Use of Lie transforms to generalize Madey's theorem for computing the gain in microwave devices

P. E. Latham, S. M. Miller, and C. D. Striffler
Laboratory for Plasma Research, University of Maryland, College Park, Maryland 20742
(Received 24 July 1991)

For near-integrable Hamiltonian systems with a nonintegrable piece of order $\epsilon \ll 1$, we use Lie transforms to derive a generalized form of Madey's theorem. Specifically, we find an expression for the average second-order (in ϵ) change of any function of momentum in terms of first-order quantities only. A formalism is given that makes gain calculations for devices like free-electron lasers and gyrotrons in complicated geometries tractable. An explicit expression is presented for the case where the nonintegrable part of the Hamiltonian is a harmonic function of the coordinates. As an example, the average change in particle kinetic energy is computed through second order in the field amplitude for gyrotrons in complicated geometries. The transform method is extended to non-Hamiltonian systems, and it is shown that there is a class of non-Hamiltonian differential equations to which Madey's theorem applies.

PACS number(s): 41.60.Cr, 42.52.+x

I. INTRODUCTION

Standard analytical techniques for computing the gain in microwave devices such as free-electron lasers (FEL's) and gyrotrons are cumbersome in simple configurations and virtually impossible to use in realistic ones. Madey's theorem [1–3] greatly simplifies this computation by expressing the gain in terms of first-order quantities. In its original form, this theorem applied only to one-dimensional systems; later it was extended to N-dimensional, near-integrable Hamiltonian systems [3]. Specifically, the theorem states that if the nonintegrable portion of a Hamiltonian system is $O(\epsilon)$, $\epsilon \ll 1$, the average change in energy to $O(\epsilon^2)$ may be expressed in terms of $O(\epsilon)$ quantities. For microwave devices the small parameter ϵ is proportional to the vector potential $\mathbf{A}$ associated with the electromagnetic field, and the Hamiltonian has the usual form

$$H = c\sqrt{m_0^2 c^2 + (\mathbf{p} - q/c\,\mathbf{A})^2}\ , \tag{1}$$

where q and m_0 are the particle's charge and mass, respectively, c is the speed of light, and $\mathbf{p}$ is the canonical momentum. We have neglected the scalar potential Φ, as space-charge forces are not usually important in microwave devices.

Here we provide a further generalization of Madey's theorem. Using Lie transforms [4–10], we first derive an expression for the time evolution of any function on phase space $f(\mathbf{z}) = f(\mathbf{q},\mathbf{p})$, where $\mathbf{q}$ and $\mathbf{p}$ are the canonical coordinates and momenta, respectively, in terms of Lie operators. Then, for the case where f is a function of momentum only and the integrable part of the Hamiltonian is a function of momentum and possibly time, we show how the average change in f can be computed through second order in the small parameter ϵ in terms of first-order quantities. In other words, we compute the quantity $\langle f(\mathbf{p}_f) - f(\mathbf{p}_0)\rangle$ to $O(\epsilon^2)$, where $\mathbf{p}_0$ and $\mathbf{p}_f$ are the values of the momentum at $\tau = \tau_0$ and $\tau = \tau_f$, respectively. The value of $\mathbf{p}_f$ is found by evolving the phase-space variable $\mathbf{z} = (\mathbf{q},\mathbf{p})$ according to

$$\frac{d\mathbf{z}}{d\tau} = \{\mathbf{z}, H\}\ , \tag{2}$$

where $\mathbf{z}_0 \equiv \mathbf{z}(\tau = 0)$ is the initial condition, $H(\mathbf{z})$ is the Hamiltonian, and $\{\ ,\ \}$ is the usual Poisson bracket [11]. The angular brackets $\langle\ \rangle$ refer to an average over the initial coordinates $\mathbf{q}_0$. Throughout our analysis we will assume that the system is autonomous, which can always be realized by introducing extended phase-space variables [9]. Thus, the Poisson bracket in Eq. (2) is the extended phase-space bracket.

Although our final result is only a minor extension of the N-dimensional form of Madey's theorem presented in Ref. [3], our derivation is radically different: it is short, simple, and physical, and it does not rely on the fortuitous cancellations that appeared necessary in previous derivations. Moreover, the method we present can easily be applied to a whole array of problems, including those that involve higher-order (in ϵ) calculations, and, with only minor modifications, to non-Hamiltonian systems. This latter observation allows us to identify the class of differential equations (broader than simply Hamiltonian differential equations) to which Madey's theorem applies.

This paper is organized as follows. In Sec. II we use Lie transforms to derive a generalized form of Madey's theorem. The expression we derive is extremely convenient for practical calculations. We illustrate this point in Sec. III, where, for simplicity, we choose a first-order Hamiltonian that is a harmonic function of its coordinates. This type of Hamiltonian applies to most microwave devices, and in Sec. IV we use the results of Sec. III to compute, through second order in the electromagnetic field amplitude, the average change in particle kinetic energy of a gyrotron beam passing through a complex cavity. Section V generalizes our results to non-Hamiltonian systems and Sec. VI contains our summary and conclusions.

II. DERIVATION OF A GENERALIZED FORM OF MADEY'S THEOREM

When analyzing a Hamiltonian system, or any system described by an ordinary differential equation, the goal is to compute the phase-space trajectories as a function of some parameter. Typically, this parameter is time, but it need not be; it can just as easily be a spatial dimension or an angle. Thus, for generality, we talk about evolution of phase-space variables as a function of the parameter τ.

There are two approaches to computing the τ evolution of phase-space variables: directly integrate the equations of motion, Eq. (2), or transform to a new frame where the equations are simple, integrate the equations in the new frame, then transform back. These two approaches are illustrated schematically in Fig. 1. The solid arrow indicates the direct integration; the dashed arrows show the indirect route (transformation to the primed frame, integration, and transformation back to the unprimed frame). In the first approach, the integration forward in τ is represented by the time development operator $S_H(\tau)$:

$$f(\mathbf{z}(\tau))=S_H(\tau)f(\mathbf{z}_0) \tag{3}$$

for any function $f(\mathbf{z})$ with $\mathbf{z}$ evolving according to the Hamiltonian H. In general, the trajectories generated by the Hamiltonian H are complicated, and S_H is difficult to compute.

An alternative method for finding $f(\mathbf{z}(\tau))$ is to transform to a frame where the time development operator is simple. This is the second approach described above. The change of variables that represents the transformation between frames is contained in the operator $T(\mathbf{z})$,

$$\mathbf{z}'=T(\mathbf{z})\mathbf{z} \ . \tag{4}$$

In the primed frame there is a new Hamiltonian, which we call K, so the time development operator there is S_K rather than S_H. The second, indirect approach is useful if the equations of motion in the primed frame are integrable; i.e., $S_K(\tau)$ is a known operator. A typical choice is to have the Hamiltonian in the primed frame, $K(\mathbf{z}')$, be independent of the coordinate $\mathbf{q}'$, so that the momenta are constant and $\mathbf{q}'$ evolves linearly in τ. For that case,

$$f(\mathbf{z}'(\tau))=S_K(\tau)f(\mathbf{z}_0')=f(\mathbf{q}_0'+\mathbf{v}_0'\tau,\mathbf{p}_0') \ ,$$

where $\mathbf{z}_0'=(\mathbf{q}_0',\mathbf{p}_0')$ and $\mathbf{v}_0'=\partial K(\mathbf{p}_0')/\partial\mathbf{p}_0'$.

