CANONICAL TREATMENT OF ACCELERATED ORBITS IN SECTOR-FOCUSED CYCLOTRONS*

M.M. GORDON

National Superconducting Cyclotron Laboratory
Michigan State University
East Lansing, MI 48824

December 1982

*Work supported by the National Science Foundation under Grant No. Phy 78-22696.

Abstract

With the aim of improving orbit computation programs as well as deriving useful invariants, we present a fairly general formulation of the Hamiltonian which is then applied both to cyclotrons having conventional radial gaps and to those (like the new superconducting cyclotrons) having spiral electric gaps. In transforming to the radial displacement variables (x,p_x), it is then found that the time t is transformed into a new variable u which is conjugate to the energy E and which is, moreover, an invariant to first order. This invariant allows us to define more accurately a phase ϕ which is then quite properly a constant of the motion for a monoenergetic group of ions executing linear oscillations about the equilibrium orbit.

When applied to the longitudinal motion itself, the analysis leads directly to Joho's Hamiltonian whose invariance not only produces the standard phase versus energy relationship, but which also includes the phase compression effect associated with a change in gap voltage with radius. Moreover, we automatically obtain the same result for spiral electric gaps as for radial gaps.

I. INTRODUCTION

It has now become fairly standard procedure to carry out detailed orbit computations as part of the process of designing cyclotron components or as an aid in diagnosing beam characteristics. Since the interpretation of the results often depends on the theorems of phase space dynamics, the orbit computation programs should be based on equations of motion which can be derived from a Hamiltonian, and as far as is feasible, an accurate Hamiltonian. The design of such programs has been of interest to us for a long time, and we present here some results (both old and new) of our research on this subject.

In addition to these considerations, a canonical treatment of the motion reveals quite naturally the existence of certain invariants which are not only very useful, but which are otherwise somewhat difficult to discover.

In reviewing the application of Hamiltonian and canonical techniques to accelerator dynamics, we generally find that the analysis of accelerated orbits follows a different and less rigorous approach than that used for nonaccelerated orbits. For example, the basic Hamiltonian was formulated in Serret-Frenet coordinates both by Courant and Snyder[1] and by Kolomensky and Lebedev[2] (following almost identical procedures), but neither of them chose to show that this same Hamiltonian could be used to analyze synchroton oscillations

as well as betatron oscillations. Such an analysis was sketched by Mills[3] in a brief paper, but a complete treatment has not as yet been published.[4]

There is, of course, an extensive use of phase space techniques in the analysis of accelerated orbits carried out by Symon and Sessler,[5] and by those who followed their lead.[6] Such analyses start, however, from the differential equations for the longitudinal motion and then, after making suitable approximations, construct a Hamiltonian to represent the approximate results.

This situation has developed perhaps from a general reluctance to use the basic Hamiltonian because of its highly nonlinear dependence on the vector potential components, and hence the fields. That is, the Lagrangian and the equations of motion themselves are completely linear in the fields, but the Hamiltonian is not.

Turning now more specifically to cyclotrons, we find that an elaborate analysis has recently been carried out by Schulte and Hagedoorn[7] with an extensive use of canonical transformations. Their work is, however, based on an oversimplified Hamiltonian so that the results, although interesting, are only qualitatively valid.

In the present paper, we start by describing a somewhat unusual but fairly rigorous formulation of the problem based on a relativistically covariant form of Hamilton's Principle

which, because of its inherent space-time symmetry, facilitates transforming the azimuth θ into the independent variable. Then, after constructing an appropriate Hamiltonian, we discuss the technique of gauge transformations and thereby obtain a suitable representation for the rf fields.

This formulation allows us to separate the Hamiltonian into a term describing the nonaccelerated orbits, and a second term which is responsible for the acceleration effects. Moreover, the resultant Hamiltonian is, to a good approximation, linear in the rf fields. For the sake of generality, this analysis is carried out not only for the radial gap geometry common to most cyclotrons, but also for the spiral gap geometry characteristic of the new superconducting cyclotrons.

In the present paper, our attention is focused almost exclusively on median plane motion and following general practice, we expand the Hamiltonian in powers of x and p_x, the deviations of r and p_r from their values on the instantaneous equilibrium orbit. The shift from (r,p_r) to the pair (x,p_x) is brought about by a specific canonical transformation and, as a result, the variable conjugate to the energy E is changed from the time t to a quantity u which we call the "commencement".

Perhaps the most interesting consequence of this transformation is that the commencement u turns out to be an invariant quantity for nonaccelerated orbits, at least to

first order. Thus, the longitudinal motion associated with free linear oscillations is characterized by two constants, E and u. Moreover, the use of u rather than t as an orbit coordinate has the advantage of eliminating the fluctuations observed in the variation of t with θ as a result of the linear oscillations.

The quantity u also leads quite naturally to a proper definition of the phase ϕ characterizing an accelerated orbit. As discussed at length in a previous paper,[8] the phase ϕ should be defined so as to be a constant of the motion for a monoenergetic group of ions executing linear oscillations. Since u is just such a constant, it seems appropriate to set $\phi = \omega_{rf}u$, where ω_{rf} is the rf angular frequency.

Actually, this ϕ definition is exact only for an isochronous field, but it requires only a minor change to make it quite general. We should also note that the difference between our phase definition and the usual one tends to disappear (as it should) when a long term average is considered.

An important feature of our analysis is that for the longitudinal motion itself, it leads directly to the Hamiltonian derived by Joho[9] following a more elementary approach. This Hamiltonian is an approximate invariant which yields the familiar relation between sin ϕ and E, and which also exhibits the phase compression effect produced by a radial

variation of the gap voltage. Moreover, our derivation shows that these results are valid for machines having spiral electric gaps as well as those with a conventional radial gap geometry.

In a future paper, we plan to extend the analysis so as to provide a more complete treatment of the effects arising from the coupling of the acceleration process with the linear radial and vertical oscillations.

