfrac{1}{2}k_2l^2xp - \tfrac{1}{8}k_2l^3p^2\end{aligned} \tag{65}$$

which is second order in both the x and p mappings. It is clear that for a sextupole, each 'kick' increases the order of the mapping by two. Thus, the next factorisation, which would correspond to two sextupole kicks along the length of the magnet, would lead to a fourth-order mapping.

In general, a symmetric factorisation attempts to express a Lie transform in the form

$$\exp(\varepsilon{:}A{:}) = \exp(\varepsilon{:}A_1{:})\exp(\varepsilon{:}A_2{:})\cdots\exp(\varepsilon{:}A_n{:})\cdots\exp(\varepsilon{:}A_2{:})\exp(\varepsilon{:}A_1{:}). \tag{66}$$

A number of people have devised techniques for producing factorisations of the form (66) for increasingly large values of n. Yoshida has produced methods that are relatively straightforward to apply, and these are described in reference [3].

It is worth remarking that when we construct a solvable map approximation to our lattice, then all the parameters of the map (including magnet lengths and field strengths) should be regarded as free parameters, that may be varied to fit the observed dynamics of the real lattice. This may seem disappointing if we are hoping to build models that can predict accurately the properties of a lattice at the design stage, when we have no experimental data against which to match our model. However, we are really in no worse position than if we were using any other modelling technique. In practice, we need a theory for which we understand the effects of each of the parameters, and this is exactly what the Lie operator approach gives us. In a Newtonian approach to nonlinear dynamics, for example, we quickly become lost in trying to solve complicated differential equations, and the effects of varying the parameters of the model can appear unpredictable. The significance of particular variables quickly becomes blurred as soon as we resort to numerical techniques. As we shall see, by using Lie operators, we arrive at a model that is simple enough to provide powerful insights, while retaining enough structure for us to be able to reproduce many of the significant features of the system. Furthermore, if we compare the results of different modelling techniques, we find that it is not always necessary to adjust the parameters of our model to fit expected results.

4.4 When is a Map Solvable?

We have so far assumed that each of the factors in the factorisation of our map produces a solvable map. Our aim is always to take a non-solvable map, and manipulate it into a solvable approximation. It would therefore be helpful to have a 'catalogue' of solvable maps we can use in producing our approximations. We present the following results without proof.

Expressed very simply, we find that:

$\exp(:f:)$ is solvable where

$$\left.\begin{aligned} f(x,p) &= \varepsilon x^a \\ f(x,p) &= \varepsilon p^a \end{aligned}\right\} \text{ for integer } a \text{, and constant } \varepsilon;$$

$$f(x,p) = \sum_{i=1}^{\infty} \mu_n J^n \quad \text{where } J = \tfrac{1}{2}\left(p^2 + kx^2\right). \tag{67}$$

Note that the mapping $\exp(-\mu{:}J{:})$ corresponds to a quadrupole map, with phase advance $\sqrt{k}\mu$. Maps of the form (67) are of particular importance for the analysis of nonlinear dynamics, as we now see.

5. Nonlinear Beam Dynamics

Central to the understanding of the effects of nonlinear lattice elements is the idea of the *nonlinear rotation*. We have seen how a linear quadrupole, for example, leads to a rotation of states in phase space around the origin. The angle of rotation (phase advance) is independent of the initial state. In a nonlinear rotation, the phase advance is a function of the distance of the state from the origin in phase space. Nonlinear rotations are useful, because they can be represented easily by solvable maps. In this

section, we show how the effects of various nonlinear elements can be approximated by nonlinear rotations. Although more complex (and non-solvable) maps are required to reproduce the dynamics of a real system with any accuracy, simple nonlinear rotations are sufficient to allow an understanding of many features of the behaviour.

Related to the idea of a nonlinear rotation is the process of *normalisation*. We have seen how, in a linear lattice, the most general motion of a particle under the action of a single-turn map is an orbit in the shape of an ellipse. In normalisation, we carry out a canonical transformation to turn the ellipse into a circle. Since a canonical transformation amounts to a change of variables, a number of important features of the system, such as the linear phase advance, are preserved. We begin by illustrating normalisation in the linear case, using a method that is sufficiently general to work for nonlinear systems as well.

5.1 The Courant-Snyder Transformation: Normalisation of a Linear Map

The general transfer map for a linear symplectic map was given in (37):

$$\mathbf{M} = \begin{pmatrix} \cos(\mu) + \alpha\sin(\mu) & \beta\sin(\mu) \\ -\gamma\sin(\mu) & \cos(\mu) - \alpha\sin(\mu) \end{pmatrix}. \tag{37}$$

As we have seen, the orbit in phase space for a particle moving under the action of this map is an ellipse:

$$\gamma x^2 + 2\alpha xp + \beta p^2 = \varepsilon. \tag{33}$$

Our aim in normalising the map is to find a canonical transformation which turns $\mathbf{M}$ into a circular rotation, rather than an 'elliptical rotation', i.e. we wish to find a matrix $\mathbf{A}$ such that:

$$\mathbf{A}^{-1}\mathbf{M}\mathbf{A} = \mathbf{R} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{68}$$

(this is the transfer map analogue of equation (30) for Koopman maps). In fact, there are an infinite number of transformations $\mathbf{A}$ that satisfy (68); the conventional choice is the Courant-Snyder transformation given by [3]:

$$\mathbf{A} = \begin{pmatrix} \sqrt{\beta} & 0 \\ -\dfrac{\alpha}{\sqrt{\beta}} & \dfrac{1}{\sqrt{\beta}} \end{pmatrix}. \tag{69}$$

It is straightforward to verify that (69) satisfies (68), with $\theta = \mu$.

Finally, we note that a circular rotation is generated by a function of the form (33) with $\gamma = \beta = 1$ and $\alpha = 0$; thus the Lie transform is given by:

$$\mathcal{M} = e^{:-\mu J:}$$

where

$$J = \tfrac{1}{2}\left(x^2 + p^2\right).$$

5.2 Physical Observables and Gauge Invariance

We mentioned before equation (69) that the canonical transformation to turn the linear single-turn map into a rotation is not unique. This means that if we hope to extract physically measurable quantities (which we call physical observables) from the transformed map, the values we obtain cannot depend on

the precise transformation we use. In other words, choosing a transformation amounts to a choice of gauge, and physical observables must be gauge invariant.

For example, the multiple turn average $\langle f \rangle$ of some function $f(x,p)$ is clearly a physical observable. In many cases, it is convenient to obtain a value for $\langle f \rangle$ from the normalised map, and by our argument above, the value we obtain must be gauge invariant. We shall proceed by first deriving an expression for $\langle f \rangle$, and then examining its gauge invariance.

If the single-turn map is $\mathcal{M}$, then we have

$$\langle f \rangle = \lim_{p\to\infty} \frac{1}{p} \sum_{k=1}^{p} \mathcal{M}^k f \,. \tag{70}$$

If $\mathcal{M}$ is normalised by a map $\mathcal{A}$, we have

$$\mathcal{R} = \mathcal{A}\mathcal{M}\mathcal{A}^{-1}$$
$$\mathcal{M} = \mathcal{A}^{-1}\mathcal{R}\mathcal{A}$$

thus

$$\mathcal{M}^k f = \left(\mathcal{A}^{-1}\mathcal{R}\mathcal{A}\right)^k f = \mathcal{A}^{-1}\mathcal{R}^k \mathcal{A} f \,.$$

Note that since $\mathcal{R}$ is a pure rotation, we can write

$$\mathcal{R} = e^{:-\mu J:}.$$

To proceed, it is helpful to introduce the eigenfunctions of the operator $e^{:-\mu J:}$. It is easy to check that these are the functions

$$h_{\pm} = x \pm ip \,,$$

which have the properties

$$e^{:-\mu J:}\left(x \pm ip\right) = e^{\mp i\mu}\left(x \pm ip\right)$$
$$e^{:-\mu J:}\left(x \pm ip\right)^m = e^{\mp im\mu}\left(x \pm ip\right)$$
$$e^{:-\mu J:}\left(x + ip\right)^m\left(x - ip\right)^n = e^{-i(m-n)\mu}\left(x + ip\right)^m\left(x - ip\right)^n \tag{71}$$

We then expand the transformed function $\mathcal{A}f$ as a series in h_+ and h_-:

$$\mathcal{A}f = \sum_{m,n=0}^{\infty} C\left(\mathcal{A}f; m, n\right) h_+^m h_-^n \,,$$

then

$$\mathcal{M}^k f = \mathcal{A}^{-1}\mathcal{R}^k \sum_{m,n=0}^{\infty} C\left(\mathcal{A}f; m, n\right) h_+^m h_-^n$$
$$= \mathcal{A}^{-1} \sum_{m,n=0}^{\infty} C\left(\mathcal{A}f; m, n\right) \exp\left(-ik(m-n)\mu\right) h_+^m h_-^n \tag{72}$$

If we substitute (72) into (70) and take the limit $p \to \infty$, only terms in (72) for which $m = n$ survive, and thus we are left with:

$$\langle f \rangle = \mathcal{A}^{-1} \sum_{n=0}^{\infty} C(\mathcal{A}f; n, n) h_+^n h_-^n$$

which we can write as:

$$\langle f \rangle = \sum_{n=0}^{\infty} 2^n C(\mathcal{A}f; n, n)(\mathcal{A}^{-1} J). \tag{73}$$