With the indirect approach, the problem shifts from solving a complicated differential equation to computing a transformation T that makes the equations of motion integrable. An especially convenient choice for T is the Lie transform operator, which has the form

$$T=e^{-L} \ , \tag{5}$$

where $Lf=\{\omega,f\}$ for any function $f(\mathbf{z})$. The quantity $\omega(\mathbf{z})$ is the Lie generator. The inverse operator is given simply by $T^{-1}=e^{L}$. The Lie transform operator given in Eq. (5) has two important properties: it generates canonical transformations and it commutes with functions. The second property implies that

$$f(T\mathbf{z})=Tf(\mathbf{z}) \ . \tag{6}$$

These two properties follow immediately from the observation that

$$\frac{d\mathbf{z}}{d\lambda}=-L\mathbf{z} \implies \mathbf{z}(\lambda)=e^{-\lambda L}\mathbf{z}(\lambda=0) \ . \tag{7}$$

In other words, T simply pushes a phase-space point forward in "time" λ by 1 unit along the vector field specified by L. Since Eq. (7) applies with $\mathbf{z}$ replaced by $f(\mathbf{z})$, T commutes with functions whether or not L represents a Hamiltonian vector field. However, when L does represent a Hamiltonian flow (with Hamiltonian ω: $L=\{\omega,\ \}$), then we are guaranteed that T generates canonical transformations.

The Lie transform operator T and time development operator S can be used to express $f(\mathbf{z}(\tau))$ in terms of $f(\mathbf{z}_0)$. Following the dashed arrows in Fig. 1, we first use Eqs. (4) and (6) to write

$$f(\mathbf{z})=f(T^{-1}(\mathbf{z}')\mathbf{z}')=T^{-1}(\mathbf{z}')f(\mathbf{z}') \ .$$

We then apply Eq. (3) (with H replaced by K because we are in the primed frame), which yields

$$T^{-1}(\mathbf{z}')f(\mathbf{z}')=S_K(\tau)T^{-1}(\mathbf{z}_0')f(\mathbf{z}_0') \ .$$

Finally, we again use Eqs. (4) and (6), producing

$$S_K(\tau)T^{-1}(\mathbf{z}_0')f(\mathbf{z}_0')=T(\mathbf{z}_0)S_K(\tau)T^{-1}(\mathbf{z}_0)f(\mathbf{z}_0) \ .$$

Thus,

$$f(\mathbf{z}(\tau))=T(\mathbf{z}_0)S_K(\tau)T^{-1}(\mathbf{z}_0)f(\mathbf{z}_0) \ . \tag{8}$$

Equation (8) represents a compact formula for $f(\mathbf{z}(\tau))$ in terms of $f(\mathbf{z}_0)$. This formula is useful only if $S_K(\tau)$ is simple. The problem of computing $f(\mathbf{z}(\tau))$ has been reduced to finding the transformation $T(\mathbf{z})$ that makes the equations of motion in the primed frame integrable.

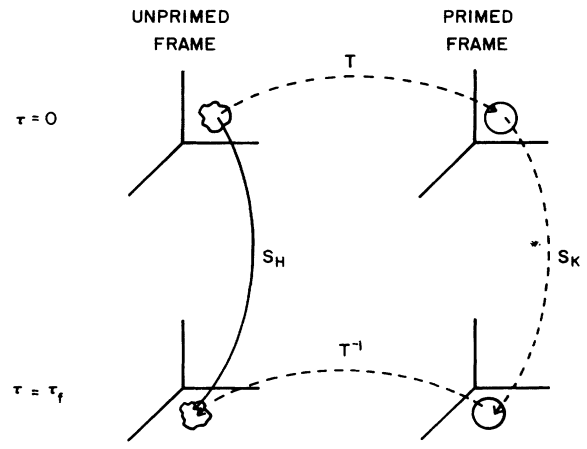

FIG. 1. Two approaches to computing the τ evolution of phase-space variables. The solid arrow is the direct approach: integrate the equations of motion in the original (unprimed) frame. The dashed arrows indicate the indirect approach: transform to the primed frame where the equations are simpler, integrate the equations of motion in the primed frame, then transform back.

For a near-integrable Hamiltonian with a nonintegrable part of order $\epsilon \ll 1$, there is a well-defined algorithm for computing T as a power series in ϵ [7–10]. This algorithm may be summarized as follows. Consider a Hamiltonian H written in the form

$$H(\mathbf{z})=H_0(\mathbf{z})+\sum_{n=1}^{\infty} H_n(\mathbf{z}) , \tag{9}$$

where H_0 is integrable. We use subscripts to denote order; quantities with subscript n are $O(\epsilon^n)$. Our goal is to find a transformation $\mathbf{z}'=T\mathbf{z}$ such that the new Hamiltonian $K(\mathbf{z}')=H(\mathbf{z})$ is simpler than H. Specifically, we wish to push the nonintegrable part of H to higher and higher order. To do this, expand K as

$$K(\mathbf{z}')=K_0(\mathbf{z}')+\sum_{n=1}^{\infty} K_n(\mathbf{z}') . \tag{10}$$

Using Eqs. (4) and (6), we have

$$K(\mathbf{z}')=H(\mathbf{z})=H(T^{-1}(\mathbf{z}')\mathbf{z}')=T^{-1}(\mathbf{z}')H(\mathbf{z}') . \tag{11}$$

Writing

$$L=\sum_{n=1}^{\infty}\frac{L_n}{n!}\equiv\sum_{n=1}^{\infty}\frac{\{\omega_n,\}}{n!} \tag{12}$$

inserting Eqs. (9) and (10) for H and K into Eq. (11), using Eq. (5) for T and (12) for L, and equating like orders, we find that through $n=2$, K_n and H_n are related by the transform equations

$$K_0=H_0 , \tag{13a}$$

$$K_1=L_1H_0+H_1 , \tag{13b}$$

$$K_2=\tfrac{1}{2}L_2H_0+H_2+L_1H_1+\tfrac{1}{2}L_1^2H_0 . \tag{13c}$$

The nth-order transform equation is given by

$$K_n=\frac{1}{n!}L_nH_0+H_n+R_n , \tag{14}$$

where R_n depends only on terms of order $n-1$ or less.

The simplest possible form for K is to have $K_n=0$ for $n\geq 1$. However, this is not in general possible; the Lie operators L_n can be chosen to eliminate only the nonessential terms. To understand which terms are essential and which are not, use Eq. (12) to write $L_nH_0=\{\omega_n,H_0\}$. Then, note that $\{\omega_n,H_0\}=d\omega_n/d\tau_0$, where the derivative $d/d\tau_0$ is to be taken along the trajectories generated by H_0, the so-called unperturbed orbits. Thus, $L_nH_0=d\omega_n/d\tau_0$, and Eq. (14) becomes

$$K_n=\frac{1}{n!}\frac{d\omega_n}{d\tau_0}+H_n+R_n .$$

Formally, the right-hand side of this equation can always be set to zero by integrating H_n+R_n along unperturbed orbits. However, the resulting expression for ω_n is not necessarily physically meaningful. For instance, ω_n may not respect the periodicity of the original system or it may not remain bounded. Thus, the essential terms come from those portions of H_n and R_n that, when integrated along unperturbed orbits, yield nonphysical expressions for ω_n. These terms represent an essential part of the dynamics and must be incorporated into K_n. Although almost all realistic problems generate essential terms, i.e., it is impossible to force all the K_n to zero, fortunately the K_n are often of a form that leaves the new Hamiltonian K integrable.