II. HAMILTONIAN CONSTRUCTION

The motion of an ion having charge q and mass m in an electromagnetic field is treated at length by Goldstein[10] and by Jackson,[11] and although their approaches are considerably different, they both establish the existence of a canonical momentum 4-vector given by:

$$P_\alpha = mu_\alpha + q'A_\alpha \ , \tag{1}$$

where $q' = q/c$ in gaussian units. When expressed in traditional Minkowsky notation, $u_\alpha = (\gamma\vec{v}, i\gamma c)$ is the velocity 4-vector, while $A_\alpha = (\vec{A}, i\phi)$ is the 4-vector constructed from the vector potential $\vec{A}$ and the scalar potential ϕ.

If the ion's position is expressed in cylindrical polar coordinates (r,θ,z), then the corresponding canonical momenta are given by:

$$P_r = m\gamma v_r + q'A_r \ , \tag{2a}$$

$$P_\theta = r(m\gamma v_\theta + q'A_\theta) \ , \tag{2b}$$

and $$P_z = m\gamma v_z + q'A_z \ . \tag{2c}$$

This P_θ can be recognized as the canonical angular momentum which is a constant of the motion when the fields are axially symmetric.

Since $u_4 = i\gamma c$ and $A_4 = i\phi$, the canonical momentum P_4 conjugate to $x_4 = ict$ is then given by:

$$P_4 = i(\gamma mc^2 + q\phi)/c = iW/c \ , \tag{3}$$

where W here is evidently the total relativistic energy of the ion. Moreover, since $P_4 dx_4 = -Wdt$, we also have $P_t = -W$ as the canonical momentum conjugate to the time t itself.

The square of the 4-vector u_α is the invariant $-c^2$, and if we identify $\vec{p} = m\gamma\vec{v}$ as the ordinary momentum vector, it then follows that:

$$p^2 = (\gamma^2 - 1)m^2c^2 \ , \tag{4a}$$

which can be recognized as a form of the momentum energy relationship for a free particle. Alternatively, if (3) is used to eliminate γ, we than find:

$$p^2 = \frac{1}{c^2}(W - q\phi)^2 - m^2c^2 \ , \tag{4b}$$

and if (2) is then used to eliminate p^2, we finally obtain:

$$(P_r - q'A_r)^2 + (\frac{P_\theta}{r} - q'A_\theta)^2 + (P_z - q'A_z)^2$$

$$= \frac{1}{c^2}(W - q\phi)^2 - m^2c^2 \ . \qquad (5)$$

This equation provides an important relationship between the four canonical momenta $(P_r,P_\theta,P_z,P_t = -W)$ and their corresponding coordinates (r,θ,z,t).

As is well known, the canonical equations of motion can be derived from Hamilton's Principle,[10] and in keeping with the First Postulate, this principle can be rewritten in a manifestly covariant form as follows:

$$\delta \int P_\alpha dQ^\alpha = 0 \ , \qquad (6)$$

where the integral is taken along the "world line" of the particle. Here, with $P_\alpha = (P_r,P_\theta,P_z,P_4)$ as given in (2,3) above, then $dQ^\alpha = (dr,d\theta,dz,icdt)$, and since $P_4 = iW/c$ and $P_t = -W$, this principle becomes:

$$\delta \int (P_r dr + P_\theta d\theta + P_z dz + P_t dt) = 0 \ , \qquad (7)$$

in our case.

As is true for all relativistically covariant equations, this equation involves the time and space coordinates in a completely symmetric way, and therefore provides an obvious basis for a formulation of Hamiltonian mechanics with any

one of them as the independent variable. Thus, if the time t is chosen for this variable, as is usually done, then we set $H = -P_t = W$ in (5), and thereby obtain the conventional Hamiltonian.

Here, however, we want θ to be the independent variable, and hence by symmetry, we set $H = -P_\theta$ in (5) to obtain the desired Hamiltonian. For a Linac, on the other hand, the z coordinate would be the appropriate choice, and this special case has been treated by Schnizer[12] following a somewhat similar approach, but using predominantly nonrelativistic considerations.

With θ as the independent variable, $Q_i = (r,z,t)$ become the relevant coordinates, and $P_i = (P_r,P_z,P_t = -W)$ the corresponding canonical momenta. We then set $P_\theta = -H$ in (5) and rewrite the result in the following simplified form:

$$H = -r(p^2 - p_r^2 - p_z^2)^{\frac{1}{2}} - q'rA_\theta \ , \qquad (8)$$

with the understanding that p, p_r, and p_z are to be expressed in terms of Q_i, P_i, and θ. That is, from (2), we have:

$$p_r = P_r - q'A_r, \ p_z = P_z - q'A_z \ , \qquad (9)$$

while (4b) gives the appropriate expression for p^2.

The resultant canonical equations of motion then take on their standard form, but with t replaced by θ; that is,

$$\frac{dQ_i}{d\theta} = \frac{\partial H}{\partial P_i} \ , \ \frac{dP_i}{d\theta} = -\frac{\partial H}{\partial Q_i} \ . \qquad (10)$$

Thus, for example, if H is independent of t, it follows immediately that P_t, and hence the total energy W, is a constant of the motion.

Before proceeding, we should call attention to another consequence of the covariant form of Hamilton's Principle given in (6). That is, the generating function for a canonical transformation should be a Lorentz scalar (or invariant) since this will preserve the covariance.

For example, the canonical momenta P_α given in (1) can be applied directly only to cartesian components and they are therefore conjugate to the position 4-vector X^α = (x,y,z,ict) expressed in cartesian coordinates. In order to find the momenta P'_β which are conjugate to the generalized coordinates Q^β = (r,θ,z,t), we can utilize the following (Lorentz scalar) generating function:

$$\Psi = P_\alpha X^\alpha(Q^\beta) \ , \tag{11}$$

which then yields:

$$P'_\beta = \partial\Psi/\partial Q^\beta = P_\alpha(\partial X^\alpha/\partial Q^\beta) \ . \tag{12}$$

For β = 1,2,3, this equation gives P_r, P_θ, and P_z shown in (2), while for β = 4, we find:

$$P_t = icP_4 = -W \ , \tag{13}$$

which confirms our previous assertion.