If the single-turn map is

$$\mathcal{M} = e^{:-\mu F:}$$

then

$$\begin{aligned} \exp(:-\mu F:) &= \mathcal{A}^{-1} \mathcal{R} \mathcal{A} \\ &= \mathcal{A}^{-1} e^{:-\mu J:} \mathcal{A} \\ &= \exp(:-\mu \mathcal{A}^{-1} J:) \end{aligned}$$

where we have used (30) in the last line. Thus:

$$\mathcal{A}^{-1} J = F \tag{74}$$

and so finally, from (73), we have:

$$\langle f \rangle = \sum_{n=0}^{\infty} 2^n C(\mathcal{A}f; n, n) F \tag{75}$$

Equation (75) is not explicitly gauge invariant, as it still involves the normalising transformation. To see that we can obtain a gauge invariant result, let us consider the case $f = x^2$, where the invariant of the single-turn map is given by

$$F = \tfrac{1}{2}(\gamma x^2 + 2\alpha x p + \beta p^2) = \tfrac{1}{2}\varepsilon. \tag{76}$$

The matrix representation of the transfer map corresponding to $\mathcal{A}$ is:

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Since $\mathbf{A}$ must be symplectic[6], we have

$$\mathrm{Det}(\mathbf{A}) = A_{11} A_{22} - A_{12} A_{21} = 1$$

$$\mathbf{A}^{-1} = \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

Then

$$\mathbf{A}^{-1} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} A_{22} x - A_{12} p \\ -A_{21} x + A_{11} p \end{pmatrix}$$

so

[6] See section 1.3.

$$\begin{aligned}\mathcal{A}^{-1}J &= \tfrac{1}{2}\mathcal{A}^{-1}\left(x^2+p^2\right)\\ &= \tfrac{1}{2}\left(A_{22}^2+A_{21}^2\right)x^2+\left(-A_{22}A_{12}-A_{21}A_{11}\right)xp+\tfrac{1}{2}\left(A_{11}^2+A_{12}^2\right)p^2\end{aligned} \tag{77}$$

Using (74), (76) and (77), we find:

$$\begin{aligned}\gamma &= A_{22}^2+A_{21}^2\\ \alpha &= -A_{22}A_{12}-A_{21}A_{11}\\ \beta &= A_{11}^2+A_{12}^2\end{aligned}$$

We also have that

$$\mathbf{A}\begin{pmatrix}x\\p\end{pmatrix}=\begin{pmatrix}A_{11}x+A_{12}p\\A_{21}x+A_{22}p\end{pmatrix}$$

so

$$\begin{aligned}\mathcal{A}x^2 &= \left(A_{11}x+A_{12}p\right)^2\\ &= \left(A_{11}\cdot\tfrac{1}{2}\left(h_+ + h_-\right)+A_{12}\cdot\tfrac{1}{2i}\left(h_+ - h_-\right)\right)^2\end{aligned}$$

from which we see that

$$C\left(\mathcal{A}x^2;1,1\right)=\tfrac{1}{2}\left(A_{11}^2+A_{12}^2\right)=\tfrac{1}{2}\beta \tag{78}$$

and all other $C\left(\mathcal{A}x^2;n,n\right)$ appearing in (75) are zero. Substituting (76) and (78) into (75) yields the final result:

$$\left\langle x^2\right\rangle=\tfrac{1}{2}\beta\varepsilon\,. \tag{79}$$

Since β and ε are parameters of the single-turn map, we see that the expression for $\left\langle x^2\right\rangle$ given in (79) is independent of the normalising transformation $\mathcal{A}$, and is thus explicitly gauge invariant. For this case, it could be argued that it is hardly necessary to consider normalisation to obtain an expression for $\left\langle x^2\right\rangle$, but there are more complex cases where an expression of the form (75) could be very useful. If we are careful to ask for quantities that are physical observables, then (75) should always yield a gauge invariant result.

5.3 Normalisation of a Linear Map by Lie Operator Methods

Suppose we have a linear, normalised map, to which we add a quadrupole 'kick' as a perturbation of order ε. This map may be written as

$$\mathcal{M}=\exp\left(:-\mu J:\right)\exp\left(:-\tfrac{1}{2}\varepsilon x^2:\right). \tag{80}$$

Note that the 'unperturbed' phase advance is μ: it is easy to check that the transfer map corresponding to the operator $\exp\left(:-\mu J:\right)$ is just a circular rotation through μ. The effect of the perturbation is to deform the circular map into an ellipse. Our aim is to find a canonical transformation $\mathbf{A}$ such that the perturbed map is restored to a circular rotation:

$$\mathcal{R}=\exp\left(:-\mu' J:\right)=\mathcal{A}\mathcal{M}\mathcal{A}^{-1}, \tag{81}$$

c.f. equation (27). Since the perturbation is linear, we expect to be able to solve (81) exactly, though this will not be the case for a nonlinear perturbation, e.g. a sextupole instead of a quadrupole kick. Indeed, there are a number of methods for identifying an appropriate operator $\mathcal{A}$ for equation (81).

For example, moving to a matrix representation of the corresponding transfer maps, if we construct a matrix $\mathbf{V}$ from the eigenvectors of $\mathbf{R}$, then

$$\mathbf{VRV}^{-1} = \mathbf{VAMA}^{-1}\mathbf{V}^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ is diagonal,} \tag{82}$$

where λ_1 and λ_2 are the eigenvalues of $\mathbf{R}$. But then, if we construct a matrix $\mathbf{D}$ from the eigenvectors of $\mathbf{M}$, we can normalise $\mathbf{D}$ such that

$$\mathbf{DMD}^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{83}$$

We then have, by comparison of (82) and (83),

$$\mathbf{D} = \mathbf{VA}$$

$$\mathbf{A} = \mathbf{V}^{-1}\mathbf{D}$$

Since the construction of eigenvectors and eigenvalues of square matrices is a straightforward procedure, we have an effective way of solving (81). Unfortunately, this method only works in the linear case. A more general method, which generalises to nonlinear cases, is as follows.

First of all, we rewrite (81), using the Lie transform representation for the (Koopman) maps instead of a (transfer map) matrix representation. Thus:

$$e^{:F:}\mathcal{M}e^{:-F:} \approx \mathcal{R} = e^{-:G:}e^{:-\mu J:} \tag{84}$$

Notice that we write the full rotation as a product of the original rotation, $e^{:-\mu J:}$, and an additional rotation $e^{:-G:}$. But $e^{:-G:}$ is only a rotation if

$$G = \delta\mu \cdot J \tag{85}$$

where $\delta\mu$ is an additional phase shift arising from the perturbing element; equation (85) therefore fixes[7] the form of the function G. Let us now write[8]

$$\mathcal{M} = e^{:-\mu J:}e^{:-f:}$$

and manipulate the left hand side of (84) to make it look explicitly like the right hand side. We find:

$$\begin{aligned} e^{:F:}\mathcal{M}e^{:-F:} &= e^{:F:}e^{:-\mu J:}e^{:-f:}e^{:-F:} \\ &= e^{:F:}e^{:-\mu J:}e^{:-f:}e^{:\mu J:}e^{:-\mu J:}e^{:-F:}e^{:\mu J:}e^{:-\mu J:} \\ &= e^{:F:}\exp\left(:-e^{:-\mu J:}f:\right)\exp\left(:-e^{:-\mu J:}F:\right)e^{:-\mu J:} \end{aligned} \tag{86}$$

where we have used equation (30) twice to achieve the last line. Using the BCH formula *to first order* on (86), we have:

$$e^{:F:}\mathcal{M}e^{:-F:} \approx \exp\left(:\left(1-e^{:-\mu J:}\right)F - e^{:-\mu J:}f:\right)e^{:-\mu J:} \tag{87}$$

Comparing this with (84) and (85), we see that

$$e^{:-\mu J:}f - \left(1-e^{:-\mu J:}\right)F \approx G = \delta\mu \cdot J \tag{88}$$

---

[7] In the nonlinear case, we shall loosen the condition expressed by (85) to make G a more general function of J. We find that $\mathcal{R}$ is still a (kind of) rotation.

[8] For a quadrupole kick, as in (80), we should use $f = \frac{1}{2}\varepsilon x^2$.

Equation (88) looks formidable, but in fact it is open to analysis. The aim is to find a form for F, such that the left hand side of (88) is proportional to J (or more generally for the nonlinear case, a function of J only).

