SOME USEFUL INVARIANTS AND A TRANSFER MATRIX FOR THE LONGITUDINAL MOTION*

M.M. Gordon

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

Abstract

Using a novel approach, the invariance of phase with changes in electric gap orientation is demonstrated again. This is followed by a simplified deduction of an invariant Hamiltonian (called here the SSJ invariant) along with the resultant equations of motion for the energy and phase. This then leads to another invariant governing the phase width of a group of ions having the same energy. In the second half of the paper, the SSJ invariant is used to construct analytically a complete transfer matrix for calculating changes in the distribution of energy and phase values within the beam between injection and extraction. The results are used to discuss the general conditions for minimizing the energy spread and/or phase width of the extracted beam. The transfer matrix is then extended to include important second order effects. Further applications of the results are summarized in some concluding notes.

1. Introduction

Our discussion here will be concerned mainly with the longitudinal motion of the ions in a cyclotron, and will consequently be confined to certain general features of the acceleration process. This work is aimed at bringing together in one place several different aspects of the subject, most of which have been treated before in other ways. Our treatment therefore represents an extension and reinterpretation of previous developments along with some new results.

The first half of this paper presents a straightforward derivation or justification of four useful invariants which govern the general changes in the energy and phase values during the acceleration process. These results are then used in the second half to develop a transfer matrix for analysing the evolution of the energy and phase values within the beam between injection and extraction.

This development grew out of a conversation with Pierre Lapostolle at the particle accelerator conference in Santa Fe last year. He mentioned at that time results of a calculation which showed that one could achieve a factor of two in phase compression by injecting the beam with an average phase of 60° and then bringing this value down to zero by a suitable alteration of the magnetic field.[1] My first thought was that this phase compression was simply the consequence of an old invariance relation (the one discussed at the end of Sec. 3 below), but I later decided that for a cyclotron with an injected beam, a proper treatment required the calculation of a complete transfer matrix.

As far as the longitudinal motion is concerned, there is a widespread belief that the best possible operating condition for a cyclotron is one having a perfectly isochronous field with a matched rf frequency. This is, however, not generally true because, as we shall see, the nature of a truly optimized cyclotron system depends very much on the distribution of energy and phase values within the injected beam, and on whether it is desired to minimize the phase width or energy spread of the extracted beam, or perhaps some combination of both. We should also recognize, of course, that with this optimization will be hampered when vertical focusing and beam extraction requirements severely limit the choice of trim coil currents and other parameters.

2. Invariance of Phase with Changes of Electric Gap Orientation

Let us consider first a special invariance principle for cyclotrons which states that the ion's phase history does not depend on the orientation of the electric gaps provided the voltage gain per turn is fixed. Our purpose here is to bury once and for all a misconception which has arisen many times during the history of cyclotrons, namely, that the ion's phase will be shifted if its orbit is not perpendicular to the electric gaps.

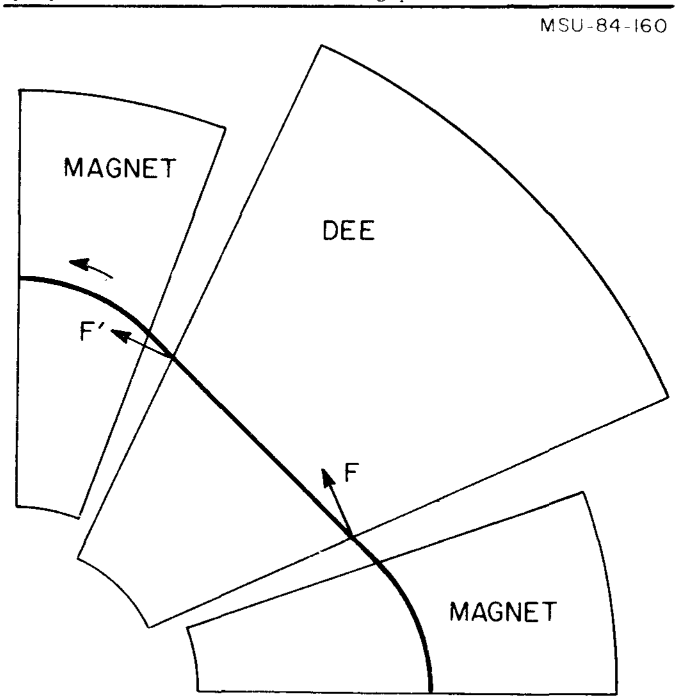


Fig. 1. Schematic diagram of one quadrant of the Indiana Cyclotron showing one dee with radial gaps 40° apart which lies in a valley between two magnet sectors. An ion moving in the orbit shown here experiences the force F as it enters the dee, and a different force F' as it exits. The resultant of these forces has a net transverse component when the ion's phase ϕ differs from zero.

*Work supported by the National Science Foundation under Grant No. PHY-83-12245.

As an example, consider the orbit shown in Fig. 1 which represents schematically the situation in one sector of the Indiana cyclotron. The dee shown here is assumed to be in a field free region and the ion's orbit therefore crosses the dee in a straight line. Since the electric gaps are radial, the ion experiences first a force $\vec{F}$ which has an outward component as it enters the dee, and then a force $\vec{F}'$ which has an inward component as it exits. If the ion's phase differs from zero, these components will be unbalanced, and the ion will therefore receive a net transverse impulse on passing through each dee.

As another example, consider the orbit shown in Fig. 2 which represents the gap geometry in our K500 superconducting cyclotron where the three 60° dees are situated in the valleys so that the electric gaps have a spiral shape which matches that of the pole tips. The orbit here is practically circular and when the ion crosses one of the gaps, it experiences a force $\vec{F}$ which evidently has a large outward component.