III. MEDIAN PLANE MOTION

As is customary, we expand H in powers of z and p_z, and assume that the expansion converges rapidly in the region of interest. We also assume the absence of imperfections which might disturb median plane symmetry so that the Hamiltonian H, and hence its expansion, is a function of z^2 as well as p_z^2.

For the present at least, we restrict the discussion to median plane motion and therefore set $z = p_z = 0$. In this case, B_z is the only nonvanishing component of the guide field, and the corresponding vector potential can then be reduced to the single component A_θ given by:

$$rA_\theta = -\int rB(r,\theta)dr \ , \tag{14}$$

where we define $B_z = -B(r,\theta)$ on the assumption that the function $B(r,\theta)$ is positive, at least on the average.

We consider first the case of nonaccelerated orbits. Since B_z is then the only nonzero field component, we can take $A_r = 0$ in (9) so that $P_r = p_r$, the ordinary radial momentum component. The Hamiltonian (8) then reduces to:

$$H^0 = -r(p^2 - p_r^2)^{\frac{1}{2}} + q'\int rB(r,\theta)dr \ , \tag{15}$$

where the superscript "0" indicates that the electric field is zero. This Hamiltonian is essentially the same as the one used by Hagedoorn and Verster[13] in their analysis of

nonaccelerated orbits in the median plane of a sector-focused cyclotron.

Since $\phi = 0$ also, eq. (4b) for p^2 reduces to:

$$p^2 = (W/c)^2 - m^2c^2 \ , \tag{16}$$

with $P_t = -W$ as before. Instead of P_t itself, it seems preferable to use W as the variable, or even better, the energy $E = \gamma mc^2$ to which it reduces in this case.

However, we could equally well choose E to be the kinetic energy, $(\gamma - 1)mc^2$, since only the momentum p and the differential $dP_t = -dE$ actually occur in the equations of motion. This choice for E seems better suited to cyclotrons, and is therefore the one we shall adopt here.

Thus, when there is no acceleration, the equations of motion (10) in the median plane then become:

$$\frac{dr}{d\theta} = \frac{\partial H^0}{\partial P_r} = \frac{rp_r}{\sqrt{(p^2 - p_r^2)}} \ , \tag{17a}$$

$$\frac{dp_r}{d\theta} = -\frac{\partial H^0}{\partial r} = (p^2 - p_r^2)^{\frac{1}{2}} - q'rB(r,\theta) \ , \tag{17b}$$

$$\frac{dt}{d\theta} = -\frac{\partial H^0}{\partial E} = \frac{\gamma mr}{\sqrt{(p^2 - p_r^2)}} \ . \tag{17c}$$

Furthermore, since H^0 is independent of t, the energy E is a constant of the motion, as it should be.

These equations of motion (or variants thereof) are ideally suited for computer programs designed to calculate

median plane orbits covering most of the cyclotron. Except for a small region near the center, the effect of the rf fields can be represented quite well by impulsive changes in E and p_r at a discrete set of "gap crossings". Such computer programs have been widely used at many laboratories since their early development at Oak Ridge.[14]

IV. RF FIELD REPRESENTATION

It is worth keeping in mind that the vector and scalar potentials are far from unique since any set of potentials can be changed by an arbitrary gauge transformation without changing the resultant fields or the equations of motion (gauge invariance).[11] Moreover, we are not restricted to the Lorentz gauge or any other special gauge since our job here is not to calculate the potentials from a given distribution of charges, currents, and boundary conditions (which is the business of electrodynamics), but rather to determine suitable potentials to represent a given set of electromagnetic fields.

We should also recognize that gauge transformations do modify the canonical momenta and Hamiltonian as is clearly evident from eq. (1) for P_α. However, such modifications are completely acceptable since they can also be produced by canonical transformations.

To see this, consider a general gauge transformation from the given A_α to a new 4-vector potential A'_α which can be expressed in covariant form as follows:

$$A'_\alpha = A_\alpha + (\partial\Lambda/\partial X^\alpha) \ , \tag{18}$$

where Λ is some (Lorentz scalar) function of the coordinates and time. This transformation changes P_α in (1) to P'_α in a way which is identical to the change produced by the following generating function:

$$\Psi = P_\alpha X^\alpha + q'\Lambda \ . \tag{19}$$

That is, the resultant canonical transformation yields:

$$P'_\alpha = \partial\Psi/\partial X^\alpha = P_\alpha + q'(\partial\Lambda/\partial X^\alpha) \ , \tag{20}$$

which therefore proves the point in question.

Turning now to the representation of the RF fields, we choose a special set of potentials by requiring $\Phi = 0$ and then determining $\vec{A}$ accordingly. For this choice of gauge, the total vector potential can be written as follows:

$$\vec{A} = \vec{A}^0 - c\int\vec{E}dt \ , \tag{21}$$

where $\vec{A}^0$ is the potential representing the static guide field, and where $\vec{E}$ is the RF electric field. Since $\Phi = 0$, the resultant fields are then given by:

$$\vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} = \vec{E} \ , \tag{22a}$$

and

$$\vec{B} = \nabla\times\vec{A} = \vec{B}^0 - c\int(\nabla\times\vec{E})dt \ , \tag{22b}$$

where the equation for E is evidently an identity, and where $\vec{B}^0$ is the static guide field. The second term in $\vec{B}$ reduces, with the aid of Faraday's law, to the RF magnetic field, as it should.

One advantage of having $\Phi = 0$ is that we can continue to set $P_t = -E$, thereby maintaining the ion's energy as one of the canonical variables. Indeed, we shall henceforth take E to be one of the coordinates and $P_e = +t$ to be the corresponding canonical momentum. Such an interchange of roles (with an appropriate change of sign) does not change the equations of motion and can therefore be brought about by a canonical transformation.