Once L_1 and L_2 have been chosen, it is a straightforward matter to find $f(\mathbf{z}(\tau))$ through second order. Inserting Eq. (5) into (8) and expanding the exponent with L given by Eq. (12), we find that through second order,

$$f(\mathbf{z}(\tau))=[S_K+S_KL_1-L_1S_K+\tfrac{1}{2}(S_KL_2-L_2S_K)+\tfrac{1}{2}(L_1^2S_K-2L_1S_KL_1+S_KL_1^2)]f(\mathbf{z}_0) . \tag{15}$$

It is convenient to define $f_\tau(\mathbf{z}_0)\equiv S_K(\tau)f(\mathbf{z}_0)$ for any function $f(\mathbf{z})$. Then, using the Jacobi identity [11],

$$\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\}=0$$

for any functions f, g, and h, and replacing L_n by $\{\omega_n,\}$, Eq. (15) becomes

$$\begin{aligned}(\Delta f)_2=&[f_\tau(\mathbf{z}_0)-f(\mathbf{z}_0)]+\{\omega_{1\tau}-\omega_1,f_\tau\}\\&+\tfrac{1}{2}\{\omega_{2\tau}-\omega_2,f_\tau\}\\&+\tfrac{1}{2}\{\omega_{1\tau}-\omega_1,\{\omega_{1\tau}-\omega_1,f_\tau\}\}\\&+\tfrac{1}{2}\{\{\omega_{1\tau},\omega_1\},f_\tau\} ,\end{aligned} \tag{16}$$

where $(\Delta f)_2\equiv f(\mathbf{z}(\tau))-f(\mathbf{z}_0)$ through second order.

Equation (16) gives us a method for finding $(\Delta f)_2$ by simply computing a set of Poisson brackets. This expression is valid regardless of the form of the operator $S_K(\tau)$ or the function $f(\mathbf{z}_0)$. Generally, however, it is not necessary to know $(\Delta f)_2$ for every initial condition and for arbitrary functions $f(\mathbf{z}_0)$. More typical is the case where the quantity of interest is $\langle(\Delta f)_2\rangle$, the average value of $(\Delta f)_2$ (defined explicitly below), f and H_0 are functions of the momentum only, and H_1 and H_2 are such that K_1 and K_2 are also functions of momentum only. The average referred to here is over initial position $\mathbf{q}_0$; with V the volume of integration, for any function $g(\mathbf{q}_0,\mathbf{p}_0)$,

$$\langle g(\mathbf{q}_0,\mathbf{p}_0)\rangle\equiv\int d^N\mathbf{q}_0\frac{1}{V}g(\mathbf{q}_0,\mathbf{p}_0) .$$

For this case we can show that the expression for $\langle(\Delta f)_2\rangle$ simplifies dramatically. Since H_0, K_1, and K_2 are all functions of momentum only, so is K through second order, and

$$f_\tau(\mathbf{z}_0)=f_\tau(\mathbf{p}_0)=f(\mathbf{p}_0)=f(\mathbf{z}_0) ,$$

where the first equality follows from the assumption that f is a function of momentum only. Consequently, the first term on the right-hand side of Eq. (16) is identically zero. In addition, the Poisson bracket of any function $\Omega(\mathbf{z}_0)$ with $f(\mathbf{p}_0)$ can be written

$$\{\Omega(\mathbf{z}_0),f(\mathbf{p}_0)\}=\partial_{\mathbf{q}_0}\Omega(\mathbf{q}_0,\mathbf{p}_0)\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0) .$$

Using this relation, and integrating by parts and assum-

ing that the boundary terms vanish, it is straightforward to show that the second, third, and fifth terms on the right-hand side of Eq. (16) are also zero. The only contribution to $\langle(\Delta f)_2\rangle$ is from the fourth term, which yields

$$\langle(\Delta f)_2\rangle=\tfrac{1}{2}\partial_{\mathbf{p}_0}\cdot\langle(\partial_{\mathbf{q}_0}\Delta\omega_1)(\partial_{\mathbf{q}_0}\Delta\omega_1)\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0)\rangle \quad (17)$$

where $\Delta\omega_1\equiv\omega_{1\tau}(\mathbf{z}_0)-\omega_1(\mathbf{z}_0)$. Note that Eq. (17) contains only first-order quantities even though Eq. (16) had second-order terms in it.

The condition that H_0 be a function of the momentum only is more restrictive than necessary; an important case arises when H_0 also depends on time. Then, since the momentum conjugate to time, which we denote p_t, is not physical, both f and ω_n are independent of p_t. In that case Eq. (17) is valid as long as the average is over all coordinates except time.

Although Eq. (17) is not the usual form of Madey's theorem, it is the most useful form for calculation: $\langle(\Delta f)_2\rangle$ can be found by computing $\Delta\omega_1$ from Eq. (13b) and taking derivatives. In addition, $\Delta\omega_1$ can be written in a form that has a natural physical interpretation. Assuming that we can set $K_1=0$ and still have a valid expression for ω_1, Eq. (13b) yields

$$\Delta\omega_1=-\int_0^{\tau}d\tau_0' H_1(\mathbf{z}_0(\tau_0')) , \quad (18)$$

where the integral is over unperturbed orbits. Thus, $\Delta\omega_1$ is simply the first-order action, and $\langle(\Delta f)_2\rangle$ is expressed solely in terms of derivatives of this quantity.

Equation (17) can be cast in a more recognizable form by setting $f(\mathbf{z})=\mathbf{p}$ in Eq. (16), so that $\partial_{\mathbf{q}_0}\Delta\omega_1=\Delta\mathbf{p}_1$, where $\Delta\mathbf{p}_1$ is the first-order change in momentum; $\Delta\mathbf{p}_1=\mathbf{p}(\tau)-\mathbf{p}_0$ to first order in ϵ. Then, Eq. (17) becomes

$$\langle(\Delta f)_2\rangle=\tfrac{1}{2}\partial_{\mathbf{p}_0}\cdot\langle\Delta\mathbf{p}_1\Delta\mathbf{p}_1\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0)\rangle . \quad (19)$$

In a one-dimensional system the inverse Hamiltonian [12], $-P_z$, is used rather than H, evolution is in the spatial coordinate z rather than time, and t and γ (the relativistic factor) are, up to a constant, canonically conjugate variables. (The scalar variable z, which denotes position in Euclidean space, should not be confused with the phase-space vector $\mathbf{z}$.) In this case, if $f=mc^2\gamma$, the relativistic energy of a particle, Eq. (19) reduces to the original form of Madey's theorem:

$$\langle(\Delta\gamma)_2\rangle=\tfrac{1}{2}\partial_{\gamma_0}\langle\Delta\gamma_1^2\rangle_{t_0} ,$$

where the average is over t_0, the entrance times of the particles. Note that there is no average over z_0; this is because in the inverse Hamiltonian formalism z plays the role of time and its conjugate momentum is nonphysical.

Before concluding this section we review the conditions necessary for the validity of the generalized form of Madey's theorem, as states in Eq. (17) or (19).

(i) H_0, the zeroth-order Hamiltonian, must depend only on momentum and time, and f, the function of interest, must depend only on momentum.

(ii) The first- and second-order transformed Hamiltonians K_1 and K_2 must also depend only on momentum and time.

(iii) When averaging Eq. (16) over coordinates, the boundary terms that arise when integrating by parts must vanish.