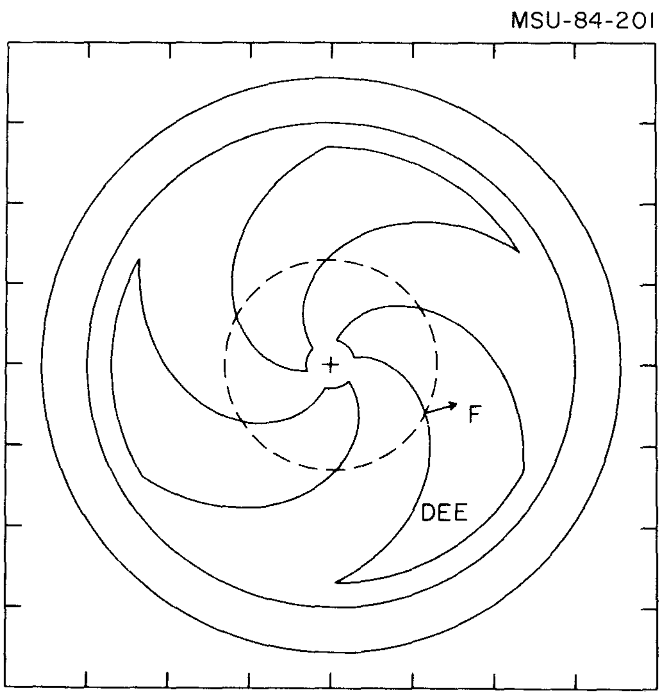


Fig. 2. Schematic diagram of the NSCL K500 superconducting cyclotron showing the three dees with spiral electric gaps 60° apart. An ion moving in the circular orbit shown here experiences a force F having a large outward component as it enters or exits each dee. The two outer circles indicate the boundaries of the superconducting coils.

In all such cases, a careful analysis shows that there are two different effects to be considered, each of which by itself causes a phase shift.[2,3] Instead of discussing these effects again, we merely point out that they cancel each other so that the phase remains invariant. This cancellation is most remarkable since it does not depend on the magnetic field shape and occurs whether or not the field is isochronous. Such a result strongly suggests that the transverse force components are quite irrelevant, and to make this more obvious, we present here an alternative view of the phenomenon.

Consider first a simple model which I used some time ago to help explain the longitudinal space charge effect,[4] and which, by coincidence, has recently been used for this purpose by Chasman and Baltz.[5] Let us assume a uniform magnetic field $\vec{B}$ and non-relativistic conditions so that the ions move in circular orbits with angular velocity $\omega = qB/m$ at all energies. The left side of Fig. 3 shows two such orbits as they would appear in the laboratory where they rotate counterclockwise. Both orbits have the same radius R and one is centered at the origin while the other is off-center by an amount $\underline{a}$.

The same two orbits are shown on the right as they would appear in a reference frame rotating counterclockwise with the angular velocity ω. The centered orbit now becomes a fixed point at a distance R from the origin, while the off-center orbit now appears as a circle of radius $\underline{a}$ in which the ion again rotates with angular velocity ω but now in the clockwise direction.

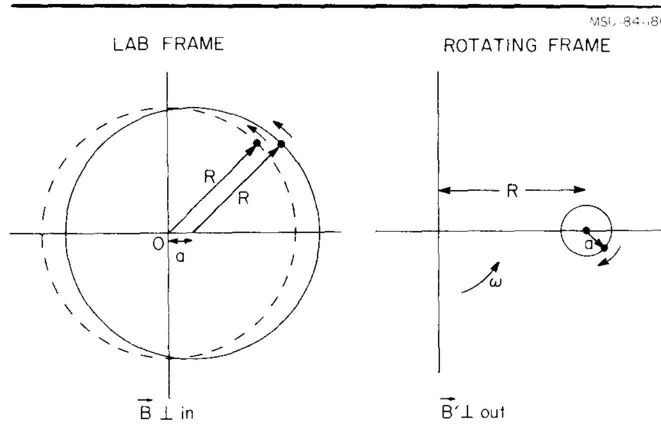


Fig. 3. In the lab frame situation shown on the left, there are two ions rotating counterclockwise with angular velocity $\omega = qB/m$ in a uniform field B which is perpendicular into the plane. The ions move in circular orbits of radius R, one centered at O and the other off-center by a distance $\underline{a}$. The same two orbits are shown on the right as they would appear in a reference frame rotating counterclockwise with the same ω. Because of the additional Coriolis and centrifugal forces, the ions now seem to move in a field $B' = -B$ which is perpendicular out of the plane. In this frame, the ions are at rest on the average, with $r' = R$ and $\theta' = $ const.

We may therefore conclude that in the rotating frame, the ions appear to move in a uniform field $\vec{B}'$ having the same magnitude as the laboratory field, but with the direction reversed, i.e., $\vec{B}' = -\vec{B}$. This reversal actually comes about because the force acting on the ion in the rotating frame is the resultant of three forces, namely, the magnetic force, the Coriolis force, and the centrifugal force. Thus, if $\vec{F} = q\vec{v} \times \vec{B}$ is the force in the lab frame, then we find that in the rotating frame, the resultant force is $\vec{F}' = q\vec{v}' \times \vec{B}'$, with $\vec{B}' = -\vec{B}$.

Referring again to the right side of Fig. 3, we find that the off-center orbit corresponds to a radial oscillation with amplitude $\Delta r = a$ combined with an angular oscillation of amplitude $\Delta\theta = a/R$. Thus, this picture also provides a simple model for the coupling of the radial and longitudinal motion as it was analysed by Schulte and Hagedoorn.[6] For the present, we assume that a/R is negligibly small so that r and θ can be replaced by their average values:

$$r = R, \qquad \theta = \omega t + \text{const.}, \tag{1a}$$

corresponding to a centered orbit. Now since $r' = r$ and $\theta' = \theta - \omega t$ in the rotating frame, we therefore find:

$$r' = R, \qquad \theta' = \text{const.} \tag{1b}$$

As might be anticipated, the constant value of θ' is related to the acceleration phase ϕ for these orbits.

In order to consider the effect of the rf electric field, we start with this field $\vec{E} = \vec{E}(r,\theta,t)$ as it appears in the laboratory. If ω_{rf} is the rf angular frequency, then $\vec{E}$ can be written as:

$$\vec{E} = \vec{E}_c(r,\theta)\cos\omega_{rf}t + \vec{E}_s(r,\theta)\sin\omega_{rf}t. \tag{2}$$

Moreover, the coefficients $\vec{E}_c$ and $\vec{E}_s$ are periodic in θ and can therefore be expressed as Fourier series. When this is carried out, the final result can be written:

$$\vec{E} = \sum_n \vec{E}_n(r) \cos\left[\omega_{rf}t - n\theta - \kappa_n(r)\right], \tag{3}$$

where the sum extends over $n = \pm1, \pm2$, etc., but with certain n values ruled out by symmetry.