As in section 5.2, we use the eigenfunctions $h_{\pm}$ of $e^{:-\mu J:}$ as a basis for series expansions of functions of x and p, thus

$$J = \tfrac{1}{2}\left(x^2 + p^2\right) = \tfrac{1}{2} h_+ h_- \tag{89}$$

$$f = \tfrac{1}{2}\varepsilon x^2 = \tfrac{1}{8}\varepsilon(h_+ + h_-)^2 = \tfrac{1}{8}\varepsilon h_+^2 + \tfrac{1}{4}\varepsilon h_+ h_- + \tfrac{1}{8}\varepsilon h_-^2 \tag{90}$$

$$F = \sum_{m,n=0}^{\infty} c_{mn} h_+^m h_-^n \tag{91}$$

where we have used a general expansion for F in terms of coefficients c_{mn}. Note that for F to be real, we need to impose the condition

$$c_{mn} = c_{nm}^* .$$

Looking at (88), we need to apply the operators $\left(1 - e^{:-\mu J:}\right)$ and $e^{:-\mu J:}$ to the functions F and f respectively. Using equations (71) and (89), we find:

$$\begin{aligned} e^{:-\mu J:} f &= e^{:-\mu J:}\left(\tfrac{1}{8}\varepsilon h_+^2 + \tfrac{1}{4}\varepsilon h_+ h_- + \tfrac{1}{8}\varepsilon h_-^2\right) \\ &= \tfrac{1}{8}\varepsilon\left(e^{-2i\mu} h_+^2 + 2h_+ h_- + e^{2i\mu} h_-^2\right) \end{aligned} \tag{92}$$

and

$$\begin{aligned} \left(1 - e^{:-\mu J:}\right)F &= \sum_{m,n=0}^{\infty}\left(1 - e^{i(n-m)\mu}\right) c_{mn} h_+^m h_-^n \\ &= 2i \sum_{m,n=0}^{\infty} e^{-\frac{1}{2}i(m-n)\mu} \sin\left(\tfrac{1}{2}(m-n)\mu\right) c_{mn} h_+^m h_-^n \end{aligned} \tag{93}$$

Inspecting the form of equation (93), we notice that we can choose the c_{mn} so that terms from $\left(1 - e^{:-\mu J:}\right)F$ cancel terms in h_+^2 and h_-^2 from $e^{:-\mu J:} f$ in equation (88). We cannot do anything about the term in $h_+ h_-$ from (92), since the corresponding coefficient in the expansion of $\left(1 - e^{:-\mu J:}\right)F$ is zero, but this does not matter, because from (89), $h_+ h_-$ is proportional to J, and this is the term we wish to leave. By inspection of (93), if we choose:

$$c_{20} = \tfrac{1}{16} i\varepsilon e^{-i\mu} \operatorname{cosec}(\mu)$$

$$c_{02} = -\tfrac{1}{16} i\varepsilon e^{i\mu} \operatorname{cosec}(\mu)$$

$$\text{all other } c_{mn} = 0$$

we achieve the desired cancellation. Inserting these coefficients into (91), we then find, after a little algebra,

$$F = \tfrac{1}{8}\varepsilon\left(x^2 - 2xp\cot(\mu) - p^2\right). \tag{94}$$

Equation (94) for F is almost incidental; the result we are really interested in is obtained simply by taking the terms in (92), which are not cancelled by (93) when used in (88). By this means we find:

$$\delta\mu \cdot J = G \approx \tfrac{1}{4}\varepsilon h_+ h_- = \tfrac{1}{2}\varepsilon J \tag{95}$$

and thus, finally:

$$\delta\mu \approx \tfrac{1}{2}\varepsilon\,. \tag{96}$$

The approximation in (95) comes from the fact that we have treated the perturbation $e^{:-f:}$ to first order using the Baker-Campbell-Hausdorff formula. Equation (96) thus gives the first order perturbation shift in tune from a quadrupole kick of strength ε. Returning to (80) and (83), we see that we can write:

$$e^{:-\mu J:}e^{:-\frac{1}{2}\varepsilon x^2:} \approx e^{:-F:}e^{:-(\mu+\frac{1}{2}\varepsilon)J:}e^{:F:} = \exp\left(:-\left(\mu+\tfrac{1}{2}\varepsilon\right)e^{:-F:}J:\right) \tag{97}$$

where F is given by (94). Writing the perturbation in the form (97) makes it clear that F generates a canonical transformation which turns the perturbed map into a rotation through an angle $\mu+\frac{1}{2}\varepsilon$. As we shall now see, the effect of a nonlinear perturbation is to make the angle of rotation dependent on the 'amplitude' of the state of the particle in phase space.

5.4 Normalisation of a Nonlinear Map

Generalising the procedure of the previous section to deal with nonlinear perturbations is surprisingly straightforward[9]. The trick is always to expand the perturbation function f in terms of the eigenfunctions $h_\pm$ of J:

$$f = \sum_{m,n=0}^{\infty} C(f;m,n)h_+^m h_-^n\,. \tag{98}$$

We then look for a function F such that, from (88),

$$e^{:-\mu J:}f - \left(1-e^{:-\mu J:}\right)F = G(J) \tag{99}$$

where G is a (nonlinear) function of J. In fact, from the work of the previous section, we can see that if we expand F as before:

$$F = \sum_{m,n=0}^{\infty} c_{mn}h_+^m h_-^n \tag{91}$$

then the results we need are:

$$\left.\begin{aligned} c_{mn} &= \tfrac{1}{2}ie^{\frac{1}{2}i(m-n)\mu}\operatorname{cosec}\left(\tfrac{1}{2}(m-n)\mu\right)C(f;m,n), && m\neq n\\ c_{mn} &= 0, && m=n\end{aligned}\right\} \tag{100}$$

$$G = \sum_{m=1}^{s} 2^m C(f:m,m)J^m \tag{101}$$

where s is the maximum order of h_+h_- occurring in (98). Note that we have replaced the approximation in (88) by an equality in (99). This is achieved by making (99) a definition of $G(J)$; the approximation enters when we relate the perturbed map to the normalised map expressed in terms of G, c.f. (84):

$$e^{:F:}e^{:-\mu J:}e^{:-f:}e^{:-F:} \approx e^{:-G(J):}e^{:-\mu J:}\,. \tag{102}$$

Equations (91) and (98) to (102) summarise the results of the normalisation of a map to first order in a nonlinear perturbation. For convenience, we refer to this as *first order perturbation theory*. To go to

[9] Actually, the procedure has been designed so that it *can* be easily generalised to nonlinear cases.

second order perturbation theory would involve including higher order terms from the Baker-Campbell-Hausdorff formula in equation (87); this makes the analysis significantly more involved, though results can be obtained. We do not attempt to go into the second order theory here, but refer the interested reader to [3]. We mention that Forest has produced a code [9] that can carry out normalisation to arbitrary order; there are instructive notes accompanying the program.

5.5 Nonlinear Rotations

A normalisation of the general form of equation (102) is a powerful tool, because it allows a simple interpretation of the normalised map. Since G is a function of J only, we have

$$\{G, J\} = 0$$

and so the BCH formula applied to (102) gives:

$$e^{:F:} e^{:-\mu J:} e^{:-f:} e^{:-F:} \approx e^{:-\mu J - G(J):}.$$

Furthermore, it can be shown that if

$$G = \sum_{n=0}^{\infty} \mu_n J^{n+1} \tag{103}$$

then

$$e^{:-G:} = \mathcal{R}_\theta \tag{104}$$

where $\mathcal{R}_\theta$ is the operator for a rotation through an angle θ, and

$$\theta = \sum_{n=0}^{\infty} \frac{n+1}{2^n} \mu_n (x^2 + p^2)^n . \tag{105}$$

Equations (103), (104) and (105) define a *nonlinear rotation*, in which the angle of rotation is a function of the phase space amplitude. Thus, the effect of a nonlinear perturbation on a linear mapping is to introduce an amplitude dependence of the tune, i.e. of the phase advance per mapping operation.

As an example, let us consider an octupole (quartic) perturbation. In our usual notation, the generating function of the perturbation can be written

$$\begin{aligned} f &= \tfrac{1}{4} \varepsilon x^4 \\ &= \tfrac{1}{4} \varepsilon \left(\frac{h_+ + h_-}{2} \right)^4 \\ &= \tfrac{1}{64} \varepsilon \left(h_+^4 + 4h_+^3 h_- + 6h_+^2 h_-^2 + 4h_+ h_-^3 + h_-^4 \right) \end{aligned}$$

From (101) this gives immediately:

$$G = \tfrac{3}{8} \varepsilon J^2 , \tag{106}$$

hence

$$e^{:-\mu J:} e^{:f:} \approx e^{:-F:} \exp\left(:-\mu J - \tfrac{3}{8} \varepsilon J^2 :\right) e^{:F:} .$$

So the phase advance in the presence of the perturbation is

$$\theta_{\text{octupole}} \approx \mu + \tfrac{3}{8} \varepsilon \left(x^2 + p^2 \right). \tag{107}$$

If required, we can readily find the generator F of the canonical transformation that normalises the map, using equations (91) and (100). In the present case, we find:

$$F = \tfrac{1}{64}\varepsilon\left(5x^4 - \left(8\cot(\mu) + 4\cot(2\mu)\right)x^3p - 6x^2p^2 - \left(8\cot(\mu) - 4\cot(2\mu)\right)xp^3 - 3p^4\right).$$

Equation (106) generalises for the case of a *2n*-pole kick, where *n* is even (i.e. quadrupole, octupole, dodecapole etc.) . If

$$f = \tfrac{1}{n}\varepsilon x^n$$

then

$$G = \frac{\varepsilon}{2^n}\cdot\frac{(n-1)!}{\frac{n}{2}!^2}J^{n/2}$$

For the case of a *2n*-pole kick where *n* is odd (sextupole, decapole…), then *to first order in perturbation theory there is no change in the tune* of the map. In these cases we need to go to at least second order to find a tune shift. Note that there is a tune shift which is first order in the sextupole strength, but we need to go to second order in perturbation theory to find it. Unfortunately, these are often cases that we are interested in, e.g. when studying the effects of chromatic sextupoles.