V. EXTENDED HAMILTONIAN

Returning now to our discussion of motion in the median plane, we have $E_z = 0$, assuming again the absence of imperfections. The RF fields are therefore determined entirely by the components E_r and E_θ which are generally functions of r, θ, and t.

Thus, A_r is no longer zero, and instead of (14), A_θ is now given by:

$$rA_\theta = -\int rB(r,\theta)dr - c\int rE_\theta dt \ , \tag{23}$$

as follows from (14, 21).

As a result of these considerations, the extended version of the Hamiltonian (8) including acceleration effects now becomes:

$$H = -r(p^2 - p_r^2)^{\frac{1}{2}} + q'\int rB(r,\theta)dr + q\int rE_\theta dt \ , \qquad (24)$$

with p_r given by:

$$p_r = P_r + q\int E_r dt \ , \qquad (25)$$

as follows from (9, 21).

Almost all of the cyclotrons now in operation have been designed with radial electric gaps, and for such machines, we can take $E_r = 0$ as a good approximation. However, spiral electric gaps form an essential part of the RF systems in the new superconducting cyclotrons and for these machines, E_r becomes quite significant and actually exceeds E_θ in magnitude at large radii.

Considering first the case of radial electric gaps with $E_r = 0$, we then have $p_r = P_r$, and the above Hamiltonian reduces to:

$$H = H^0 + H' \ , \qquad (26)$$

with H^0 given in (15) and,

$$H' = q\int rE_\theta dt \ . \qquad (27)$$

That is, the entire effect of the RF field is completely described by the added term H'.

We should note here that even when $E_r = 0$, there may still be a radial electric force in the median plane. That

is, in addition to E_θ, we may have the RF magnetic field component

$$\delta B_z = -c\int[\frac{1}{r}\frac{\partial}{\partial r}(rE_\theta)]dt \ , \qquad (28)$$

as follows from (22b). This formula implies that a radial electric force will occur whenever the gap voltage, and hence the product rE_θ, changes with radius.

For RF systems (such as those having spiral electric gaps) where E_r differs from zero, we can expand H in powers of this component and hence write:

$$H = H^0 + H' \ , \qquad (29)$$

just as in (26), but now with:

$$H' = q\int(rE_\theta + r'E_r)dt \ , \qquad (30)$$

to first order in E_r. The factor $r' = dr/d\theta$ is to be replaced here by its zero order value given in (17a), and this factor therefore depends on p_r and E as well as r.

Since H^0 is independent of t, the rate of energy gain becomes:

$$\frac{dE}{d\theta} = \frac{\partial H'}{\partial t} = q(rE_\theta + r'E_r) \ , \qquad (31)$$

which is entirely equivalent to $q\vec{E}\cdot d\vec{s}/d\theta$, as it should be.

Indeed, this equation is completely correct if the exact value of r' (rather than the zero order value) is

inserted. Thus, the expression (30) for H' implicitly contains the higher order terms in E_r.

VI. GAP CROSSINGS

As noted at the end of Section III, the effect of a gap crossing is very often approximated by impulsive changes in E and p_r. Given a set of gaps located around the machine at $\theta = \theta_n$ (with n = 1, 2, ...), we therefore assume that the variation of the electric field across each gap can be represented by a delta function, $\delta(\theta - \theta_n)$.

We first consider radial electric gaps for which $E_r = 0$, and then assume as a model that the rate of energy gain is given by:

$$dE/d\theta = \sum_n qV_n(r)\sin(\omega_{rf}t - k_n)\delta(\theta - \theta_n) \ , \qquad (32)$$

where the sum extends over all of the gaps, with the gap at $\theta = \theta_n$ having a voltage of amplitude $V_n(r)$ and phase k_n. For simplicity, the RF angular frequency ω_{rf} is assumed to be the same for all gaps, but this assumption is not essential.

Integrating over θ, we obtain the following energy gain at the n^{th} gap considered by itself:

$$\delta E = qV_n(r)\sin(\omega_{rf}t - k_n) \ , \qquad (33)$$

where the values of r and t here must correspond to the ion's coordinates at $\theta = \theta_n$ as derived, for example, by integrating the differential equations (17) between successive gaps.

The value of H' can be derived by working backwards from eq. (31). In this way we obtain:

$$H' = -(q/\omega_{rf}) \sum_n V_n(r)\cos(\omega_{rf}t - k_n)\delta(\theta - \theta_n) \ . \qquad (34)$$

Because the gap voltage and hence H' depends on r, the ion will experience a radial impulse due to the gap crossing at $\theta = \theta_n$ which, since $p_r = P_r$ here, is given by:

$$\delta p_r = -\int \frac{\partial H'}{\partial r}\, d\theta = \frac{q}{\omega_{rf}} \left(\frac{dV_n}{dr}\right)\cos(\omega_{rf}t - k_n) \ , \qquad (35)$$

and this impulse must be evaluated simultaneously with the energy gain δE found above.

We note in passing that this radial impulse corresponds to the force produced by the RF magnetic field component δB_z given in (28), and this effect (at least for cyclotrons) was first pointed out by Müller and Mahrt,[15] and then more fully explained by Joho.[9] We should also note that the above expressions for δE and δp_r were incorporated into the computer program "Indigo",[16] which was designed to calculate accelerated orbits in the Indiana cyclotron where the gap voltage depends quite strongly on radius.

The RF systems of most cyclotrons contain dees rather than cavities, which for our purpose simply means that the electric gaps occur in pairs with correlated voltages. To treat this situation in more detail, we assume an idealized yet fairly general dee geometry just like the one treated in a previous paper.[8]

Suppose there are N_d identical dees which are numbered $i = 1, 2, \ldots, N_d$ in the order in which the ions traverse them. Suppose further that the ions enter the i^{th} dee at $\theta = \theta_{i1}$ and exit at $\theta = \theta_{i2}$ given by:

$$\theta_{ij} = (i - 1)2\pi/N_d + (-1)^j D/2 + \theta_c \ , \qquad (36)$$

with $j = 1$ or 2. Here, $2\pi/N_d$ is the angular interval between corresponding points on successive dees, and D is the angular width of each dee with $D < 2\pi/N_d$. The value $\theta = \theta_c$ specifies the center line of dee #1, and it should be noted that for spiral electric gaps, θ_c is a function of r.