Here, $\vec{E}_n(r)$ is the vector amplitude of the n^{th} Fourier harmonic and $\kappa_n(r)$ is its phase angle. For radial electric gaps, the κ_n are all constants. For spiral gaps specifically, we assume that: $\kappa_n(r) = n\,\zeta(r) + \text{const.}$, which corresponds to the situation shown in Fig. 2 where all of the gaps follow the same spiral curve, namely, $\theta + \zeta(r) = \text{const.}$ For either radial or spiral gaps, we therefore write:

$$\vec{E} = \sum_n \vec{E}_n(r) \cos\left[\omega_{rf}t - n(\theta + \zeta(r)) - \kappa_n\right], \tag{4}$$

with κ_n now being constant in either case.

For $n > 0$, the terms in the above sum correspond to waves traveling in the $+\theta$ direction, while for $n<0$, they travel in the $-\theta$ direction, and in either case their angular velocity is ω_{rf}/n. Now if $h = \omega_{rf}/\omega$ is the integral harmonic number, then in this case, the wave with $n = +h$ will have the same angular velocity and direction as the ions. It therefore follows that this term in the above series is the only one with real significance since it corresponds to an electric wave which rotates in perfect synchronism with the ions. That is, it is this term and only this term which produces on the average a net energy gain per turn.

With this in mind, we assume all of the terms can be discarded except the one with $n = +h$. The electric field therefore reduces to:

$$\vec{E} = \vec{E}_h(r) \cos\left[\omega_{rf}t - h(\theta + \zeta(r)) - \kappa_h\right], \tag{5}$$

and we can now recognize that the quantity in the brackets is the acceleration phase ϕ for the ion as it is usually defined, i.e.,

$$\phi = \omega_{rf}t - h(\theta + \zeta(r)) - \kappa_h. \tag{6}$$

In order to transform this field into the rotating frame, we next set $r = r'$ and $\theta = \omega t + \theta'$. In addition, we set $\omega_{rf} = h\omega$, and therefore obtain:

$$\vec{E}' = \vec{E}_h(r') \cos\left[h(\theta' + \zeta(r')) + \kappa_h\right], \tag{7}$$

and we note that the time dependence drops out. Thus, the resultant electric field $\vec{E}'$ in the rotating frame appears as a quasi-static field to the ions considered above, which, in the absence of this field, are all at rest on the average.

A polar map of this quasi-static field $\vec{E}'$ is shown at the top of Fig. 4 for the case of radial gaps, and a similar map for spiral gaps is shown on the bottom. For simplicity, only maps for h=1 are given, and it should be kept in mind that these maps provide only a qualitative picture of the field. For both kinds of gaps, the curves $\phi = \text{const.}$ are shown radiating out from the center and as can be seen, these curves have exactly the same form as those for the gaps, i.e., $\theta + \zeta(r) = \text{const.}$ Consequently, the arrows representing the field $\vec{E}'$ are drawn perpendicular to these curves with their lengths proportional to $\cos\phi$. For simplicity again, it is assumed that $|\vec{E}'_h|$ is independent of r', which is not realistic. For example, if the voltage gain per turn V is constant , then $|\vec{E}'_h| = V/2\pi r'$, in the case of radial gaps.

Let us return again to the motion of the ions in the rotating frame, and note that in addition to the electric field $\vec{E}'$ shown here, there is a uniform magnetic field $\vec{B}' = -\vec{B}$, which is perpendicular outward from the plane of this figure. Now as you may recall, for such a combination of mutually orthogonal electric and magnetic fields, the motion of an ion is, at least on the average, in a direction perpendicular to both with a "drift velocity" given by:

$$\vec{v}_d = (\vec{E}' \times \vec{B}')/B^2. \tag{8}$$

By using the usual right hand rule for vector products, we find that this velocity vector is always directed along the gap line following one of the curves with $\phi = \text{const.}$ shown in Fig. 4. In addition, the direction of this motion will be radially outward or inward depending on whether $\cos\phi$ is positive or negative, which is completely in accord with the energy gain.

We have implicity assumed here that adiabatic conditions prevail which excludes the central region, and under these conditions, the rf electric field can be treated as a small perturbation acting on the orbits. In addition, we have neglected the possible effect of the rf magnetic field which is treated later. Furthermore, we have considered only the time average or secular motion of the ions, and we should note that the actual motion will show small oscillations or fluctuations about this average.

To summarize now, we find that this simple model of the ion's motion shows that the orbits remain isochronous, and hence that the phase ϕ is invariant, whether or not the orbits cross the gaps normal to the gap line. We should also note, however, that the transverse component of the electric force does indeed modify the orbits, and may even influence the radial focusing.[7] But these effects are generally small outside the central region, and what is important to

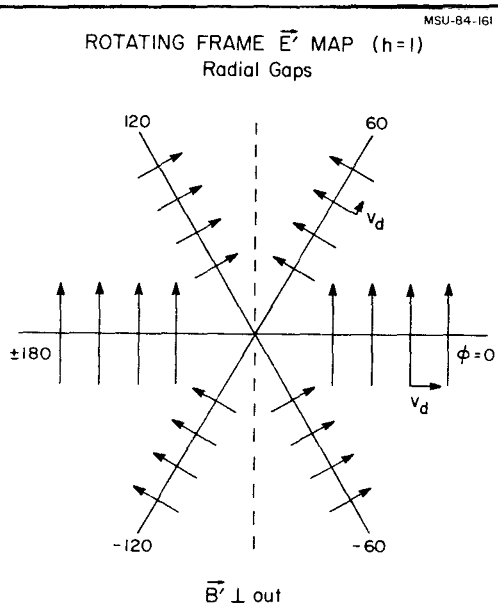
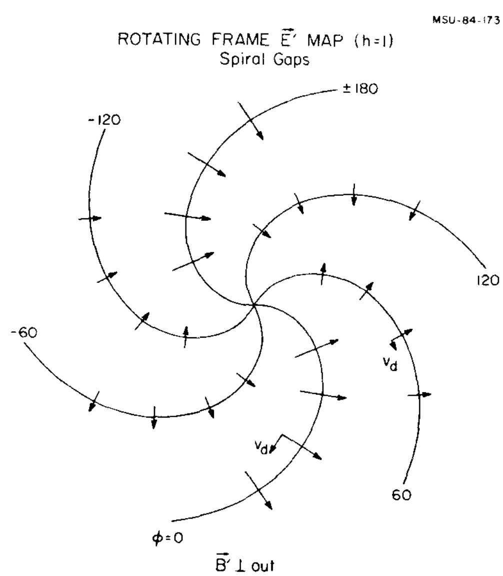


