CERN 87-03
21 April 1987

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CAS CERN ACCELERATOR SCHOOL
ADVANCED ACCELERATOR PHYSICS

The Queen's College, Oxford, England
16–27 September 1985

PROCEEDINGS

Editor: S. Turner

Vol. I

GENEVA
1987

ABSTRACT

This advanced course on general accelerator physics is the second of the biennial series given by the CERN Accelerator School and follows on from the first basic course given at Gif-sur-Yvette, Paris, in 1984 (CERN Yellow Report 85-19). Stress is placed on the mathematical tools of Hamiltonian mechanics and the Vlasov and Fokker-Planck equations, which are widely used in accelerator theory. The main topics treated in this present work include: nonlinear resonances, chromaticity, motion in longitudinal phase space, growth and control of longitudinal and transverse beam emittance, space-charge effects and polarization. The seminar programme treats some specific accelerator techniques, devices, projects and future possibilities.

ADVANCED ACCELERATOR PHYSICS COURSE, OXFORD SEPTEMBER 1985

Time	Monday 16 September	Tuesday 17 September	Wednesday 18 September	Thursday 19 September	Friday 20 September	Saturday 21 September	Sunday 22 September	Monday 23 September	Tuesday 24 September	Wednesday 25 September	Thursday 26 September	Friday 27 September
09.00	B R E A K F A S T (served 08.00 - 08.30)						B R E A K F A S T					
09.30	Registration Welcome and Introduction J. Mulvey K. Johnsen	Non-Linear Resonances E. Wilson	Non-Linear Resonances E. Wilson	Landau Damping H. Hereward	Waves in Plasmas C. Lashmore-Davies	FREE DAY	FREE DAY	Wake Fields V.G. Vaccaro	Instabilities J.-L. Laclare	Instabilities J.-L. Laclare	Synchro-Betatron Resonances A. Piwinski	Diffusion due to R.F. Noise G. Dôme
10.00	C O F F E E							C O F F E E				
10.30	Hamiltonian Dynamics J. Bell	Chromaticity B. Montague	Chromaticity G. Guignard	Beam-Beam Effect L. Evans	Beam-Beam Effect and Beam Disruption L. Evans (Presented by J. Gareyte)			Betatron Coupling with Radiation G. Guignard	Space-Charge Dominated Beam Transport I. Hofmann	Beam Break-up J. LeDuff	Schottky Noise and Beam Transfer Function Diagnostics D. Boussard	Beam Loading D. Boussard
11.30	Hamiltonian Dynamics J. Bell	Hamiltonian Treatment of Longitudinal Motion G. Dôme	Kinetic Theory and Vlasov Equation C. Lashmore-Davies	Hamiltonian Treatment of Longitudinal Motion G. Dôme	Dynamics with Radiation and Non-linear Wigglers J. Jowett			Stochastic Cooling D. Möhl	Stochastic Cooling D. Möhl	Electron Cooling H. Poth	Intra-beam and Multiple Scattering A. Piwinski	Seminar High Intensity Neutron Beams G. Rees
12.30	L U N C H							L U N C H				
14.00	S T U D Y S E S S I O N											
15.00	Seminar RFQ M. Puglisi	Discussion	Discussion	Discussion	Discussion			V I S I T TO RUTHERFORD APPLETON LABORATORY	Study Session and Discussion	Study Session and Discussion	Film on Polarization and Discussion	Closing Remarks K. Johnsen
16.00		T E A							T E A			
16.30	TEA	Seminar Particle Tracking H. Mais	Seminar S.C. R.F. Systems H. Piel	Seminar High-Field Linacs J. LeDuff	Seminar Heavy Ions; The Present and Future H. Gutbrod				Seminar Free-electron Laser A. Renieri	Polarisation in Electron and Proton Beams J. Buon	Seminar Linear Colliders versus Storage Rings U. Amaldi	
17.00												
17.30	Welcome Reception											
18.00	E V E N I N G M E A L (start 19.00 prompt)							E V E N I N G M E A L				
20.00				Organ Recital The Queen's College Chapel M.S. Gautrey					Bingham String Quartet Holywell Music Rooms		Cocktails + Banquet	

CERN ACCELERATOR SCHOOL

Rutherford Appleton Laboratory &
Department of Nuclear Physics, Oxford
will jointly organize a course on

ADVANCED ACCELERATOR PHYSICS

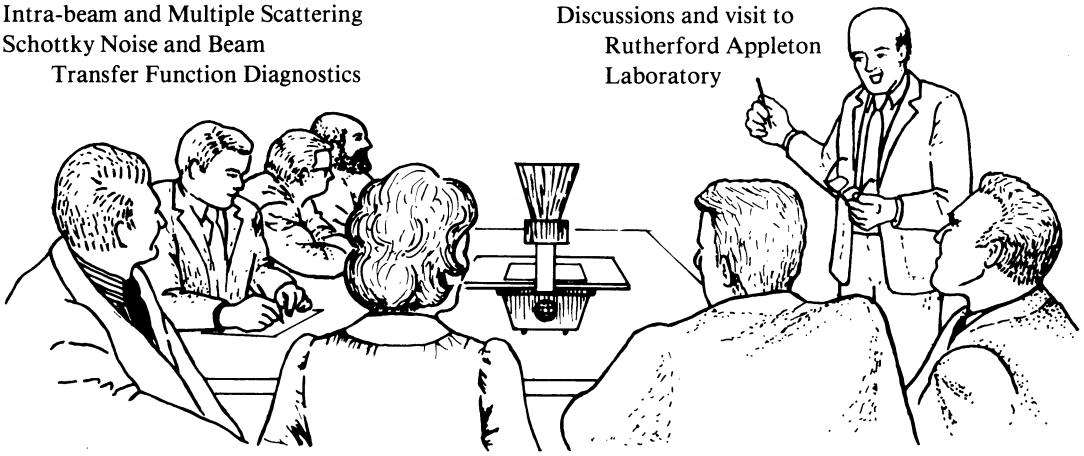
16–27 SEPTEMBER 1985
at The Queen's College, Oxford, England

Lectures:

Hamiltonian Dynamics
Non-linear Resonances
Chromaticity
Betatron Coupling with Radiation
Synchro-Betatron Resonances
Hamiltonian Treatment of Longitudinal Motion
Diffusion due to R.F. Noise
Beam Loading
Beam–Beam Effect
Beam Disruption
Beam Break-up
Space-Charge Dominated Beam Transport
Wake Fields
Instabilities
Stochastic Cooling
Electron Cooling
Intra-beam and Multiple Scattering
Schottky Noise and Beam
Transfer Function Diagnostics
Kinetic Theory and Vlasov Equation
Waves in Plasma
Landau Damping
Dynamics with Radiation and Non-linear Wigglers
Polarization in Electron and Proton Beams

Seminars:

RFQ
Particle Tracking
Superconducting Radiofrequency Systems
High-Field Linacs
Heavy Ions
Free-electron Laser
Linear Colliders versus Storage Rings
High Intensity Neutron Beams
Discussions and visit to
Rutherford Appleton
Laboratory



General information Students should have a basic knowledge of accelerators. Persons wishing to attend this course can obtain further information and application forms from the CERN Accelerator School, Mrs. S. von Wartburg, LEP Division, CH 1211 Geneva 23. Application forms must be returned no later than 1st June 1985. The registration fee of 970.- FS **includes full board and lodging** in The Queen's College, Oxford, England. All participants will receive a copy of the proceedings. The number of participants is limited to 130.

Administration Head of School: K. Johnsen; Programme Committee: P. Bryant, J. Buon, K. Johnsen, W. Joho, R. Kohaupt, J. Lawson, K. Nielsen, S. Tazzari;
Local Organizing Committee: P. Bryant, S. Collier, K. Johnsen, J. Maidment, J. Mulvey, S. von Wartburg.

CONTENTS

*) Contribution not received

FOREWORD

The CERN Accelerator School was established in 1983 with the main mission to preserve and disseminate the knowledge accumulated at CERN and elsewhere on particle accelerators and storage rings of all kinds. This is being achieved by means of a biennial programme of basic and advanced courses on general accelerator physics, supplemented by topical courses organised jointly with the US Particle Accelerator School, and specialised courses as needs arise. The basic and advanced courses together bridge the gap between a science or engineering degree and the level of knowledge appropriate for starting accelerator research work.

In 1984 the first basic course on general accelerator physics was held at Gif-sur-Yvette, Paris and the proceedings subsequently appeared as CERN Yellow Report 85-19, Volumes I and II. In September of the following year the first complementary advanced course was held at Oxford and the present proceedings bring together the lectures and seminars presented there. In alternate future years these courses will be repeated with some modifications for new material and current developments. Since this is an advanced course primarily meant for those interested in going deeply into accelerator theory, a certain stress is placed on the mathematical tools of Hamiltonian mechanics, the Vlasov equation and the Fokker-Planck equation, which are widely applied in the lectures.

With the publication of these proceedings I should like to take the opportunity to thank, on behalf of the School, the CERN Directorate and the School's Advisory and Programme Committees for their continued support and effort. The support of the Department of Nuclear Physics, Oxford, in organising the course, and the help and sponsorship of the Rutherford Appleton Laboratory are gratefully acknowledged. Particular thanks are also due to the lecturers who not only prepared and presented the different topics but also completed the exacting task of writing their chapters for the proceedings. I am also very grateful to the many people in the various CERN services who have given invaluable help in producing these proceedings. Finally, and most important, I would like to thank the 114 participants in the course who made it all so worthwhile.

This course was both the first of the advanced general accelerator physics courses to be run by the CERN Accelerator School and the last to be directed by our founder Head, Professor Kjell Johnsen. Since CAS was formed in 1983, Professor Johnsen has guided us through four courses and one workshop, and has set the school onto a firm foundation with clear objectives. He now hands over to Dr. P. Bryant and in the brief eight months before he retires he will take on a new role as Chairman of an Advisory Panel on new ideas for electron-positron Colliders for CERN as part of Carlo Rubbia's Working Group on the Scientific and Technological Longterm Future of CERN.

On behalf of all who have participated in the CAS activities, I should like to extend our thanks to Professor Johnsen and to convey our best wishes to him for his new work and his retirement.

S. Turner
Head of Administration
CERN Accelerator School

OPENING ADDRESS

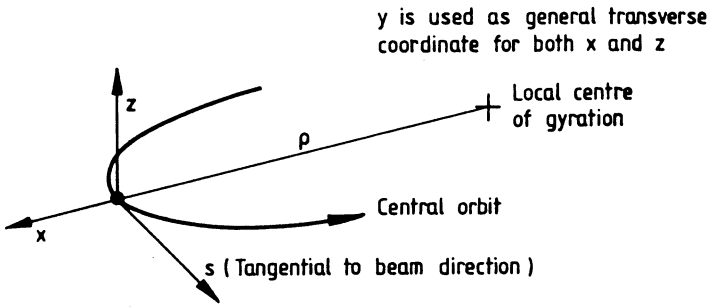
J. Mulvey
University of Oxford, Department of Nuclear Physics, Oxford, England

First of all of course, a very warm welcome to Oxford. Actually it seems a little chillier today but those of you who were here yesterday saw that the sun was shining and there is every expectation that it will shine at least once again before you go. I welcome you to Oxford, to this ancient University. The Queen's College, where you are staying, was founded in the year 1341, not one of the first colleges, Merton usually claims to be the first foundation about a hundred years earlier. Perhaps some of you who have had the opportunity of spending a first night in the college will appreciate the feelings of an American lady who came here with her husband, an eminent geneticist, spending a year here as a visiting professor, and after she left she wrote a book about her experiences, with the title "These ruins are inhabited". However things have moved on. It is generally assumed that the natural period or relaxation time for this University is about 250 years, but it was not so very long ago that one visitor to the university being shown around, was suddenly struck with a question which he put to his guide. "Tell me" he said , "the students who live in these colleges, where do they wash?" The guide had to think for a few moments and then showed the visitor the pump standing in the quadrangle. When the visitor expressed some surprise that that was the only such facility the guide was again at a bit of a loss until he realised that there really was not any problem, as he said, "Sir, the young gentlemen only spend 8 weeks here at a time". After these calumnies, which I am sure you realise are repeated in front of every conference that visits Oxford, I should provide you with some further information which may be of assistance.

The first thing that lecturers should notice is that if they put their lecture notes on this side of the table the fan here will blow them off! You will have registered in the school office so you know where it is. Well, regrettably, the school office will move. As from this afternoon it will be at another location in the college which I think is best described to you by saying that we will put a map on the door of the present office to tell you how to get to the new one. However, one advantage of the new one is that it has a telephone in it, and if at any time you want somebody to telephone you to leave a message this is the telephone number. If you want any information about the college, where to find the gents toilets or the ladies toilets or to find a piano or anything that you can think of like that, please in the first instance ask the porter at the lodge. If he cannot satisfy you, then come to the office and we will try to solve it. In the wallet that you all have received, you will find a certain amount of information about restaurants and entertainments in the town, as well as the school programme. You will note that for the weekend, Saturday and Sunday we have departed from what I think is normal tradition in not arranging an excursion. Our excuse for this is the following. There are very many excursions that can be made, you can tour around the Cotswolds, you can visit ancient castles, you can punt on the river here - if the weather is fine enough - or you may wish to visit London. With all these possibilities we thought you should choose your own pleasures. We will, of course, be glad to give you help and advice in making arrangements.

Finally, may I again welcome you to Oxford and wish you all a very enjoyable, and instructive stay.

LOCAL COORDINATES FOR THE BEAM AND FREQUENTLY USED SYMBOLS



(MKS units)

$\beta_y(s)$	transverse betatron amplitude function [m] in plane (y,s)
$\alpha_y(s) = -\frac{1}{2}\frac{d\beta y}{ds}$	
$\phi_y(s) = \int_0^s \frac{ds}{\beta_y(s)}$	phase of betatron oscillation
Q_y	number of betatron oscillations per revolution in transverse plane (y,s)
$\eta(s) = \frac{y(s)}{\sqrt{\beta_y(s)}}$	normalised amplitude of betatron oscillation
$\phi(s) = \int_0^s \frac{ds}{Q_y \beta_y(s)}$	normalised betatron phase
p	particle momentum [GeV/c]
$D(s) = \frac{y(s)}{\Delta p/p}$	transverse dispersion or momentum compaction [m](local y transverse co-ordinate of an off-axis closed orbit normalised by the fractional momentum deviation)
ε_y	transverse emittance [π m.rad] (defn. using 2 standard deviations half beam width, $\varepsilon = 4\sigma_y^2/\beta_y$).
ρ	local radius of bending [m]
R	average machine radius [m]
t	time [s]
I	beam current [A]
e	electronic charge [A.s]
c	velocity of the light [$m.s^{-1}$]
β	ratio of particle velocity to that of light

γ	ratio of total energy of particle to its rest energy
B	magnetic induction [T]
E	electric field strength [V/m]
$B\rho$	magnetic rigidity [T.m]
$\eta = \frac{1}{\gamma^2} - \frac{1}{\gamma_t^2}$	revolution frequency spread per unit of momentum spread, γ_t being the value of γ at the transition energy
ϕ, ϕ_s	radio frequency phase seen by particle crossing a cavity and phase seen by synchronous particle
Ω_s	synchrotron oscillation angular frequency [s^{-1}]
r_e, r_p	classical radii of the electron and proton respect. [m]

$'$	denotes d/ds	$\dot{}$	denotes d/dt		
$\bar{}$	average value	$\hat{}$	maximum value		
$\check{}$	minimum value	$\langle\rangle$	average over distribution	$\lvert\rvert$	modulus

HAMILTONIAN MECHANICS

J.S. Bell
CERN, Geneva, Switzerland.

ABSTRACT

For the bright gift of poetry was his;
And in lone walks and sweetly pensive musings
He would create new worlds and people them
With fond hearts and sweet sounds and sights of Beauty.
He had been gifted, too, with sterner powers.
Even while a child he laid his daring hand
On Science' golden key; and ere the tastes
Or sports of boyhood yet had passed away
Oft would he hold communion with the mind
Of Newton, and with awed enthusiasm learn
The eternal Laws which bind the Universe
And which the Stars obey.

W.R. Hamilton (c. 1830)

1. INTRODUCTION

William Rowan Hamilton (1805-1865) was a mathematician as well as a poet. His first publications were concerned with geometrical optics. An analogy between geometrical optics and mechanics led to his distinctive formulation of classical mechanics. For Hamilton the relation between optics and mechanics was one of analogy only. But in the meantime the analogy has acquired physical substance with the appearance, around 1925, of 'quantum mechanics'. It is now thought that the motion of particles is guided, somehow, by associated waves. And that it is only when the evolution of these waves is well approximated by geometrical optics, i.e. when wavelengths are small and frequencies high, that classical mechanics is good.

Consider the evolution of a 'wave packet', i.e. a wave train confined to a small region of space, but nevertheless containing many (necessarily very short) wavelengths. If the wavelength is $(2\pi/k)$, and the corresponding time period $(2\pi/\omega)$, the mean position q of the wave packet moves with the group velocity $d\omega/dk$. This is very familiar in the case of a homogeneous medium constant in time, when ω is a function of k only. But in the case of q and k large, it remains true (for short wave trains) in the case of an inhomogeneous inconstant medium, when ω has to be regarded as depending on position q and time t as well as on wave number k:

$$\frac{dq}{dt} = \frac{\partial}{\partial k}\omega(q,k,t). \qquad (1)$$

When ω depends on q and t, the mean wave number k is not constant during the propagation. In analogy with (1), it can be shown that

$$\frac{dk}{dt} = -\frac{\partial}{\partial q}\omega(q,k,t). \tag{2}$$

Defining

$$p = \hbar k, \qquad H(q,p,t) = \hbar\omega(q,k,t), \tag{3}$$

where $2\pi\hbar$ is Planck's constant, we have finally

$$\left.\begin{aligned}\frac{dq}{dt} &= \frac{\partial}{\partial p}H(q,p,t),\\ \frac{dp}{dt} &= -\frac{\partial}{\partial q}H(q,p,t),\end{aligned}\right\} \tag{4}$$

i.e. *Hamilton's equations* for classical particle motion; p is called the *momentum conjugate* to coordinate q, and H(q, p, t) is called the *Hamiltonian* function.

For simplicity we consider explicitly above, and often below, the case of just one coordinate q. For a particle moving in three dimensions we have actually three coordinates, q_1, q_2, and q_3 (and for N particles 3N q's), and corresponding momenta. Equations (4) hold for each coordinate q_n and momentum p_n:

$$\left.\begin{aligned}\frac{dq_n}{dt} &= \frac{\partial}{\partial p_n}H(q,p,t),\\ \frac{dp_n}{dt} &= -\frac{\partial}{\partial q_n}H(q,p,t).\end{aligned}\right\} \tag{5}$$

The Hamiltonian for a particle of mass m and charge e moving in an electromagnetic field is

$$H = e\phi + c\sqrt{(\vec{p}-e\vec{A})^2 + m^2c^2}, \tag{6}$$

where $\varphi(\vec{q}, t)$ and $\vec{A}(\vec{q}, t)$ are scalar and vector potentials, corresponding to electric and magnetic fields:

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial\vec{A}}{\partial t}, \qquad \vec{B} = \vec{\nabla}\times\vec{A}. \tag{7}$$

One can arrive at (6) just by writing down the simplest Lorentz- and gauge-invariant wave equation, which defines $\omega(q, k, t)$, and using (3). One can also check that (6) yields the familiar equations of motion. From (5)

$$\frac{dq_n}{dt} = c\left(p - eA_n\right)\Big/\sqrt{\left(\vec{p} - e\vec{A}\right)^2 + m^2c^2}\ , \tag{8}$$

$$\frac{dp_n}{dt} = -e\frac{\partial\phi}{\partial q_n} + ec\left[\sum_m \left(p_m - eA_m\right)\frac{\partial A_m}{\partial q_n}\right]\Big/\sqrt{\left(\vec{p} - e\vec{A}\right)^2 + m^2c^2}\ . \tag{9}$$

Using

$$\frac{dA_n}{dt} = \frac{\partial A_n}{\partial t} + \sum_m \frac{dq_m}{dt}\ \frac{\partial A_n}{\partial q_m} \tag{10}$$

and (8), equation (9) becomes

$$\begin{aligned}\frac{d}{dt}\left(p_n - eA_n\right) = &-e\frac{\partial\phi}{\partial q_n} - e\frac{\partial A_n}{\partial t} \\ &+ e\sum_m \frac{dq_m}{dt}\left(\frac{\partial A_m}{\partial q_n} - \frac{\partial A_n}{\partial q_m}\right) \\ = &\ e\vec{E}\left(\vec{q}, t\right) + e\frac{d\vec{q}}{dt} \times \vec{B}\left(\vec{q}, t\right)\ .\end{aligned} \tag{11}$$

From (8)

$$m\dot{\vec{q}}\Big/\sqrt{1 - \dot{q}^2/c^2} = \vec{p} - e\vec{A}$$

(where $\dot{\vec{q}} \equiv d\vec{q}/dt$). Using this in (11),

$$\frac{d}{dt}\left(\frac{m\dot{\vec{q}}}{\sqrt{1 - \dot{q}^2/c^2}}\right) = e\vec{E} + e\,\dot{\vec{q}} \times \vec{B} \tag{12}$$

as required.