Condition (i) is easy to check; (ii) and (iii) must be examined on a case by case basis. For most microwave devices, however, the Hamiltonian is a periodic function of the coordinates; i.e., $H_n=\sum_{\mathbf{k}}H_{n\mathbf{k}}(\mathbf{p})e^{i\mathbf{k}\cdot\mathbf{q}}$. In this case, as long as $\mathbf{k}\cdot\mathbf{v}_0\neq0$ for all relevant values of $\mathbf{k}$ and $\mathbf{p}$ [recall that $\mathbf{v}_0=\partial H_0(\mathbf{p})/\partial\mathbf{p}$], then (ii) and (iii) are automatically satisfied.

Implicit in the whole analysis, of course, is the assumption that the H_n are sufficiently small. For H_n periodic, as above, "sufficiently small" means

$$|H_{1\mathbf{k}}|\ll\frac{(\mathbf{k}\cdot\mathbf{v}_0)^2}{\left|\mathbf{k}\cdot\dfrac{\partial^2H_0}{\partial\mathbf{p}\,\partial\mathbf{p}}\cdot\mathbf{k}\right|} .$$

Physically, this expression is equivalent to saying that the particles are not trapped. In a linear analysis, of course, this can always be satisfied *unless* some symmetry of the Hamiltonian forces $\mathbf{k}\cdot\mathbf{v}_0$ to be identically zero (e.g., $\mathbf{k}=k_x\hat{\mathbf{x}}$ and H_0 independent of p_x). As we shall see in the next section, for periodic systems the expression for the second-order change in any function of momentum is valid even as $\mathbf{k}\cdot\mathbf{v}_0\to0$, so the condition $\mathbf{k}\cdot\mathbf{v}_0\neq0$ must be checked separately.

III. PERIODIC SYSTEMS

We now illustrate this method for a particle of mass m_0 and charge q interacting with an electromagnetic field characterized by a vector potential $\mathbf{A}$ (again we neglect the scalar potential Φ). The Hamiltonian for such a system is given in Eq. (1). Expanding the Hamiltonian in powers of $\mathbf{A}$, the first-order Hamiltonian is $H_1=-(q/c)\mathbf{v}_0\cdot\mathbf{A}$, where again $\mathbf{v}_0=\partial H_0/\partial\mathbf{p}$. For a plane wave, $H_1\sim e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, and H_1 is oscillatory and periodic in the coordinates. Even when there are external fields present, H_1 is typically oscillatory and periodic. For instance, in a gyrotron [13],

$$H_1=\sum_{n,s,l}G_{nsl}(L,I,h)e^{i(k_nz-\omega t+s\psi+l\theta_g)}+\text{c.c.} , \quad (20)$$

where c.c. denotes complex conjugate, ψ and θ_g are, respectively, the gyrophase and guiding center angle, and L, I, and h are the momenta conjugate to ψ, θ_g, and t. In fact, in some gyrotron geometries H_1 is even more complicated: both G_{nsl} and k_n can depend discontinuously on z (we consider this case in the next section). It was such a geometry that led us to the Lie transform technique, as all other perturbation theories appeared intractable. In a free-electron laser [14] (FEL), the first-order Hamiltonian is similar to Eq. (20), except typically FEL's are analyzed in one or two dimensions, so θ_g and possibly ψ are not present.

Here we choose a Hamiltonian simpler than that given in Eq. (20), but one that still captures the essential physics. Let $H=H_0(\mathbf{p})+H_1(\mathbf{q},\mathbf{p})+\cdots$ with

$$H_1(\mathbf{q},\mathbf{p})=\sum_{\mathbf{k}}H_{1\mathbf{k}}(\mathbf{p})e^{i\mathbf{k}\cdot\mathbf{q}}+\text{c.c.} , \quad (21)$$

where the wave vectors $\mathbf{k}$ lie on a rectangular lattice; i.e., $\mathbf{k}=(n_1k_1,n_2k_2\cdots)$ with k_i fixed and n_i positive integers. To compute the second-order change in any function of momentum, we use Eq. (17) with $\Delta\omega_1$ given by Eq. (18). For this Hamiltonian, the unperturbed orbits are simply $\mathbf{p}=\mathbf{p}_0$ and $\mathbf{q}=\mathbf{q}_0+\mathbf{v}_0\tau$, where $\mathbf{v}_0=\partial H_0(\mathbf{p}_0)/\partial\mathbf{p}_0$. Thus, Eq. (18) yields

$$\Delta\omega_1=-\int_0^\tau d\tau_0'\sum_{\mathbf{k}}H_{1\mathbf{k}}(\mathbf{p}_0)e^{i\mathbf{k}\cdot(\mathbf{q}_0+\mathbf{v}_0\tau_0')}+\text{c.c.}$$

The integral is straightforward, and we arrive at

$$\Delta\omega_1=-\tau\sum_{\mathbf{k}}H_{1\mathbf{k}}(\mathbf{p}_0)e^{i[\mathbf{k}\cdot\mathbf{q}_0+\mathbf{k}\cdot\mathbf{v}_0\tau/2]}\times\operatorname{sinc}(\mathbf{k}\cdot\mathbf{v}_0\tau/2)+\text{c.c.}\ , \tag{22}$$

where

$$\operatorname{sinc}(x)\equiv\frac{\sin x}{x}\ .$$

Inserting Eq. (22) into Eq. (17) and performing a small amount of straightforward algebra, we arrive at

$$\langle(\Delta f(\mathbf{p}_0))_2\rangle=\tau^2\sum_{\mathbf{k}}\mathbf{k}\cdot\partial_{\mathbf{p}_0}[|H_{1\mathbf{k}}(\mathbf{p}_0)|^2\operatorname{sinc}^2(\mathbf{k}\cdot\mathbf{v}_0\tau/2)\mathbf{k}\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0)]\ . \tag{23}$$

From this equation we see that the average second-order change in any function of momentum can be found rather simply in terms of the Fourier transform of the first-order Hamiltonian. Note that the form of the gain function depends weakly on the Hamiltonian. This is especially true in the limit of large τ (τ is essentially the length of the system): when $\tau\mathbf{k}\cdot\mathbf{v}_0\gg v_0$, Eq. (23) becomes

$$\langle(\Delta f)_2\rangle\approx\frac{\tau^3}{2}\sum_{\mathbf{k}}|H_{1\mathbf{k}}(\mathbf{p}_0)|^2(\mathbf{k}\cdot\partial_{\mathbf{p}_0}f_0)\times[(\mathbf{k}\cdot\partial_{\mathbf{p}_0})^2H_0]\frac{\partial}{\partial x}\operatorname{sinc}^2(x)$$

where $x\equiv\mathbf{k}\cdot\mathbf{v}_0\tau/2$. Thus, for large τ the second-order change in all functions of momentum has essentially the same form, regardless of the dimensionality of the system or the details of H_0 and H_1. Since the gain in a device is proportional to the second-order change in energy, we have immediately that emittance growth, changes in the moments of the perpendicular momentum, etc. are proportional to gain. This formalism is easily extended to first-order Hamiltonians of a form more complicated than that given by Eq. (21).