We now assume that the rate of energy gain for the ion is, by analogy to (32), given by:

$$dE/d\theta = \sum_i \sum_j (-1)^j qV_o(r)\sin(\omega_{rf}t - k_i)\delta(\theta - \theta_{ij}) \ , \qquad (37)$$

where all dee gaps are assumed to have the same voltage amplitude $V_o(r)$. If $h = \omega_{rf}/\omega_o$ is the integral harmonic ratio of the RF frequency to the "ideal" frequency, then the RF voltage phases should satisfy:

$$k_{i+1} = k_i + 2\pi h/N_d \ , \qquad (38)$$

for operation on this harmonic.

Following the same procedure as above, we obtain the following expression for H':

$$H' = -\frac{qV_n(r)}{\omega_{rf}} \sum_i \sum_j (-1)^j \cos(\omega_{rf}t - k_i)\delta(\theta - \theta_{ij}) \ , \qquad (39)$$

where the sum again extends over all values of i and j. Since this H' depends on r as well as t, there will also be a radial impulse δp_r accompanying the energy gain δE at each gap crossing directly analogous to the values given in (33, 35) above.

Turning now to the case of spiral electric gaps, we find that the changes required in the foregoing results are surprisingly small. That is, the form of $dE/d\theta$ is generally assumed to be the same for both spiral and radial gaps, and this assumption will therefore yield the same functions for δE and H' as those found above.

The principal change is in the radial impulse δp_r at the gap crossing which, in accordance with (25), is now given by:

$$\delta p_r = \delta P_r + qE_r\delta t \ , \qquad (40)$$

where δP_r here is the part resulting from the change in gap voltage with radius given, for example, in (35). The term $qE_r\delta t$ is evidently the radial impulse produced by the "slant" of the spiral gaps and the effects of this term have been discussed at length in a previous paper[17] which also includes a description of the "Spiral Gap" program used for orbit computations in superconducting cyclotrons.

VII. HAMILTONIAN EXPANSION

It has become general practice to use an equilibrium orbit code to compute as a function of the ion's energy E all of the required properties of the equilibrium orbit and the associated linear oscillations for a given field $B(r,\theta)$. This information, or relevant parts of it, can be stored for reference purposes in computer programs designed to calculate accelerated orbits in the same field, and it is quite advantageous to do so. We shall therefore assume that such information is readily available.

Suppose that the coordinates r, p_r, and t for an ion moving in the EO (equilibrium orbit) at a given energy E are specified by:

$$r_{eo} = R(E,\theta), \quad (p_r)_{eo} = Q(E,\theta) , \tag{41a}$$

and

$$t_{eo} = (\theta/\omega) + \psi(E,\theta) + u , \tag{41b}$$

where R, Q, and ψ are all periodic functions of θ having the same periodicity as the field. Here, $\omega = 2\pi/\tau$, where $\tau = \tau(E)$ is the ion's rotation period. The function ψ arises from the noncircularity of the EO, while u is simply a constant characterizing the ion's "starting time".

The EO provides an ideal reference orbit at each energy, and we therefore replace the pair (r,p_r) by the radial displacement variables (x,p_x) defined by:

$$x = r - R(E,\theta), \qquad p_x = p_r - Q(E,\theta) , \tag{42}$$

with R and Q given above.

We now assume that the Hamiltonian H in (24) is expanded in powers of x and p_x, and that this expansion is represented schematically by:

$$H = H_0 + H_1 + H_2 + \cdots , \tag{43}$$

where the subscript n indicates the part of H which is of n^{th} order in x and p_x. Since H has been separated into two parts, $H = H^0 + H'$, as in (29), we shall likewise use the same subscript notation for the expansion terms of each part. Furthermore, since we will not be concerned here with nonlinear effects, all terms beyond n = 2 will henceforth be dropped.

Since H^0 is the part of the Hamiltonian describing non-accelerated orbits, the zero order term H_0^0 describes the EO itself, while the term H_2^0 applies to the free linear oscillations about the EO. Moreover, because the EO is an actual orbit, the first order term H_1^0 can be disregarded, as will be shown below. We should also note that all of these terms were first considered for cyclotrons in the analysis of Hagedoorn and Verster.[13]

Since H' is the part of the Hamiltonian describing all of the acceleration effects, the zero order term H'_0 describes

the effect on the EO itself, while the second order term H'_2 concerns changes produced in the free linear oscillations. These changes correspond to the radial electric focusing effects discussed in a previous paper.[8]

The first order term H'_1 changes the free linear oscillations to forced oscillations driven by the acceleration process, and it is this term which is responsible for an actual displacement of the EO from $x = p_x = 0$. This displaced orbit, which is also known as the "accelerated equilibrium orbit", plays an important role in the analysis of accelerated orbits, and its properties have been recognized for a long time. Indeed, a computer program "Disport"[18] for calculating this orbit came into use at this laboratory as early as 1965.

We should note in passing that the expansion of the total Hamiltonian in powers of z and p_z described at the beginning of Section III can be combined with the one indicated by (43) above, and in this combined expansion, the second order term H_2 will be the lowest order term affecting the vertical motion or the coupling of the vertical and longitudinal motion. This is a consequence of median plane symmetry, and the resultant absence of lower order terms in z and p_z provides a qualitative explanation of why a cyclotron beam having an extremely narrow energy width can have such a large ratio of vertical to radial emittance.

VIII. INVARIANT COMMENCEMENT

If a canonical transformation is used to change variables from (r,p_r) to (x,p_x), and if this transformation includes the pair (E,t), as it should, then additional conditions are required to make the transformation unique.

For this purpose, we first specify that the ion's energy E should be retained as a coordinate, while its conjugate momentum changes from $P_e = t$ to a new value $P'_e = u$. If we also require that u reduces to the starting time defined in (41b) when $x = p_x = 0$, then the resultant transformation will be unique.