The change in tune with amplitude is an important concept, that, as we shall see, can be used to explore some of the most significant features of nonlinear maps, including the stability of the system. For example, suppose that the tune of a linear lattice is slightly below $2\pi n + \frac{1}{4}$, where n is an integer. If a short octupole is added to the lattice, to give the effect of an octupole kick, then from (107), we see that at certain amplitudes, the tune can be increased by the perturbation to reach exactly $2\pi n + \frac{1}{4}$. The precise amplitude at which this happens depends on the strength of the perturbation, but when it does happen, we find that on every fourth turn, the particle returns exactly to its starting point. In other words, the perturbation pushes some particles at large enough amplitudes onto a fourth order resonance. This in itself does not necessarily make the map unstable. Points in phase space which are mapped back to themselves after m turns (operations of the map) are fixed points of the mth iterate of the map. In simple terms, the fixed points can be either elliptic or hyperbolic. If elliptic, then points near the fixed points follow (nearly) elliptical orbits in phase space around the fixed points; if hyperbolic, then the 'amplitude' of nearby points will increase without limit as the map continues to be applied[10]. The latter case is clearly disastrous for the stability of a beam of particles in an accelerator. On the other hand, the former case is undesirable if we are attempting resonant extraction.

Since any nonlinear perturbation leads to a shift in tune with amplitude, any nonlinear element can potentially excite any order of resonance, if the linear tune is close to enough to the resonance and on the right side of it. For example, it is perfectly possible for a sextupole to excite a fourth order resonance. This is in disagreement with the conventional wisdom on nonlinear effects in accelerators, derived from an attempt to analyse such systems with an approach rooted in Newtonian mechanics. It is common for introductory texts to state a relationship between the order of the nonlinearity of the magnetic field and the order of the resonance driven. For example, one author [10] states:

> "Higher order resonances are driven by magnetic fields with higher order nonlinearities. For instance, octupole fields…drive fourth order resonances. Decapole fields drive fifth order resonances, and so on."

The author of [11] makes a similar assertion, though giving more mathematical detail. However, the perturbation theory based on the properties of Lie transforms suggests that such statements are misleading, and as we shall see, the results of simulations using tracking codes clearly support the theory developed in works such as [3], and summarised in the present note.

[10] There are of course a variety of other possibilities - see any text on nonlinear dynamics, e.g. [8].

5.6 Normalisation of a Sequence of Elements

Having seen the procedure for the normalisation of a perturbed rotation, we continue by considering the normalisation of a sequence of lattice elements. For this section, we draw heavily on reference [3]; in fact we attempt to give here only a brief sketch of the more rigorous argument that Forest uses to obtain the significant results.

Consider the map given by a sequence of four elements, namely a dipole, quadrupole, octupole and a second dipole:

$$\mathcal{M} = \mathcal{D}\mathcal{Q}\mathcal{O}\mathcal{D}$$

This is the example used in reference [3], though any sequence of linear elements with a single nonlinear element at some position would suit our purposes. It is possible, of course, construct different maps starting from any point in this sequence, thus,

$$\mathcal{M}_0 = \mathcal{D}\mathcal{Q}\mathcal{O}\mathcal{D}$$
$$\mathcal{M}_1 = \mathcal{Q}\mathcal{O}\mathcal{D}\mathcal{D}$$
$$\mathcal{M}_2 = \mathcal{O}\mathcal{D}\mathcal{D}\mathcal{Q}$$
$$\mathcal{M}_3 = \mathcal{D}\mathcal{D}\mathcal{Q}\mathcal{O}$$

We can construct Courant-Snyder transformations $\mathcal{A}_i$ to normalise the linear parts of each of these maps (i.e. neglecting nonlinear terms):

$$\mathcal{R}_i = \mathcal{A}_i \mathcal{M}_i \mathcal{A}_i^{-1}$$

If we now consider the linearly normalised map $\mathcal{A}_0 \mathcal{M}_0 \mathcal{A}_0^{-1}$, we can write

$$\begin{aligned} \mathcal{A}_0 \mathcal{M}_0 \mathcal{A}_0^{-1} &= \mathcal{A}_0 \mathcal{D}\mathcal{Q}\mathcal{O}\mathcal{D} \mathcal{A}_0^{-1} \\ &= \mathcal{A}_0 \mathcal{D} \mathcal{A}_1^{-1} \mathcal{A}_1 \mathcal{Q} \mathcal{A}_2^{-1} \mathcal{A}_2 \mathcal{A}_3^{-1} \mathcal{A}_3 \mathcal{O} \mathcal{A}_3^{-1} \mathcal{A}_3 \mathcal{D} \mathcal{A}_0^{-1} \end{aligned} \tag{108}$$

Note that in the last line of (108), by inserting a set of operators $\mathcal{I} = \mathcal{A}_i^{-1}\mathcal{A}_i$, we have introduced transformations on each element in the sequence. The remarkable result is that, although these transformations are not strictly of a kind we have seen so far, they normalise the linear elements into rotations, e.g.

$$\mathcal{A}_0 \mathcal{D} \mathcal{A}_1^{-1} = \mathcal{R}_{0\to 1}$$
etc.

Thus:

$$\mathcal{A}_0 \mathcal{M}_0 \mathcal{A}_0^{-1} = \mathcal{R}_{0\to 1}\mathcal{R}_{1\to 2}\mathcal{R}_{2\to 3}\tilde{\mathcal{O}}\mathcal{R}_{3\to 4}$$

where

$$\tilde{\mathcal{O}} = \mathcal{A}_3 \mathcal{O} \mathcal{A}_3^{-1}. \tag{109}$$

Note that the map has now been transformed into a rotation with a nonlinear element at one location. Hopefully, the effect of the nonlinear element is small. If this is the case, a further remarkable result is that we can carry out a nonlinear normalisation on the linearly normalised map, and find that *to first order*, the effect of the perturbation is localised at the nonlinear element. To second or higher order in the perturbation, the nonlinear element has more complicated effects.

Our conclusion is that to first order, a seemingly complex sequence of elements can be modelled very simply, in this case as a linear phase advance equal to the tune of the lattice, with an amplitude-

dependent phase advance resulting from the presence of a nonlinear element. The nonlinear phase advance is not exactly as would be expected, for example from (107), because of the transformation (109). As mentioned at the end of section 4.3, however, the parameters of any model should not be regarded as fixed. We shall see that with a pragmatic approach, even a very simple model can accurately reproduce significant features of an apparently complex map.

6. Results of Tracking Experiments

The emphasis in this note has been on the use of Lie operator methods to gain an insight into the behaviour of linear and nonlinear maps. We have developed techniques that enable us to construct simple models, that we hope would reproduce significant features of more detailed models that can be investigated using conventional tracking codes. A possible drawback with using conventional codes is that they can easily be seen as 'black boxes', where a particular lattice structure is fed in, and tracking results are obtained, with the user gaining very little understanding over how the results depend on particular features of the lattice. Our aim is to show how the user can control the dynamics, in particular nonlinear effects, by identifying relevant parameters in the lattice, and understanding how these will affect the map. For simplicity, we work in one degree of freedom; this avoids the effects of coupling, which could obscure the behaviour we are principally interested in at this stage. The reader should note, however, that the theory extends elegantly to more than one degree of freedom.

The present section is added for completeness, and attempts no more than a qualitative comparison of the results of simple models with the results of tracking in a specific lattice design using a conventional code. For ease of discussion, we refer to a comparison between a *simple model* and a *full lattice*, although it should be understood that all the results presented here are obtained by some method of modelling. Note that in the simple model, we use a mapping procedure, by applying a single-turn (or, rather, single-cell) map to the phase space co-ordinates, whereas in the full lattice, the tracking involves applying maps for individual elements. We begin by outlining the models that we use.

6.1 The Full Lattice

Our 'reference' model is a double-bend achromat structure, using a minimal number of elements. The aim was to produce a lattice that was not optimised for any particular purpose, but was flexible enough to cover a wide range of situations for comparison with a simple model. The investigations involved a substantial amount of tracking; clearly a very simple lattice structure was needed, so that tracking large numbers of particles over many turns did not take more than a few minutes.

One cell of the lattice we arrived at is shown in Figure 5, together with the beta functions and dispersion through the lattice. A circular machine would be composed of 15 identical cells. In the case shown, the horizontal phase advance per cell is $0.752 \times 2\pi$ radians, although this was one of the parameters which was adjusted as required for different comparisons. The length of the cell was fixed at 18.76 m.

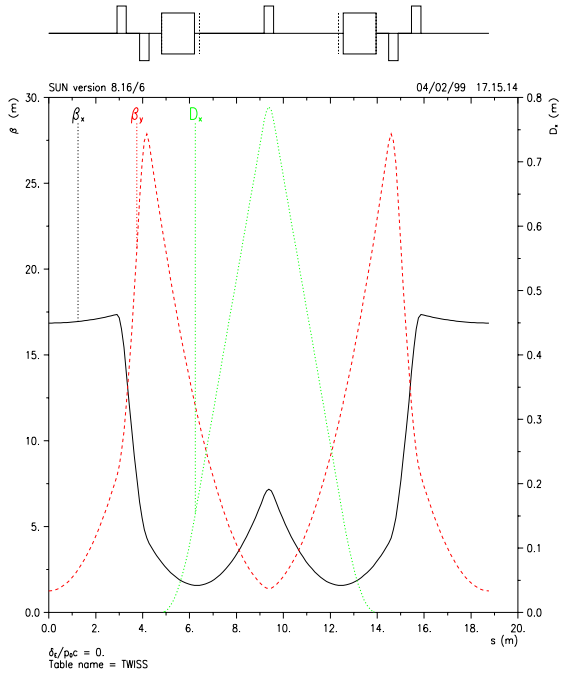

Figure 5

Beta functions and dispersion in a 'reference' DBA lattice.

Nonlinear elements (usually a single sextupole or a single octupole) were added as required, in the dispersive section of the cell, with the drift space being reduced by an equal amount so as to leave the total length of the cell unchanged.