IV. APPLICATION TO GYROTRONS IN COMPLEX CAVITIES

The first-order Hamiltonian given in Eq. (21) may appear restrictive, but it turns out to apply to a wide range of microwave devices. In particular, it is applicable to a gyrotron in which the electron beam interacts with the fields in a finite length complex cavity made up of sections of uniform cylindrical waveguides. One such cavity is illustrated in Fig. 2. There is an applied magnetic field $B_0\hat{\mathbf{z}}$ which, in the absence of electromagnetic fields, generates helical trajectories. (Again, the scalar variable z should not be confused with the phase-space vector $\mathbf{z}$.) For clarity, the cavity illustrated in Fig. 2 is oversimplified, although such cavities have been used successfully in realistic, high-power gyroklystron amplifiers [15,16]. Usually, significantly more complex configurations are employed. For instance, in gyroklystron amplifiers the drift section between cavities may be loaded with rings of lossy dielectric [16] or the output cavity may have a complex shape to enhance mode selectivity [17]. In addition, tapered cavities and cavities whose radius changes smoothly with axial position may be modeled by a series of steps. Thus, in this section we consider a complex cavity consisting of an arbitrary number of sections of different radii.

The fields in a cavity of this type can be represented as a superposition of empty waveguide modes in each section, with the amplitudes of the modes chosen to enforce continuity of the fields at the interface between sections. In practice the amplitudes may be computed using the scattering matrix formalism [18], so here we will assume that they are known. Shortly we will write down explicit expressions of the fields. First, however, we will review the inverse Hamiltonian formalism needed to describe this system.

When there is an external uniform magnetic field present, the relativistic Hamiltonian is written

$$H=c\sqrt{m_0^2c^2+(\mathbf{p}-q/c\,\mathbf{A}_0-q/c\,\mathbf{A})^2}\ , \tag{24}$$

where

$$\mathbf{A}_0=\frac{B_0}{2}\hat{\mathbf{z}}\times\mathbf{x}_\perp$$

is the vector potential corresponding to the uniform field $B_0\hat{\mathbf{z}}$ and $\mathbf{A}$ is the vector potential for the electromagnetic fields [see Eq. (31)]. If we were interested in the temporal evolution of the system we would use the above Hamiltonian. However, we are interested in the spatial evolution, so we use instead the inverse Hamiltonian [12],

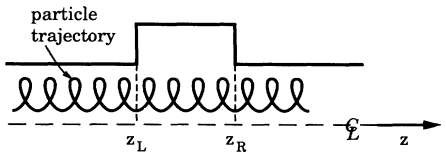

FIG. 2. Schematic of gyrotron beam passing through a complex cavity. z_L and z_R label the left and right edges of the section.

which we denote $\mathcal{P}$. Essentially, $\mathcal{P}$ is minus the z component of the canonical momentum $\mathbf{p}$ expressed in terms of the Hamiltonian H and the remaining canonical variables. With this formalism,

$$h \equiv -H$$

becomes the "momentum" variable conjugate to time. In addition, we must work in extended phase space so our inverse Hamiltonian $\mathcal{P}$ is autonomous. To this end we introduce the (nonphysical) variable ρ_z which is canonically conjugate to z. Then, solving Eq. (24) for p_z and setting $\mathcal{P}=\rho_z - p_z$, we arrive at

$$\mathcal{P}=\rho_z - \sqrt{(h/c)^2 - m_0^2 c^2 - (\mathbf{p}_\perp - q/c\,\mathbf{A}_0 - q/c\,\mathbf{A}_\perp)^2} - \frac{q}{c}A_z \ . \tag{25}$$

Expanding Eq. (25) to first order in the electromagnetic vector potential $\mathbf{A}$, we arrive at the zeroth- and first-order Hamiltonians

$$\mathcal{P}_0=\rho_z - \sqrt{(h/c)^2 - m_0^2 c^2 - (\mathbf{p}_\perp - q/c\,\mathbf{A}_0)^2} \ , \tag{26a}$$

$$\mathcal{P}_1 = -\frac{(\mathbf{p}_\perp - q/c\,\mathbf{A}_0)\cdot q/c\,\mathbf{A}_\perp}{\sqrt{(h/c)^2 - m_0^2 c^2 - (\mathbf{p}_\perp - q/c\,\mathbf{A}_0)^2}} - \frac{q}{c}A_z \ . \tag{26b}$$

To express the unperturbed Hamiltonian $\mathcal{P}_0$ in terms of momentum only, we need to change from Cartesian coordinates to guiding center variables. Defining the nonrelativistic cyclotron frequency Ω_0 as

$$\Omega_0 \equiv \frac{qB_0}{m_0 c} \ ,$$

the Larmor radius $\mathbf{r}_L$ and the guiding center radius $\mathbf{r}_g$ are written

$$\mathbf{r}_L = \frac{1}{m_0\Omega_0}\hat{\mathbf{z}}\times\left[\mathbf{p}_\perp - \frac{q}{c}\mathbf{A}_0\right] \ , \tag{27a}$$

$$\mathbf{r}_g = \frac{1}{m_0\Omega_0}\hat{\mathbf{z}}\times\left[-\mathbf{p}_\perp - \frac{q}{c}\mathbf{A}_0\right] \ . \tag{27b}$$

The canonical variables are the angular momentum L, the generalized guiding center momentum I, and their conjugate pairs ψ and θ_g; these are expressed in terms of the Larmor radius and guiding center radius as

$$L = -m_0\Omega_0\frac{r_L^2}{2} \ , \tag{28a}$$

$$I = m_0\Omega_0\frac{r_g^2}{2} \ , \tag{28b}$$

$$\psi = \tan^{-1}(r_{Ly}/r_{Lx}) \ , \tag{28c}$$

$$\theta_g = \tan^{-1}(r_{gy}/r_{gx}) \ . \tag{28d}$$

With these definitions it is straightforward to show that $\{\psi, L\} = \{\theta_g, I\} = 1$ and all cross brackets vanish. The eight-dimensional (four degree of freedom) phase space is now given by

$$\mathbf{z} = (\psi, \theta_g, t, z, L, I, h, \rho_z) \ .$$

In terms of this new set of canonical variables, the zeroth- and first-order Hamiltonians in Eq. (26) become

$$\mathcal{P}_0 = \rho_z - \sqrt{(h/c)^2 - m_0^2 c^2 + 2m_0\Omega_0 L} \tag{29a}$$

$$\mathcal{P}_1 = \frac{q}{c}\left[\frac{m_0\Omega_0\hat{\mathbf{z}}\times\mathbf{r}_L\cdot\mathbf{A}_\perp}{p_{z0}(h,L)} - A_z\right] \ , \tag{29b}$$

where

$$p_{z0}(h,L) \equiv \sqrt{(h/c)^2 - m_0^2 c^2 + 2m_0\Omega_0 L} \tag{30}$$

is the zeroth-order axial momentum.

Our next step is to express $\mathcal{P}_1$ in terms of guiding center variables. To do this we need the vector potential for the electromagnetic field. We will assume that the fields consist of a single azimuthal mode number m and a superposition of radial mode numbers, labeled by n. (The restriction to a single m is not really important; to include many azimuthal mode numbers we could simply take our final expression and sum on m.) Let us consider the region in Fig. 2 between z_L and z_R, the left and right ends of a section. The vector potential $\mathbf{A}$ in that section can be written

$$\mathbf{A} = \frac{m_0 c^2}{q}\mathrm{Re}\left[ae^{-i\omega t}\sum_n [F_n^-(z)e_n(\mathbf{x}_\perp)\hat{\mathbf{z}} + iF_n^+(z)\mathbf{E}_n(\mathbf{x}_\perp)]\right] \ . \tag{31}$$