In order to carry out the transformation, we use a generating function F which is a function of the old coordinates and new momenta (as well as θ), and which has the following form:[10]

$$F = Eu + (r - R)p_x + rQ + g(E,\theta) \ , \tag{44}$$

where $R = R(E,\theta)$ and $Q = Q(E,\theta)$ are the EO coordinates defined in (41a), and where the function $g(E,\theta)$ is to be determined.

To test whether this generating function works properly, we first calculate:

$$x = \partial F/\partial p_x = r - R(E,\theta) \ , \tag{45a}$$

$$p_r = \partial F/\partial r = p_x + Q(E,\theta), \tag{45b}$$

and

$$E = \partial F/\partial u = E \ , \tag{45c}$$

which agrees with our basic definitions of x and $p_{\dot{x}}$ in (42), and which satisfies our requirement that E remain unaltered.

However, there is a fourth transformation equation, namely,

$$t = \frac{\partial F}{\partial E} = u - \frac{\partial R}{\partial E} p_x + \frac{\partial Q}{\partial E} (R + x) + \frac{\partial g}{\partial E} , \qquad (46)$$

where we have set r = R + x. This t is now required to reduce to the t_{eo} in (41b) when $x = p_x = 0$, and this determines the function g. As a result, we finally obtain:

$$g(E,\theta) = \int (\frac{\theta}{\omega} + \psi - R \frac{\partial Q}{\partial E}) dE , \qquad (47)$$

which completes the specification of F in (44).

This g also yields from (46) the following relation between t and u:

$$t = u + \frac{\theta}{\omega} + \psi - \frac{\partial R}{\partial E} p_x + \frac{\partial Q}{\partial E} x , \qquad (48)$$

which now holds true for arbitrary x and p_x. Because of its connection with the starting time defined in (41b), the canonical variable u will henceforth be called the "commencement".

Since the generating function F also depends on θ, there will be a concurrent transformation of the Hamiltonian from H to K given by:

$$K = H + \partial F/\partial \theta , \qquad (49)$$

where this $K = K(x,p_x,E,u;\theta)$ is a function of the new variables. This change actually affects only H^0, the part of H describing nonaccelerated orbits, and since $H = H^0 + H'$, we therefore write:

$$K = K^0 + H' , \qquad (50a)$$

with

$$K^0 = H^0 + \partial F/\partial \theta . \qquad (50b)$$

Furthermore, if K and K^0 are expanded in powers of x and p_x, just like H in (43), we then discover that the zero and first order terms in K^0 vanish; that is,

$$K^0_0 = 0, \text{ and } K^0_1 = 0 . \qquad (51)$$

Thus, in the expansion of K the only zero and first order terms are those resulting from the acceleration process, and we can therefore write this expansion as follows:

$$K = H'_0 + H'_1 + H'_2 + H^0_2 , \qquad (52)$$

correct to second order.

This interesting result could have been anticipated since in terms of the new variables, the equilibrium orbit itself corresponds to $x = p_x = 0$ identically, with E and u both being constants. Moreover, these defining conditions would no longer hold if K^0 contained a zero or first order term.

An important consequence of the expansion (52) is that the commencement u is a constant of the motion for any non-accelerated orbit executing strictly linear oscillations about the EO. That is, when H' = 0, then $du/d\theta = 0$ except for second and higher order terms in x and p_x. This conclusion is true not only for cyclotrons but for any FFAG accelerator.

As a simple example of where this result could be applied, consider the usual 2 x 2 transfer matrix which carries the vector (x,p_x) from one gap crossing to the next at a fixed energy E. Since u is also constant for this orbit segment, the change in the time t can be calculated directly from (48) relating t and u. That is, we need not enlarge the transfer matrix to find time changes between gap crossings.

Before proceeding, we first obtain a simple approximate formula relating t and u by considering circular orbits, or what is nearly the same, by averaging over θ. By either approach, we find: $\psi = 0$, $Q = 0$, and $R = R_0(E)$ independent of θ. Moreover, we can take:

$$\frac{\partial R}{\partial E} = \frac{dR_0}{dE} = \frac{R_0}{v p \nu_r^2} \tag{53}$$

in accordance with the usual approximation for the momentum compaction (or dispersion). Thus, eq. (48) for t now becomes:

$$t = u + (\theta/\omega) - R_0 p_x/v p \nu_r^2 \ , \tag{54}$$

and since $v = \omega R_0$, we then find:

$$\omega t - \theta + p_x/p\nu_r^2 = \omega u = \text{const.} \ , \tag{55}$$

a result previously obtained by more elementary methods.[8]

IX. PROPER PHASE DEFINITION

In order for the ion to be accelerated efficiently, it must have an orbital frequency ω which almost matches a subharmonic h of the RF angular frequency ω_{rf}, that is,

$$\omega \simeq \omega_0 = \omega_{rf}/h \ , \tag{56}$$

where $\omega = \omega(E)$ is defined in (41b), and where the ideal frequency ω_0 was introduced in connection with eq. (38).

Even though the difference between ω_0 and $\omega(E)$ must necessarily remain very small, it is nonetheless quite significant. In order to take this difference into account, we change from u to a new variable u' defined by:

$$u' = u + \Omega(E)\theta/\omega_0 \ , \tag{57}$$

with

$$\Omega(E) = (\omega_0/\omega) - 1 \ , \tag{58}$$

so that $\Omega(E)$ measures the frequency error as a function of E. Thus, when the magnetic field is perfectly isochronous, then $\Omega(E) = 0$, and $u' = u$.

We are now ready to define what we shall call the "proper" phase, $\phi = \phi(\theta)$. To do so, we simply set $\phi = \omega_{rf}u'$, and combining (57) above with eq. (48) for t, we then have

$$\phi(\theta) = \omega_{rf}t - h\theta - \omega_{rf}\left[\psi - \frac{\partial R}{\partial E}p_x + \frac{\partial Q}{\partial E}x\right] . \qquad (59)$$

This definition, though more complicated than the usual one, has certain advantageous properties.