Those who would not be bothered with quantum mechanics can simply start from the observation that with definition (6), equations (5) and (12) are equivalent.

2. POISSON BRACKETS

The rate of change of any function F(q, p, t) along a dynamical orbit is

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \sum_n \left(\frac{\partial F}{\partial q_n} \frac{dq_n}{dt} + \frac{\partial F}{\partial p_n} \frac{dp_n}{dt} \right)$$

$$= \frac{\partial F}{\partial t} + \sum_n \left(\frac{\partial F}{\partial q_n} \frac{\partial H}{\partial p_n} - \frac{\partial F}{\partial p_n} \frac{\partial H}{\partial q_n} \right) \qquad (13)$$

[using (5)]. For any two functions F and G we define their

Poisson Bracket:

$$\sum_n \left(\frac{\partial F}{\partial q_n} \frac{\partial G}{\partial p_n} - \frac{\partial F}{\partial p_n} \frac{\partial G}{\partial q_n} \right) \equiv (F, G). \qquad (14)$$

Then (13) becomes

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + (F, H). \qquad (15)$$

When F and G are identical the Poisson bracket (14) vanishes. In particular, (H, H) ≡ 0. So taking F in (15) to be H,

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} . \qquad (16)$$

In particular, if H does not depend directly on t, (∂H/∂t = 0), it is unchanged by the variation of q and p along a dynamical orbit:

$$dH / dt = 0. \qquad (17)$$

The value of H is called the *energy*, and (17) expresses *conservation of energy*.

3. STATIONARY AND VARYING ACTION

Consider a finite path in (q, p, t) space -- not necessarily a dynamical orbit (a solution of Hamilton's equations). We define the corresponding *action* S to be the integral along the path,

$$S(\text{path}) = \int \left(\sum_n p_n dq_n - H\, dt \right) . \qquad (18)$$

[Warning: sometimes the name 'action' is used for a quantity which omits the dt contribution in (18).] We will consider how S changes when the path is changed slightly.

It is convenient to label the points along the path by a parameter λ, which can be supposed to begin at 0 and end at 1. Then (q, p, t) are functions of λ, and considering different paths means considering different functions of λ. The expression (18) can be written

$$S = \int_0^1 d\lambda \left(p \frac{dq}{d\lambda} - H \frac{dt}{d\lambda} \right), \tag{19}$$

where for simplicity we omit the summation over n.

When the path is varied slightly, the increment in S is

$$\delta S = \int d\lambda \left(\delta p \frac{dq}{d\lambda} - \delta H \frac{dt}{d\lambda} + p \frac{d(\delta q)}{d\lambda} - H \frac{d(\delta t)}{d\lambda} \right)$$

$$= \int d\lambda \left[\delta p \frac{dq}{d\lambda} - \delta H \frac{dt}{d\lambda} - \frac{dp}{d\lambda} \delta q + \frac{dH}{d\lambda} \delta t + \frac{d}{d\lambda} \left(p \delta q - H \delta t \right) \right]. \tag{20}$$

Remember that H is a function of (q, p, t):

$$\delta H = \frac{\partial H_e}{\partial q} \delta q + \frac{\partial H_e}{\partial t} \delta t + \frac{\partial H_e}{\partial p} \delta p \, .$$

Using this,

$$\delta S = \int d\lambda \left\{ \delta p \left(\frac{dq}{d\lambda} - \frac{\partial H_e}{\partial p} \frac{dt}{d\lambda} \right) + \delta q \left(- \frac{dp}{d\lambda} - \frac{\partial H_e}{\partial q} \frac{dt}{d\lambda} \right) \right.$$

$$\left. + \delta t \left(\frac{dH}{d\lambda} - \frac{\partial H_e}{\partial t} \frac{dt}{d\lambda} \right) + \left[p \delta q - H \delta t \right]_{\lambda=0}^{\lambda=1} \right\}$$

$$= \int dt \left\{ \delta p \left(\frac{dq}{dt} - \frac{\partial H_e}{\partial p} \right) + \delta q \left(- \frac{dp}{dt} - \frac{\partial H_e}{\partial q} \right) \right.$$

$$\left. + \delta t \left(\frac{dH}{dt} - \frac{\partial H_e}{\partial t} \right) + \left[p \delta q - H \delta t \right]_{\lambda=0}^{\lambda=1} \right\}. \tag{21}$$

The coefficients of δp, δq, δt, are just the expressions required by Hamilton's equations, (4) and (16), to vanish. And so we have

<u>Hamilton's principle of varying action</u>: if, and only if,

$$\delta S = \left[p \delta q - H \delta t \right]_0^1 \tag{22}$$

for any small variation from a given path, is that path a dynamical orbit (i.e. a solution of Hamilton's equations).

If we consider only variations with $\delta p = \delta q = \delta t = 0$ at the ends, we have from Eq. (21):

Hamilton's principle of stationary action: if, and only if,

$$\delta S = 0 \tag{23}$$

for any small variation of path in which q, p, and t are held fixed at the end-points, is the given path a dynamical orbit.

Requiring (23) for fixed end-points requires Hamilton's equations everywhere in between, because δq, δp, and δt can be arbitrarily varied arbitrarily close to the end-points. It is striking that the restricted requirement, (23), for fixed end-points, by requiring Hamilton's equations requires (22) for unfixed end-points.

4. POINCARE INVARIANT

Consider a family of dynamical orbits

$$\left.\begin{aligned} q(\lambda, u, v, \ldots) \\ p(\lambda, u, v, \ldots) \\ t(\lambda, u, v, \ldots) \end{aligned}\right\} . \tag{24}$$

For fixed u, v, ..., as the parameter λ varies from 0 to 1 these functions trace out a solution of Hamilton's equations. The parameters u, v, ... serve to distinguish different members of the family of solutions considered. For each member we have the action (18):

$$S(u, v, \ldots) = \int (p\,dq - H\,dt) . \tag{25}$$

The variation formula (22) is valid for variation from a dynamical orbit to any nearby path, and so in particular for variation to a nearby dynamical orbit

$$\delta S = (p\,\delta q - H\,\delta t)_{\lambda=1} - (p\,\delta q - H\,\delta t)_{\lambda=0} . \tag{26}$$

Integrating this, we have for variation from an orbit u, v, ... to another u', v', ...:

$$\begin{aligned} S(u', v', \ldots) - S(u, v, \ldots) \\ = \int_1 (p\,dq - H\,dt) - \int_0 (p\,dq - H\,dt) . \end{aligned} \tag{27}$$

The two integrals are over the paths traced out, at $\lambda = 1$ and $\lambda = 0$ respectively, by varying u, v, ... in (24). If we return finally to the original orbit, i.e. if u' = u, v' = v, ..., the left-hand side of (27) vanishes, and we have

$$\oint_1 (p\,dq - H\,dt) = \oint_0 (p\,dq - H\,dt) . \tag{28}$$

The notation $\oint$ indicates that we now have integrals over paths which return to where they start, i.e. *closed* paths. We find then, making explicit the sum over degrees of freedom, that

$$\oint \left(\sum_n p_n dq_n - H\, dt \right) \tag{29}$$

is the same for any two closed paths obtained one from another by motion along dynamical orbits. This is the *Poincaré invariant* (or 'integral invariant', or 'relative integral invariant').

5. LAGRANGE INVARIANT

Consider again the family of orbits (24). From (25) and (26)

$$\frac{\partial S}{\partial u} = \left[p \frac{\partial q}{\partial u} - H \frac{\partial t}{\partial u} \right]_0^1 , \tag{30}$$

whence

$$\frac{\partial^2 S}{\partial v \partial u} = \left[\frac{\partial q}{\partial u} \frac{\partial p}{\partial v} + p \frac{\partial^2 q}{\partial v \partial u} + \cdots \right]_0^1 . \tag{31}$$

Equally

$$\frac{\partial^2 S}{\partial u \partial v} = \left[\frac{\partial q}{\partial v} \frac{\partial p}{\partial u} + p \frac{\partial^2 q}{\partial u \partial v} + \cdots \right]_0^1 . \tag{32}$$

Subtracting (32) from (31),

$$0 = \left[\frac{\partial q}{\partial u} \frac{\partial p}{\partial v} - \frac{\partial p}{\partial u} \frac{\partial q}{\partial v} - \frac{\partial t}{\partial u} \frac{\partial H}{\partial v} + \frac{\partial H}{\partial u} \frac{\partial t}{\partial v} \right]_0^1 . \tag{33}$$

So we have that the *Lagrange bracket*

$$[u,v] = \sum_n \left(\frac{\partial q_n}{\partial u} \frac{\partial p_n}{\partial v} - \frac{\partial p_n}{\partial u} \frac{\partial q_n}{\partial v} \right) - \left(\frac{\partial t}{\partial u} \frac{\partial H}{\partial v} - \frac{\partial H}{\partial u} \frac{\partial t}{\partial v} \right) \tag{34}$$

is constant along dynamical orbits when u and v are constant along dynamical orbits.

Consider a small variation from one dynamical orbit to another,

$$\delta_1 q = \frac{\partial q}{\partial u} \delta u \quad , \quad \delta_1 p = \frac{\partial p}{\partial u} \delta u \quad , \quad \cdots \quad , \tag{35}$$

and a second such variation

$$\delta_2 q = \frac{\partial q}{\partial v}\delta v \ , \ \delta_2 p = \frac{\partial p}{\partial v}\delta v \ , \ \cdots \tag{36}$$

It follows from the constancy of (34) that

$$\sum_n \left(\delta_1 q_n \delta_2 p_n - \delta_1 p_n \delta_2 q_n \right) - \left(\delta_1 t \delta_2 H - \delta_1 H \delta_2 t \right) \tag{37}$$

for variations from one dynamical orbit to any two nearby dynamical orbits, is constant. This is the *Lagrange invariant*.

There is a close relation between Lagrange and Poincaré invariants in the case of one degree of freedom when time t is held constant in relating one orbit to another. From (29) the integral

$$\oint p\, dq \tag{38}$$

is then constant as the points (p, q) of the integration path move in time t in accordance with Hamilton's equations. The integral (38) is just the area of the (p, q) plane contained by the curve (Fig. 1). So this area is a constant of the motion (a special case also of *Liouville's theorem*, Section 7).

Consider in particular a small parallelogram in (q, p) space with for sides $(\delta_1 q, \delta_1 p)$ and $(\delta_2 q, \delta_2 p)$ (Fig. 2). This has an area

$$\delta_1 q \, \delta_2 p - \delta_1 p \, \delta_2 q \, . \tag{39}$$

It evolves in the course of the motion, to the approximation linear in small quantities, into another small parallelogram -- which must have the same area. From the Poincaré invariance of (38) we arrive therefore at the Lagrange invariance of (39), a special case of (37).

Conversely, a large area can be seen as made up of little parallelograms (Fig. 3). The Lagrange invariance of the small component areas then implies the Liouville invariance of the total area, and the Poincaré invariance of the corresponding line integral (38).

6. SYMPLECTICS

Let δ_1 and δ_2 again denote small variations from one dynamical orbit to two others, for fixed values of the independent variable t:

$$\delta_1 t = \delta_2 t = 0 .$$

Let x_1 and x_2 denote the small variations of q and p arranged as column matrices:

$$X_1 \equiv \begin{bmatrix} \delta_1 q_1 \\ \delta_1 p_1 \\ \delta_1 q_2 \\ \delta_1 p_2 \\ - \\ - \\ - \end{bmatrix} \qquad X_2 \equiv \begin{bmatrix} \delta_2 q_1 \\ \delta_2 p_1 \\ \delta_2 q_2 \\ \delta_2 p_2 \\ - \\ - \\ - \end{bmatrix} \tag{40}$$

The Lagrange invariance of (37) then takes the form

$$\tilde{X}_1 \, \eta \, X_2 = \text{constant} , \tag{41}$$

where η is the square matrix

$$\eta \equiv \left[\begin{array}{cc|cc|cc} 0 & 1 & 0 & 0 & - & - \\ -1 & 0 & 0 & 0 & - & - \\ \hline 0 & 0 & 0 & 1 & - & - \\ 0 & 0 & -1 & 0 & - & - \\ \hline - & - & - & - & - & - \\ - & - & - & - & - & - \end{array}\right] \qquad \text{i.e.} \qquad \eta \begin{bmatrix} \delta q_1 \\ \delta p_1 \\ \delta q_2 \\ \delta p_2 \\ - \\ - \end{bmatrix} = \begin{bmatrix} \delta p_1 \\ -\delta q_1 \\ \delta p_2 \\ -\delta q_2 \\ - \\ - \end{bmatrix} \tag{42}$$

The tilde (~) denotes, as usual, transposition of rows and columns. In particular, $\tilde{X}_1$ is the row matrix

$$\left[\delta_1 q_1 , \delta_1 p_1 , \delta_1 q_2 , \delta_1 p_2 , - - - \right] .$$

In the course of dynamical evolution, from time t to time t', the small variations will undergo (to linear approximation) a linear transformation,

$$x' = S x \tag{43}$$

where the square matrix S depends on t and t'. From (41)

$$\tilde{X}_1' \eta X_2' = \tilde{X}_1 \eta X_2 \tag{44}$$

or

$$\tilde{X}_1 \tilde{S} \eta S X_2 = \tilde{X}_1 \eta X_2 . \tag{45}$$

This is so for arbitrary X_1 and X_2 if, and only if,

$$\tilde{S} \eta S = \eta \quad , \quad \tilde{S} \eta = \eta S^{-1} . \tag{46}$$

Such a matrix S is said to be *symplectic*. Note the partial analogy with *orthogonal* matrices O:

$$\tilde{O} O = 1 , \quad \tilde{O} = O^{-1} .$$

7. LIOUVILLE'S THEOREM

An equation such as (46) for matrices holds also for their determinants, so

$$(\det \tilde{S})(\det \eta)(\det S) = \det \eta .$$

Since

$$\det \tilde{S} = \det S , \quad \det \eta = 1 \neq 0 ,$$

we have then

$$\det S = \pm 1. \tag{47}$$

For a small time interval

$$S \approx 1 , \quad \det S = +1 .$$

It follows then by continuity that

$$\det S = 1 \tag{48}$$

always. This is *Liouville's theorem*. As we have derived it here it is a theorem about small deviations from *any* dynamical orbit to any neighbouring orbit. It can be expressed in two

other ways. The determinant of S is just the Jacobian of the transformation between q_1, p_1 ... and q_1', p_1', ... So this Jacobian is unity:

$$\frac{\partial(q_1', p_1', q_2', p_2', \text{---})}{\partial(q_1, p_1, q_2, p_2, \text{---})} = 1 \; . \tag{49}$$

And since the Jacobian gives the ratio of corresponding small volume-elements in (q, p) and (q', p'):

$$\int_V dq_1 dp_1 dq_2 dp_2 \text{---} = \int_{V'} dq_1', dp_1', dq_2', dp_2' \text{---} \tag{50}$$

where V is some region in (q, p) space; and V' is the region occupied, at a later time, by the points into which those of V evolve in accordance with Hamilton's equations. We will refer to (49) (or 48) and (50), equally, as Liouville's theorem.

8. CONSERVED QUADRATIC FORM

The bilinear form

$$\tilde{X}_1 \, \eta \, X_2$$

is conserved by *all* symplectic transformations S. The *quadratic* form

$$U = \tilde{X} \, \eta \, T X \tag{51}$$

is conserved by the *particular* symplectic transformation T from which it is constructed. Thus:

$$\tilde{X}' \eta T X = \tilde{X} (\tilde{T} \eta T) T X$$

$$= \tilde{X} \eta T X$$

from the assumed symplectic nature of T,

$$\tilde{T} \eta T = \eta \; .$$

Let T be the transformation, for one turn, of small deviations from a closed orbit. Suppose that T is constant from turn to turn, so that the evolution of X is given by repeated application of the *same* transformation T. Then the quadratic (51) remains constant throughout the motion -- always in the approximation linear in small deviations. It may be that the quadratic (41) is positive- or negative-definite, i.e. not zero for any X. Then its constancy guarantees that the motion is *bounded* (always in the linear approximation).

In the case of only one degree of freedom (or of one degree of freedom decoupled from others) η and T are 2 x 2 matrices:

$$\eta \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad T \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{52}$$

$$U = \tilde{X} \eta T X = \delta q (c \delta q + d \delta p) - \delta p (a \delta q + b \delta p)$$

$$= c (\delta q)^2 + (d-a) \delta q \delta p - b (\delta p)^2 . \tag{53}$$

This is definite (given det T = ad - bc = 1) if

$$(a + d)^2 / 4 \quad < 1 \tag{54}$$

and then has always the same sign as c (or -b).

In the case of several *uncoupled* degrees of freedom, U is a sum of similar contributions from each:

$$\left.\begin{aligned} U &= U_1 + U_2 + \cdots \\ &= c_1 (\delta q_1)^2 + (d_1 - a_1) \delta q_1 \delta p_1 - b_1 (\delta p_1)^2 \\ &\quad + c_2 (\delta q_2)^2 + (d_2 - a_2) \delta q_2 \delta p_2 - b_2 (\delta p_2)^2 + \cdots \end{aligned}\right\} \tag{55}$$

If *all* the U_n are positive definite (or *all* negative definite) then U also is positive definite (or negative definite). *Small* changes in the coefficients cannot immediately change a definite form into an indefinite one. So the stability (in linear approximation) of the uncoupled case survives the introduction of small enough coupling terms. This has some significance in particular for the coupling of horizontal and vertical betatron oscillations.

(For analysis of other conserved quadratics, $\tilde{X} \eta T^n X$, see J.S. Bell, AERE T/R1114, January 1953, and AERE T/R1383, March 1954.)

9. CHARACTERISTIC EXPONENTS

Consider again the problem of following small deviations, from a closed orbit, around many turns. If the transformation for one turn is T, and is constant from turn to turn, after n turns

$$X' = T^n X . \tag{56}$$

To evaluate this it is convenient to introduce the eigenvalues λ_k, and the corresponding eigenvectors $X^{(k)}$, of the matrix T. They are defined by

$$T X^{(k)} = \lambda_k X^{(k)} , \tag{57}$$

i.e. transformation of $X^{(k)}$ by T is equivalent to multiplying every component by the same factor λ_k. The λ_k satisfy

$$\det [T - \lambda I] = 0 , \tag{58}$$

where I is the unit matrix. This has as many roots λ_k as the matrix T has rows (or columns). When all these roots are different, the $X^{(k)}$ are all linearly independent. An arbitrary vector X can be expanded:

$$X = \sum_k a_k X^{(k)} . \tag{59}$$

From (56) and (57) then

$$X' = \sum_k (\lambda_k)^n a_k X^{(k)} . \tag{60}$$

This remains bounded for all n if

$$|\lambda_k| \leqslant 1 \tag{61}$$

for all k. We then have stability -- in the linear approximation in small deviations.