Fig. 4. Two polar maps of the quasi-static component of the rf electric field $\vec{E}$ as seen in the rotating frame used in Fig. 3. Only maps for harmonic h=1 are shown, and the one on top is for radial gaps like those in Fig. 1, while the map at bottom is for spiral gaps like those in Fig. 2. The curves for ϕ = const. are shown radiating outward, and the geometry of these curves matches that for the gap lines. Arrows representing $\vec{E}'$ are drawn perpendicular to each of these curves and with lengths proportional to $\cos\phi$. For such mutually perpendicular $\vec{E}'$ and $\vec{B}'$ fields, the motion of the ion is, on the average, normal to both of them with a drift velocity $\vec{v}_d = (\vec{E}' \times \vec{B}')/B^2$, as indicated by arrows marked v_d. Because of this drifting, each ion tends to follow the same curve ϕ = const. throughout its motion, and its average phase ϕ is therefore invariant.



us here, they do not influence the phase, at least to first order.

3. The SSJ Invariant

Let us consider now another approach to the longitudinal motion which is similar to the one pioneered by Symon and Sessler for synchrotrons,[8] and which leads to a result found by Joho for cyclotrons.[9] Within the limits of certain approximations, this approach is quite general and results in the SSJ invariant, named after Symon, Sessler, and Joho.

To start, we divide the motion into two parts, namely, longitudinal motion in the closed equilibrium orbit (EO) characterized by the coordinates E and ϕ, and the transverse motion about this orbit which is described by the (x,p_x) and (z,p_z) coordinates. Thus, we divide the Hamiltonian into three terms,

$$H = H_{long} + H_{osc} + H_{coup}, \qquad (9)$$

where H_{long} describes the motion in the EO's including acceleration effects, H_{osc} describes the oscillations about these orbits, while the third term H_{coup} contains all of the coupling effects. This picture is somewhat analogous to the model in nuclear physics where the Hamiltonian is again divided into three terms, one describing the rotational and vibrational motion of the nucleus as a whole, a second one describing the motion of individual nucleons in shell model orbits, and a third one describing the coupling of these motions.

We consider here only the longitudinal motion and focus our attention on the secular part of this motion, that is, the part that describes the slow systematic changes in E and ϕ. To eliminate the fluctuations in these variables, we average the motion over θ (which also removes the coupling effects to first order). The result of this averaging on H_{long} can be written as follows:

$$\langle H_{long} \rangle = K(E,\phi), \qquad (10)$$

where $K = K(E,\phi)$ is therefore the Hamiltonian for the secular variations in E and ϕ, and is independent of θ because of the averaging. Hence, the corresponding equations of motion are

$$\frac{dE}{dn} = \frac{\partial K}{\partial \phi}, \qquad \frac{d\phi}{dn} = -\frac{\partial K}{\partial E}, \qquad (11)$$

where we have chosen E to be the coordinate with $p_e = \phi$ its conjugate momentum, and where $n = \theta/2\pi$ is the turn number. (For simplicity, a factor $2\pi\omega_{rf}$ was omitted from Eq. (10).)

Now since K is independent of θ (or n), it is therefore a constant of the motion. That is, this K is an approximate invariant for the longitudinal motion, and it is this quantity which we have named the SSJ invariant.

Assuming again that adiabatic conditions prevail, we can equate the energy gain per turn to the work done by the rf electric field on an ion moving around the closed EO, that is,

$$\frac{dE}{dn} = \oint q\vec{E}\cdot d\vec{s}, \qquad (12)$$

where the electric field $\vec{E} = \vec{E}(r,\theta,t)$ here is evaluated along the closed EO, and it is clear that only the tangential component is involved. In accordance with (6), we also set

$$\omega_{rf}t = \phi + h(\theta + \zeta(r)) + \kappa_h, \qquad (13)$$

when evaluating this integral, and the result will then be proportional to cos ϕ. Thus, we can write

$$\frac{dE}{dn} = qV(E)\cos\phi, \qquad (14)$$

where V(E), the peak voltage gain per turn, is a function of E whenever the gap voltage depends on radius, or when the shape of the EO relative to the gap geometry changes with E.

Combining this result with (11) above, we have

$$\frac{dE}{dn} = \frac{\partial K}{\partial \phi} = qV(E)\cos\phi, \qquad (15)$$

and after integrating with respect to ϕ, the resultant expression for K is

$$K = qV(E)\sin\phi + K_o(E), \qquad (16)$$

where $K_o(E)$ is the value of K when there is no acceleration. In this case, we know that the average phase slip per turn is given by

$$\left(\frac{d\phi}{dn}\right)_o = h(\omega_o\tau(E) - 2\pi), \qquad (17a)$$

where $\omega_{rf} = h\omega_o$ here, and where $\tau(E)$ is the orbital period in the EO. When this is combined with (11) above, we obtain

$$\left(\frac{d\phi}{dn}\right)_o = -\frac{\partial K_o}{\partial E} = h(\omega_o\tau(E) - 2\pi), \qquad (17b)$$

and after integrating with respect to E, the result is

$$K_o = -hF(E), \qquad (18)$$

with

$$F(E) = \int (\omega_o\tau(E) - 2\pi)dE, \qquad (19)$$

so that F(E) measures the deviation of the magnetic field from isochronism.