First, we recognize that for an isochronous field this phase reduces to $\phi = \omega_{rf}u$ which is therefore a constant of the motion for any monoenergetic group of ions executing strictly linear oscillations about the EO. As discussed at some length in a previous paper,[8] this feature of the phase definition is particularly important for cyclotrons.

We also recognize that the bracketed quantity in (59) averages to zero in the long run, and we can therefore write:

$$\bar{\phi} = \langle\omega_{rf}t - h\theta\rangle , \qquad (60)$$

which, aside from some additive constant, is the customary definition for ϕ used in cyclotron work.

Returning again to the approximate formulas discussed at the end of Section VIII, and introducing the same approximations into (59), we then obtain:

$$\phi(\theta) \simeq \omega_{rf}t - h\theta + hp_x/pv_r^2 , \qquad (61)$$

which is the same as the phase definition derived in the previous paper referred to above.

Strictly speaking, the foregoing proper phase definitions apply only to radial gaps and require some modification

for spiral gaps, or any other nonradial gaps. To see this we need only recall the rule of thumb which says that the ions which cross the center line of a dee just when the voltage reverses sign will receive the maximum energy gain and should therefore be assigned the phase $\phi = 0$.

Suppose that $\theta = \theta_c(r)$ defines the curve for the center line of dee #1 as noted, for example, after eq. (36). Suppose further that this curve intersects the EO radius $R(E,\theta)$ at the particular angle $\theta_0 = \theta_0(E)$ given by:

$$\theta_0 = \theta_c(R(E,\theta_0)) . \qquad (62)$$

This θ_0 now provides at each energy value the appropriate reference angle for defining the phase $\phi(\theta)$.

We first return to the eq. (41b) for t_{eo} and change it to the following:

$$t_{eo} = (\theta'/\omega) + \psi'(E,\theta) + u , \qquad (63)$$

where u is now the starting time not at $\theta = 0$, but rather at $\theta = \theta_0(E)$. That is, we define θ' and ψ' as follows:

$$\theta' = \theta - \theta_0 , \qquad (64a)$$

and

$$\psi' = \psi(E,\theta) - \psi(E,\theta_0) , \qquad (64b)$$

so that $\theta' = 0$ and $\psi' = 0$ when $\theta = \theta_0$.

With regard to the canonical transformation in Section VIII and the generating function F in (44), the only change required is in the function $g(E,\theta)$ and in this function, we need only change θ to θ' and ψ to ψ' within the integrand in (47). As a result, the same two changes will also occur in the relation (48) between t and u. However, these relatively minor changes will not alter the important conclusion that the commencement u is an invariant to first order for nonaccelerated orbits.

Returning now to the discussion leading up to the proper phase definition, we first replace θ by θ' in the relation (57) to obtain:

$$u' = u + \Omega(E)\theta'/\omega_0 \ , \tag{65}$$

so that $u' = u$ when $\theta = \theta_0$. Thus, the proper phase definition (59) now becomes:

$$\phi(\theta) = \omega_{rf}t - h\theta' - \omega_{rf}[\psi' - \frac{\partial R}{\partial E} p_x + \frac{\partial Q}{\partial E} x] \ , \tag{66}$$

and these changes serve to generalize the definition without altering the advantages described above.

We should also note that this generalization to spiral gaps can be extended still further by removing the restriction imposed on the definition of $\theta_0(E)$ in (62). That is, given the geometry of the rf gaps and given the amplitudes and phases of the gap voltages, we can determine $\theta_0(E)$ simply by considering an ion moving in the EO and requiring that

its energy gain per turn be proportional to cos ϕ, as will be described more fully below.

X. LONGITUDINAL MOTION

The change of variables from u to u' in (57) or (65) can be brought about by a canonical transformation which retains E as the coordinate and also leaves x and p_x unaltered. There is, however, a concurrent change in the zero-order term of the Hamiltonian K in (52), and the revised form of this term now becomes:

$$K_0 = H_0' - \int \Omega(E)dE/\omega_0 \ , \tag{67}$$

where $\Omega(E)$ is the frequency error defined in (58). The revised form of the Hamiltonian is therefore given by:

$$K = K_0 + H_1' + H_2' + H_2^0 \ , \tag{68}$$

correct to second order in x and p_x as in (52).

We can just as easily change variables from u' to $\phi = \omega_{rf}u'$ by a "scale" transformation. This simply requires changing the Hamiltonian from K to $\omega_{rf}K$.

The K_0 in (68) describes the longitudinal motion itself since it concerns variations of the energy and phase assuming the ion moves continuously in an EO. As noted before, the term H_2^0 deals with the free linear oscillations so that the other terms, $H_1' + H_2'$, are responsible for the coupling of these oscillations with the longitudinal motion.

We restrict our attention for the present to the longitudinal motion by itself and assume a quite general form for H' rather than one of the more specialized forms given in Section VI. Since $x = p_x = 0$ for now, we set $r = R(E,\theta)$, $p_r = Q(E,\theta)$, and:

$$\omega_{rf}t = h\theta' + \phi + \omega_{rf}\psi'(E,\theta) \ , \tag{69}$$

as follows from (66).

With these changes, $H' \rightarrow H_0'(E,\phi,\theta)$, and we then have from (67);

$$K_0 = H_0'(E,\phi,\theta) - \int\Omega(E)dE/\omega_0 \ , \tag{70}$$

which is therefore the pertinent Hamiltonian. Note that for fixed values of E and ϕ, this K_0 is periodic in θ.

Making use of the scale transformation noted above, the zero-order equations of motion are now given by:

$$\frac{dE}{d\theta} = \omega_{rf}\frac{\partial K_0}{\partial\phi} = \omega_{rf}\frac{\partial H_0'}{\partial\phi} \ , \tag{71a}$$

$$\frac{d\phi}{d\theta} = -\omega_{rf}\frac{\partial K_0}{\partial E} = h\Omega(E) - \omega_{rf}\frac{\partial H_0'}{\partial E} \ . \tag{71b}$$

These equations show incidentally that the change of variables from u' to ϕ is indeed brought about by a simple scale transformation.