The left-hand side of (58) is a polynomial in λ with zeros at $\lambda = \lambda_k$, and with leading term

$$(-\lambda)^{2n} ,$$

where n is the number of degrees of freedom of the system, 2n the number of rows (or columns) of T. Then

$$\det [T - \lambda I] = \prod_k (\lambda_k - \lambda) . \tag{62}$$

With $\lambda = 0$ this gives

$$\prod_k \lambda_k = \det T = 1 \quad . \tag{63}$$

So the 2n eigenvalues multiply to unity. In particular: no eigenvalue is zero.

A theorem of Poincaré goes further. The reciprocal λ^{-1} of any eigenvalue λ is also an eigenvalue. To see this, first multiply (57) by

$$\lambda_k^{-1} \; T^{-1}$$

and interchange the two sides:

$$T^{-1} X^{(k)} = \lambda_k^{-1} X^{(k)} \; . \tag{64}$$

So the reciprocal of an eigenvalue of T is an eigenvalue of T^{-1}. But an eigenvalue of T^{-1} is also an eigenvalue of T when T is symplectic. For, using (46) and $\det \eta = 1$,

$$\begin{aligned} \det [T^{-1} - \lambda I] &= \det \eta [T^{-1} - \lambda I] \\ &= \det [\tilde{T} - \lambda I] \eta \\ &= \det [\tilde{T} - \lambda I] \; , \end{aligned}$$

so

$$\det [T^{-1} - \lambda I] = \det [T - \lambda I] \tag{65}$$

(since interchange of rows and columns does not change a determinant, and does not change the unit matrix I). Equation (65) shows that T^{-1} and T have the same eigenvalues. So: the reciprocal of an eigenvalue of T is also an eigenvalue.

Because T is real, the complex conjugate λ^* of any root λ of (58) is also a root. So when

$$\lambda = e^{\alpha + i\beta} \tag{66}$$

is an eigenvalue of T,

$$\lambda^* = e^{\alpha - i\beta} \tag{67}$$

$$\lambda^{-1} = e^{-\alpha - i\beta} \tag{68}$$

$$(\lambda^{-1})^* = e^{-\alpha + i\beta} \tag{69}$$

are also eigenvalues. For instability (in the linear approximation in small deviations) we must have some α non-zero. If the system is destabilized (in linear approximation) by some linear perturbation, then as that perturbation is diminished the vanishing of α brings two different roots together:

$$\lambda \to (\lambda^{-1})^* \to e^{i\beta} . \tag{70}$$

Conversely, *if a system is stable (in linear approximation) and all roots λ are different, it cannot be destabilized by an arbitrarily small linear perturbation.* For instability can set in only when the perturbation is strong enough first to bring two previously different eigenvalues together. (One zero of a polynomial cannot suddenly turn into two when the coefficients are varied continuously.)

The logarithms, $\alpha + i\beta$, of the eigenvalues λ are called the *characteristic exponents.* For stability they must all be pure imaginary. For while a negative α in (66) gives a decreasing term in (60), the corresponding reciprocal eigenvalue (68) gives an increasing term.

10. CANONICAL TRANSFORMATIONS

The variational principle (23) greatly facilitates change of variables. Let the old variables (q, p, t) be functions of new variables ($\bar{q}$, $\bar{p}$, $\bar{t}$) (and vice versa), and let $\bar{H}(\bar{q}, \bar{p}, \bar{t})$ be some function of the new variables, just as the Hamiltonian H is some function of the old (q, p, t). The transformation is said to be *canonical* if for arbitrary small changes in the variables

$$\left(\sum p\,\delta q - H\,\delta t\right) - \left(\sum \bar{p}\,\delta\bar{q} - \bar{H}\,\delta\bar{t}\right) = \delta F \tag{71}$$

where F is some function of the old (or new) variables. Define for any path

$$W = \int \left(p\,dq - H\,dt\right) \tag{72}$$

$$\bar{W} = \int \left(\bar{p}\,d\bar{q} - \bar{H}\,d\bar{t}\right) = W + [F]_0^1 . \tag{73}$$

This will be the first of a series of canonical transformations. At each step the new variables of the old transformation become the old variables of the new transformation. We will then drop the bars (-) in their notation.

12. CHANGE OF INDEPENDENT VARIABLE

Introduce the notation

$$t \equiv q_0 \quad , \quad -H \equiv p_0 \quad . \tag{88}$$

Then the basic stationary principle

$$\delta \int \left(\sum_1^N p\, \delta q - H\, \delta t \right) = 0$$

becomes

$$\delta \int \sum_0^N p_n\, \delta q_n = 0 \quad . \tag{89}$$

The apparent asymmetry between t and the coordinates q has disappeared. Clearly we can take some coordinate other than q_0, say $q_N \equiv \bar{t}$, as the independent variable. Then $-p_N$ becomes the Hamiltonian. Solving

$$-p_0 = H(q_0 \cdots q_N, p_1 \cdots p_N)$$

for p_N

$$-p_N = \bar{H}(q_0 \cdots q_N, p_0 \cdots p_{N-1}) \tag{90}$$

defines the new Hamiltonian function $\bar{H}$. Then from (89) we have Hamilton's equations

$$\frac{dq_n}{d\bar{t}} = \frac{\partial \bar{H}}{\partial p_n} \quad , \quad \frac{dp_n}{d\bar{t}} = -\frac{\partial \bar{H}}{\partial q_n} \quad , \qquad n = 0 \cdots N-1 \; , \tag{91}$$

with $\bar{t} \equiv q_N$ as the independent variable.

For example, from the Hamiltonian (87),

$$H = c\left[m^2c^2 + (p_s - eA_s)^2 (1 + x/\rho)^{-2} + (p_x - eA_x)^2 + (p_z - eA_z)^2 \right]^{1/2} + e\phi \tag{92}$$

taking s as the independent variable, and solving for $-p_s$, we have the new Hamiltonian

$$\bar{H} = -eA_s \mp \left(1+\frac{x}{\rho}\right)\left[\left(\frac{p_0+e\phi}{c}\right)^2 - m^2c^2 - (p_x-eA_x)^2 - (p_z-eA_z)^2\right]^{\frac{1}{2}} \tag{93}$$

and the Hamilton equations

$$\left.\begin{aligned} \frac{dt}{ds} &= \frac{\partial \bar{H}}{\partial p_0}\,, & \frac{dp_0}{ds} &= -\frac{\partial \bar{H}}{\partial t}\,, \\ \frac{dx}{ds} &= \frac{\partial \bar{H}}{\partial p_x}\,, & \frac{dp_x}{ds} &= -\frac{\partial \bar{H}}{\partial x}\,, \\ \frac{dz}{ds} &= \frac{\partial \bar{H}}{\partial p_z}\,, & \frac{dp_z}{ds} &= -\frac{\partial \bar{H}}{\partial z}\,. \end{aligned}\right\} \tag{94}$$

This general method of changing the independent variable is strongly recommended as compared with *ad hoc* approaches (for more detail, see M. Bell, AERE T/M 125, June 1955).

13. SCALING

Consider a change of variables such that

$$\sum p\,\delta q - H\,\delta t = C\left(\sum \bar{p}\,\delta\bar{q} - H\,\delta t\right) \tag{95}$$

where C is a constant. When C = 1 this is a canonical transformation [with F ≡ 0 in (71)] but not otherwise. But even with C ≠ 1 the Hamilton form of the equations of motion is preserved. This is so because

$$\int p\,dq - H\,dt = C\int \bar{p}\,d\bar{q} - \bar{H}\,d\bar{t} \tag{96}$$

and the stationary principle

$$\delta\int p\,dq - H\,dt = 0 \tag{97}$$

(with q, p, t fixed at the ends) is equivalent to

$$\delta\int \bar{p}\,d\bar{q} - \bar{H}\,d\bar{t} = 0 \tag{98}$$

(with $\bar{q}$, $\bar{p}$, $\bar{t}$ fixed at the ends).

And from (98),

$$\frac{d\bar{q}}{d\bar{t}} = \frac{\partial \bar{H}}{\partial \bar{p}}\,, \qquad \frac{d\bar{p}}{d\bar{t}} = -\frac{\partial \bar{H}}{\partial \bar{q}}\,. \tag{99}$$

For a canonical transformation

$$\sum_0^N p\,\delta q - \sum_0^N \bar{p}\,\delta\bar{q} = \delta F .$$

In particular then

$$\sum_0^N \left(p\frac{\partial q}{\partial u} - \bar{p}\frac{\partial \bar{q}}{\partial u} \right) = \frac{\partial F}{\partial u} , \tag{112}$$

$$\sum_0^N \left(p\frac{\partial q}{\partial v} - \bar{p}\frac{\partial \bar{q}}{\partial v} \right) = \frac{\partial F}{\partial v} . \tag{113}$$

Differentiating (112) with respect to u and (113) with respect to v, and remembering

$$\partial^2 F / \partial u\,\partial v = \partial^2 F / \partial v\,\partial u$$

etc., we find

$$\sum_0^N \left(\frac{\partial p}{\partial u}\frac{\partial q}{\partial v} - \frac{\partial p}{\partial v}\frac{\partial q}{\partial u} \right) = \sum_0^N \left(\frac{\partial \bar{p}}{\partial u}\frac{\partial \bar{q}}{\partial v} - \frac{\partial \bar{p}}{\partial v}\frac{\partial \bar{q}}{\partial u} \right) . \tag{114}$$

That is, the Lagrange bracket

$$[u,v] = \sum_0^N \left(\frac{\partial q_n}{\partial u}\frac{\partial p_n}{\partial v} - \frac{\partial p_n}{\partial u}\frac{\partial q_n}{\partial v} \right) \tag{115}$$

is unchanged when the variables (q, p) are replaced by others related by canonical transformation to the originals.

Consider now a small variation from the point (u, v, ...) in which only u is varied, by δu, and v etc. are held constant. The variations of the canonical variables are

$$\left.\begin{aligned} \delta_1 q &= \frac{\partial q}{\partial u}\delta u , & \delta_1 \bar{q} &= \frac{\partial \bar{q}}{\partial u}\delta u , \\ \delta_1 p &= \frac{\partial p}{\partial u}\delta u , & \delta_1 \bar{p} &= \frac{\partial \bar{p}}{\partial u}\delta u . \end{aligned}\right\} \tag{116}$$

Consider also a second small variation in which only v is varied, by δv, and u etc. are held constant. The variations of the canonical variables are

$$\left.\begin{aligned} \delta_2 q &= \frac{\partial q}{\partial v}\,\delta v\ , \qquad \delta_2 \bar{q} = \frac{\partial \bar{q}}{\partial v}\,\delta v\ , \\ \delta_2 p &= \frac{\partial p}{\partial v}\,\delta v\ , \qquad \delta_2 \bar{p} = \frac{\partial \bar{p}}{\partial v}\,\delta v\ . \end{aligned}\right\} \tag{117}$$

Multiplying (114) by (δu) (δv),

$$\sum_0^N \left(\delta_1 p\, \delta_2 q - \delta_1 q\, \delta_2 p \right) = \sum_0^N \left(\delta_1 \bar{p}\, \delta_2 \bar{q} - \delta_1 \bar{q}\, \delta_2 \bar{p} \right). \tag{118}$$

That is, for any two variations δ_1 and δ_2 from a given state, the Lagrange bilinear

$$\sum_0^N \left(\delta_1 p\, \delta_2 q - \delta_1 q\, \delta_2 p \right) \tag{119}$$

is unchanged by canonical change of variables. As a particular case, it is unchanged by dynamical evolution [Eq. (37)].

17. LIOUVILLE INVARIANT

Consider now transformations in which the independent variable is unaltered:

$$t \equiv \bar{t}\ . \tag{120}$$

And consider small variations in which the independent variable is not varied:

$$\delta t \equiv \delta \bar{t} = 0\ . \tag{121}$$

Let the other small variations be arranged into column matrices:

$$\begin{bmatrix} \delta_1 q_1 \\ \delta_1 p_1 \\ \delta_1 q_2 \\ \delta_1 p_2 \\ \vdots \end{bmatrix} \equiv X_1 \qquad \begin{bmatrix} \delta_2 q_1 \\ \delta_2 p_1 \\ \delta_2 q_2 \\ \delta_2 p_2 \\ \vdots \end{bmatrix} \equiv X_2$$

The Lagrange bilinear (119), omitting the n = 0 term by virtue of (121), is then

$$\tilde{X}_1 \, \eta \, X_2 \quad , \tag{122}$$

where η is defined in Eq. (42). The expression (112) is unaltered by canonical transformation, because of the invariance of (119). That is,

$$\tilde{\bar{X}}_1 \, \eta \, \bar{X}_2 = \tilde{X}_1 \, \eta \, X_2 \quad , \tag{123}$$

where

$$\bar{X}_1 \equiv \begin{bmatrix} \delta_1 \bar{q}_1 \\ \delta_1 \bar{p}_1 \\ \delta_1 \bar{q}_2 \\ \delta_1 \bar{p}_2 \\ \vdots \end{bmatrix} \qquad \text{etc.}$$

The small variations are related, in linear approximation, by a linear transformation

$$\bar{X} = L X \quad . \tag{124}$$

From (123),

$$\tilde{X}_1 \, \tilde{L} \, \eta \, L \, X_2 = \tilde{X}_1 \, \eta \, X_2$$

for arbitrary x_1 and x_2. Then

$$\tilde{L} \, \eta \, L = \eta \quad . \tag{125}$$

The matrix L of a canonical transformation is symplectic. It follows immediately from (125) that

$$(\det L)^2 = 1 \quad . \tag{126}$$

In fact

$$\det L = 1 \quad . \tag{127}$$

[See, for example, Hammermesh: Group Theory (Addison-Wesley Publ. Co., Inc., Reading, Mass., 1962).]

Equivalently, the Jacobian of a canonical transformation (which does not change t) is unity:

$$\frac{\partial(\bar{q}, \bar{p})}{\partial(q, p)} = 1 \quad . \tag{128}$$

Equivalently, the volume of a given region in (q, p) space, at given $t = \bar{t}$, is the same when measured in terms of canonically transformed variables $\bar{q}$ and $\bar{p}$:

$$\int_R \prod_1^N dq \prod_1^N dp = \int_{\bar{R}} \prod_1^N d\bar{q} \prod_1^N d\bar{p} . \tag{129}$$

Remember again that $q_0 (\equiv t)$ and p_0 are not included in (128) and (129), and the canonical transformation here has $t = \bar{t}$.

18. ONE DEGREE OF FREEDOM

Consider a system with one degree of freedom only, i.e. one p and one q (not counting p_0 and $q_0 \equiv t$). Consider a canonical transformation with $t \equiv \bar{t}$. Applying (110) to a closed curve along which $dt = d\bar{t} = 0$,

$$\oint p\,dq = \oint \bar{p}\,d\bar{q} . \tag{130}$$

This line integral is equal to the area enclosed by the curve. So this Poincaré invariance (130) is equivalent to the Liouville invariance (128) or (129). It says that the area of a given region in phase space is the same whether measured in the original q, p or in the new $\bar{q}$, $\bar{p}$. It will be shown here that for one degree of freedom this area invariance is not only *necessary* for a canonical transformation (with $t \equiv \bar{t}$) but also *sufficient*. That is, given (130), there are functions F and $\bar{H}$ such that

$$(p\,\delta q - H\,\delta t) - (\bar{p}\,\delta\bar{q} - \bar{H}\,\delta t) = \delta F . \tag{131}$$

From (130) the integral

$$\int_{q'',p''}^{q,p} (p\,dq - \bar{p}\,d\bar{q}) \tag{132}$$

depends only on the end-points, (q", p") and (q, p), and not on the choice of path joining them -- for the change in (132) on going from one path to another is just

$$\oint (p\,dq - \bar{p}\,d\bar{q})$$

where the closed curve is formed by going out along the second path and back along the first. This is zero by (130). So for arbitrary fixed (q", p"), the integral (132) defines a function of q and p

$$\int_{q'',p''}^{q,p} (p\,dq - \bar{p}\,d\bar{q}) = F(q,p) \tag{133}$$

and when q and p are varied,

$$\delta F = p\,\delta q - \bar{p}\,\delta\bar{q} \quad . \tag{134}$$

All this is for fixed $t \equiv \bar{t}$. In general, F will depend on t, if the transformation is time-dependent, and when t also is varied

$$\delta F = p\,\delta q - \bar{p}\,\delta\bar{q} + A(q,p,t)\,\delta t \tag{135}$$

where A is some function of the variables indicated. *Define* now

$$\bar{H}(\bar{q},\bar{p},t) = H(q,p,t) - A(q,p,t). \tag{136}$$

Then (135) reduces to (131). That is, if a transformation from (q, p) to ($\bar{q}$, $\bar{p}$) has the area-preserving property, it is part of a canonical transformation.

From (135) note that

$$A = \bar{H} - H = \frac{\partial F}{\partial t}\Big|_{q,\bar{q}} \quad . \tag{137}$$

It is important here, in the partial derivative, to be explicit about the variables (q, $\bar{q}$) that are held constant, as well as the variable t that is varied. For in (133) we indicated that F be regarded as a function of (q, p), and the conventions implicit elsewhere in these notes would imply for the notation

$$\partial F / \partial t$$

the explicit significance

$$\partial F / \partial t \,|_{q,p}$$

which is *not* what is needed in (137).

19. ACTION AND ANGLE VARIABLES

Often we are interested in an oscillatory motion characterized by an amplitude or a phase. Suppose the motion traces a closed curve in the (q, p) plane (Fig. 5) as the independent variable t increases. The 'amplitude' of the oscillation might be defined by the maximum value of q, or the maximum value of p, or in some other way -- for example by the area of the closed curve. This latter definition gives what is called the action variable J:

$$J = \frac{\text{area}}{2\pi} = \frac{1}{2\pi}\oint p\,dq \quad . \tag{138}$$

The area between two curves of slightly different J (Fig. 6) is just

$$2\pi \quad \Delta J \quad . \tag{139}$$

The *angle variable* ψ of a point P (Fig. 6) is then defined as the fraction of the area (138) which lies between P and some reference point -- for example P_0, where q takes its maximum for the given J -- multiplied by 2π:

$$\psi = \lim_{\Delta J \to 0} \left(\frac{\text{area between P and } P_0, \text{ J and J}+\Delta J}{\Delta J} \right) \qquad (140)$$

Clearly, from (138), ψ increases by 2π as P makes a complete oscillation from P_0, around clockwise back to P_0 with J fixed. The definition of J and ψ in terms of areas in the (q, p) plane is such that the area of any region in that plane is given equally by

$$\iint dq\, dp$$

and

$$\iint dJ\, d\psi \,.$$

That is, the transformation from (q, p) to (J, ψ) preserves areas, or has unit Jacobian

$$\partial(q, p) / \partial(\psi, J) = 1 \qquad (141)$$

and is therefore a canonical transformation (Section 18).

The reader should, as an exercise, construct the canonical transformation explicitly for harmonic and anharmonic oscillators. Here we will treat the slightly more complicated case of betatron oscillations in a storage ring. (The treatment of phase oscillations can be done in a similar way.) The complication is that the Hamiltonian, say (104), depends on the independent variable s. However, it does so in a periodic way, and then in the linear approximation the motion does trace out repeatedly a given curve, as in Fig. 5 (an ellipse in this case), provided attention is fixed on a given s and equivalent points s + c, s + 2c, ... , where c is the circumference. The transformation from (q, p) to (ψ, J) is then s-dependent -- but periodic with period c.

20. SMALL DEVIATIONS FROM CLOSED ORBIT

Let the transformation T of Section 8, for some degree of freedom which decouples from the others, be written

$$T \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\equiv \begin{bmatrix} \cos\mu + \alpha \sin\mu & \beta \sin\mu \\ -\gamma \sin\mu & \cos\mu - \alpha \sin\mu \end{bmatrix} \qquad (142)$$

so defining the parameters $\alpha(s)$, $\beta(s)$, $\gamma(s)$, μ [as in Courant and Snyder, Annals of Physics 3 (1958) 1].