The quantities in this expression are defined as follows: a is the dimensionless amplitude of the electromagnetic vector potential, $F_n^\pm(z)$ are the axial field profiles associated with the normal mode n,

$$F_n^\pm(z) = f_n e^{ik_n(z-z_L)} \pm b_n e^{-ik_n(z-z_R)} \ , \tag{32}$$

where k_n is the axial wave number and f_n and b_n are the forward and backward wave amplitudes, $\mathbf{E}_n(\mathbf{x}_\perp)$ is the perpendicular waveguide electric field, which is expressed in terms of the axial waveguide electric and magnetic fields $e_n(\mathbf{x}_\perp)$ and $h_n(\mathbf{x}_\perp)$ as

$$\mathbf{E}_n(\mathbf{x}_\perp) = \gamma_n^{-2}[k_n\nabla_\perp e_n - (\omega/c)\hat{\mathbf{z}}\times\nabla_\perp h_n] \ , \tag{33}$$

where γ_n is the perpendicular wave vector and the axial fields have the usual form

$$h_n(\mathbf{x}_\perp) = C_{Hn}J_m(\gamma_n r)e^{im\theta} \ , \tag{34a}$$

$$e_n(\mathbf{x}_\perp) = C_{En}J_m(\gamma_n r)e^{im\theta} \tag{34b}$$

with r and θ the radius and azimuthal angle ($r=|\mathbf{x}_\perp|$, $\tan\theta = y/x$), and J_m the Bessel function of order m. The quantities C_{Hn} and C_{En} are normalization factors for the TE and TM modes, respectively. For perfectly conducting walls, either $C_{Hn}=1$ and $C_{En}=0$ (pure TE) or $C_{En}=1$ and $C_{Hn}=0$ (pure TM). In this case the quantities C_{En} and C_{Hn} are essentially irrelevant. However, when $m\neq 0$ and there is dielectric loading in the cavities or the walls have a finite resistivity, both C_{En} and C_{Hn} are simultaneously nonzero; i.e., pure TE and TM modes do not exist. Here we consider C_{Hn} and C_{En} to be known, but essen-

tially arbitrary, constants to allow for general configurations. Finally, the frequency ω is taken to be the real part of the complex cold cavity frequency ω_0, which in turn is related to the wave vectors γ_n and k_n by

$$\frac{\omega_0^2}{c^2}=\gamma_n^2+k_n^2 . \tag{35}$$

For perfectly conducting walls of radius r_w, γ_n takes on an especially simple form: $\gamma_n=p_{mn}/r_w$ for TM modes and $\gamma_n=p'_{mn}/r_w$ for TE modes, where p_{mn} (p'_{mn}) is the nth zero of J_m (J'_m)

We assume that the quantities in Eqs. (31)–(35), i.e., the frequency ω, the wave vectors γ_n and k_n, the field amplitudes f_n and b_n, and the renormalization factors C_{En} and C_{Hn} are known. In practice, these can be computed from the cold cavity fields using the scattering matrix formalism, as in Ref. [18]. Of course, because the cold cavity fields are modified to some extent by the electron beam, the theory is not totally self-consistent. However, for practical gyrotron beams the modification is small enough that it can be neglected.

It is now a straightforward, although somewhat tedious, matter to work out the first-order Hamiltonian $\mathcal{P}_1$: simply insert the vector potential **A**, Eq. (31), into the expression for $\mathcal{P}_1$, Eq. (29b), and use Eqs. (27) and (28) to express the Cartesian momenta and coordinates in terms of the canonical variables. The final result is reasonably compact:

$$\mathcal{P}_1^{(l)}=\frac{m_0c}{2}\sum_{n,s}ae^{i[s\psi-\omega t+(m-s)\theta_g]} \times[G_{sn}^{+}(L,I,h)f_ne^{ik_n(z-z_L)} +G_{sn}^{-}(L,I,h)b_ne^{-ik_n(z-z_R)}]+\text{c.c.}, \tag{36}$$

where $G_{sn}^{\pm}$ is given by

$$G_{sn}^{\pm}(L,I,h)=-iJ_{m-s}(\gamma_n r_g)\times\left[C_{Hn}\frac{\omega}{\gamma_n c}\frac{\Omega_0 r_L}{c}\frac{m_0c}{p_{z0}}J'_s(\gamma_n r_L) -iC_{En}\left[\frac{k_n}{\gamma_n}\frac{s\Omega_0}{\gamma_n c}\frac{m_0c}{p_{z0}}\pm1\right]J_s(\gamma_n r_L)\right] .$$

Recall that r_L, r_g, and p_{z0} are functions of the canonical momenta L, I, and h as given in Eqs. (28) and (30).

In Eq. (36) we have given $\mathcal{P}_1$ a superscript (l); this denotes section. In fact, γ_n, k_n, z_L, z_R, C_{En}, and C_{Hn} also depend on section, so these quantities too should have a label attached to them. We do not include such a label for clarity, although when performing actual calculations we need to be aware that it exists.

The expression for $\mathcal{P}_1^{(l)}$ in Eq. (36) has almost the same form as the first-order Hamiltonian (21); the only difference is that $\mathcal{P}_1^{(l)}$ has a fairly complicated dependence on z. However, since we average over ψ, θ_g, and t but not z, most of the analysis in Sec. III is applicable here. The primary difference is that the computation of $\Delta\omega_1$ must be done section by section; i.e.,

$$\Delta\omega_1=\sum_l\Delta\omega_1^{(l)} , \tag{37}$$

where

$$\Delta\omega_1^{(l)}=-\int_{\tau_L}^{\tau_R}d\tau'_0\mathcal{P}_1^{(l)}(\tau'_0) , \tag{38}$$

τ_L and τ_R correspond to $z=z_L$ and $z=z_R$, respectively, and $\mathcal{P}_1(\tau'_0)$ is to be evaluated along unperturbed orbits. As usual, the momenta are constant along unperturbed orbits and the coordinates evolve according to

$$\psi(\tau_0)=\psi_0+\tau_0\partial_L\mathcal{P}_0 , \tag{39a}$$

$$\theta_g(\tau_0)=\theta_{g0}+\tau_0\partial_I\mathcal{P}_0 , \tag{39b}$$

$$t(\tau_0)=t_0+\tau_0\partial_h\mathcal{P}_0 , \tag{39c}$$

$$z(\tau_0)=z_0+\tau_0\partial_{p_z}\mathcal{P}_0 . \tag{39d}$$

Previously we used a subscript 0 to denote initial value of the momenta. Here we drop that subscript for clarity, but we should keep in mind that the momenta L, I, and h are to be evaluated at their initial values L_0, I_0, and h_0.

To compute $\Delta\omega_1^{(l)}$, all we need to do is insert the expressions for ψ, θ_g, t, and z given in Eq. (39) into the expression for $\mathcal{P}_1^{(l)}$ given in Eq. (36), then perform the integral over τ'_0 in Eq. (38). To simplify the resulting expression we define the unperturbed axial velocity,

$$v_{z0}\equiv-\frac{c^2p_{z0}}{h} ,$$

the unperturbed relativistic cyclotron frequency,

$$\Omega_{c0}\equiv-\frac{m_0c^2}{h}\Omega_0 ,$$

the unperturbed transit time through section l,

$$T^{(l)}\equiv\frac{z_R-z_L}{v_{z0}} ,$$

and the unperturbed transit time from $z=0$ to the center

of section l,

$$\bar{T}^{(l)} \equiv \frac{z_R + z_L}{2v_{z0}} .$$

Since in the absence of electromagnetic fields, p_{z0} is the axial mechanical momentum and $h = -H = -m_0c^2\gamma$, where γ is the relativistic factor, the above quantities correspond to the usual definitions. Finally, using $\partial_{p_z}\mathcal{P}_0 = 1$, we arrive at the expression for $\Delta\omega_1^{(l)}$,

$$\Delta\omega_1^{(l)} = -\frac{m_0c}{2} T^{(l)} v_{z0} \sum_{n,s} a e^{i[s\psi_0 - \omega t_0 + (m-s)\theta_{g0}]} \times \Delta\omega_{ns}^{(l)}(L,I,h) + \text{c.c.} ,$$

where

$$\Delta\omega_{ns}^{(l)}(L,I,h) \equiv e^{ik_n v_{z0} T^{(l)}/2 - i(\omega + s\Omega_{c0})\bar{T}^{(l)}} \{ G_{ns}^{+}(L,I,h) f_n \operatorname{sinc}[(\omega + s\Omega_{c0} - k_n v_{z0}) T^{(l)}/2] + G_{ns}^{-}(L,I,h) b_n \operatorname{sinc}[(\omega + s\Omega_{c0} + k_n v_{z0}) T^{(l)}/2] \} .$$