The variation of E and ϕ with θ can be divided into two parts, a rapidly oscillating part, and a slowly varying

"secular" part. Generally speaking, the secular part is much more important and is very often the only part considered. Following this custom, we eliminate the rapidly oscillating part by an averaging process.

Rather than averaging H_0', we use instead the quite general equation (31) and thereby obtain:

$$\langle dE/d\theta\rangle = (q/2\pi)\S(rE_\theta + r'E_r)d\theta = (q/2\pi)\S\vec{E}\cdot d\vec{s} \ , \tag{72}$$

where the integral is taken around the closed EO. The result of the integration can depend only on ϕ and E, and following convention, we take:

$$\S\vec{E}\cdot d\vec{s} = V(E)\cos\phi \ , \tag{73}$$

which serves quite generally to determine $\theta_0(E)$ as noted at the end of Section IX.

This result also defines the peak voltage gain per turn, V(E). This parameter will depend on E if the gap voltage depends on radius or if the shape of the EO relative to the gap locations changes with energy. For the idealized but fairly general dee geometry used at the end of Section VI, the result is:

$$V(E) = 2N_dV_0(E)\left|\sin(hD/2)\right| \ , \tag{74}$$

where $V_0(r)$ has been changed to $V_0(E)$ through the process indicated in (73).

After averaging eq. (71a) and making use of the foregoing results, we can then integrate with respect to ϕ and thereby obtain:

$$\omega_{rf}\langle H_0'\rangle = (q/2\pi)V(E)\sin\phi \ . \tag{75}$$

We now define the secular part of the (zero-order) Hamiltonian as follows:

$$\overline{K}_0 = \omega_{rf}\langle K_0\rangle = \omega_{rf}\langle H_0'\rangle - h\int\Omega(E)dE \ , \tag{76}$$

where, as described above, the factor ω_{rf} has been inserted to make E and $P_e = \phi$ the new canonical variables.

Combining these results, we finally obtain:

$$\overline{K}_0 = (q/2\pi)V(E)\sin\phi - h\int\Omega(E)dE \ . \tag{77}$$

This secular Hamiltonian is essentially the same as the Hamiltonian found by Joho.[9] Actually, Joho considered rf systems including harmonics of the base frequency ω_{rf}, such as those used for "flat-topping" of the voltage wave form. Our analysis could, however, be readily expanded to include such cases.

An important consequence of the analysis is that $\overline{K}_0$ is a constant of the motion (at least approximately). Instead of (77), this result is usually expressed as an equation giving ϕ as a function of E:

$$\sin\phi = \frac{V(E_i)}{V(E)}\sin\phi_i + \frac{2\pi h}{qV(E)}\int\Omega(E)dE \ , \tag{78}$$

where the integration extends from some initial energy E_i where $\phi = \phi_i$ to the variable energy E. This simple equation forms an essential part of computer programs designed to calculate operational values for trim coil currents. It is therefore important to recognize that ϕ here should more properly be identified with $\overline{\phi}$, its average value, as in (60).

For most cyclotrons, the variation of V(E) turns out to be quite small, so that the phase compression effect is practically unnoticeable. This effect is, however, quite significant in the Indiana cyclotron, as noted before, and is also appreciable in the large SIN cyclotron.

We should emphasize that the results of this section apply equally well to RF systems with spiral electric gaps (or other nonradial geometries). In this connection, there is an obvious but interesting difference between the two effects noted in Section VI which produce a radial impulse at the gap crossings. The one resulting from a radial dependence of the gap voltage leads to a phase compression effect, while the one resulting from spiral electric gaps produces zero net phase shift (at least on the average). The latter result emerges automatically from the canonical treatment, while the more elementary derivation requires considerable analysis.[17,19]

REFERENCES

1. E.D. Courant and H.S. Snyder, Ann. of Phys. 3, 1 (1958). See Appendix B.

2. A.A. Kolomensky and A.N. Lebedev, Theory of Cyclic Accelerators (J. Wiley, New York, 1966). See Section 2.5.

3. F.E. Mills, Nucl. Instrum. Methods 23, 197 (1963).

4. See, however, C.J.A. Corsten, thesis, Technische Hogeschool Eindhoven, 1982.

5. K.R. Symon and A.M. Sessler, High Energy Accelerators and Pion Physics I (CERN, Geneva, 1956), pp. 44-58.

6. A.J. Lichtenberg, Phase Space Dynamics of Particles (J. Wiley, New York, 1969). See Section 4.2.

7. W.M. Schulte and H.L. Hagedoorn, Nucl. Instrum. Methods 171, 409 (1980).

8. M.M. Gordon and F. Marti, Particle Accelerators 12, 13 (1982).

9. W. Joho, Particle Accelerators 6, 41 (1974).

10. H. Goldstein, Classical Mechanics, 2nd. ed. (Addison-Wesley, Reading, 1980). See Sections 7-9, 8-5, 9-2, and 11-2.

11. J.D. Jackson, Classical Electrodynamics, 2nd. ed. (J. Wiley, New York, 1975). See Sections 12.1 and 6.5.

12. B. Schnizer, CERN Report No. 70-77, ISR Div. (March 1970).

13. H.L. Hagedoorn and N.F. Verster, Nucl. Instrum. Methods 18, 19, 201 (1962).

14. M.M. Gordon and T.A. Welton, ORNL Report No. 2765 (1959).

15. R.W. Müller and R. Mahrt, Nucl. Instrum. Methods 86, 241 (1970).

16. M.M. Gordon, Bull. APS 18, 1422 (1973).

17. M.M. Gordon, Nucl. Instrum. Methods 169, 327 (1980).

18. M.M. Gordon, IEEE Trans. Nucl. Sci. 13, 48 (1966).

19. S. Adam, Proceedings Ninth International Conf. on Cyclotrons. Les Editions de Physique, Les Ulis, France, 1981, p. 431.