Combining (18) and (16), we finally obtain

$$K = qV(E)\sin\phi - hF(E), \qquad (20)$$

and in addition to (15) for dE/dn, we now have

$$\frac{d\phi}{dn} = -\frac{\partial K}{\partial E} = h(\omega_o\tau(E) - 2\pi) - q\left(\frac{dV}{dE}\right)\sin\phi. \qquad (21)$$

Thus, when acceleration is present, it contributes to the average phase slip per turn only through the variation of V with energy. As shown by Joho,[9] this variation is produced by the effect of the rf magnetic field during the gap crossings, an effect first pointed out by Mueller and Mahrt.[10]

The final equations of motion (15) and (21), together with the invariant K in (20), provide the most basic information about the longitudinal motion in cyclotrons. For example, the invariance of K leads directly to the standard phase-energy relation which determines ϕ vs. E at any energy in terms of its value ϕ_i at the initial energy E_i. Thus, we have

$$qV(E)\sin\phi(E) = qV_i\sin\phi_i + h[F(E) - F_i], \qquad (22)$$

where the subscript i indicates an evaluation at $E=E_i$.

We are assuming, of course, that V(E) and F(E) are known functions of the energy. It should be noted, however, that values of F(E) are part of the standard output from either the Equilibrium Orbit Code or the Cyclops program.[11]

There is an auxilliary invariance relation which follows directly from the above equation. Consider two particles, a and b, and let $\phi_a(E)$ and $\phi_b(E)$ be their respective phase values at a given E. It then follows from (22) that the difference

$$V(E)[\sin\phi_b(E) - \sin\phi_a(E)] = \text{const.}, \qquad (23)$$

that is, it is independent of E.

Suppose now that $\phi_a < \phi < \phi_b$ for all the ions in the beam at each E, and let

$$\bar{\phi} = \frac{1}{2}(\phi_a + \phi_b), \text{ and } \Delta\phi = \phi_b - \phi_a, \qquad (24)$$

so the $\bar{\phi}$ is the average value and $\Delta\phi$ is the width of the phase distribution. If the corresponding values of ϕ_a and ϕ_b are inserted in (23), we then find

$$V(E)\cos\bar{\phi}\,\sin\tfrac{1}{2}\Delta\phi = \text{invariant}, \qquad (25)$$

which shows that the minimum width occurs where $V(E)\cos\bar{\phi}$ is a maximum, and vice versa.

Suppose now we compare values at the initial energy E_i and the final energy E_f, using subscripts i and f to indicate corresponding values for other quantities. We then find

$$\sin\tfrac{1}{2}\Delta\phi_f = \frac{V_i\cos\bar{\phi}_i}{V_f\cos\bar{\phi}_f}\sin\tfrac{1}{2}\Delta\phi_i, \qquad (26)$$

which shows how the final width $\Delta\phi_f$ is related to the initial value $\Delta\phi_i$. Although this phase width relation is an exact one, it requires $\Delta E_f=0$ as well as $\Delta E_i=0$, and it is therefore most useful for those cyclotrons which operate without separated turns since in these machines, the ions are extracted solely on the basis of their final energy (or radius). A very good (but untypical) example is the TRIUMF cyclotron in which H^- ions are accelerated and the resultant proton beam is extracted with a stripping foil.

The above equation for $\Delta\phi_f$ will not apply in general to those cyclotrons which start with a narrow phase width $\Delta\phi_i$ (obtained either by using internal phase slits or by injecting a bunched beam) and thereby achieve single turn extraction. For such machines, it seems much more appropriate to use a transfer matrix technique like the one discussed below. As we shall show later (in Sec. 6), the above equation for $\Delta\phi_f$ represents a special case of a more general relationship.

4. Transfer Matrix

The number of cyclotrons which use an injected beam, rather than an internal ion source, has grown rapidly in recent years. In most cases, the injected beam is bunched beforehand and therefore starts out occupying a fairly well defined (six dimensional) volume in phase space.

According to the Liouville-Poincare theorem, this volume will remain invariant with respect to changes in θ (assuming that space-charge effects are sufficiently small). If we again assume adiabatic conditions so that coupling effects can be neglected, then the longitudinal part of the phase-space volume will also remain invariant. For our purposes here, this means that the area $\iint dEd\phi$ occupied by the ions will be independent of n.

Under these conditions, we can also use a transfer matrix to describe the change in the values of (E,ϕ) for each ion between injection and extraction. To do so, we must first specify the coordinates of a suitably chosen central ray to serve as a reference from which the deviations of all other trajectories can be measured. If we let (E,ϕ) vs. n be the coordinates of the central ray, and let $(E+\delta E,\phi+\delta\phi)$ vs. n be those for any other trajectory, then we can write

$$\begin{bmatrix} \delta E_f \\ \delta\phi_f \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} \delta E_i \\ \delta\phi_i \end{bmatrix}, \qquad (27)$$

where subscripts i and f denote the initial and final values and where the R_{jk} are the elements of the transfer matrix. In general, this matrix should also include nonlinear effects.

To see how this works out in a simple case, consider an isochronous field so that τ is constant, and to make it more interesting, assume that ω_o does not quite match $2\pi/\tau$. In this case, we introduce a parameter λ to specify the frequency error, and hence write

$$\lambda = h(\omega_o\tau - 2\pi), \qquad (28)$$

so that λ is the phase slip per turn given in (21), and is therefore generally small.

If we also assume that the voltage gain per turn V is constant, then Eqs. (15,21) can be integrated directly with the following results:

$$\phi_f = \phi_i + \lambda n, \qquad (29a)$$

$$E_f = E_i + nqV\mu\cos(\phi_i + \tfrac{1}{2}\lambda n), \qquad (29b)$$

where

$$\mu = (\sin\tfrac{1}{2}\lambda n)/(\tfrac{1}{2}\lambda n) = 1 - \tfrac{1}{6}(\lambda n/2)^2 + \ldots, \qquad (29c)$$

so that μ represents an efficiency factor.