We assume

$$\sin \mu \neq 0 \tag{143}$$

and make the convention

$$\beta \geqslant 0 \tag{144}$$

and note that

$$\det T = \beta\gamma - \alpha^2 = 1 \ . \tag{145}$$

The conserved quadric form (53) becomes, denoting the small deviations δq and δp simply by q and p,

$$U = c q^2 + (d-a) q p - b p^2 \tag{146}$$

$$= -\sin\mu \left(\gamma q^2 + 2\alpha p q + \beta p^2\right) \tag{147}$$

$$= -\sin\mu \left(\frac{q^2}{\beta} + \beta\left(p + \frac{\alpha}{\beta} q\right)^2 \right) . \tag{148}$$

The constancy of U (when s is increased by c, 2c, ...) defines the closed curve of Fig. 5 for given s and the given oscillation amplitude. The action variable J is then defined by the area enclosed by this curve, or equivalently the integral around it

$$J = \frac{1}{2\pi} \oint p \, dq \ . \tag{149}$$

The integration is simplified noting that (148) allows the introduction of an angle ψ such that

$$\left.\begin{aligned} q &= \sqrt{-U/\sin\mu}\ \sqrt{\beta}\ \cos\psi \\ p + \frac{\alpha}{\beta} q &= \sqrt{-U/\sin\mu}\ \sqrt{1/\beta}\ (-\sin\psi) \end{aligned}\right\} . \tag{150}$$

Then from (149)

$$J = -U/2\sin\mu \tag{151}$$

whence from (150)

$$\left.\begin{aligned} q &= \sqrt{2J}\,\sqrt{\beta}\,\cos\psi \\ p+\tfrac{\alpha}{\beta}q &= -\sqrt{2J}\,\sqrt{1/\beta}\,\sin\psi \end{aligned}\right\} . \tag{152}$$

Moreover, by elementary inspection of areas, or by verification that the Jacobian (141) is unity, ψ can be identified with the *angle variable* of the last section.

The methods of Section 18 could be used to construct the generating function F of the canonical transformation, from (q, p) to (ψ, J), and the new Hamiltonian $\bar{H}(\psi, J)$. However, the generating function is of little interest when we already have the transformation explicitly in (152), and know that it is canonical. The Hamitonian $\bar{H}$ can be obtained more quickly as follows. From Poincaré invariance (Section 4), the integral (149) is independent of s:

$$\frac{dJ}{ds} = -\frac{\partial \bar{H}}{\partial \psi} = 0 . \tag{153}$$

In the linear approximation, an increase of oscillation amplitude simply means increasing q and p by constant factors; that is, increasing J in (152) without changing ψ. It follows that $d\psi/ds$ is independent of J:

$$\frac{d\psi}{ds} = \frac{\partial \bar{H}}{\partial J} = \omega(s) , \tag{154}$$

where ω is some function of s [it cannot depend on ψ, for $\bar{H}$ does not so depend, from (153)]. From (153) and (154)

$$\bar{H} = \omega(s)\ J \tag{155}$$

apart from an unimportant constant.

From (152) the transformation for one turn can be found in the form

$$\begin{bmatrix} \bar{q} \\ \bar{p} \end{bmatrix} = \begin{bmatrix} \cos\Delta + \alpha\sin\Delta & \beta\sin\Delta \\ -\gamma\sin\Delta & \cos\Delta - \alpha\sin\Delta \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix} , \tag{156}$$

where

$$\Delta = \psi(s+c) - \psi(s) . \tag{157}$$

Identifying the matrix in (156) with that of (142) gives, using (154),

$$\mu = \Delta = \int_s^{s+c} \omega(s)\, ds \ . \tag{158}$$

Since ω is periodic in s with period c, this shows that μ is independent of s. Then (151) shows that U as well as J is a constant of the motion as s changes continuously (i.e. not only in jumps of c). Of course these two results are readily obtained by more elementary methods.

Because ω is a function of s, ψ increases non-uniformly. One can define a related angle variable φ which does increase uniformly:

$$\phi = \psi - \int_0^s \omega(s')\, ds' + \mu s/c \ . \tag{159}$$

In fact

$$d\phi / ds = \mu / c \ . \tag{160}$$

The transformation from (q, p) to (ψ, J) is still canonical, for

$$\frac{\partial(q,p)}{\partial(\phi,J)} = \frac{\partial(q,p)}{\partial(\psi,J)} \frac{\partial(\psi,J)}{\partial(\phi,J)} = 1 \times 1 \ . \tag{161}$$

From (160) and the constancy of J the new Hamiltonian is

$$\bar{\bar{H}}(\phi, J) = (\mu/c) J \ . \tag{162}$$

An explicit expression for $\omega(s)$ is easily found in the betatron oscillation case when

$$p = dq / ds \ . \tag{163}$$

Solving (152) for p,

$$p = -\sqrt{2J}\,(\sin\psi + \alpha\cos\psi)/\sqrt{\beta} \ . \tag{164}$$

Differentiating the first equation of (152),

$$dq/ds = -\sqrt{2J}\left[\sqrt{\beta}\,\omega(s)\sin\psi + \left(\frac{d\sqrt{\beta}}{ds}\right)\cos\psi\right] . \tag{165}$$

Comparing (164) and (165),

$$\omega(s) = 1/\beta(s) \quad . \tag{166}$$

Also, by the way,

$$\beta'(s) = 2\alpha(s) \quad . \tag{167}$$

21. ADIABATIC INVARIANCE OF J

Let the Hamiltonian be

$$H(q, p, s, \alpha) \quad , \tag{168}$$

where α is a parameter (or several parameters), such as mean bending field, which is constant in a first approximation. In this approximation suppose that an action variable has been found:

$$J(q, p, s, \alpha) \quad . \tag{169}$$

The rate of change of J is

$$\frac{dJ}{ds} = \frac{\partial J}{\partial q}\frac{dq}{ds} + \frac{\partial J}{\partial p}\frac{dp}{ds} + \frac{\partial J}{\partial s} + \frac{\partial J}{\partial \alpha}\frac{d\alpha}{ds}$$

$$= \frac{\partial J}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial J}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial J}{\partial s} + \frac{\partial J}{\partial \alpha}\frac{d\alpha}{ds} \quad . \tag{170}$$

The first three terms cancel together, for J is constant when α is constant. Then

$$\frac{dJ}{ds} = \frac{\partial J}{\partial \alpha}\frac{d\alpha}{ds} \quad . \tag{171}$$

In general $\partial J/\partial\alpha$ will be a function of the oscillation phase ψ. Suppose now that α varies very slowly and smoothly. Then J changes appreciably only over many oscillations, and in calculating this change $\partial J/\partial\alpha$ can be replaced by its average over ψ -- which is independent of ψ. So

$$\delta J = F(J, \alpha)\,\delta\alpha \quad . \tag{172}$$

That is, the final value of J is determined entirely by the initial J, independently of the initial ψ. It follows that all the particles lying initially on a closed curve of given J in the (q, p) plane again lie finally on a different curve of given J. But then, by Poincaré or

Liouville invariance, the areas enclosed by these curves must be equal. The final and initial J values are the same. *The action variable J is an adiabatic invariant.*

22. SMALL CANONICAL TRANSFORMATION

For perturbation theory a form of canonical transformation is required in which old and new variables are only a little different. Take $t = \bar{t}$, and the defining equation for canonical transformation

$$\delta F = p\,\delta q - \bar{p}\,\delta\bar{q} + (\bar{H} - H)\,\delta t . \tag{173}$$

From this,

$$\delta\left(F + \bar{p}\bar{q} - \bar{p}q\right) = (p - \bar{p})\,\delta q + (\bar{q} - q)\,\delta\bar{p} + (\bar{H} - H)\,\delta t . \tag{174}$$

Now since $\bar{p}$ will be almost the same as p, $(q, \bar{p}, t)$ will be a set of independent variables, and we may regard

$$F + \bar{p}\bar{q} - \bar{p}q \equiv \lambda X(q, \bar{p}, t) \tag{175}$$

as a function of them. From (174) then, and generalizing to many degrees of freedom,

$$\left.\begin{aligned} p_n &= \bar{p}_n + \lambda \frac{\partial X}{\partial q_n} , \\ \bar{q}_n &= q_n + \lambda \frac{\partial X}{\partial \bar{p}_n} , \\ \bar{H} &= H + \lambda \frac{\partial X}{\partial t} . \end{aligned}\right\} \tag{176}$$

When λ is small, the new and old variables are almost the same. We have the desired form of canonical transformation.

There is no approximation in (176). If we work only to first order in λ, the difference between p and $\bar{p}$ can be neglected in the coefficients of λ in (176):

$$\left.\begin{aligned} \bar{p} - p &= -\lambda\,\partial X/\partial q , \\ \bar{q} - q &= \lambda\,\partial X/\partial p , \\ \bar{H} - H &= \lambda\,\partial X/\partial t . \end{aligned}\right\} \tag{177}$$

Compare these with the infinitesimal form of Hamilton's equations

$$\left.\begin{aligned} \delta p &= -\delta t\, \partial H/\partial q \\ \delta q &= \delta t\, \partial H/\partial p \\ \delta H &= \delta t\, \partial H/\partial t . \end{aligned}\right\} \tag{178}$$

That is, Hamilton's equations give an infinitesimal canonical transformation. We see again that dynamical evolution generates a canonical transformation.

23. CANONICAL PERTURBATION THEORY

Suppose that when some part of the Hamiltonian is neglected there is a canonical transformation, to angle and action variables (ψ, J), which brings the approximate Hamiltonian to the form

$$H_0(J) \quad . \tag{179}$$

The *same* transformation from (q, p) to (ψ, J) will still be canonical when used with the complete Hamiltonian, transforming it to

$$H_0(J) + U(\psi, J, s) \, . \tag{180}$$

We look for a further canonical transformation to remove, as far as possible, the dependence of the Hamiltonian on ψ and s. To the extent that this can be done, the new Hamiltonian has the trivial form

$$\bar{H}(\bar{J}) \, .$$

Then

$$\left.\begin{aligned} d\bar{J}/ds &= -\partial \bar{H}/\partial \psi = 0 \\ d\psi/ds &= \partial \bar{H}/\partial \bar{J} = \omega(\bar{J}) \end{aligned}\right\} . \tag{181}$$

Because ψ is an angle, and because we deal with deviations from a closed orbit of circumference c,

$$\left.\begin{aligned} U(\psi + 2\pi, J, s) &= U(\psi, J, s) \\ \psi(\psi, J, s + c) &= U(\psi, J, s) \end{aligned}\right\} . \tag{182}$$

Because of these periodicity properties, U can be developed in a double Fourier series

$$\left.\begin{aligned} U &= U_{00} + U' \\ U' &= \sum{}' U_{\ell m}(J)\, e^{i\ell\psi + im\Omega s} \end{aligned}\right\}, \tag{183}$$

where

$$\Omega = 2\pi / c\,. \tag{184}$$

The summation Σ' is over all positive and negative integers ℓ and m, omitting the term ℓ = m = 0, which has been separated out. To allow explicitly for more than one degree of freedom it must be understood that

$$\ell\psi \equiv \sum_n \ell_n \psi_n \tag{185}$$

with as many angles ψ_n, and integers ℓ_n, as there are degrees of freedom.

Consider now a canonical transformation of the form (175), absorbing however the factor λ into the definition of X:

$$J = \bar{J} + (\partial/\partial\psi) X(\psi, \bar{J}, s)\,, \tag{186}$$

$$\bar{\psi} = \psi + (\partial/\partial\bar{J}) X(\psi, \bar{J}, s)\,, \tag{187}$$

$$\begin{aligned} \bar{H} &= H + (\partial/\partial s) X(\psi, \bar{J}, s)\,, \qquad (188) \\ &= H_0(\bar{J}) + (\partial H_0(\bar{J})/\partial\bar{J})(J - \bar{J}) \\ &\quad + U_{00}(\bar{J}) + U'(\psi, \bar{J}, s) \\ &\quad + \partial X(\psi, \bar{J}, s)/\partial s + R\,, \qquad (189) \end{aligned}$$

where R is of second order in U and X. Calculating (J - $\bar{J}$) from (186), and defining

$$\omega(J) = \partial H_0 / \partial J\,, \tag{190}$$

we have

$$\begin{aligned} \bar{H} &= H_0(\bar{J}) + U_{00}(\bar{J}) \\ &\quad + Y(\psi, \bar{J}, s) + R \end{aligned} \tag{191}$$

where

$$Y = \frac{\partial}{\partial s} X(\psi, \bar{J}, s) + \omega(\bar{J}) \frac{\partial}{\partial\psi} X(\psi, \bar{J}, s) + U'(\psi, \bar{J}, s)\,. \tag{192}$$

We can make Y zero, so that $\bar{H}$ is independent of $\bar{\psi}$ and s in first order, by taking

$$X = \sum{}' \frac{i\, U_{\ell m}(\bar{J})\, e^{i\ell\psi + im\Omega s}}{m\Omega + \ell\omega(\bar{J})} \tag{193}$$

provided there are no integers m, ℓ_1, ℓ_2, ..., other than 0, 0, 0, ..., such that

$$U_{\ell m} \neq 0, \qquad m\Omega + \sum_n{}' \ell_n \omega_n(\bar{J}) = 0. \tag{194}$$

The new Hamiltonian will have $\bar{\psi}$ and s dependence, and corrections to (181), only in the second order of small quantities.

The process can be repeated, pushing ψ and s dependence to higher and higher order, subject at each stage to (194). The study of whether this process, repeated indefinitely, converges, leads to the famous KAM theorem. Roughly, there is convergence at 'most' points in $\omega(\bar{J})$ space when U is small enough. But U must be *very* small for the proof, and arbitrarily small changes in parameters cause infinitely many changes from convergence to non-convergence and back again. So the relevance of the very beautiful theorem to real machines is obscure.

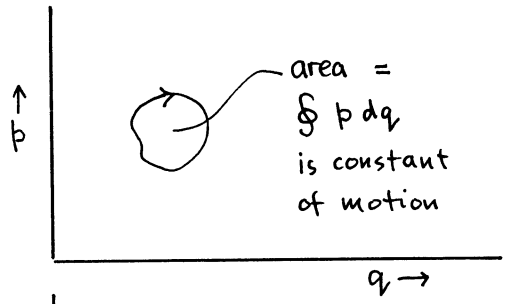


Fig. 1 Poincaré invariant.

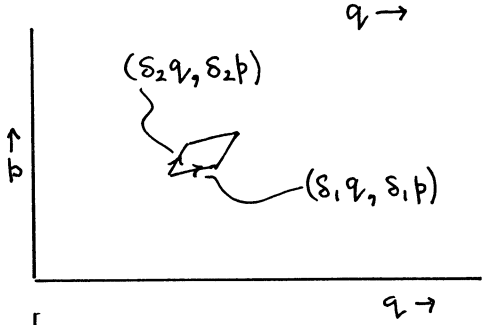


Fig. 2 Lagrange invariant.

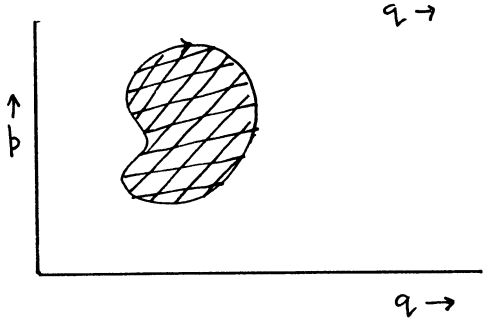


Fig. 3 Poincaré invariant as sum of Lagrange invariants.

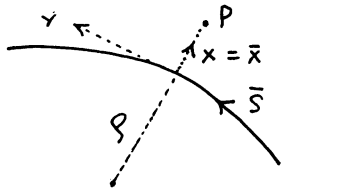


Fig. 4 Curvilinear coordinates for point P off reference orbit.

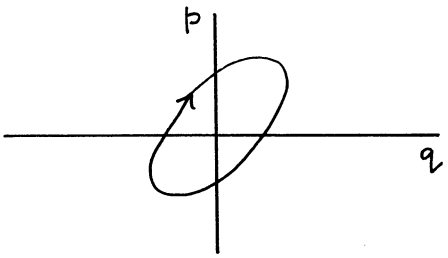


Fig. 5 Area is 2πJ where J is action variable.

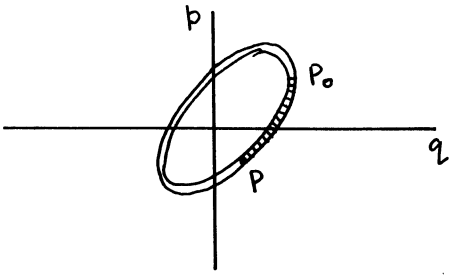


Fig. 6 The shaded area is ψΔJ where ψ is angle variable.

NONLINEAR RESONANCES

E.J.N. Wilson
CERN, Geneva, Switzerland

ABSTRACT

Building on classical Hamiltonian dynamics, this paper shows how a number of transformations can isolate the perturbation due to nonlinear terms in the guide field of a synchrotron. The concept of a resonance in transverse phase space is extended to include islands in phase space and the catalytic influence of synchrotron motion on beam diffusion.

1. INTRODUCTION

Hamiltonian mechanics is not everyone's cup of tea. Even experienced synchrotron designers often avoid the use of canonical transformation exercises to solve the simpler problems like the influence of sextupole magnets on betatron motion. It is true that for these problems a simple physical model can be used in which a sextupole is imagined as an element whose focusing strengths increase linearly with radius[1]. Such a tangible model related to the familiar concepts of optics is a valuable key to the door of nonlinear theory and may be used with great effect to predict resonant behaviour near a third integer resonance. One may extend this model to higher-order multipoles and into two degrees of freedom but at some point the handwaving becomes too vigorous to be credible.

Unfortunately, the topical problems of nonlinear theory applied to today's accelerator and storage rings include high-order resonances in three dimensions which can cause an inexorable beam loss over periods of hours. Typical of these problems is the influence of the beam-beam potential seen by a particle traversing an oncoming bunch in a collider. One can tune the Q values to avoid all low-order resonant conditions so that it is only the higher terms in the polynomial expansion of this potential which are important. The intuitive model is difficult to apply to polynomial terms which are typically of tenth order and it is worth taking the trouble to become adept at the Hamiltonian formalism to solve such a problem.

In the contribution which precedes this one[2], J. Bell has laid a firm foundation for the understanding of this formalism. In this contribution I shall first take his expression for the general Hamiltonian of a particle in an electromagnetic field and judiciously simplify it by approximations which are valid for large synchrotrons. This will reveal one of the advantages of the formalism, that each term in the multipole expansion of the field has a one-to-one relationship with a term in Hamiltonian.

The next procedure is to apply a number of canonical transformations which remove the lower-order linear terms in the Hamiltonian. To the uninitiated these transformations are unnerving in that they bring one further from the everyday world in which our familiar optics can be applied. However, their power lies in the fact that the nonlinearity is isolated and we are able to plot its effect as a trajectory in a phase space in which ordinary dynamics is reduced to circular trajectory. The perturbation of the nonlinearity becomes immediately visible.

To help the reader chart our course we list the canonical transformations to be applied in the form of a table. Each one is to be found in Ref. 2.

Table 1

Canonical Transformations

	Transformation	Purpose	Final Coordinates
1	Change of independent variable	Express as function of s, not t.	x, x', s
2	To action angle variables (J, ψ)	Remove variation of focusing with azimuth.	J, ψ, y
3	To the coordinates of a harmonic oscillator		J_1, ψ_1, s
4	Infinitesimal point transformation	Non-linear terms become first order perturbation.	J_2, ψ_2, s
5	Change independent variable to θ	Opens the way to finding a periodic solution.	J_3, ψ_3, θ
6	Transformation to a rotating coordinate system	Freezes the trajectory in a stroboscopic picture of closed curves, islands and separatrices.	J_4, ψ_4, θ

When this series of transformations is complete we shall have travelled a long road. Many of the steps on the journey will have to be taken on trust by those not yet skilled in imagining the new canonical transformations necessary to recast variables in a simpler form. It may help to think of each transformation as an imaginative leap like solving a hard integral, a difficult step for the uninitiated but easily checked once it is identified. In the end we shall find a simple geometrical trajectory which exhibits the physical features of resonant behaviour including unstable and stable fixed points, islands of stable motion, separatrices and stochastic layers which can lead particles to ever growing amplitude and cause the diffusion which it is our aim to explain. The position and separation of these features in phase space are precisely related to the coefficients in the field expansion of the real world.