6.2 The Simple Model

The argument in Section 5.6 attempts to show that a sequence of elements can be normalised to first order, with a nonlinear perturbation localised at one point in the sequence. For this reason, our simple model for a lattice with a single sextupole is based on the map

$$\mathcal{M} = \mathcal{R}_\mu \mathcal{S}$$

where

$$\mathcal{R}_\mu = e^{:-\mu J:} \tag{110}$$

is the operator for a phase advance (rotation) of μ, and

$$\mathcal{S} = \exp\left(:-\tfrac{1}{4} l p^2:\right) \exp\left(:-\tfrac{1}{6} k l x^3:\right) \exp\left(:-\tfrac{1}{4} l p^2:\right) \tag{111}$$

is a second-order solvable map approximation to a sextupole of length l and strength k in one degree of freedom, see equation (63).

For programming purposes, we need polynomial representations of the maps (110) and (111). These are readily found by direct calculation:

$$\mathcal{R}_\mu x = x\cos(\mu) + p\sin(\mu)$$
$$\mathcal{R}_\mu p = -x\sin(\mu) + p\cos(\mu)$$

for the linear map, and for the sextupole map, from equations (65):

$$\begin{aligned}\mathcal{S}x &= x + lp - \tfrac{1}{4}k_2 l^2 x^2 - \tfrac{1}{4}k_2 l^3 xp - \tfrac{1}{16}k_2 l^4 p^2 \\ \mathcal{S}p &= p - \tfrac{1}{2}k_2 l x^2 - \tfrac{1}{2}k_2 l^2 xp - \tfrac{1}{8}k_2 l^3 p^2\end{aligned} \tag{112}$$

These maps allow us quickly to map a particle through a cell containing a single sextupole, no matter how complex the linear lattice.

6.3 Comparison of TRANSPORT, LIE3 and LIE4 Tracking Methods

MAD allows tracking by three different methods, namely TRANSPORT, LIE3 and LIE4. The differences between these methods are detailed in [7], but we discuss them briefly here. A TRANSPORT map is a Taylor series for the exact transfer map, truncated at order two. In general, the truncation means that the map will not be symplectic. For linear elements, the transfer map should only contain terms to first order, so the truncation should have no effect on the symplecticity. However, careful treatments of dipoles and quadrupoles show that even these elements are not exactly linear, so the TRANSPORT map for a lattice composed only of these 'linear' elements will not be symplectic.

LIE3 and LIE4 maps are both derived from solvable Lie transforms, and so are symplectic maps. LIE3 maps include terms up to third order in the exponent of the Lie transform, so include quadrupoles and sextupoles, but neglect octupole and higher order fields. LIE4 maps include terms up to fourth order in the exponent; octupole fields are included in LIE4 maps, but fields of higher order than octupoles are neglected.

To illustrate these differences, consider Figure 6 and Figure 7. Figure 6 shows particles tracked through the linear lattice of Figure 5, using the TRANSPORT method, and Figure 7 shows exactly the same situation, but tracking using the LIE3 method. The non-symplectic nature of the TRANSPORT map is clearly visible, with particles at large amplitudes tending to spiral inwards in phase space. The LIE3 map, by contrast, is clearly symplectic, with all particles shown travelling round closed ellipses in phase space.

Although the non-symplectic effects of truncating the Taylor series in the TRANSPORT method are apparently small, and only really visible in this situation at unrealistically large amplitudes, it is clear that if the user is particularly concerned about symplecticity, then the maps derived from Lie operator techniques are the more reliable. For example, if the user is interested in the effects of synchrotron radiation, then it is important that there are no artificial non-symplectic effects from approximations made in the dynamical model which might obscure the effects of the radiation itself.

6.4 Nonlinear Rotation Caused by an Octupole Field

The effect of a nonlinear rotation can clearly be demonstrated by tracking a set of particles, initially spread along the x co-ordinate axis, through a small number of cells. For reasons that will become apparent, we first adjust the strengths of the quadrupoles in the lattice so that the phase advance per three cells is close to $2.25\times 2\pi$ radians. Note that this is near a fourth order resonance. We then place a single octupole per set of three cells, effectively giving the lattices a five-fold symmetry.

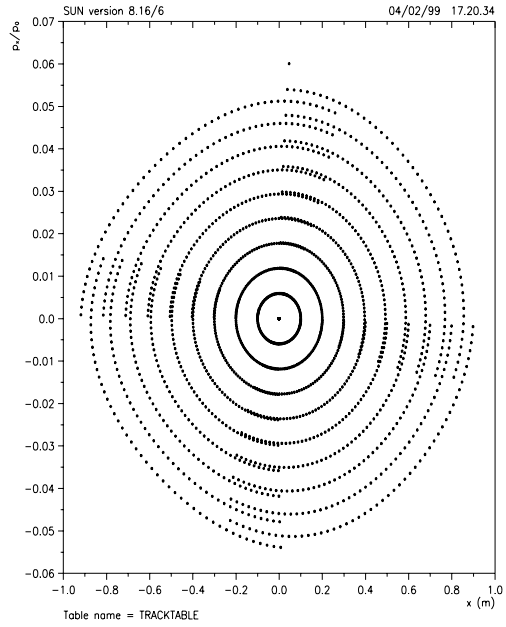

Figure 6

Tracking particles through a 'linear' lattice using the TRANSPORT method.

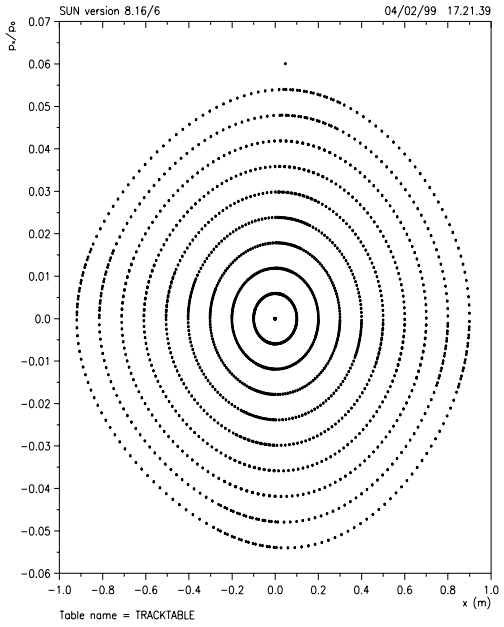

Figure 7

Tracking particles through a 'linear' lattice using the LIE3 method.

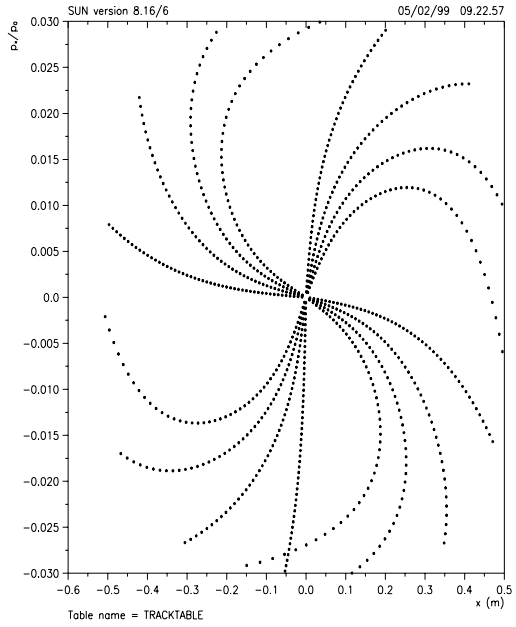

Figure 8

Results of tracking (LIE4) with an octupole perturbation in a five-fold symmetric lattice. Phase advance per set of three cells is $2.255\times 2\pi$ radians.

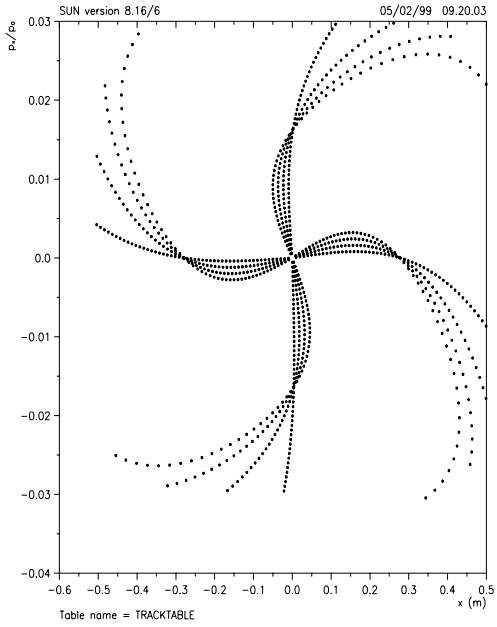

Figure 9

Results of tracking (LIE4) with an octupole perturbation in a five-fold symmetric lattice. Phase advance per set of three cells is $2.245\times 2\pi$ radians.

Plotting the positions of the particles every fifth of a turn gives the results shown in Figure 8 and Figure 9. In both these figures, 50 particles are tracked through 16 sets of three cells, with the positions of the particles in phase space plotted after each set. The octupole has strength $k = 20$ and length $l = 0.2$ m in each case.