These equations apply both to the central ray and to any other trajectory as well. Thus, by expanding these equations while holding n fixed, we can obtain the deviations δE_f and $\delta\phi_f$ defined above, namely,

$$\delta\phi_f = \delta\phi_i, \qquad (30a)$$

$$\delta E_f = \delta E_i - nqV\mu\left[\sin(\phi_i+\tfrac{1}{2}\lambda n)\delta\phi_i + \tfrac{1}{2}\cos(\phi_i+\tfrac{1}{2}\lambda n)(\delta\phi_i)^2\right], \qquad (30b)$$

where the important second order term has also been included.

It now follows from (27) that the transfer matrix R resembles that for a straight section, and we therefore write

$$\underline{R} = \begin{bmatrix} 1 & L \\ 0 & 1 \end{bmatrix}, \qquad (31)$$

where the "length" L is split into two parts,

$$L = L_1 + L_2, \qquad (32)$$

with L_1 and L_2 representing the linear and quadratic contributions, respectively. When use is made of (29), these can be put in the following useful form:

$$L_1 = -(E_f - E_i)\tan(\phi_i+\tfrac{1}{2}\lambda n), \qquad (33a)$$

$$L_2 = -\tfrac{1}{2}(E_f - E_i)\delta\phi_i. \qquad (33b)$$

The effects of these two terms separately and in combination are shown in Fig. 5 in a case where the initial $(\delta E_i,\delta\phi_i)$ distribution is rectangular.

As can be seen, the resultant final energy spread will be minimized when $L_1 = 0$, which corresponds to adjusting the parameter λ in (29) so that $\phi_f = -\phi_i$. For other cases, specifically those with "tilted" distributions in $(\delta E_i,\delta\phi_i)$ values, some other choice for λ would very likely be optimal.

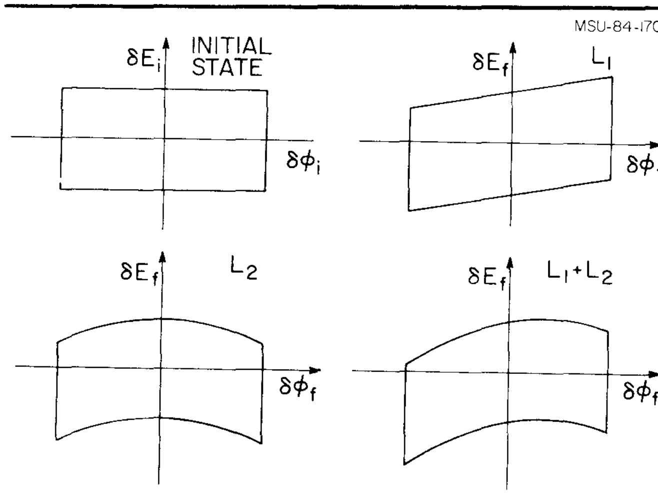


Fig. 5. Transformation of the longitudinal phase-space distribution produced by the transfer matrix for an isochronous field with a mismatched rf frequency. The initial δE vs. $\delta\phi$ distribution, which is assumed to be rectangular, is shown at top left. The final distribution shown at top right results from the linear term L_1 produced by a frequency shift. The final distribution at bottom left results from the quadratic term L_2 which is always present. The resultant final distribution at bottom right shows the combined effect of L_1 and L_2.

5. General Case

Turning now to the general case of a non-isochronous field, we assume that the function $\phi(E)$ vs. E for the central ray is known over the energy range appropriate to the given ion and operating condition, i.e., $E_i \leq E \leq E_f$. At our own laboratory, this curve is determined in each case by the Fielder program or one of its descendants.[12] We should also note that in most cyclotrons, phase probes are usually installed to measure the ϕ vs. E curve once the machine is running.

First, the invariant K of (20) is written in two alternate forms which serve to relate the final state and the initial state,

$$K = qV_i \sin\phi_i - hF_i, \quad = qV_f \sin\phi_f - hF_f, \tag{34}$$

and these equations apply to any trajectory. Thus, for a displaced ray, we find

$$\delta K = qV_i \cos\phi_i \delta\phi_i - \phi_i' \delta E_i, \quad = qV_f \cos\phi_f \delta\phi_f - \phi_f' \delta E_f, \tag{35}$$

where $\phi' = d\phi/dn$ is the change in ϕ per turn given by (21). The second order terms in the expansion for δK have been omitted since they seem much less important in general than the one derived below. Hence, the above δK expressions provide one of the two relations needed to find the transfer matrix.

The second relation is derived from (15) for dE/dn which, when inverted and integrated, yields

$$n = \int \frac{dE}{qV\cos\phi}, \tag{36}$$

where the integration runs from E_i to E_f. Using (20) for K, this integral can be rewritten in a more useful form, namely,

$$n = \int \left[(qV)^2 - (K+hF)^2\right]^{-\frac{1}{2}} dE, \tag{37}$$

which shows that $n = n(E_f, E_i, K)$. This expression for n also applies to any trajectory.

Since the ions in each bunch under consideration all start with n=0 and are extracted on the same turn number $n = n_f$, it follows that $\delta n_f = 0$ for the final state of all these ions. Thus, expanding δn_f, we have:

$$\delta n_f = \frac{\partial n}{\partial E_f} \delta E_f + \frac{\partial n}{\partial E_i} \delta E_i + \frac{\partial n}{\partial K} \delta K + \frac{1}{2} \frac{\partial^2 n}{\partial K^2} (\delta K)^2, \tag{38}$$

where here, we have retained the one second order term which turns out to be important in all cases.

We will return to the second order term later, but for now, we consider only the three first order terms in δn_f above. Evaluating the derivatives and setting $\delta n_f = 0$, we then obtain

$$\frac{\delta E_f}{qV_f \cos\phi_f} = \frac{\delta E_i}{qV_i \cos\phi_i} - I\left(\frac{\delta K}{qV_i qV_f}\right), \tag{39}$$

where the integral I is given by:

$$I = \int \left[\frac{\sin\phi}{\cos^3\phi}\right] g(E) dE, \tag{40}$$

which includes a factor $g(E) = V_f V_i / V^2(E)$ arising from the possible variation of V with E.