In the remainder of this report we shall explain each of these transformations in some detail and then discuss the effect of nonlinearity in creating unstable regions of phase space, on the width of resonances and on the growth of amplitude when crossing resonances. All this will be treated for motion in one degree of freedom. We shall then see how by adding a further degree of freedom, synchrotron motion, the topology of phase space is altered and sketch out how one may lead to the Cherikov limit beyond which slow but continuous beam growth occurs. A criterion for purity of magnetic field then emerges.

We leave the details of discussion of the consequences of adding a second transverse degree of freedom and a third dimension to the field shape to other authors[3] but indicate the qualitative consequences of these extensions to the one dimensional model.

In all sections of this report, we shall refer to J. Bell's contribution to the 1985 CERN Advanced Accelerator School[2] which the reader is well advised to have at hand. Another useful reference is E.D. Courant, R.D. Ruth and W.T. Weng[4].

The next section describes the link between the rigorous and general treatment of Ref. 2 and the simplified Hamiltonian which is the starting point of our discussion. The reader may choose to omit it on first reading and pass to Eq. (8).

2. THE GENERAL FORM OF THE HAMILTONIAN

It is a good idea to start from one of the fundamental expressions to be found in text books on dynamics in an electromagnetic field. We shall rather soon introduce approximations which are justified in the context of a modern synchrotron. Watch these carefully. If your particular synchrotron is a small one you may wish to review their validity.

Reference 2 gives the general Hamiltonian for a charged particle of mass, m, and charge, e, in a magnetic vector potential, $\bar{A}$, and electric potential, φ:

$$H = e\varphi + c\sqrt{(\bar{p} - e\bar{A})^2 + m^2c^2}\ , \qquad (1)$$

where

c is the velocity of light,
$\bar{A}$ and φ are functions of space and time,
$\bar{p}$ is the momentum conjugate to the space coordinate.

We shall choose the coordinates shown in Fig. 1. The transverse displacements are x and z while s is measured along the particle's trajectory. The local radius of curvature, ϱ, is dependent on the local magnetic field and is therefore a function of s. We can express the Hamiltonian in these coordinates as:

$$H = c\left\{m^2c^2 + (p_s - eA_s)^2\left(1 + \frac{x}{\varrho}\right)^{-2} + (p_x - eA_x)^2 + (p_z - eA_z)^2\right\}^{1/2} + e\varphi \ . \quad (2)$$

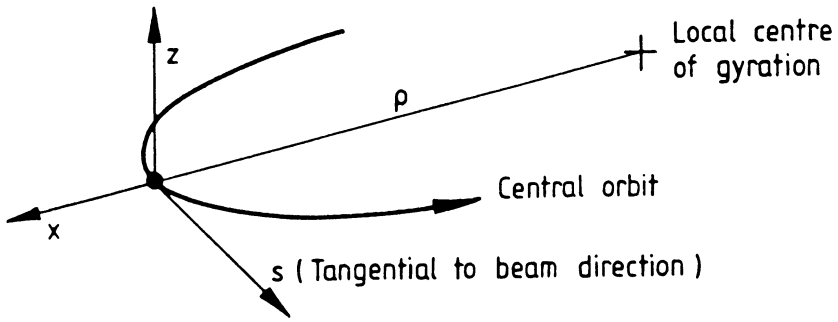


Fig. 1 Transverse coordinate system.

From now on we shall drop φ which is assumed to be constant. We are only considering magnetic and not electric fields.

The independent coordinate in (2) is t but it would be more convenient to use s since the machine turns out to be periodic in s. In his treatment of scaling Bell[2] gives explicitly the canonical transformation for this change of independent variable. The reader may remember that since H is conjugate to t and p_s to s, one may turn the Hamiltonian (2) inside out by writing a new Hamiltonian, K, which is really $-p_s$ expressed as a function of the other canonical variables together with s, the new independent variable:

$$K = -eA_s \mp \left(1 + \frac{x}{\varrho}\right)\left\{\left(\frac{p_0}{c}\right)^2 - m^2c^2 - (p_x - eA_x)^2 - (p_z - eA_z)^2\right\}^{1/2} . \quad (3)$$

We can further apply simple scaling to new coordinates (with a bar):

$$\bar{q} = q, \quad \bar{s} = s, \quad \bar{p} = \frac{p}{p}, \quad \bar{K} = \frac{K}{p} , \quad (4)$$

where

$$pc = \sqrt{p_0^2 - mc^2}. \quad (5)$$

So that

$$\bar{K} = - \frac{eA_s}{p} - \left(1 + \frac{x}{\rho}\right)\left\{1 - \left(\bar{p}_x - \frac{eA_x}{p}\right)^2 - \left(\bar{p}_z - \frac{eA_z}{p}\right)^2\right\}^{1/2} . \qquad (6)$$

Finally, in our attempt to recast the Hamiltonian in a more tractable form, we shall ignore the vertical plane and approximate by assuming $\rho = \infty$. Furthermore, we shall assume the magnetic field is only transverse so that $A_x = A_z = 0$.

These drastic measures should not be forgotten. They are completely justified to study horizontal motion in a large synchrotron comprising long magnets but would have to be re-examined in the case of a small machine in which the end fields of lattice elements are of importance. With these reservations we obtain a new Hamiltonian which we call H because we are running out of alphabet:

$$H \approx - \frac{eA_s}{p} - (1 - \bar{p}_x^2)^{1/2} , \qquad (7)$$

and if $\bar{p}_x \ll 1$,

$$H \approx - \frac{eA_s}{p} + \frac{\bar{p}_x^2}{2} . \qquad (8)$$

3. THE MAGNETIC VECTOR POTENTIAL FOR MULTIPOLES

It would be difficult to unravel the problem of nonlinear motion in a synchrotron if we were not able to analyse magnetic fields in to a series in which each term corresponds to a magnet with a certain number of poles.

We have already announced our intention to ignore the ends of magnets where there can be transverse components of the magnetic vector potential and restrict our analysis to the body of long magnets where only A_s is finite and there are only the transverse components of field, B_x and B_z . This has the virtue that the vector potential can be expressed by a series:

$$A_s = \sum_n A_n f_n(x, z) \qquad (9)$$

$$= \sum_n A_n (x + iz)^n . \qquad (10)$$

In this expansion f_n corresponds to a multipole with 2 n poles. The real

terms form a series for "normal" multipoles for which the field is normal to a horizontal mid-plane and the imaginary terms correspond to skew multipoles.

We can derive the vertical field component for a 2n pole magnet:

$$B_z(z=0) = \frac{\partial A_s}{\partial x} = n\, A_n\, x^{(n-1)} \quad , \tag{11}$$

and we can write this as a term in the Taylor series:

$$B_z(z=0) = \frac{1}{(n-1)!} \frac{\partial^{(n-1)} B_z}{\partial x^{(n-1)}} x^{(n-1)} \quad . \tag{12}$$

We can also write the first term in H (Equation 8) as:

$$\frac{eA_s}{p} = \frac{1}{B\varrho} \sum \frac{1}{n!} \frac{\partial^{(n-1)} B_z}{\partial x^{(n-1)}} x^n \quad , \tag{13}$$

and our Hamiltonian becomes:

$$H = \frac{p_x^2}{2} + \sum_{n=0}^{\infty} \frac{1}{(B\varrho)} \frac{1}{n!} \frac{\partial^{(n-1)} B_z}{\partial x^{(n-1)}} x^n \quad . \tag{14}$$

We immediately see that each order of multipole will contribute a term to the Hamiltonian. It is hardly a difficult step to see that if we were to allow both transverse degrees of freedom a normal multipole with 2n poles can contribute a set of terms x^n, $x^{n-2}z^2$, $x^{n-4}z^4$... while a skew multipole, corresponding to the imaginary terms in the expansion of Eq. 10, would introduce the missing terms $x^{n-1}n$, $x^{n-3}z^3$... in the homogeneous polynomial. It can be shown that each term corresponds to a line in the Q diagram, so this is helpful in identifying the kind of error which may cause resonant loss at a particular working point.

4. LINEAR DYNAMICS IN ACTION ANGLE VARIABLES

Readers who are already familiar with the theory of transverse dynamics will remember that linear motion is described locally by Hill's equation:

$$\frac{d^2x}{ds^2} + k(s)\, x = 0 \quad . \tag{15}$$

This is a differential equation with a periodic coefficient. The focusing strength, k(s), repeats every turn of the ring, or even every superperiod, if the lattice has a symmetry within a turn. The periodic variation of k(s) distinguishes the solution from simple harmonic motion giving rise to an amplitude function, $\sqrt{\beta}(s)$, that varies periodically with s reflecting the focusing pattern, k(s). The motion can be converted into that of a harmonic oscillator with a simple sine like solution by a change of variables[5]:

$$\eta = \frac{x}{\sqrt{\beta}} \ . \tag{16}$$

We can see that the Hamiltonian (if we take only linear terms) has the same periodic coefficient:

$$H = \frac{p^2}{2} + \frac{k(s)x^2}{2} \ , \tag{17}$$

where:

$$k = \frac{1}{(B\varrho)} \frac{\partial B_z}{\partial x} \ . \tag{18}$$

Incidentally it is an excellent exercise for newcomers to apply Hamilton's equations to (17) and thus derive Hill's equation.

Expressed in the jargon of Hamiltonian mechanics the difficulty with Eq. 17 is that H is not time-invariant. (Remember that earlier in Section 2 we used s instead of t as an independent variable). The difficulty is really the same as that with Hill's equation because if the Hamiltonian were to be time invariant it would generate the differential equation of a harmonic oscillator, and the trajectory of the particle in phase space would be the same closed ellipse independent of the observer's position, s. The trajectory could then be labelled with a numerical value, a constant of the motion. The momentum conjugate to any coordinate which is not in the expression for H is invariant, this applies equally to H itself which is conjugate to s.

Perhaps this is the first inkling for the diffident recipient of Hamilton's legacy that it might be of some practical value. He or she will be pleased to learn that the canonical transformation to action and angle variables discussed by Bell[2] is Hamilton's solution to this difficulty.

If we were to derive Hill's equation from the Hamiltonian (17) and solve it we would find:

$$y = E^{1/2}\beta^{1/2}(s) \cos[\varphi(s) + \delta] \tag{19}$$

$$\varphi(s) = \int_0^s \frac{ds}{\beta(s)} \quad , \tag{20}$$

where

y is the general transverse coordinate,
Q is the betatron wave number,
$\beta(s)$ is an amplitude function periodic in azimuth,
E is an emittance defining an azimuthal invariant,
δ is an arbitrary phase.

We must find the momentum which is conjugate to y and for this we use Hamilton's Equation:

$$\frac{dy}{ds} = \frac{\partial H(y,s)}{\partial p} = p \quad . \tag{21}$$

By differentiating Eq. 19 we find:

$$p = -E^{1/2}\beta^{-1/2}(s)\left\{\sin[\varphi(s) + \delta] - \frac{\beta'}{2}\cos[\varphi(s) + \delta]\right\}. \tag{22}$$

We are now in a position of having physical coordinates (p and y) which are conjugate according to the rules of Hamilton. Neither of them is a constant of the motion. In our desire to freeze out the linear terms it would be an advantage to transform (p and y) to new coordinates (J and ψ) in which J becomes independent of time. In Ref. 2, Bell achieved this for the case of a harmonic oscillator by using a canonical transformation into action angle variables.

The mathematical procedure for performing this type of transformation is to construct a function of a pair of old (p,q) and new (P,Q) coordinates which must have one of the four forms:

$$F_1(q,Q,t),\ F_2(q,P,t),\ F_3(p,Q,t) \text{ or } F_4(p,P,t) \ . \tag{23}$$

In general F_1 can be used for interchange of momentum with displacement or mixing them together to form invariants, while F_2 can be used for rotations or small perturbations in displacement or momentum space.

For the purposes of transformation into action and angle coordinates we select the first type. There is a prescription for performing the change of coordinates:

$$p_i = \frac{\partial F_1}{\partial q_i} \tag{24}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i} \tag{25}$$

$$K = H + \frac{\partial F_1}{\partial t} \ . \tag{26}$$

The way in which this works may be tried out by using the function:

$$F_1(q,Q,t) = qQ \tag{27}$$

to achieve an interchange of initial displacement and momentum.

By using the less obvious:

$$F_1(q,\psi,s) = \frac{1}{2}\sqrt{\frac{m}{k}}\, q^2 \cot\psi \ , \tag{28}$$

one may reproduce the transformation of the harmonic oscillator into action, angle variables and obtain Bell's result[2]:

$$q = \sqrt{2J}\ (k/m)^{1/4} \cos\psi \tag{29}$$

$$p = -\sqrt{2J}\ (m/k)^{1/4} \sin\psi \tag{30}$$

$$J = \text{constant} \tag{31}$$

$$\dot{\psi} = \partial H/\partial J = \omega = 1/\sqrt{mk} \ . \tag{32}$$

Having flexed our muscles in this case we can move to the real problem of transforming Hill's equation. This is a general form of a harmonic oscillator and it is not surprising therefore to find in Ref. 4 Eq. 4.51 that the generating function has a form not unlike Eq. 28:

$$F_1(y,\psi,s) = -\frac{y^2}{2\beta}\left[\tan\psi - \frac{\beta'}{2}\right] \ . \tag{33}$$

We remember that β is a function of s.

The rules for applying this function are:

$$p = \frac{\partial F_1}{\partial y} \ , \quad J = -\frac{\partial F_1}{\partial \psi} \ , \tag{34}$$

and the new Hamiltonian

$$K = H + \frac{\partial F_1}{\partial s} \quad . \tag{35}$$

The full transformation requires a number of clever substitutions to obtain J, ψ as explicit functions of the old coordinates and K as a function of the new ones. The student may find it instructive to work out the Eqs. 34. The first yields an expression for a $\tan\psi$ which can be substituted in the second to obtain:

$$2J = \frac{1}{\beta}\left[y^2 + \left(\beta y' - \frac{\beta' y}{2}\right)^2\right] . \tag{36}$$

This will be recognized by anyone familiar with Courant and Snyder[5] as an invariant of particle motion around a synchrotron. Even though β is a function of s, J is constant. If J is the invariant for the largest amplitude particle in a beam 2J is just the emittance, E, which includes the beam. πE or $2\pi J$ is the area of the phase space ellipse anywhere in the ring.

We remember that the action and angle transformation[2] of a harmonic oscillator also gave an invariant $2\pi J$ which was the area of the phase contour.

If we were to go through the full transformation procedure (see Ref. 2 again) we would find that the new Hamiltonian is:

$$K = J/\beta(s) \ , \tag{37}$$

which should be compared with

$$K = J\omega \tag{38}$$

for an harmonic oscillator. The analogy between frequency and $1/\beta(s)$ is apparent. Hamilton's Equations give:

$$\frac{d\psi}{ds} = \frac{\partial K}{\partial J} = \frac{1}{\beta(s)} \quad . \tag{39}$$

ψ is none other than the familiar betatron phase advance and we are not surprised to find:

$$y = \sqrt{2J\beta}\,\cos\psi \tag{40}$$

$$y' = -\sqrt{2J/\beta}\,[\sin\psi - (\beta'/2)\cos\psi] \ . \tag{41}$$

We have arrived at an invariant momentum, J, conjugate to ψ but the Hamiltonian K still varies like $1/\beta(s)$ around the ring. To finish the job of freezing out the linear motion, we need to transform into a new coordinate system (J_1,ψ_1). (Note we use suffix 1 to indicate new coordinates. We shall continue the policy of adding one to this suffix each time we transform). The new and old coordinates are related by a generating function of the second kind. This kind of generating function includes rotations in space and does not mix momenta and coordinates as does F_1:

$$F_2(\psi,J_1) = J_1\left[\frac{2\pi Qs}{C} - \int_o^s \frac{ds'}{\beta}\right] + \psi J_1 \tag{42}$$

$$\Psi_1 = \frac{\partial F_2}{\partial J_1} = \psi + \frac{2\pi Qs}{C} - \int_o^s \frac{ds'}{\beta(s')} \tag{43}$$

$$J = \frac{\partial F_2}{\partial \psi} = J_1 \tag{44}$$

where

$$2\pi Q = \int_o^{2\pi R} \frac{ds}{\beta} \quad \text{and} \quad C = 2\pi R \ . \tag{45}$$

The new Hamiltonian is related to the old one of Equation 37 by:

$$K_1 = K + \frac{\partial F_2}{\partial s} = \frac{2\pi Q}{2\pi R} \cdot J_1 = \frac{Q}{R} \cdot J = \text{constant} \ . \tag{46}$$

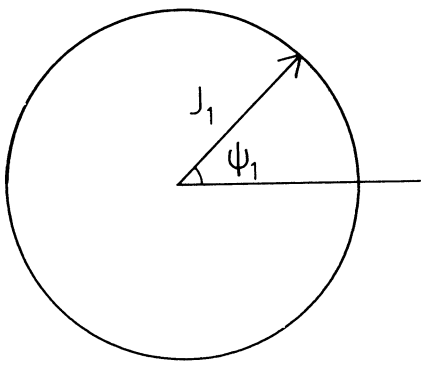


Fig. 2 Phase space trajectory of a linear system in action angle coordinates.

We have now reached our aim insofar as the linear dynamics is concerned. In the new coordinates (J_1,ψ_1) the trajectory of the particle is a circle of radius J which is numerically equal to half the emittance of the beam it includes. The variable ψ goes from 0 to 2π in one betatron oscillation. The Hamiltonian is independent of s (or time). In the next section we shall see how this leaves the way free to express non-linear motion as a perturbation of this circle. The circle will be distorted, its circumference will develop waves and an archipelago of islands will form.

5. PERTURBATION THEORY

With the linear problem behind us we are ready to use perturbation theory to treat the nonlinear effects. The terms in the Hamiltonian (Eq. 14) with a cubic or higher transverse dependence are nonlinear. They arise from sextupole and higher multipole fields distributed in azimuth, s, around the focusing lattice. Unlike the linear terms they retain their s dependence even when we apply the canonical transformation to action-angle variables, J_1 and ψ_1. In order to reveal the way they distort the perfect circle of Fig. 2 we must first apply a transformation which removes the s dependence of the Hamiltonian so that H becomes a new invariant of the motion. The transformation which does this will tell us how much the circle is distorted.

The prescription for finding the canonical transformation which irons out the s dependence is described in Ref. 2 in the Section on Canonical Perturbation Theory. To help the reader relate to Ref. 2 we should point out that we add one to the suffix rather than use a bar to indicate a new set of coordinates .

At the centre of canonical perturbation theory is a generating function of the second type which is an infinitesimal perturbation of the function which generates the identity transformation:

$$F_2(J_2,\psi_1,s) = J_2\psi_1 + \chi(J_1,\psi_1,s) \ , \tag{47}$$

where χ is small.

The new and old coordinates are simply related by the first derivatives of this small quantity, χ.

$$J_1 = J_2 + \frac{\partial}{\partial\psi_1}\,\chi(\psi_1,J_2,s) \tag{48}$$

$$\psi_2 = \psi_1 + \frac{\partial}{\partial J_2}\,\chi(\psi_1,J_2,s) \tag{49}$$

$$H_2 = H_1 + \frac{\partial}{\partial s}\,\chi(\psi_1, J_2, s) \quad . \tag{50}$$

The first term in Eq. 50, H_1, will contain the unperturbed Hamiltonian (Eq. 17) plus a small nonlinear perturbation due to one or more of the cubic or higher terms in Eq. 14. Of course, we must transform this nonlinear term into (J_1,ψ_1) coordinates using substitutions from Eqs. 37, 40, 41, 43, 44 and 46. The nonlinear terms will give it an s dependence.