In Figure 8, the phase advance is just above the fourth order resonance. The effect of the octupole, from equation (107) with[11] $\varepsilon = \frac{1}{24} kl$, is to increase the phase advance with the amplitude of the particle. This can be seen in the increasing curvature of the line of particles in phase space, with distance from the origin. In Figure 9, the phase advance is just below the fourth order resonance. Again, the effect of the octupole is to increase the phase advance with amplitude. But now, because the increase is starting from *below* a fourth order resonance, at a particular amplitude, there is a particle which lies almost exactly on the resonance. This can be seen clearly in Figure 9, with the appearance of four additional fixed points of the fourth order iterate of the map[12]. A particle at this amplitude moves sequentially from one of these fixed points to the next with each pass through a set of three cells, because the phase advance per set of three cells for this particle is exactly $2.25 \times 2\pi$ radians.

Note that it appears from Figure 9 that the presence of the fourth order resonance does not affect the stability of the map. This can be seen more clearly by tracking the particles for many turns of the lattice. The results are shown in Figure 10.

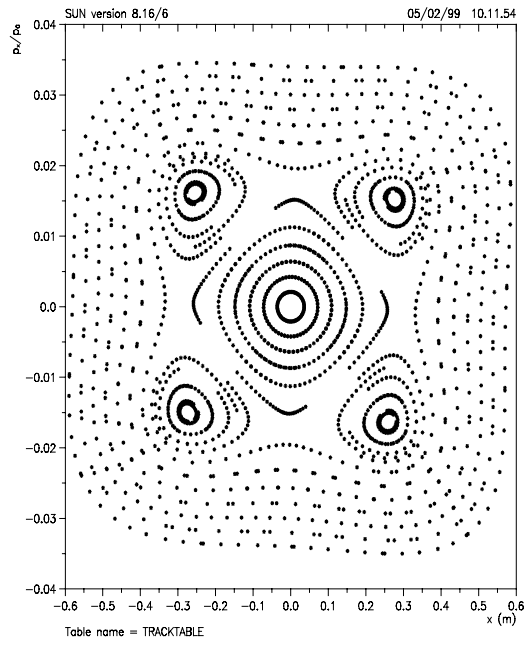

Figure 10

Results of tracking (LIE4) thirty turns with an octupole perturbation in a five-fold symmetric lattice. Phase advance per set of three cells is $2.245 \times 2\pi$ radians.

[11] k is the octupole strength, and l is the length of the magnet; the expression for ε follows from a symmetric factorisation of the octupole Hamiltonian.

[12] Strictly, the fourth order iterate of the three-cell map.

An interesting feature of Figure 10 is the appearance of 'islands' located near the fixed points of the iterated map. This is clearly to do with the presence of a resonance. Nevertheless, there is an apparently stable region extending outside the islands.

We now have sufficient knowledge to be able to control topological features of tracking plots. For example, if we retune our lattice to lie just below a fifth order resonance, we can easily produce a phase space tracking plot with five islands shown in Figure 11. Note that in this case, though, the 'islands' have shrunk to points.

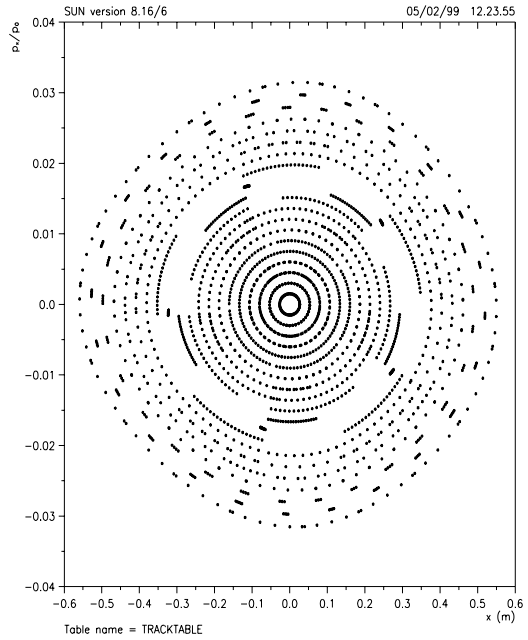

Figure 11

Results of tracking (LIE4) thirty turns with an octupole perturbation in a five-fold symmetric lattice. Phase advance per set of three cells is $2.195 \times 2\pi$ radians.

6.5 Normalised Map with an Octupole Kick

Let us compare the phase space plots produced by tracking through the full lattice with those from a simple model. From equation (107), we can write the first order normalised map in the presence of an octupole kick as:

$$x \to x\cos(\theta) + p\sin(\theta)$$
$$p \to p\cos(\theta) - x\sin(\theta)$$

where

$$\theta = \mu + \tfrac{3}{8} \cdot \tfrac{1}{24} kl\left(x^2 + p^2\right)$$

Applying this map 15 times to a set of points initially lying along the x axis in phase space produces the plots shown in Figure 12 and Figure 13[13].

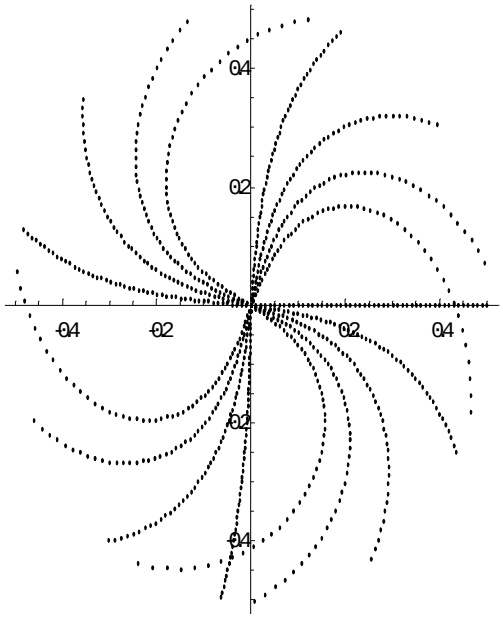

Figure 12

Plot of nonlinearly normalised map for a linear lattice with an octupole perturbation. The parameters for the model were taken directly from values used to produce Figure 8. (Tune = 0.255)

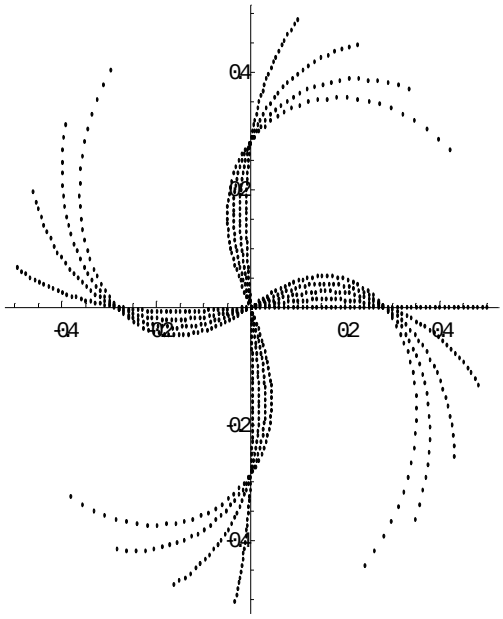

Figure 13

Plot of nonlinearly normalised map for a linear lattice with an octupole perturbation. The parameters for the model were taken directly from values used to produce Figure 9. (Tune = 0.245)

The plots in Figure 12 and Figure 13 show a striking similarity to the plots in Figure 8 and Figure 9, the main difference being a scaling of the momentum axis. Since we are interested mainly in qualitative comparisons between the full lattice and a simple model, we leave investigation of this point to a later note.

The model does have its limitations. For example, if we try to reproduce Figure 10 using the simple model, we obtain a plot shown in Figure 14. We see that in Figure 14 there are none of the islands that we see in Figure 10. This should not surprise us; in Figure 10, there is something clearly more complex than a rotation taking place, and our normalisation procedure reduces every map to a (possibly nonlinear) rotation. Forest [3] suggests that to produce islands, we need to include terms in the exponent of the Lie transform that look like $\cos(n\phi)$, where $\tan(\phi) = -\frac{p}{x}$. Unfortunately, such maps are not solvable, and their investigation is beyond the scope of the present note.

[13] The calculations were performed and the plots in Figure 12 and Figure 13 were produced using *Mathematica*.

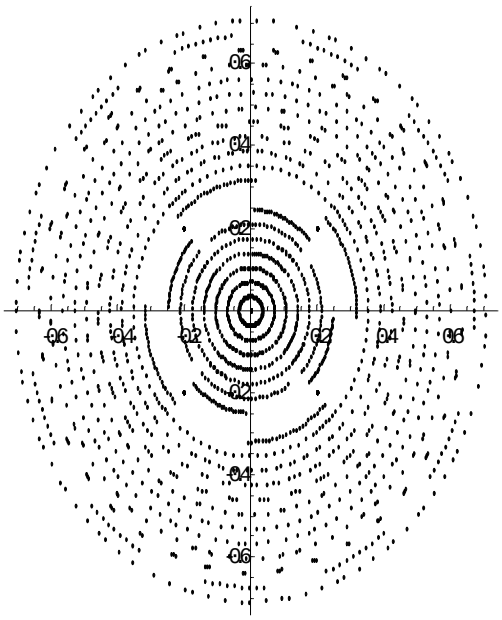

Figure 14

Plot of nonlinearly normalised map for a linear lattice with an octupole perturbation, map applied 100 times. The parameters for the model were taken directly from values used to produce Figure 10. (Tune = 0.245)

6.6 Sextupole Models

We have so far worked with octupole maps because there is a simple expression for the first order perturbation phase shift with amplitude resulting from an octupole. For a sextupole, to first order, the phase shift with amplitude is zero, and we need to go to second order in the perturbation theory to find a phase shift. This makes constructing a map expressed as a single Lie transform more difficult. For the present work, where we are interested mainly in qualitative comparisons, we therefore resort to applying a map of the form

$$\mathcal{M} = \exp(:-\mu J:)\mathcal{S} \tag{113}$$

where $\mathcal{S}$ is given by (112). To assist investigation of this model, the author has produced a simple program, DynAp, which will show the results of iterating the map on any starting point in phase space, and will search for a maximal boundary of stable points. We refer to the region enclosed by such a boundary the 'dynamic aperture' of the map. This terminology requires careful qualification; for a discussion, see Appendix B, where the algorithm for DynAp is described.