It is now completely straightforward to perform the average over ψ, θ_g, and t in Eq. (17). Keeping in mind that $\Delta\omega_1$ is a sum over sections of $\Delta\omega_1^{(l)}$ [Eq. (37)], we find that

$$\langle (\Delta f(L,I,h))_2 \rangle = \frac{m_0^2c^2|a|^2}{4} \sum_s [s\partial_L + (m-s)\partial_I - \omega\partial_h] \left| \sum_{n,l} v_{z0} T^{(l)} \Delta\omega_{ns}^{(l)} \right|^2 \times [s\partial_L + (m-s)\partial_I - \omega\partial_h] f(L,I,h) . \quad (40)$$

Equation (40) represents our final result. Typically the quantity of interest is the change in particle kinetic energy, $m_0c^2\langle(\Delta\gamma)_2\rangle$, for which $f = -h$ (since $h = -m_0c^2\gamma$). This simplifies the last term in Eq. (40):

$$[s\partial_L + (m-s)\partial_I - \omega\partial_h](-h) = \omega .$$

Once $\langle(\Delta\gamma)_2\rangle$ is known the gain can be computed from power balance. Defining the linear efficiency η by

$$\eta = -\frac{\langle(\Delta\gamma)_2\rangle}{\gamma - 1}$$

(there is a minus sign in the definition because $\Delta\gamma \equiv \gamma_{\text{final}} - \gamma_{\text{initial}}$), we see that the power going into the cavity is ηP_b, where P_b is the beam power: $P_b =$ (beam voltage)$\times$(beam current). The power leaving the cavity is $\omega W_{Em}/Q$, where W_{Em} is the field energy and Q is the cavity quality factor, including both diffractive and absorptive losses. In terms of these quantities, power balance is written

$$\frac{dW_{Em}}{dt} = \eta P_b - \frac{\omega W_{Em}}{Q} .$$

Letting W_{Em} have an exponential dependence on time $W_{Em} \sim e^{\Gamma t}$, the gain Γ becomes

$$\Gamma = \frac{\eta P_b}{W_{Em}} - \frac{\omega}{Q} .$$

Since η is the linear efficiency, it is proportional to W_{Em} and our expression for gain is independent of the amplitude of the electromagnetic field. Thus, calculating $\langle(\Delta\gamma)_2\rangle$ from Eq. (40) leads directly to an expression for the gain.

V. NON-HAMILTONIAN SYSTEMS

The transform approach need not be confined to Hamiltonian systems; any differential equation can be solved perturbatively by looking for a change of variables that simplifies the equations. For instance, consider the differential equation

$$\frac{d\mathbf{z}}{d\tau} = \mathcal{H}\mathbf{z} \equiv \sum_{n=0}^{\infty} \mathcal{H}_n \mathbf{z} , \quad (41)$$

where the $\mathcal{H}_n$ are operators of the form $\mathcal{H}_n = \mathbf{H}_n(\mathbf{z})\cdot\partial_{\mathbf{z}}$. As with the Hamiltonian equations we can look for a change of variables $\mathbf{z}' = T\mathbf{z} = e^{-L}\mathbf{z}$, but now L has the general form

$$L = \sum_{n=1}^{\infty} \frac{1}{n!} \mathbf{w}_n(\mathbf{z})\cdot\partial_{\mathbf{z}} \equiv \sum_{n=1}^{\infty} \frac{L_n}{n!} .$$

The transform operator T no longer generates canonical transformations but it still commutes with functions. [That T commutes with functions follows from the same arguments as in the Hamiltonian case: Eq. (7) and the comments following it apply even though the vector field generated by L is non-Hamiltonian.] With this form for L and $\mathcal{H}$, in the primed frame the differential equation is

$$\frac{d\mathbf{z}'}{d\tau} = e^{L(\mathbf{z}')} \mathcal{H}(\mathbf{z}') e^{-L(\mathbf{z}')} \mathbf{z}' .$$

Writing $\mathcal{K} = e^{L}\mathcal{H}e^{-L}$ leads to a set of transform equations analogous to Eq. (13), except now $L_n \rightarrow [L_n, \,]$ where $[\,,\,]$ is the usual quantum-mechanical commutator: $[A,B] = AB - BA$. For example, the transform equations for Eq. (41) are

$$\mathcal{K}_0=\mathcal{H}_0 ,$$

$$\mathcal{K}_1=[L_1,\mathcal{H}_0]+\mathcal{H}_1 ,$$

$$\mathcal{K}_2=\tfrac{1}{2}[L_2,\mathcal{H}_0]+\mathcal{H}_2+[L_1,\mathcal{H}_1]+\tfrac{1}{2}[L_1,[L_1,\mathcal{H}_0]] .$$

This approach is especially useful for normal form calculations [19], in which the problem is to determine which nonlinearities in an ordinary differential equation cannot be eliminated by a change of variables.

It is straightforward to show that Eq. (15), which gives the second-order change in any function of z, is valid even for non-Hamiltonian systems; the only difference is that S_K is replaced by $S_{\mathcal{H}}$, the time development operator associated with Eq. (41). Equation (16) is also valid as long as $\{\omega, \}$ is replaced by $\mathbf{w}\cdot\partial_z$ and the Poisson bracket $\{\omega_{1\tau},\omega_1\}$ is replaced by the commutator $[\mathbf{w}_{1\tau}\cdot\partial_z\mathbf{w}_{1\tau}\cdot\partial_z]$: for the non-Hamiltonian differential equation (41),

$$\begin{aligned}(\Delta f)_2=&[f_\tau(\mathbf{z}_0)-f(\mathbf{z}_0)]+\Delta\mathbf{w}_1\cdot\partial_z f_\tau\\&+\tfrac{1}{2}\Delta\mathbf{w}_2\cdot\partial_z f_\tau+\tfrac{1}{2}\Delta\mathbf{w}_1\cdot\partial_z\Delta\mathbf{w}_1\cdot\partial_z f_\tau\\&+\tfrac{1}{2}[\mathbf{w}_{1\tau}\cdot\partial_z,\mathbf{w}_\tau\cdot\partial_z]f_\tau ,\end{aligned}$$

where, analogous to previous definitions, $\Delta\mathbf{w}_n\equiv\mathbf{w}_{n\tau}-\mathbf{w}_n$ and $f_\tau(\mathbf{z}_0)\equiv S_{\mathcal{H}}(\tau)f(\mathbf{z}_0)$. Thus, as with Eq. (16), the τ evolution of a function can be determined merely by taking a set of derivatives. Again, this approach is useful only if the operator $S_{\mathcal{H}}(\tau)$ is simple.