Our plan of campaign is to choose the function χ in such a way that its differential, $\partial\chi/\partial s$, exactly cancels out the s-dependence making H_2 a constant of the motion.

But how do we choose the exact form of χ? At the moment it is arbitrary. We can first rewrite Eq. 50 to explicitly include the linear part of the Hamiltonian, H_0, and the perturbation U. Both are functions of J_1 but we apply a trick and will write them as functions of $J_2 + \partial\chi/\partial\psi_1$ rather than J_1. This is to make it easier to identify first and second-order terms when we expand the terms:

$$H_2 = H_0\left(J_2 + \frac{\partial\chi}{\partial\psi_1}\right) + U\left(\psi_1,\ J_2 + \frac{\partial\chi}{\partial\psi_1},\ s\right) + \frac{\partial\chi}{\partial s} \quad . \tag{51}$$

We next must eliminate second-order terms and to do this we shall use another trick and introduce three new pairs of terms which cancel. This will help us group all the terms which are either of second order in U or are independent of J, ψ or s in rectangular brackets:

$$\begin{aligned} H_2 = H_0(J_2) &+ \left[H_0\left(J_2 + \frac{\partial\chi}{\partial\psi_1}\right) - H_0(J_2) - \frac{Q}{R}\,(J_2)\,\frac{\partial\chi}{\partial\psi_1}\right] \\ &+ \left[U\left(\psi_1,\ J_2 + \frac{\partial\chi}{\partial\psi_1}, s\right) - U(\psi_1, J_2, s)\right] + \frac{Q}{R}\,(J_2)\,\frac{\partial\chi}{\partial\psi_1} + \frac{\partial\chi}{\partial s} + U(\psi_1, J_2, s). \end{aligned} \tag{52}$$

If we examine the contents of the rectangular brackets very carefully and remember that (Eq. 46):

$$\frac{Q}{R}\,(J_1) = \frac{\partial H_0}{\partial J_1} \quad , \tag{53}$$

we find that the brackets reduce to:

$$\frac{1}{2}\,\frac{\partial QR}{\partial J_2}\left\{\frac{\partial\chi}{\partial\psi_1}\right\}^2 + \frac{\partial U}{\partial J_2}\,\frac{\partial\chi}{\partial\psi_1} \quad , \tag{54}$$

which are purely second order and can be left out to give:

$$H_2 = H_0(J_2) + \frac{Q}{R}(J_2)\frac{\partial \chi}{\partial \psi_1} + \frac{\partial \chi}{\partial s} + U(\psi_1, J_2, s) \ . \qquad (55)$$

At this point we shall digress to attend to a small detail. The independent variable is still s but we shall find it much more convenient as we move into Fourier analysis to use the azimuthal angle $\theta = s/R$. In Ref. 2 it was shown that one may scale a pair of conjugate variables. To preserve the Poincarré invariant we must multiply H_2 the conjugate to s by R:

$$H_3 = RH_2, \quad \theta = s/R \qquad (56)$$

$$\psi_3 = \psi_2 \ ,$$

$$J_3 = J_2 \ .$$

This gives the Hamiltonian of Eq. 56 as a function of θ:

$$H_3(\psi_3, J_3, \theta) = H_0(J_3) + Q(J_3)\frac{\partial \chi}{\partial \psi_1} + R\frac{\partial \chi}{d\theta} + RU(\psi_1, J_3, \theta) \ . \qquad (57)$$

To return to our aim of removing the s dependence in the last three terms of Eq. 57 we can simply impose the condition:

$$Q(J_3)\frac{\partial \chi}{\partial \psi_1} + \frac{\partial \chi}{\partial \theta} + RU(\psi_1, J_3, \theta) = 0 \ , \qquad (58)$$

and H_3 will be independent of s.

If we know the driving perturbation $U(\psi_1, J_3, \theta)$ we can solve the differential equation 58 to find χ.

To find this solution it is convenient, and physically revealing, to analyse U into a Fourier expansion, either a single series in ψ, or a double series in ψ and θ. The single series method is simpler to start with and is appropriate if one is far from a particular resonance condition. One assumes the function χ can be also expanded as a Fourier series and then one solves the differential equation for χ term by term, rather in the manner of solving electrical circuit problems.

6. EFFECT OF NONLINEARITIES FAR FROM A RESONANCE

To first flex our muscles in perturbation theory we solve the problem far from resonance. By far from resonance we mean that the quantity, mQ-n, is large compared with the width of the nearest resonance. We express the perturbation as a single Fourier series in ψ, the "betatron" variable. The index of the summation is the multipole number, 3 for sextupole, 4 for octupole, etc.

$$U = \sum_{n=0}^{\infty} U_n(J_3,\theta)e^{in\psi_3} \quad . \tag{59}$$

The solution to (58), χ, will also be a Fourier series:

$$\chi = \sum_{n=0}^{\infty} \chi_n(J_3,\theta)e^{in\psi_3} \quad , \tag{60}$$

and by substitution in (58) we find:

$$\left[inQ(J_3) + \frac{\partial}{\partial\theta}\right] \chi_n = -RU_n \quad . \tag{61}$$

This has the solution:

$$\chi_n = \frac{i}{2\ \sin\pi nQ} \int_{\theta}^{\theta+2\pi} e^{inQ(\theta'-\theta-\pi)}U_n(\theta')d\theta' \quad . \tag{62}$$

The function χ formed by summation (Eq. 60) will then remove the s dependence of the Hamiltonian and make the new Hamiltonian a constant of the motion:

$$H_3 = H_0(J_3) + \sum_n RU_n(J_3,\theta)e^{in\psi_3} \quad . \tag{63}$$

How do we use this function χ which magically renders the Hamiltonian invariant in order to deduce the distortions it makes to the simple linear circle, $J_1(\psi_1)$? We can show that not only is the new H_3 independent of θ but that J_3 is invariant to first order. Equation 53 reminds us that the rate of change of ψ_1 is just Q/R, a constant. It can be shown that the first-order effect of the perturbation is simply to add a small constant term to Q so that, to first order, J_3 is an invariant. We must now look carefully at Eq. 48. The new J is invariant and so the old J_1 is defined as a function of azimuth by the derivative of χ. If we know χ from computation of Eq. 62 we

can find this derivative as a function of ψ_1 and trace out the perturbed $J_1\psi_1$ circle.

7. RESONANCES

Close to a resonance we must expect to have to take into account particular harmonics of the error distribution which act coherently with betatron motion. As a first step we assume that the perturbation is periodic in both ψ and θ and expand as a double Fourier series:

$$U = \Sigma\, U_{nm}(J_2)e^{i(n\psi_2 - m\theta)} \tag{64}$$

in which the coefficients are:

$$U_{nm} = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_{\psi_2}^{2\pi+\psi_2} U(\psi J_2, \theta) e^{i(n\psi - m\theta)} d\theta d\psi \quad . \tag{65}$$

The treatment follows a parallel path to that of Section 6 and we arrive at the expression for χ and H:

$$\chi = i \sum_{m,n} \frac{RU_{nm}}{nQ - m} (J_2)\ e^{i(n\psi_2 - m\theta)} \tag{66}$$

$$H_3 = H_0(J_3) + RU_{nm}\ e^{i(n\psi - m\theta)} \quad . \tag{67}$$

Again we may use χ to compute the nonlinear distortion of the circular trajectory but close to the resonant condition $nQ = m$ the dominator becomes small. Therefore, we begin to suspect the validity of our first-order approximation and in particular our argument that H_3 was essentially independent of explicit J_3 dependence.

Fortunately, if we zoom in on the one resonant term (which we are justified in doing when we are close to $mQ = n$) there is a transformation we can apply which gives an H_4 which is the exact invariant of the motion and which yields frozen contours in phase space. We no longer need to apply Eq. 48 assuming J_3 = constant since the equation of H_4 itself defines a contour or trajectory in phase space. This alternative technique of finding the trajectory only works if we drop all the Fourier terms except the resonant combination, n,m.

Starting from Eq. 67 we apply an F_2 type of generating function

$$F_2(\psi_3, J_4, \theta) = \left(\psi_3 - \frac{m}{n}\,\theta\, J_4\right) \tag{68}$$

to find

$$J_3 = J_4 \tag{69}$$

$$\varphi = \left(\psi_3 - \frac{m}{n}\,\theta\right) \tag{70}$$

$$H_4 = \left(Q - \frac{m}{n}\right)J_4 + RU_{00}(J_4) + RU_{mn}\, e^{in\varphi} \quad . \tag{71}$$

The new angle coordinate means that our reference frame rotates so that $\varphi = 0$ advances n turns while ψ advances m oscillations. This has the effect of "freezing" the resonance in our new phase space. The result is a Hamiltonian which is independent of time and which defines a closed contour in phase space. We shall use this to explore the examples of third and fourth-order resonance before generalizing the theory and moving on to explain invariants and islands in phase space.

We have finished the bulk of the Hamiltonian formalism so a little recapitulation is perhaps in order.

There are two approaches to finding out how nonlinear terms distort the simple circular trajectory due to linear focusing in (H_1, ψ_1) space. The first, outlined in Section 5 uses first order perturbation theory to find a generating function, χ assumed to be a single Fourier series related to the azimuthal pattern of the perturbation. Once found, the generating function tells us the distortion of the circle provided it is small. The second approach is to use a double Fourier expansion and then change the coordinate ψ so that it keeps pace with the resonant oscillations at a stopband: $m\psi = n\theta$. This gives an exact invariant Hamiltonian which itself defines the path in phase space but ignores all but the resonant perturbation.

8. THE THIRD-INTEGER RESONANCE

One of the confusing aspects of a generalized description involving a number of transformations is that we tend to lose track of the numerical relation between the final coordinates and the initial physical system. In this example we try to give the reader the link between these coordinate frames.

Let us now take a practical example of a sextupole-driven resonance. We suppose that the Hamiltonian expressed in "normal" coordinates includes a sextupole field:

$$H = \frac{p^2}{2} + \frac{k(s)}{2} x^2 + \frac{1}{3!(B\varrho)} \frac{\partial^2 B_z}{\partial x^2} x^3 \quad . \tag{72}$$

First we must transform into action angle coordinates (J,Ψ) applying Eqs. 37, 40 and 41 and then, to arrive in a coordinate system in which the Hamiltonian of a linear system would be a constant of the motion, we apply Eqs. 43, 44 and 46. The final coordinates are (J_1,ψ_1) and the Hamiltonian, H_1:

$$H_1 = \frac{Q}{R} J_1 + \frac{[2J_1\beta(s)]^{3/2}}{3!(B\varrho)} \frac{\partial^2 B_z}{\partial x^2} \cos^3\psi_1 \quad . \tag{73}$$

The new angle variable, ψ_1, is defined:

$$\left.\begin{aligned} \psi_1 &= \psi - \left[\int_0^s \frac{ds'}{\beta(s)} - \frac{Q}{R} s \right] \\ \text{where} \qquad \psi &= \int_0^s \frac{ds}{\beta(s)} \quad . \end{aligned}\right\} \tag{74}$$

We recall that the old coordinates are x and x' related to J and ψ via:

$$\left.\begin{aligned} x &= \sqrt{2J\beta(s)}\, \cos\psi \\ p = x' &= -\sqrt{2J/\beta(s)}\, \sin\psi \quad . \end{aligned}\right\} \tag{75}$$

We can now make use of the purely trigonometrical relation:

$$\cos^3\psi_1 = \frac{1}{2^{(3-1)}} (\cos 3\psi_1 + 3 \cos\psi_1) \quad . \tag{76}$$

Ignoring the second term which in general does not drive a one-third-integer resonance:

$$H_1 = \frac{Q}{R} J_1 + \left[\frac{2^{3/2}}{2^{3-1}} \cdot \frac{J^{3/2}\beta(s)^{3/2}}{3!\ (B\varrho)} \cdot \frac{\partial^2 B_z(s)}{\partial x^2} \right] \cos 3\psi_1 \quad . \tag{77}$$

The second term in this equation is the perturbation, U, in Eqs. 55 and 58 and, applying directly the result of perturbation theory at a resonance (Eq. 71) we have:

$$H_4 = \left(Q - \frac{m}{3} \right) J_4 + RU_{3m} \cos 3\varphi \quad . \tag{78}$$

The U_{00} term in Eq. 71 is zero for m = odd multipoles. Here U_{3m} is the double Fourier coefficient (with respect to ψ and θ) of $U(\psi,\theta)$ and written explicitly is:

$$U_{3m} = \frac{J^{3/2}2^{3/2}}{3!(B\varrho)2^{3-1}} \left\{ \frac{1}{2\pi R} \int_0^{2\pi R} \beta^{3/2} B'' e^{i\int_0^s 3[(1/\beta)-(Q/R)]ds'} e^{-ims/R} ds \right\}. \quad (79)$$

In addition φ is given by Eq. 70.

The integral exponent merely takes care of the relative phase slip of betatron motion with respect to $Q\theta$ which occurs between quadrupoles of a regular lattice. In many cases it may be ignored.

Note too that the content of the curly brackets is none other than Guignard's[6] d_p. We have left the powers of 2 and factorials in their crude state from which one may generalise for any multipole with 2m poles:

$$U_{nm} = \frac{J^{n/2}}{n!(B\varrho)2^{n-1}} \left\{ \frac{1}{2\pi R} \int_0^{2\pi R} \beta^{n/2} B^{(n-1)} e^{i\int_0^s n[(1/\beta)-(Q/R)]ds'} e^{-ims/R} ds \right\}. \quad (80)$$

Having computed U_{nm} numerically, we can substitute back in Eq. 71 to find the Hamiltonian for any desired order of resonance. It is worth remembering that the expression for U_{nm} becomes much simplified when one does the Fourier analysis for a single short sextupole or a random distribution. Guignard[6] gives expressions for d_p, the curly brackets for this latter case.

9. THE TRAJECTORY OF A THIRD-INTEGER RESONANCE

All this has become very mathematical and the reader may be forgiven for a little impatience to see some phase plots showing the perturbed trajectories. In order to find the shape of these phase plots we must take a hard look at Eq. 78. We shall drop the suffixes from the notation and examine the contour defined by the Hamiltonian of Eq. 78

$$H_4 = \left(Q - \frac{m}{3}\right)J_4 + RU_{3m}\cos 3\varphi \ . \quad (81)$$

Rewriting it to show the J dependence and defining constants $\delta = Q-(m/3)$ and $\varepsilon = RU_{3m}/J^{3/2}$ to obtain a streamlined:

$$H = \delta J + \varepsilon J^{3/2}\cos 3\varphi \ . \quad (82)$$

If the coefficient of the second term is small, either because the amplitude is small or the driving strength is weak, the contour will be close to a circle (J = constant). The same will be true if the Q is distant from a third integer making, δ, the coefficient of the first term large.

If the converse is true the circle will be distorted inwards when $\varphi = 0$, $2\pi/3$, $4\pi/3$ and outwards inbetween to become somewhat triangular. Note that this triangle does not rotate but is frozen in (J,φ) space - a consequence of the change to a rotating coordinate system. Figure 3 shows constant H contours in J,ψ space. We can see this kind of distortion growing with amplitude.

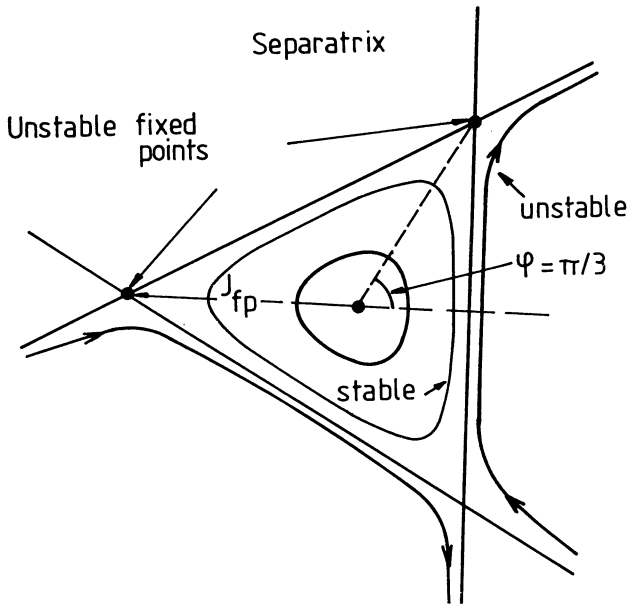


Fig. 3 Phase space plot in (J,ψ) coordinates of a third-integer resonance (adapted from Reference 3).

We remember from Hamilton's Equations that:

$$\frac{dJ}{ds} = -\frac{\partial H}{\partial \varphi} = 3\varepsilon J^{3/2} \sin 3\varphi \qquad (83)$$

$$\frac{d\varphi}{ds} = \frac{\partial H}{dJ} = \delta + \frac{3}{2}\varepsilon J^{1/2} \cos 3\varphi \ . \qquad (84)$$

When both these expressions are zero a particle will stagnate in phase space lying on a "fixed point". For this to be the case Eq. 83 demands:

$$\sin 3\varphi = 0 \ , \qquad (85)$$

so that $dJ/ds = 0$. Then we can write $\cos 3\varphi = \pm 1$ in Eq. 84 and ask how this can be zero. Above resonance, the sign of δ is positive and provided ϵ is positive $d\varphi/ds$ can only be zero if:

$$\cos 3\varphi = -1 \ . \tag{86}$$

These conditions on ψ define three fixed points at

$$\varphi = \pi/3, \ 3\pi/3 \text{ and } 5\pi/3 \ . \tag{87}$$

From Eq. 84 we can also find their amplitude:

$$J_{fp} = (2\delta/3\epsilon)^2 \ . \tag{88}$$

Inside this amplitude the triangular trajectory becomes less and less distorted and more circular as one approaches $J = 0$. Outside, the trajectories diverge in unbounded motion towards infinite amplitude.

The value of Q in Eq. 33 is the unperturbed Q for zero amplitude particles which corresponds to the centre of the diagram. The betatron wave number increases with J and becomes exactly one third integer at the fixed points. This difference in Q

$$\delta = Q - (m/3) \tag{89}$$

is just the stopband width for a particle of amplitude J. Any particle with an unperturbed Q closer to the third integer than this will be already unstable if its amplitude is J. Note that if we tune the Q closer to m/3 or increase the strength of ϵ the stable triangle where motion is bounded will shrink and expell particles. This is the principle of third-integer ejection.

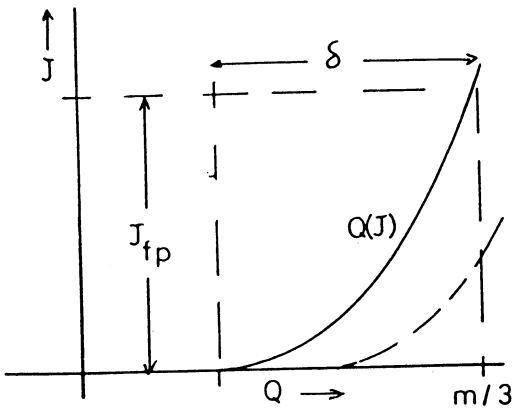


Fig. 4 The variation of Q as a function of amplitude close to a third-integer resonance.

10. THE EFFECT OF AN OCTUPOLE

This section is not merely a repetition of the previous one for a different multipole. An octupole leads to an extra term in the invariant Hamiltonian which corresponds to a nonlinear variation of frequency with amplitude. This is characteristic of "even" multipoles and defines quite a different topology of phase space.

First we must write down the Hamiltonian in "physical" coordinates:

$$H = \frac{p^2}{2} + \frac{1}{2} k(s)x^2 + \frac{1}{4!(B\rho)} \frac{\partial^3 B_z}{\partial x^3} x^4 \quad . \tag{90}$$

where k(s) is defined in Eq. 18.