Combining this equation for δE_f with (35) for δK, the exact expression for the linear part of the transfer matrix can then be found. When this is done, the result is

$$R_{11} = \frac{V_f \cos\phi_f}{V_i \cos\phi_i} + I\left(\frac{\phi_i'}{qV_i}\right) \cos\phi_f, \tag{41a}$$

$$R_{12} = -I \cos\phi_f \cos\phi_i, \tag{41b}$$

$$R_{21} = \left[\frac{\phi'_f}{qV_i \cos\phi_i} - \frac{\phi'_i}{qV_f \cos\phi_f}\right] + I\left(\frac{\phi'_i}{qV_i}\right)\left(\frac{\phi'_f}{qV_f}\right), \tag{41c}$$

$$R_{22} = \frac{V_i \cos\phi_i}{V_f \cos\phi_f} - I\left(\frac{\phi'_f}{qV_f}\right) \cos\phi_i. \tag{41d}$$

For a given E_i and E_f, the value of these matrix elements therefore depend on the values of ϕ_i and ϕ_f together with the integral I of (40). For the isochronous field case treated at the end of Sec. 4, I can be evaluated analytically and as expected, these formulas for R_{jk} reduce exactly to those given in (33).

We note that I may be positive or negative, and for example, if ϕ is sufficiently small so that $\sin\phi \approx \phi$ and $\cos\phi \approx 1$, then we obtain

$$I \approx (E_f - E_i)\langle\phi\rangle, \tag{42}$$

where the angular brackets indicate an average. Clearly, the value of I can be varied over a substantial range simply by adjusting the trim coil currents to change the average phase values.

At this laboratory, trim coil current settings are always calculated for each operating condition in advance, and as an auxilliary part of this calculation, we try to impose the condition I = 0. This requirement is indeed an option in the original Fielder program and was based on the supposition that $\Delta E_i = 0$ could be assumed for a cyclotron having an internal source, and in this case, I = 0 leads to $\Delta E_f = 0$ to first order. For this reason, the condition I = 0 was originally named "energy focusing." But for an injected beam, assuming $\Delta E_i = 0$ is no longer warranted, so that the condition on I for obtaining energy focusing is generally quite different.

6. Energy Focusing and Phase Compression

Returning once more to the $(\delta E, \delta\phi)$ distribution within the beam at E_i and E_f, we know that for a linear transfer matrix, an elliptical distribution at E_i will transform into another elliptical distribution at E_f. Following standard notation,[13] the equation for either ellipse is written as

$$\gamma(\delta E)^2 + \beta(\delta\phi)^2 + 2\alpha(\delta E)(\delta\phi) = \varepsilon, \tag{43a}$$

where $\pi\varepsilon$ is the invariant area of the ellipse, and where

$$\beta\gamma = 1 + \alpha^2. \quad (43b)$$

Moreover, the total energy width ΔE or phase width $\Delta\phi$ for this distribution, (as determined by the maximum values of δE or $\delta\phi$) is given by

$$\Delta E = 2(\varepsilon\beta)^{1/2}, \quad (44a)$$

or

$$\Delta\phi = 2(\varepsilon\gamma)^{1/2}. \quad (44b)$$

Note that ε here corresponds to the normalized emittance and could, for example, be expressed in units of MeV x millirad.

Let us assume that the initial distribution is given so that the initial values of the ellipse parameters $(\alpha_i, \beta_i, \gamma_i)$ together with ε are all known. The relation between these parameters and those for the final distribution $(\alpha_f, \beta_f, \gamma_f)$ can be expressed in terms of the linear transfer matrix elements R_{jk}. Thus, for example,

$$\beta_f = \beta_i R_{11}^2 - 2\alpha_i R_{11} R_{12} + \gamma_i R_{12}^2, \quad (45)$$

and it then follows that the minimum value of β_f, and hence the final energy spread ΔE_f, is obtained when

$$R_{12} = R_{11}\,\alpha_i/\gamma_i. \quad (46)$$

This therefore becomes the general condition for energy focusing and since the value of R_{12}, as given in (41b), is proportional to I, it also becomes the corresponding condition for I, as discussed above.

Assuming this condition is fulfilled, then (44b) and (45) lead to the following relation between the minimum final energy width and the given initial width:

$$\min(\Delta E_f) = R_{11} k_i \Delta E_i, \quad (47)$$

where $k_i = 1/(1+\alpha_i^2)^{1/2}$ so that $k_i \leq 1$. This minimum value is evidently proportional to R_{11}, which can be calculated from (41a) now that I is fixed. We should also note that if $\alpha_i = 0$ so that the initial ellipse is erect, then (46) leads to I = 0, which is exactly the old energy focusing condition described at the end of Sec. 5.

In addition to (45) for β_f, the ellipse transformation equations also yield

$$\gamma_f = \gamma_i R_{22}^2 - 2\alpha_i R_{22} R_{21} + \beta_i R_{21}^2, \quad (48)$$

and as a result, the minimum value of γ_f occurs when

$$R_{21} = R_{22}\,\alpha_i/\beta_i. \quad (49)$$

If this condition is fulfilled, then by direct analogy to (46) and (47), the minimum final phase width is found to be

$$\min(\Delta\phi_f) = R_{22} k_i \Delta\phi_i, \quad (50)$$

where $\Delta\phi_i$ is the given initial phase width. Thus, this minimum is proportional to R_{22}.

As shown by (41c), R_{21} depends on ϕ_i' and ϕ_f' as well as I, and it is therefore possible in principle to satisfy both conditions (46) and (49) simultaneously, and thereby achieve phase compression as well as energy focusing. A closer look shows, however, that these two results are somewhat counterproductive. That is, R_{11} and R_{22} are closely related, as shown by (41).

For example, if $\alpha_i = 0$ so that the initial ellipse is erect, then (46) requires I=0, as noted above, while (49) can be satisfied if $\phi_i' = \phi_f' = 0$. As a result, both $R_{12} = 0$ and $R_{21} = 0$, and in addition, we find

$$R_{22} = \frac{V_i \cos\phi_i}{V_f \cos\phi_f} = 1/R_{11}. \quad (51)$$

Thus, if we set $V_f = 3V_i$, which is approximately the case for the Indiana cyclotron, and also take $\phi_i = 60°$ with $\phi_f = 0$, as suggested by Lapostolle,[1] the result is

$$R_{22} = \frac{1}{6}, \text{ and } R_{11} = 6. \quad (52)$$

Evidently, there is a large compression of the phase width in this case, but at the same time, there is a correspondingly large dilatation of the energy spread. This conclusion is still qualitatively correct even when $\alpha_i \neq 0$.