It is natural to ask to what extent Madey's theorem applies to the non-Hamiltonian differential equations given in Eq. (41). For this question to even make sense we need to prescribe a subset of the coordinates $\mathbf{z}$ to average over. Analogously to the Hamiltonian case we write $\mathbf{z}=(\mathbf{q},\mathbf{p})$, although here we do not demand that the vectors $\mathbf{q}$ and $\mathbf{p}$ have the same number of components. For definiteness we will call $\mathbf{q}$ the coordinates and $\mathbf{p}$ the momenta, and we will average over the initial values of the coordinates. We choose the zeroth-order differential operator $\mathcal{H}_0$ to be of the form

$$\mathcal{H}_0=\mathbf{v}_0(\mathbf{p})\cdot\partial_{\mathbf{q}} \tag{42}$$

and the nth-order operators to be of the form

$$\mathcal{H}_n=\mathbf{H}_n(\mathbf{z})\cdot\partial_{\mathbf{z}} . \tag{43}$$

We will assume that $\mathcal{H}_0$ and $\mathcal{H}_n$ are chosen so that the transform equations can be solved with $\mathcal{K}_1=0$ and $\mathcal{K}_2$ of the form $\mathbf{K}_2(\mathbf{p})\cdot\partial_{\mathbf{q}}$. In addition we will consider f to be a function of momentum only: $f(\mathbf{z})=f(\mathbf{p})$. With these restrictions it is straightforward to show that

$$\begin{aligned}\langle(\Delta f)_2\rangle=&\tfrac{1}{2}\partial_{\mathbf{p}_0}\cdot\langle\Delta\mathbf{p}_1\Delta\mathbf{p}_1\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0)\rangle-\tfrac{1}{2}\langle(\partial_{\mathbf{z}_0}\cdot\Delta\mathbf{w}_1)\Delta\mathbf{w}_1+(\partial_{\mathbf{z}_0}\cdot\mathbf{w}_1)\mathbf{w}_{1\tau}-(\partial_{\mathbf{z}_0}\cdot\mathbf{w}_{1\tau})\mathbf{w}_1\rangle\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0)\\&+\tfrac{1}{2}[\partial_{\mathbf{p}_0}\cdot\langle\mathbf{w}_{1\tau}\mathbf{w}_1-\mathbf{w}_1\mathbf{w}_{1\tau}\rangle]\cdot\partial_{\mathbf{p}_0}f(\mathbf{p}_0) ,\end{aligned} \tag{44}$$

where the angular brackets indicate an average over initial coordinates $\mathbf{q}_0$, and $\Delta\mathbf{p}_1=\Delta\mathbf{w}_1\cdot\partial_{\mathbf{z}_0}\mathbf{p}_0$ is the change in momentum to first order.

The first term in this equation is identical in form to the Hamiltonian version of Madey's theorem. The others arise because of the non-Hamiltonian nature of the differential equation. Although Eq. (44) is relatively complicated, it does have the advantage that the right-hand side has no second-order quantities. This in itself simplifies calculations, and even further simplification can be achieved if we place some restriction on the differential operators in Eqs. (42) and (43). For instance, if we demand that the differential equation be volume preserving through first order, i.e.,

$$\partial_{\mathbf{z}}\cdot\mathbf{H}_0=\partial_{\mathbf{z}}\cdot\mathbf{H}_1=0 ,$$

then it is straightforward to show that the second term in angular brackets in Eq. (44) vanishes. Whether or not the third term vanishes depends on $\mathcal{H}_1$, but for periodic functions $\mathbf{H}_1=\sum_{\mathbf{k}}\mathbf{H}_{1\mathbf{k}}(\mathbf{p})e^{i\mathbf{k}\cdot\mathbf{q}}$, the third term is identically zero. Thus, some form of Madey's theorem applies to a class of differential equations more general than Hamiltonian, and for a restricted class the Hamiltonian version of Madey's theorem as expressed in Eq. (19) applies directly.

VI. SUMMARY AND CONCLUSIONS

We have shown that the Lie transform method can be used to solve perturbatively Hamiltonian differential equations by introducing a change of variables which removes the nonessential nonlinearities. In this method, a vector field is constructed whose solution gives the desired change of variables. Such an approach is both simple and powerful, it is easy to generalize to non-Hamiltonian systems, and it has a ready geometrical interpretation. For Hamiltonian equations, as well as a restricted class of non-Hamiltonian equations, we derived a generalized form of Madey's theorem. In this case, the vector field which produces the coordinate transformation is generated by the first-order action, $-\int_0^\tau d\tau_0' H_1$. Derivatives of this quantity carry all the information necessary to compute $\langle(\Delta f(\mathbf{p}))_2\rangle$, the second-order change in any function of momentum. Because of the generality of the transform method, we were able to derive a universal form for $\langle(\Delta f)_2\rangle$ in terms of the Fourier transform of the first-order Hamiltonian. Finally, for a gyrotron interacting in a complicated cavity, we computed the average change in particle kinetic energy through second order in the field amplitude, which led directly to an expression for the gain.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy.

[1] J. M. J. Madey, Nuovo Cimento **65**, 64 (1979).
[2] S. Krinsky, J. M. Wang, and P. Luchini, J. Appl. Phys. **8**, 5453 (1982).
[3] P. Luchini an dS. Solimeno, IEEE J. Quantum Electron. **QE-21**, 952 (1985).
[4] A. Deprit, Cel. Mech. **1**, 12 (1969).
[5] R. L. Dewar, J. Phys. **9**, 2043 (1976).
[6] R. L. Dewar, J. Phys. **11**, 9 (1978).
[7] J. R. Cary, Phys. Rep. **79**, 130 (1981).
[8] J. R. Cary and A. N. Kaufman, Phys. Fl. **24**, 1238 (1981).
[9] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer, New York, 1983).
[10] A. Dragt and J. M. Finn, J. Math. Phys. **17**, 2215 (1976).
[11] H. Goldstein, *Classical Mechanics,* 2nd ed. (Addison-Wesley, Reading, MA, 1980).
[12] E. D. Courant and H. S. Snyder, Ann. Phys. **3**, 1 (1958).
[13] As far as we know, there are no published results in which the Hamiltonian formalism is applied to the full set of gyrotron equations. For a discussion of guiding center variables, from which the first-order Hamiltonian is easily derived, see C. S. Kou, D. B. McDermott, N. C. Luhman, Jr., and K. R. Chu, IEEE Trans. Plasma Sci. **PS-18**, 343 (1990), and references cited therein.
[14] C. Grebogi and H. S. Uhm, Phys. Fl. **28**, 1984 (1985).
[15] K. R. Chu, V. L. Granatstein, P. E. Latham, W. Lawson, and C. D. Striffler, IEEE Trans. Plasma Sci **13**, 424 (1985).
[16] W. Lawson, J. P. Calame, B. Hogan, P. E. Latham, M. E. Read, V. L. Granatstein, M. Reiser, and C. D. Sriffler, Phys. Rev. Lett. **67**, 520 (1991).
[17] H. Guo, D. S. Wu, G. Liu, Y. M. Miao, S. Z. Qian, and W. Z. Qin, IEEE Trans. Plasma Si. **18**, 326 (1990).
[18] J. M. Nielson, P. E. Latham, M. Caplan, and W. Lawson, IEEE Trans. Microwave Theory, Tech. **37**, 1165 (1989).
[19] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields* (Springer-Verlag, New York, 1983), Sec. 3.3.