After transforming to action-angle coordinates:

$$H_1 = \frac{Q_0}{R} J + \frac{(2J\beta)^{4/2}}{4!(B\rho)} B^{(3)} \cos^4\psi \ , \tag{91}$$

where

$$B^{(3)} = \frac{\partial^3 B_z}{\partial x^3}\bigg|_{\substack{x=0 \\ z=0}} \ . \tag{92}$$

Again we expand

$$\cos^4\psi = \frac{1}{2^4}\left(\cos 4\psi + 4\cos 2\psi + \frac{4!}{[(4/2)!]^2}\right) \ . \tag{93}$$

The first term in Eq. 93 will resonate at quarter integer Q values, the second at half-integer values and the third adds a U_{00} term to the Fourier analysis of the perturbation. It is this that gives the amplitude variation of Q.

We ignore the half integer term for the purpose of this discussion, but we are interested in the U_{00} term. This was missing in the case of the third (and all odd) order resonances. Equation 71, the general form of the Hamiltonian for a resonance

$$H = \left(Q - \frac{m}{n}\right)J + RU_{00}(J) + RU_{nm} \cos n\varphi \ , \tag{94}$$

may be differentiated to find an average shift in tune:

$$\frac{d\psi}{d\theta} = \frac{\partial H}{\partial J} = \left(Q - \frac{m}{n}\right) + \frac{R\partial U'_{00}(J)}{\partial J} \quad . \tag{95}$$

For the fourth-integer case we find ΔQ_{NL}, the perturbed second term, is protortional to J:

$$\Delta Q_{NL} = \frac{4!\ R}{2^4[(4/2)!]^2} \cdot \frac{J\beta}{3!\ B\varrho}\, B^{(3)} \tag{96}$$

and for other even orders of n we can find this nonlinear time shift Q_{NL} by differentiating:

$$U_{00} = \frac{J^{n/2} n!}{2^n[(n/2)!]^2} \cdot \frac{1}{\pi R} \int_0^{2\pi R} \beta^{n/2} B^{(n-1)} ds \quad . \tag{97}$$

11. PHASE-SPACE TOPOLOGY WITH AMPLITUDE FREQUENCY VARIATION

The trajectories discussed for the third-integer case had no term equivalent to $RU_{00}(J)$ in Eq. 14. Let us explore the topology of phase space trajectories with a generalised invariant Hamiltonian of the same form as Eq. 82 but including the nonlinear shift $\alpha(J)$:

$$H = \delta J + \alpha(J) + \varepsilon J^{n/2} \cos m\varphi \quad . \tag{98}$$

We remember that the perturbed Q is:

$$Q = \frac{\partial H}{\partial J} = \delta + \alpha'(J) + \frac{n}{2} \varepsilon J^{(n/2)-1} \cos m\varphi \quad . \tag{99}$$

The first two terms cancel on average when J has a resonant value J_r defined by the condition:

$$\alpha'(J_r) = -\delta \quad . \tag{100}$$

If we look at small changes in J about J_r by a second differentiation we find:

$$(J - J_r) = -\frac{n}{2} \frac{\varepsilon}{\alpha''} J_r^{(n/2)-1} \cos m\varphi \quad . \tag{101}$$

We now have stable and unstable fixed points at $\cos m\varphi = -1$ and $+1$ respectively and separated in J by:

$$\frac{n\varepsilon}{\alpha''} J_1^{(n/2)-1} \quad . \tag{102}$$

The existence of stable fixed points in the real plane is new and is only the case for even-order resonances. We find that they are the centres of an archipelago of stable islands. Trajectories circulate around these points within the islands. The unstable fixed points are at the junction of the islands and the diagram is identical to a series of RF buckets plotted in polar coordinates with a harmonic number equal to m.

The width of these islands may be calculated and is found to be:

$$\Delta J_W = 2 \sqrt{\frac{\varepsilon (J_u)^{n/2}}{\alpha''}} \quad . \tag{103}$$

where J_u is that of an unstable fixed point.

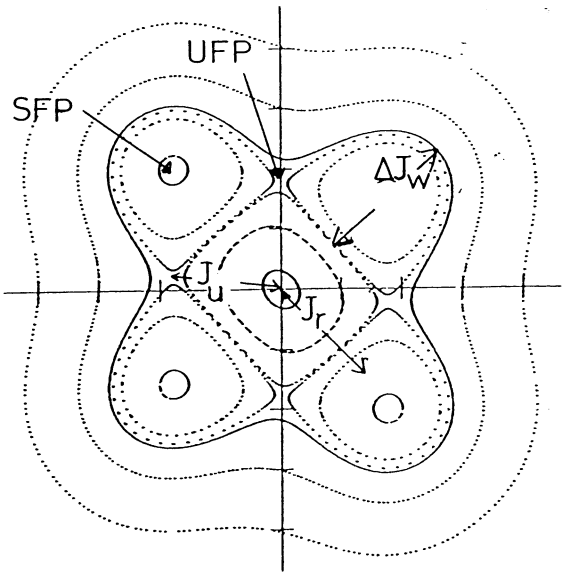


Fig. 5 The effect of an octupole near a fourth-order resonance (adapted from Reference 7).

Motion will remain contained within the islands unless there is some way in which particles may "leak" out. One mechanism which may cause leakage is the presence of another archipelago which overlaps. This is said to cause chaotic behaviour. The next set of islands of comparable order will correspond to a change of Q of the order of 1/n where n is the order of the resonance. Since α'' controls $\partial Q/\partial J$ the spacing in phase space is:

$$\Delta J_S = \frac{1}{\alpha''} \cdot \frac{1}{n} \quad . \tag{104}$$

We can expect these archipelagos to overlap and chaos to ensue when:

$$\Delta J_W = \Delta J_S \tag{105}$$

$$4\varepsilon J^n \geq \frac{1}{n^2} \cdot \frac{1}{\alpha''} \quad . \tag{106}$$

This is the Chirikov limit and normally does not happen in any machine built with reasonable care. Unfortunately, even in a clean machine, synchrotron motion or magnet ripple can produce sidebands to the main resonance and increases the number of resonance lines each of which produces its own set of islands. The chaotic condition is much easier to reach as phase space becomes much more crowded.

12. AMPLITUDE GROWTH ON CROSSING A RESONANCE

So far we have described the dynamics of a machine in which neither the magnetic field nor the momentum of the particles change with time. Many of the less welcome effects of nonlinear fields only become apparent when either the magnetic field changes because of ripple in the power supply or when the particle's momentum changes sinusoidally as it executes synchrotron oscillations in an RF bucket. It is a relatively common experience to find that an injected beam survives until the accelerating cavities are switched on and then is extracted by a nonlinear resonance. The explanation usually given is that oscillations in radial position together with the chromatic properties of the lattice cause the particle to cross and recross the resonance.

One way to see what happens in the time domain when a crossing occurs is derived by Guignard[6] . We will not attempt to give his derivation here but only his result. Suppose we cross a resonance (order, m) at a speed of ΔQ_t per turn. The equation of growth is:

$$\int_0^x \frac{dx}{x^{n-1}} = \Delta Q_e \frac{2\pi}{(n|\Delta Q_t|)^{1/2}} \quad , \tag{107}$$

where:

n = order of the resonance,
ΔQ_e = half width of the resonance.

On integration this gives (for one crossing):

$$\left(1 - \frac{x_0}{x}\right)^{n-2} = \frac{2\pi Q_e}{(n|\Delta Q_t|)^{1/2}} \quad . \tag{108}$$

One may go on and calculate what happens if the time is modulated due to synchrotron oscillations:

$$Q = Q_0 + \hat{Q} \sin(Q_S \Omega t) \ , \tag{109}$$

where Ω is the revolution frequency.

However, it is more direct at this point to move into the frequency domain and analyse synchrobetatron resonances.

13. SYNCHROTRON RESONANCES

It is a fact well known to electron accelerator physicists that each nonlinear resonance has a series of satellite lines, parallel to it and spaced by Q_S/n on either side of the resonance, where Q_S is the synchrotron wave number or number of synchrotron oscillations per turn.

If we return to the Hamiltonian for a nonlinear resonance we find the perturbation term is always proportional to:

$$\cos(n\psi - m\theta) \ , \tag{110}$$

where:

$$d\psi/d\theta = Q,$$
$$\psi \quad = \int Q \, d\theta .$$

Now if Q is modulated by synchrotron motion:

$$Q = Q_0 + \hat{Q} \sin Q_S \theta \tag{111}$$

then

$$\psi = \psi_0 - \frac{\hat{Q}}{Q_S} \cos Q_S \theta \ , \tag{112}$$

and the perturbation term in the Hamiltonian will change:

$$\cos m\psi \rightarrow \cos\left[n\psi_0 - \left(\frac{n\hat{Q}}{Q_S}\right)\cos Q_S\theta\right] \tag{113}$$

$$= \sum_K J_K\left(\frac{n\hat{Q}}{Q_S}\right)\cos[n\psi_0 + kQ_S\theta] \ .$$

Each value of k corresponds to a different satellite resonance line. The resonant condition becomes:

$$Q = \frac{m}{n} + \frac{kQ_S}{n} \quad , \quad k = 0 \pm 1, 2, 3, \ldots \tag{114}$$

The strength of each line is determined by:

$$J_k(n\hat{Q}/Q_S) \ . \tag{115}$$

Figure 6 shows how the strength will vary with the order $k = n\Delta Q/Q_S$ where k is the distance from the resonance in units of Q_S/n and with the depth of modulation $n\hat{Q}/Q_S$ expressed in the same units. In general the effect will be worst when $\Delta\hat{Q} = Q_S$, $2Q_S$, $3Q_S$, etc. and will drop to 30% when either ΔQ or $\Delta\hat{Q}$ becomes larger than $10Q_S/n$. The suffix ν in the plot is just k.

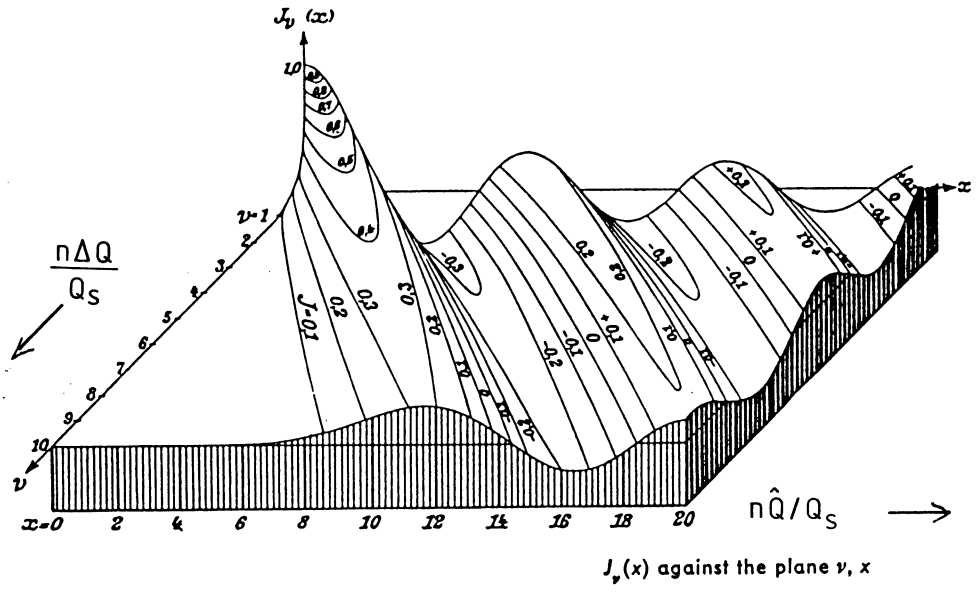


Fig. 6 The variation of the Bessel function with distance from a resonance and with depth of modulation (with acknowledgement to Jahnke, Emde and Lösch, Tables of Higher Functions, pub. B.G. Teubner, Stuttgart).

This is nevertheless a rather considerable extension of the influence of a single nonlinear resonance. In the worst case, when the depth of modulation is about $10Q_s$. The single line is split into about ten other lines of comparable strength.

Each satellite line generates its own set of islands in phase space, so that spacing can be reduced by an order of magnitude. Each resonance is narrower by $J_k(n\hat{Q}/Q_s)$ but as Fig. 6 shows this is a factor of the order of 1/3. The consequence is that the Chirikov threshold is reduced by a factor of 3 or so.

14. BEAM LIFETIME DUE TO MAGNET IMPERFECTIONS

We are at the end of a long road and in the last few miles have abandoned some of the luggage containing the rigour of mathematics to speed to our goal. Our conclusion is not to present a neat analytic expression into which one may plug field imperfections and obtain beam lifetime but to outline a method to arrive at this result. What is offered is a means to bypass the perhaps prohibitively expensive computer simulations needed to prove stability.

Previous attempts to predict lifetime on the basis of nonlinear field content have often foundered by including neither the concept of overlapping islands nor the catalytic effect of ripple or synchrotron motion. When both of these ingredients are included it turns out that the nonlinear detuning, α, plays a crucial role in determining machine performance. This detuning may be due to effects other than the nonlinearity studied like beam-beam or ion-beam effects.

The scenario then is as follows:

i) Choose an order of resonance too high to be avoided by tuning Q.
ii) Calculate the nonlinear detuning due to ALL forms of nonlinearities.
iii) Invoke Q_s to find the spacing and strength of the resonance lines.
iv) Calculate the Chirikov limit, and if it is exceeded go on to calculate diffusion rates.
v) Calculate the growth per crossing.
vi) Assume crossings are uncorrelated because of Qs.
vii) Sum the growth rate in random walk fashion to get a diffusion rate.

If this method proves to be effective one would expect a lifetime of 24 hours for a beam-beam effect in the CERN $Sp\bar{p}S$ storage ring[7], ten minutes for ion-induced resonances in the CERN Antiproton Accumulator (Ref. 8), and perhaps only a few tens of seconds for a large ring with inadequate power

supply filtering[9] . The experience found in the Refs. 7 to 9 suggests this may not be far from the truth.

Those who are already expert in the field may scoff at an attempt to describe beam diffusion ignoring the second transverse degree of freedom which changes the topology of the problem. Nevertheless, the numbers seem to fit and anyone with ambitions to include both transverse planes will find the next Section contains the prescription for this.

15. THE EFFECT OF TWO DIMENSIONS OF TRANSVERSE MOTION

The linear motion in one transverse degree of freedom is a circle in (J_1,ψ_1) space. The original Hamiltonian only included terms proportional to p^2 and to x^2 coming from the transverse distribution of A_S for a multipole (Eq. 10) remembering that $m = 2$ for a quadrupole and a normal orientation gives only the imaginary terms. We find:

$$H = \frac{p_x^2}{2} + \left(\frac{1}{\rho^2} - k\right)\frac{x^2}{2} + \frac{p_z^2}{2} + k\,\frac{z^2}{2} \quad . \tag{116}$$

We may transform these coordinates into action angle coordinates and remove the s dependence of H, J_x and J_z just as in the one dimensional case. In the case of one degree of freedom, plus s the azimuth, the particle was confined to a torus. Slicing the torus at any azimuth gave a circle. With two degrees of freedom the "slice" will produce a four dimensional surface which becomes a circle when projected on either (J_x,Ψ_x) plane or (J_z,Ψ_z). Either J_x or J_z is a horizontal plane when plotted against (Ψ_x,Ψ_z) in a "hill and dale" fashion.

Now we move on to include other multipoles. Equation 10 tells us we should include in the Hamiltonian all the terms in the homogeneous polynomial expansion of:

$$(x + iz)^n \quad . \tag{117}$$

We should select only real terms if the 2n poled magnet is in normal orientation and only imaginary terms if the magnet is skew. Thus a normal sextupole will give two terms:

$$x^3 - 3xz^2 \tag{118}$$

while a skew sextupole will give:

$$3x^2z - z^3 \quad . \tag{119}$$

Now we can apply all the tricks of perturbation theory to be found in Sections 5 to 7. We simply must write two element vectors for Q, Ψ, J and n taking dot products to form terms like $\bar{n}\cdot\bar{\Psi}$. One element is for the x plane, the other for the z plane.

When we come to Fourier analysis there will be two indices, n_1 for the phase angle Ψ_x and n_2 for Ψ_z . The denominator of Eq. 66 which defines a resonant condition becomes:

$$n_1Q_x + n_2Q_z - m = 0 \quad . \tag{120}$$

The resonances determined by this condition are the familiar mesh of lines in the working diagram (Fig. 7). If m_1 and m_2 have the same sign they are sum resonances and if they are different, they are difference resonances.

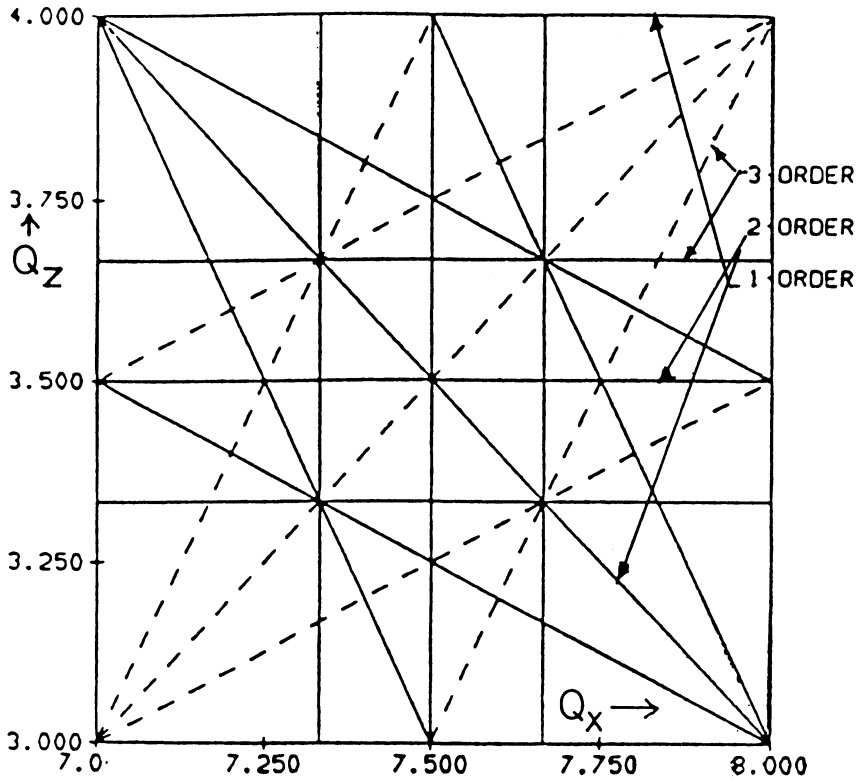


Fig. 7 Typical working diagram with sum and difference resonances to third order (Ref. 3).

The Hamiltonian, truncated to include just the resonant term (Eq. 67) will become:

$$H_3 = H_0(\bar{J}_3) + RU_{\bar{n},m}e^{i(n_1\Psi_x+n_2\Psi_z-m\theta)} \quad . \tag{121}$$

Defining:

$$n_1\varphi_x = n_1\Psi_x + n_2\Psi_z - m\theta \tag{122}$$

we can freeze the motion with a generating function reminiscent of Eq. 68 but with an extra term.

$$F_2(\bar{\Psi},\bar{J}_4,\theta) = (\bar{n}\cdot\bar{\Psi} - m\theta)J_{4,x} + \Psi_z J_{4,z} \quad . \tag{123}$$

We now find a new Hamiltonian:

$$H = (\bar{n}.\bar{Q} - m)J_x + Q_z J_z + RU_{00}(J) + RU_{\bar{n}m}\,\mathrm{cos}n\varphi_x \quad . \tag{124}$$

We also find on making the transformation from (J_3,ψ_3) to (J_4,φ) that:

$$\varphi_z = \Psi_z \tag{125}$$

$$J_{4,x} = n_1 J_{4,x} \tag{126}$$

$$J_{3,z} = n_2 J_{4,z} + J_{4,z} \quad . \tag{127}$$

The invariants of the motion whose conjugates are absent from the Hamiltonian are:

$$H_4 = \text{constant} \tag{128}$$

$$J_{4z} = J_{3z} - (n_2/n_1)J_{3x} \quad . \tag{129}$$

Only opposite signs of n_1 and n_2 will ensure stability. Thus difference resonances are stable but can couple the two planes while sum resonances are unstable.