We should, however, emphasize that these results apply only to the linear effects. In actual practice, one often finds that the second order effect on the energy spread (discussed below) predominates completely over the linear effect, and in such cases, one could indeed obtain a significant phase compression along with energy focusing without substantially increasing the total (linear plus quadratic) energy spread of the extracted beam.

7. Second Order Effect

We consider again the second order effect in the transfer matrix which is produced by the $(\delta K)^2$ term in δn_f of (38). When this term is included in the calculation, we find that the change in the matrix elements R_{jk} given in (41) can be obtained simply by making the following change in I:

$$I \rightarrow I + \frac{1}{2} I' \left(\frac{\delta K}{qV_i}\right). \quad (53)$$

where

$$I' = qV_i \frac{\partial I}{\partial K}. \quad (54)$$

with I given in (40). After evaluating this derivative and recasting the result, we finally obtain

$$I' = \int \left[\frac{1+2\sin^2\phi}{\cos^5\phi}\right] g'(E)\,dE, \quad (55)$$

where the integration again runs from E_i to E_f, and where $g'(E) = V_f V_i^2/V^3(E)$ represents the correction due to the nonconstancy of V(E).

It is clear that I' is always significant, and to obtain an approximate value, we set g' = 1, expand the integrand in powers of ϕ, and thereby obtain (to second order)

$$I' \approx (E_f - E_i)(1 + \frac{9}{2}\langle\phi^2\rangle). \quad (56)$$

Thus, the value of I' is generally quite large and increases rapidly as the mean square phase $\langle\phi^2\rangle$ increases.

In those cases where the voltage V varies significantly with E, the assumption g' = 1 is not valid. If we assume instead that V varies linearly with E from its initial value V_i at E_i to its final value V_f at E_f, then the average value of g' is given by

$$\langle g'\rangle = (V_i + V_f)/2V_f, \quad (57)$$

and this value should be inserted as a correction factor into the I' formula given above. For example in the case of the Indiana cyclotron where $V_f = 3V_i$, we obtain $\langle g'\rangle \approx 2/3$ for this factor. Thus, a rather large increase in V with energy produces a relatively small (but significant) decrease in the second order effect.

Since I occurs in all of the matrix elements R_{jk} given in (41), the simple algorithm (53) above shows that I' will likewise contribute a second order correction to each. This contribution will generally involve δE_i as well as $\delta\phi_i$ since δK involves both. That is, from (35), we have

$$\frac{\delta K}{qV_i} = \cos\phi_i \delta\phi_i - (\frac{\phi'_i}{qV_i})\delta E_i. \quad (58)$$

The second term will, however, usually be negligible since ϕ'_i is generally so small.

Judging from the results found for the isochronous field case in Sec. 4, the most important second order effect will occur in R_{12}. If we write $R_{12} = L_1 + L_2$, as in (32), then the second order "length" now becomes

$$L_2 = -\frac{1}{2}I'\cos\phi_f \cos^2\phi_i \delta\phi_i, \quad (59)$$

where it is assumed that the δE_i term in (58) can be dropped. As might be expected, this L_2 reduces to that given in (33) for the isochronous field when $\phi(E)$ is sufficiently small.

8. Conclusions

The analytical formulas for the transfer matrix elements R_{jk} given in (41), together with the second order corrections generated by the algorithm in (53), are explicit functions of quantities which can either be determined in advance or else measured for a given operating configuration of the cyclotron. Among these quantities, those of greatest importance are the distribution of ($\delta E, \delta\phi$) values within the injected beam, and the phase curve ϕ vs. E for the central ray. It is also important to know how this phase curve depends on the trim coil currents and other machine parameters since variations in this curve provide the principle mechanism for optimizing the final distribution of ($\delta E, \delta\phi$) values for the extracted beam.

This optimization could be carried out most conveniently with the aid of a computer program (somewhat like the "Transport" program[13]) which is based on the given analytical formulas for the matrix elements. Such a program would allow the user to explore the advantages of various possible injection conditions, and to evaluate the effects of changes in trim coil currents and other parameters. This program could also prove useful in assisting the cyclotron operator in the process of optimizing the entire beam acceleration and transport system from the ion source to the target.

Our entire discussion here has, of course, completely ignored the longitudinal space charge effect.[4] However, one usually begins the process of optimizing cyclotron performance with a sufficiently low beam current to avoid activation difficulties, and for such currents, the longitudinal space charge effects should be minimal. Then, as the beam current is increased, the settings of various machine parameters can be readjusted to mitigate this space charge effect and thereby re-establish optimum conditions.

References

1. P. Lapostolle, GANIL Report 83R/027/TP/01, May 1983.
2. M.M. Gordon, Nucl. Instrum. Methods 169, 327 (1980).
3. S. Adam, Proc. Ninth Intl. Conf. on Cyclotrons (Les editions de Physique, Les Ulis, France, 1981), p. 431.
4. M.M. Gordon, Proc. Fifth Intl. Cyclotrons Conf. (Butterworths, London, 1969), p. 305.
5. C. Chasman and A.J. Baltz, Nucl. Instrum. Methods, Phys. Res. 219, 279 (1984).
6. W.M. Schulte and H.L. Hagedoorn, Nucl. Instrum. Methods 171, 409 (1980).
7. M.M. Gordon and F. Marti, Particle Accelerators 12, 13 (1982).
8. K.R. Symon and A.M. Sessler, High Energy Accelerators and Pion Physics I (CERN, Geneva, 1956), p. 44.
9. W. Joho, Particle Accelerator 6, 41 (1974).
10. R.W. Mueller and R. Mahrt, Nucl. Instrum. Methods 86, 241 (1970).
11. M.M. Gordon, Particle Accelerators, in press.
12. M.M. Gordon and D.A. Johnson, Proc. Sixth Intl. Cyclotron Conf. (American Inst. Phys., New York, 1972), 298.
13. K.L. Brown, Nucl. Instrum. Methods 187, 51 (1981).