For our investigations, we used the same double-bend achromat lattice structure as before, and retained the five-fold symmetry, but replaced the octupole with a sextupole of the same length. We chose a sextupole strength $k = 50$, which would be expected to give the sextupole a much larger effect than the octupole. The reasons for this choice are indicated below.

Figures show tracking results produced by MAD, and results produced by DynAp for the map in (113), for a linear phase advance per set of three cells of $2.255 \times 2\pi$ radians. Note that the results from DynAp have been rotated by an appropriate angle to give a clearer comparison with the results from MAD. This corresponds simply to a fixed phase difference, which would correspond simply to different observation points in the two models. It is clear that there are close similarities between the

results; in particular, we observe a central region of stable orbits, with four islands of stability placed some distance from the origin. The stable islands are apparently surrounded by regions where the orbits are unstable. The presence of four islands again corresponds to a fourth order resonance, which we expect to see from the tune. This time, however, it is not clear how the presence of the islands affects the overall dynamic aperture. These islands only appear for relatively high sextupole strengths; this was the reason for our choice of k for these models.

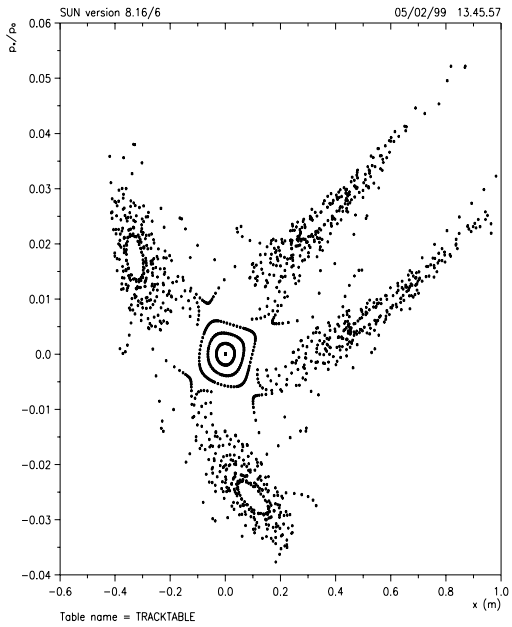

Figure 15

Tracking results from MAD for a sextupole perturbation in a five-fold symmetric lattice, phase advance between sextupoles $2.255 \times 2\pi$ radians.

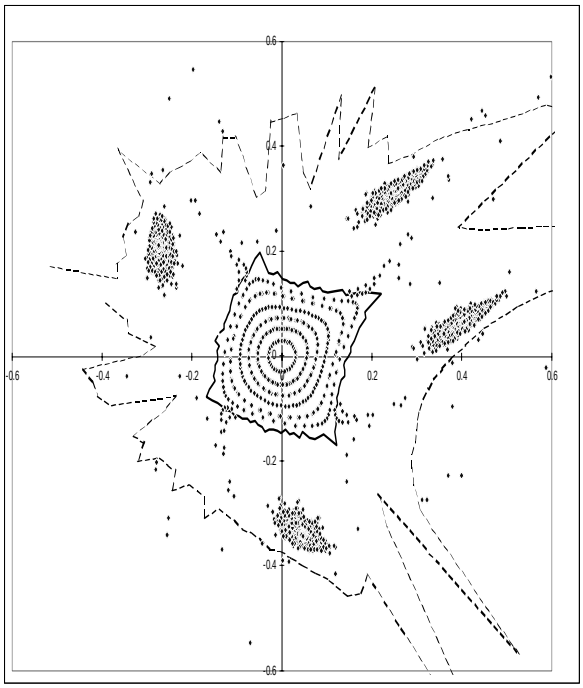

Figure 16

Mapping results from DynAp for the map (113), linear phase advance $2.255 \times 2\pi$ radians. The solid black line shows the 'minimal' dynamic aperture; the broken line shows the 'maximal' dynamic aperture.

In Figure 16, the solid line shows the boundary of the central region of stable orbits, which we refer to as the *minimal* dynamic aperture. The minimal dynamic aperture is found by searching from the origin outwards for unstable points. The broken line in Figure 16 shows the maximal dynamic aperture, which is found by searching inwards from a large radius for stable points. From the differences between the minimal and maximal boundaries, it is clear that the dynamic aperture is not a simple quantity; it could even be the case that no smooth or even well-defined boundary exists[14].

An advantage of using the program DynAp to study maps such as (113), is that in comparison to MAD, it works extremely fast, and can be used to obtain very quickly a measurement of the area of the minimal dynamic aperture. Obtaining the dynamic aperture in MAD is a rather cumbersome process, and requires several times the processing time for even a rough estimate. However, it is possible to use both techniques to find, for example, how the (minimal) dynamic aperture of a lattice varies with the linear phase advance per cell.

[14] The way in which the boundary is produced is related to the way in which fractal images are produced, the most famous example of which is the Mandelbrot set. These images have bizarre properties which are still being studied extensively by mathematicians.

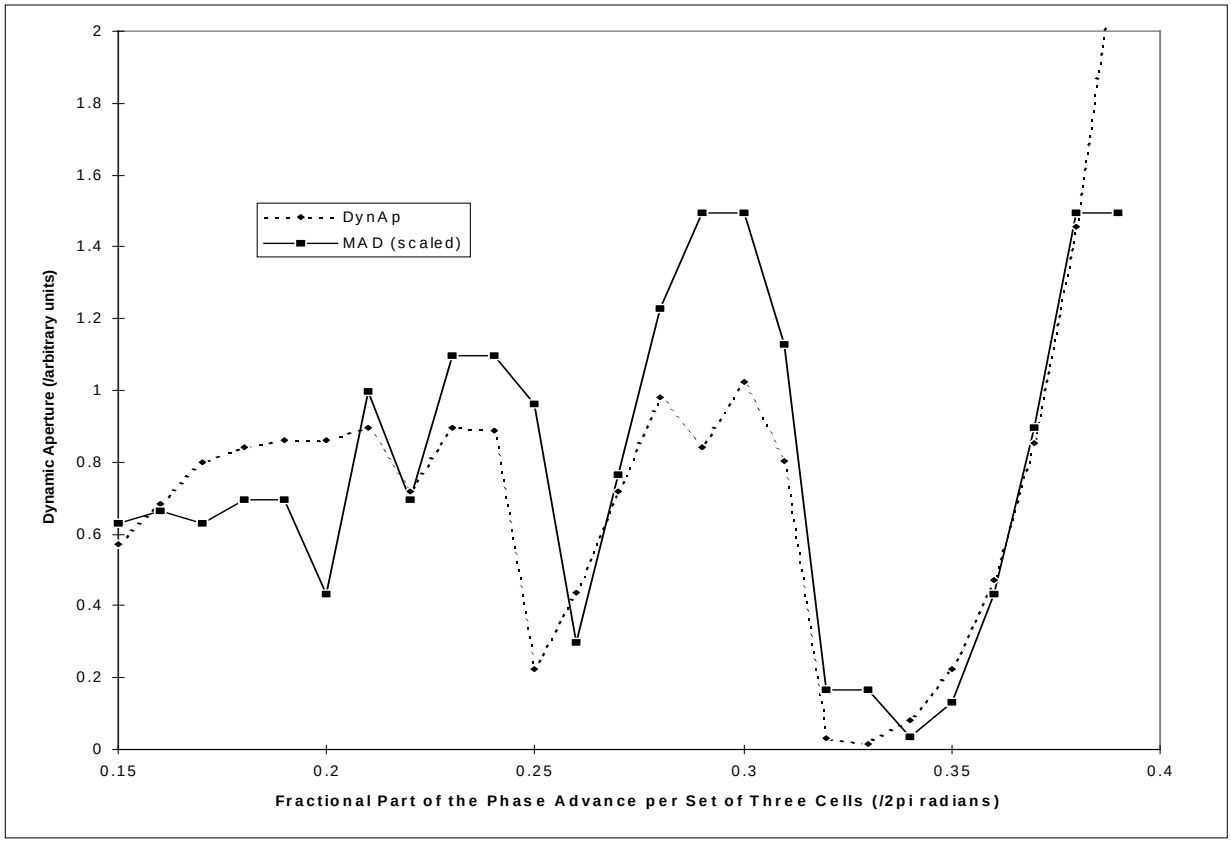

Figure 17

Comparison of the results of mapping and tracking from an investigation of variation of dynamic aperture with tune.

Figure 17 shows how the dynamic aperture of the lattice varies as the linear phase advance is varied. We see that there is a significant (but by no means) perfect correlation between tracking using MAD, and mapping using DynAp. It should be emphasised that this was by no means a rigorous investigation. Nevertheless, both models predict essentially similar behaviour for the dynamic aperture; in particular, a drop near the fourth order resonance, as well as the expected drop near the third order resonance. As we have already observed, conventional treatments of sextupole effects tend to ignore the excitation of fourth order resonances. A clear advantage of the mapping technique is that results can be obtained in a fraction of the time taken to obtain results by element-by-element tracking in MAD. This, together with the ability to investigate the map analytically, raises the possibility of real optimisation of the nonlinear elements in a lattice, rather than 'experimental optimisation' that is really a process of trial and error.

7. Conclusion

We have seen how a Hamiltonian approach to the dynamics of particles in an accelerator leads to the development of Lie transforms and single-turn maps as appropriate mathematical tools for studies in this area. Many of the results of standard linear accelerator theory can be derived very quickly and elegantly using these techniques, given a few general assumptions about the motion. A rigorous treatment can incorporate the full three degrees of freedom, given an appropriate expression for the Hamiltonian. In the present work, we concentrated on motion in a single dimension, transverse to the design orbit of the machine, so as to develop general techniques that will allow us insight into the behaviour of particles in circular accelerators.

A Lie transform provides a map that gives us the evolution of any function of the phase space co-ordinates. The map can always be expressed as a power series in the phase space co-ordinates; in some cases, the series terminates, in which case the map is said to be solvable. In most cases, however, the map generates an infinite series. If we truncate the map, we lose the symplecticity which is a fundamental feature of our system. It is therefore preferable to make an approximation to the Lie operator that gives a solvable map. An important technique for doing this is symmetric factorisation.

Given a Lie transform, we have found that it is possible to perform a canonical transformation that changes the transform into a rotation; in the case of a nonlinear map, the rotation is nonlinear, by which we mean that the angle of rotation is a function of the amplitude of the particle in phase space. It is also possible to normalise a sequence of elements containing a single nonlinear perturbation, so that, to first order, the transformed sequence is a simple rotation with a nonlinear rotation localised at the position of the perturbation. This is a useful technique, as it indicates, for example, the effects of resonances and the region of stability in phase space.

Combining all these methods of Lie transforms, single-turn maps and normalisation, we can construct a very simple model of a lattice with a nonlinear perturbation. Testing this model against a conventional tracking code in some specific cases shows that the model can reproduce many of the interesting features of the behaviour predicted by a more complete description of the lattice. The importance of this is that it suggests the possibility of a more analytical approach to the effects of nonlinear elements in a lattice than is possible by conventional means. It is not suggested that the simple models we can construct with these tools will replace, for example, element by element tracking, but they should make it possible to carry out a much more rapid characterisation of a system, and perhaps lead to a means of optimisation, that can be verified by more rigorous models.

We have not attempted in this note to give a rigorous treatment of such effects as the longitudinal motion, coupling, dispersion, chromaticity and momentum compaction, and non-symplectic effects such as synchrotron radiation; nor have we considered the construction of realistic Hamiltonians, which are necessary for accurate modelling. All these issues are discussed by Forest [3].

The author's study of Lie operator methods was initially prompted by the need to achieve effective control over nonlinear effects, introduced in an accelerator, for example, by sextupoles used to correct the chromaticity of a lattice. In particular, we hope to be able to cancel the detrimental nonlinear effects of the chromatic sextupole by using harmonic sextupoles. We hope that the ideas and techniques discussed in this note will allow some progress with this problem.

Appendix A: Hamiltonian for a Particle in an Electromagnetic Field

We show that the Hamiltonian given by (46) and (47) reproduces the familiar expression for the Lorentz force, given by (45). For convenience, we adopt the component notation A_i, where $A_1 = A_x$, $A_2 = A_y$, $A_3 = A_z$. We also use the summation convention, where a summation is to be assumed over any index repeated in a single term, thus $x_i A_i = \sum_{i=1}^{3} x_i A_i = \mathbf{x} \cdot \mathbf{A}$.

We start by writing for the canonical momentum (47),

$$p_i = m\dot{x}_i + qA_i . \tag{114}$$

Substituting this into Hamilton's equation (2) gives

$$m\ddot{x}_i + q\frac{\partial A_i}{\partial t} + q\dot{x}_j \frac{\partial A_i}{\partial x_j} = -\frac{\partial H}{\partial x_i} \tag{115}$$

where we have used

$$\frac{dA_i}{dt} = \frac{\partial A_i}{\partial t} + \dot{x}_j \frac{\partial A_i}{\partial x_j}$$

Now, from (46) we can write

$$H = \frac{1}{2m}\left(p_j p_j - 2qp_j A_j + q^2 A_j A_j \right) + q\varphi$$

and so, remembering that the p_i are (canonically) independent of the x_i, we have

$$\begin{aligned} \frac{\partial H}{\partial x_i} &= -\frac{q}{m} p_j \frac{\partial A_j}{\partial x_i} + \frac{q^2}{2m} \frac{\partial}{\partial x_i}\left(A_j A_j \right) + q \frac{\partial \varphi}{\partial x_i} \\ &= -q\dot{x}_j \frac{\partial A_j}{\partial x_i} + q \frac{\partial \varphi}{\partial x_i} \end{aligned} \tag{116}$$

where we have used (114) in the last line. Combining (115) and (116) gives

$$m\ddot{x}_i = q\dot{x}_j \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \right) - q\left(\frac{\partial A_i}{\partial t} + \frac{\partial \varphi}{\partial x_i} \right) \tag{117}$$

If we now define the field strengths

$$E_i = -\frac{\partial A_i}{\partial t} - \frac{\partial \varphi}{\partial x_i} = -\frac{\partial A_i}{\partial t} - \left(\nabla \varphi \right)_i$$

$$B_k = \varepsilon_{kij} \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \right) = \left(\nabla \times \mathbf{A} \right)_k$$

and note that

$$\begin{aligned} \left(\dot{\mathbf{x}} \times \mathbf{B} \right)_i &= \varepsilon_{ijk} \dot{x}_j B_k \\ &= \varepsilon_{ijk} \varepsilon_{klm} \dot{x}_j \left(\frac{\partial A_m}{\partial x_l} - \frac{\partial A_l}{\partial x_m} \right) \\ &= \varepsilon_{kij} \dot{x}_j \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \right) \end{aligned}$$

we see that we can write (117) as

$$m\ddot{\mathbf{x}} = q\left(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B} \right)$$

from which we deduce the Lorentz force (45).

Appendix B: The Program DynAp

The program DynAp has been written to identify the boundary of a set of stable points under the action of a polynomial mapping. The program sets up a set of n points, with each point a distance r_n from the origin. Initially, the r_n are equal, so the points are distributed evenly around a circle. For each point in turn, the program follows the following sequence:

1. Apply the map a maximum of I times to the nth point. If the distance of the point from the origin after any iteration exceeds R, then the initial point is assumed to be unstable.

2. If the point is unstable, then $r_n \to r_{n0} + 0.9\left(r_n - r_{n0}\right)$ where r_{n0} is the last known r_n for which the point was stable; if no r_{n0} has been found, then $r_n \to r_n/(1+\delta)$.

3. If the point is stable, then $r_n \to (1+\delta)r_n$.

The sequence is repeated a number of times determined by the user. *Stability* is defined by the parameters I and R, since if the point goes beyond a circle of radius R after I iterations of the map, it is unstable. Clearly, the accuracy with which the boundary of stability is determined, depends on the parameters n and δ. In practice, the sequence is followed until the area enclosed by the boundary makes small oscillations around a fixed value. Typical values for the parameters are $n = 100$, $I = 200$, $R = 100$, $\delta = 0.01$.

If the r_n are initially very small, and all the initial points are stable, then the algorithm effectively searches outwards for unstable points. If the r_n are initially very large, and all points are initially unstable, then the algorithm effectively searches inwards for stable points. The procedure works very quickly, but is only really reliable when the boundary of the stable region is smooth, which by no means covers all cases. It is possible that for some maps, the boundary of the stable region has fractal characteristics.

Acknowledgement

I am indebted to Etienne Forest, who invariably replied with patience and encouragement to my many emailed queries regarding the work set out in his impressive book.

References

[1] H. Goldstein, *Classical Mechanics* (Addison-Wesley, 2nd Ed. 1980)

[2] F. Willeke, *Tutorial on Modern Tools of Particle Tracking*, DESY Internal Report M-94-12, 1994.

[3] E. Forest, *Beam Dynamics: A New Attitude and Framework* (Harwood Academic Publishers, 1998).

[4] A. Wolski, Communication with E. Forest, available at http://bc1.lbl.gov/CBP_pages/educational/critics/part1.pdf

[5] J. Rossbach and P. Schmüser, *Basic Course on Accelerator Optics*, in Proceedings of the CERN Accelerator School Fifth General Accelerator Physics Course, 7-18 September 1992 (CERN 94-01), pp.17-88.

[6] J. D. Jackson, *Classical Electrodynamics*, (Wiley, 2nd Ed. 1975).

[7] F. C. Iselin, *The MAD Program V.8.13, Physical Methods Manual*, CERN/SL/92-??

[8] A. H. Nayfeh and B. Balachandran, *Applied Nonlinear Dynamics* (Wiley-Interscience, 1995).

[9] E. Forest, http://bc1.lbl.gov/CBP_pages/educational/main.html

[10] M. Cornacchia, *Lattices*, in *Synchrotron Radiation Sources*, ed. H. Winick (World Scientific, 1994).

[11] E. Wilson, *Non-Linearities and Resonances*, in Proceedings of the CERN Accelerator School Fifth General Accelerator Physics Course, 7-18 September 1992 (CERN 94-01), pp.239-251.