The motion is difficult to plot or imagine. The exchange of J components ensures that particle simulation plotted on a single phase plane will not have a continuous path but appear as a halo of dots as J breathes. We no longer see undulations around a closed circle. However, Ruth[3] has suggested that if one plots either J_x or J_z as a function of Ψ_x and Ψ_z to form a mountain range plot, one reveals the modes of distortion characteristic of the indices m_1 and m_2 (see Fig. 8).

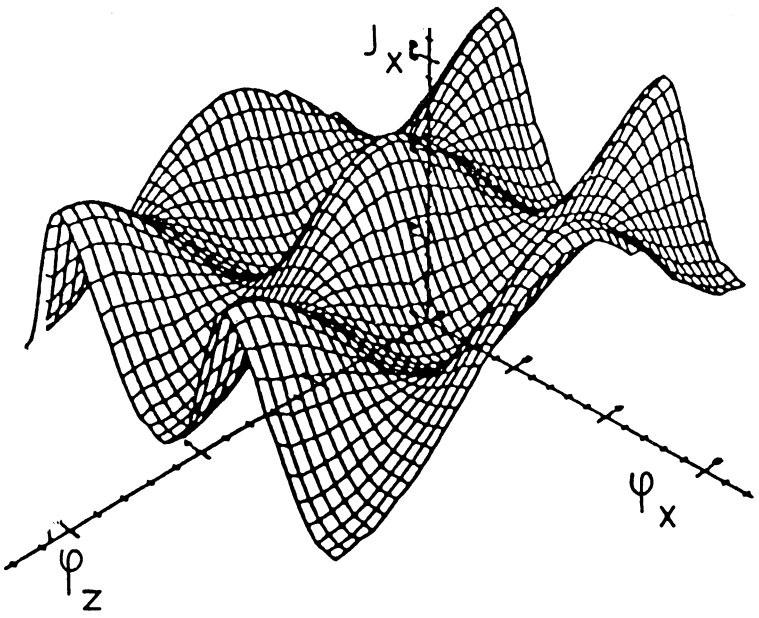


Fig. 8 A plot showing an invariant surface (H = constant) as a function of phase variables in both transverse planes (Ref. 3).

16. THREE DIMENSIONS OF MAGNETIC FIELD

At the outset I warned that some of the approximations made to the Hamiltonian in Section 2 might need to be reconsidered for small rings where the magnets are short. In this Section we discuss why this is so.

We are all taught very early in our initiation to magnets and accelerators that the field measured by a long coil passing right through the magnet

$$\int_{-\infty}^{\infty} B d\ell$$

will satisfy Laplace's equation in two dimensions and hence the particle will see an integral as if there was merely an A_s component. One would not therefore expect any terms in the Hamiltonian other than those of Eq. 10 even the magnet was short and consisted almost entirely of end field.

Two things spoil this simple picture. The first is that the suffix, s, indicates a direction along the particle trajectory and not parallel to the axis of the multipole field. Strictly

$$p_s = \bar{p} \cdot \hat{s} \left(1 + \frac{x}{\rho} \right) \tag{130}$$

$$A_S = \bar{A}\cdot\hat{s}\left(1 + \frac{x}{\varrho}\right) . \qquad (131)$$

If the particles trajectory is expected to follow a curved path, as in a bending magnet, we must somehow obtain an A_S (also A_X, A_Z) expression from the purely axial A_Z.

Secondly, the particle may well be focused (or defocused) as it passes through the end field of a magnet, so that the betatron amplitude has a chance to change. If the magnet has an aperture which is comparable to the length we should really examine A_Z at each point in the fringe field in computing the Fourier coefficients driving resonances. An example where this turns out to be important is the end field of a quadrupole. The A_Z component must have the symmetry of a quadrupole in polar coordinates, cos2θ, but it can have a radial variation normally associated with an octupole in the two dimension case. Locally

$$A_S = f(z)r^4 \cos 2\theta \quad . \qquad (132)$$

f(z) reverses sign in the fringe field but if beta has changed in the meantime this would contribute to the Fourier term driving a four order resonance.

I will not attempt a comprehensive treatment of these three-dimensional effects here but merely warn that they should not be forgotten in the design of small rings of large acceptance.

17. CONCLUSIONS

We have covered all the basic nonlinear theory related to single-particle behaviour and in so doing identified a link between the specification of magnetic field quality and beam lifetime. There are also nonlinear forces on the beam due to collisions with an oncoming "target beam" in a collider or, if the particles of the beam are negative, with the space charge field of the neutralizing ion cloud. The potentials due to these effects are usually richer in high-order multipole fields at the beam radius. However, I shall now leave other authors to take up the baton in this relay race from alternating gradient focusing to Chirikov diffusion.

* * *

REFERENCES

1. E.J.N. Wilson, Non-Linearities and Resonances, Proc. of CERN Accelerator School on General Accelerator Physics, Gif-sur-Yvette, CERN 85-19, p. 96 (1985).

2. J.S. Bell, Hamiltonian Dynamics, Proc. of this School.

3. R.D. Ruth, Non-Linear Dynamics in Circular Accelerators and Storage Rings, Proc. of the 5th US Summer School High Energy Particle Accelerators, SLAC (1985).

4. E.D. Courant, R.D. Ruth and W.T. Weng, Stability in Dynamical Systems I, SLAC-PUB-3415.

5. E.D. Courant and Snyder, Theory of the Alternating-Gradient Synchrotron, Ann. Phys. 3 (1958).

6. G. Guignard, Effets des champs magnétiques perturbateurs d'un synchrotron, CERN 70-24, Chapter 4 (1970).

7. L.R. Evans, The Beam-Beam Interaction, Proc. of CERN Accelerator School on Antiprotons for Colliding Beam Facilities, Geneva, CERN 84-15, p. 319 (1984).

8. E. Jones, F. Pedersen, A. Poncet, S. van der Meer and E.J.N. Wilson, Transverse Instabilities due to Beam-Trapped Ions and Charged Matter in the CERN Antiproton Accumulator, Proc. Particle Accel. Conf., Vancouver (1985).

9. E.D. Courant, A Computer Study of Widening of Stopbands by Ripple, CERN Lab.II-DI-PA/Int. 74-14 (1974).

10. R. Stiening and E.J.N. Wilson, Transverse Single Particle Instability in the NAL 500 GeV Accelerator, Nucl. Instrum. Methods, Vol. 121 (1974).

CHROMATIC EFFECTS AND THEIR FIRST-ORDER CORRECTION

Bryan W. Montague
CERN, Geneva, Switzerland

ABSTRACT

The momentum dependence of accelerator properties has become increasingly important with improved performance. With the advent of high-luminosity storage rings the basic concept of chromaticity has had to be generalised to take into account the local chromatic perturbations of the orbit functions. The first-order theory described here enables lattice parameters to be chosen so as to facilitate chromatic correction, and also provides insight into the origin of higher-order effects.

1. INTRODUCTION

Chromatic effects in accelerators arise because the focal properties of magnetic lens systems depend on the particle momentum, in close analogy with the corresponding effects in classical optics. In the early days of A.G. synchrotrons chromaticity, i.e. the variation of betatron tune Q with momentum, was not a matter of serious concern. However, the situation changed with the increase in intensity of accelerated beams, resulting in the appearance of collective instabilities, in particular the head-tail effect, and making necessary the correction of the natural lattice chromaticity. Considerable refinement of correction methods was introduced in the CERN ISR, in order to accommodate high-intensity proton stacks of large momentum spread within the limited space of the tune diagram free of low-order nonlinear resonances.

Up to this stage chromaticity had been considered as a single number, the average of chromatic effects and their corrections around the circumference of a machine. With the advent of high-luminosity electron-positron storage rings using low-β insertions, it became evident that the concept of global chromaticity was inadequate to describe the true situation. In contrast to a regular lattice, where the chromatic effects are reasonably well distributed, the strong quadrupoles of a low-β section generate large local chromatic perturbations of the betatron motion, as well as contributing to the overall chromaticity. Since it is not practicable to correct these chromatic effects close to their source, one is led to study chromatic effects in terms of first-order perturbation of lattice functions for off-momentum particles. First formulated by Zyngier[1], this approach as elaborated by Montague[2] is used as the basis for this lecture.

In Section 2 we discuss the basic principles of chromaticity correction and demonstrate the need for a more general approach in the presence of strong localised sources of chromatic errors. Section 3 describes a first-order chromatic perturbation

formalism which yields a clear physical picture of the effects in one betatron phase plane. In Section 4 this is extended to two dimensions and used to optimise linear parameters of the machine lattice for correcting chromatic errors. A simple concentrated correction scheme is shown to be impracticable and leads to a discussion of the criteria for smooth distributed correction.

2. BASIC IDEAS

The definition of chromaticity Q' used here is :

$$Q' = p_0 \frac{dQ}{dp} = \frac{dQ}{d\delta} \quad , \tag{1}$$

where $\delta = \Delta p/p_0$ is the relative momentum deviation from the nominal p_0. A definition sometimes found in the literature is :

$$\xi = \frac{p_0}{Q_0} \frac{dQ}{dp} = \frac{Q'}{Q_0} \quad .$$

The dispersion, i.e. the variation x_c of the closed-orbit position with momentum δ, is commonly defined in two ways, viz. :

$$D_x = \frac{\partial x_c}{\partial \delta} \tag{2}$$

or

$$\eta_x = \frac{x_c}{\delta} \quad ,$$

and to lowest order $D_x \sim \eta$. The coordinate x is taken to be the horizontal transverse direction, where the dispersion is an intrinsic machine property; vertical dispersion, normally arising only from errors, may be similarly defined with the appropriate coordinate. In earlier times the symbol α_p was used for the dispersion; it fell out of use due to the inconvenience of adding subscripts.

In classical optics, chromatic aberrations of lens systems can be corrected, at least partially, by using multiple elements of glasses with various refractive indices. Unfortunately there is no such equivalent in charged-particle focusing systems produced by Laplacian fields; magnetic quadrupoles have a fixed variation of focusing strength with momentum. From the definition of the quadrupole strength parameter :

$$K = \frac{e}{p} \frac{\partial B}{\partial x}$$

follows immediately the variation of K with momentum :

$$\frac{\Delta K}{K_0} = - \frac{\Delta p}{p_0} = - \delta \ . \tag{3}$$

Using the well-known formula[3] for the first-order tune variation ΔQ due to gradient perturbations $\Delta K(s)$:

$$\Delta Q = \frac{1}{4\pi} \int_0^{2\pi R} \beta(s) \, \Delta K(s) \, ds$$

we obtain with Eq. (1) and (3) :

$$Q' = - \frac{1}{4\pi} \int_0^{2\pi R} \beta(s) \, K_0(s) \, ds \tag{4}$$

where $\beta(s)$ is the betatron envelope function. This gives quite generally the first-order global chromaticity of an uncorrected machine.

2.1 FODO lattice

As an example we take a regular machine consisting of N periods of FODO lattice, each of length L and betatron phase advance μ. In thin-lens approximation, which is usually quite accurate, it is straightforward to calculate the transfer matrix over one period between the mid-points of the quadrupoles, and to equate corresponding elements to those of the Courant and Snyder matrix[3] between symmetry planes :

$$M = \begin{pmatrix} \cos\mu & \beta \sin\mu \\ -\frac{1}{\beta} \sin\mu & \cos\mu \end{pmatrix} \ .$$

Then with quadrupole length $\ell_q \ll L$, one finds :

$$K_0 \ell_q = \frac{4}{L} \sin \frac{\mu}{2} \tag{5}$$

$$\beta_{max} = \frac{L}{2 \tan \frac{\mu}{2} \, (1 - \sin \frac{\mu}{2})} \tag{6}$$

$$\beta_{min} = \frac{L}{2 \tan \frac{\mu}{2} \, (1 + \sin \frac{\mu}{2})} \tag{7}$$

where β_{max}, β_{min} correspond to the matrices taken from F to F and D to D quadrupoles respectively.

Now in thin-lens approximation :

$$[K_0]_F = -[K_0]_D = K_0 > 0 \ ,$$

and, neglecting the weak focusing effect of dipole bending magnets, the integral in Eq. (4) can be replaced by a sum over N identical periods, since $N = 2\pi R/L$. Hence :

$$Q' = -\frac{1}{4\pi} \sum_N (\beta_{max}[K_0]_F + \beta_{min}[K_0]_D)\, \ell_q$$

$$= -\frac{N}{4\pi} \quad (\beta_{max} - \beta_{min})\, K_0 \ell_q \ .$$

Using Eqs.(5), (6) and (7), together with $N\mu = 2\pi Q$, we have :

$$Q' = -\frac{2Q}{\mu} \tan\frac{\mu}{2} \ .$$

We see that for small phase advance $\mu \to 0$, $Q' \to -Q$, and for a typical value $\mu = \pi/2$, we have $Q' = -4Q/\pi$, not very much different. Large, high-energy accelerators and storage rings necessarily have large betatron tunes, in order to keep the aperture requirements within acceptable limits. A similar situation occurs in dedicated synchrotron-light sources, where the particularly strong focusing required to obtain low beam emittance and high brightness results in relatively large values of Q. In both types of machine the natural chromaticity is inevitably large and requires careful correction.

2.2 Local chromaticity correction

The only practical method of correcting chromatic effects makes use firstly of the variation in closed-orbit position with momentum, given by the dispersion D_x in Eq. (2), and secondly of the property of a sextupole lens that particles traversing off-axis experience a quadrupole gradient proportional to the off-axis displacement. The combination of these features results in an effective quadrupole gradient which, in principle at least, can have an arbitrary first-order momentum dependence, and can therefore be used to compensate the natural chromatic properties of the quadrupoles.

From the definition of the sextupole gradient parameter

$$K'_0 = \frac{e}{p_0} \frac{\partial^2 B}{\partial x^2} \ ,$$

one sees that the equivalent quadrupole strength for particles with momentum error δ is :

$$\Delta K = D_x\, K_0'\, \delta$$

which, by analogy with Eqs.(3) and (4), yields the chromaticity integral :

$$Q' = -\frac{1}{4\pi}\int_0^{2\pi R} \beta(s)\,\{K_0(s) - D_x(s)\,K_0'(s)\}\,ds\ . \qquad (8)$$

For the special case of local chromaticity correction we choose :

$$K'_0(s) = \frac{K_0(s)}{D_x(s)} \quad \text{for all } s,$$

which results in the integrand of Eq. (8) vanishing identically, leading to $Q' = 0$. Strictly speaking this implies that the sextupole fields must be coincident with the quadrupoles, which is not normally very convenient. In practice a very good approximation to this situation is obtained by placing a sextupole close to each quadrupole.

2.3 Low-β insertions

In order to obtain high luminosity in colliding-beam storage rings it is necessary to focus the beams down to very small cross sections, which requires very strong quadrupoles in the vicinity of the collision point. On the other hand, the quadrupoles cannot be too close, since space is required for the detector. The low value of β^* at the collision point, together with the distance to the quadrupoles, results in a large value of β at the quadrupoles. There is thus a large contribution to the integral of Eq. (8) arising from large values of both β and K_0.

The problem is compounded by the need to have $D_x = 0$ in this region, in order both to maintain a high luminosity and to avoid an important source of synchro-betatron resonances. It follows that local correction is not possible, and the chromatic contribution from the low-β interaction regions can only be corrected in the main arcs of the machine lattice, where the dispersion is non-vanishing. It is then evident from Eq. (8) that, to make $Q' = 0$, the value of $K_0'(s)$ has to be increased above that necessary for local correction of the lattice quadrupoles.

It is now easy to see that we have a machine in which the focusing strength is almost everywhere momentum-dependent, despite having $Q' = 0$. This results in a chromatic mismatch of the betatron envelope functions for off-momentum particles, extending in general all around the circumference. In large, high-luminosity storage rings this mismatch can be large and leads to a number of undesirable side effects, in particular to higher-order variations of betatron tune with momentum and reduction of the effective machine aperture for stable orbits. Considerable insight into these effects can be obtained from first-order calculation of the chromatic perturbations.

3. CHROMATIC PERTURBATION EQUATIONS

The motion of an ensemble of particles with a momentum deviation $\delta = \Delta p/p$ may be compared with that of an ensemble having zero momentum error by assigning two distinct sets of Twiss parameters β, α, and phase advance ϕ :

$$\begin{aligned} \beta_2 &= \beta(\delta), \quad \alpha_2 = \alpha(\delta), \quad \phi_2 = \phi(\delta), \\ \beta_1 &= \beta(0), \quad \alpha_1 = \alpha(0), \quad \phi_1 = \phi(0), \end{aligned}$$

in each plane. The relation between the two sets will be determined later by boundary or symmetry conditions.

Considering only one plane for the moment we define :

$$\begin{aligned} \Delta\beta &= \beta_2 - \beta_1, \quad \beta = \sqrt{\beta_2\beta_1} \\ \Delta\phi &= \phi_2 - \phi_1, \quad \phi = (\phi_2+\phi_1)/2 \end{aligned} \tag{9}$$

and use these to define the chromatic variables

$$B = \frac{\beta_2 - \beta_1}{\sqrt{\beta_2\beta_1}} = \frac{\Delta\beta}{\beta} \quad , \tag{10}$$

$$A = \frac{\alpha_2\beta_1 - \alpha_1\beta_2}{\sqrt{\beta_2\beta_1}} \quad . \tag{11}$$

Additionally we can define the difference in the local focusing gradient parameter :

$$\Delta K = K_2 - K_1 = K(\delta) - K(0). \tag{12}$$

All the above parameters are, of course, functions of the position s in the machine.

From the definitions above and using the well-known relations :

$$\frac{d\phi}{ds} = \frac{1}{\beta}, \ \frac{d\beta}{ds} = -2\alpha \text{ and } \frac{d\alpha}{ds} = K\beta - \frac{(1+\alpha^2)}{\beta} \quad ,$$

we obtain directly :

$$\frac{d}{ds}(\Delta\phi) = -\frac{\Delta\beta}{\beta^2} = -\frac{B}{\beta} \tag{13}$$

and, after a little calculation :

$$\frac{dB}{ds} = - 2A \frac{d\phi}{ds} \quad , \tag{14}$$

$$\frac{dA}{ds} = \beta \Delta K + 2B \frac{d\phi}{ds} \quad . \tag{15}$$

In an achromatic region, where $\Delta K = 0$, it follows from Eqs. (14) and (15) that :

$$\frac{d}{ds} (A^2 + B^2) = 0 \tag{16}$$

and that (A^2+B^2) is therefore invariant.

The equations so far are exact and do not involve any assumptions about the variation with δ of the quantities defined. This makes it possible, by tracking the chromatic functions through a machine for various momentum errors δ, to estimate the importance of higher-order variations in δ^2, δ^3, etc. For further discussion it is convenient, however, to introduce some approximations for small perturbations. If we define :

$$\Delta\alpha = \alpha_2 - \alpha_1$$

we obtain from Eq. (11) :

$$A = \sqrt{\frac{\beta_1}{\beta_2}} \cdot \Delta\alpha - \alpha_1 \frac{\Delta\beta}{\beta} \approx \Delta\alpha - \alpha \frac{\Delta\beta}{\beta} \tag{17}$$

if $\frac{\Delta\alpha}{\alpha}$ and $\frac{\Delta\beta}{\beta}$ are small. It is worth noting for future reference that, for $\Delta\beta = 0$, the approximation in Eq. (17) becomes exact. If, as a further approximation, we take*) $(\beta_1 + \beta_2)/2 \approx \sqrt{\beta_1\beta_2}$ it is easy to show that :

$$\frac{d}{d\phi} (\Delta\phi) \approx -B \tag{18}$$

and that :

$$\frac{dA}{d\phi} \approx \beta^2 \Delta K + 2B \quad . \tag{19}$$

If $\Delta K = 0$, Eq. (19) becomes exact. Finally, without approximation, Eq. (14) becomes :

$$\frac{dB}{d\phi} = -2A \ . \tag{20}$$

It will become apparent later that ϕ is a more suitable independent variable than s in the present context.

---

* The arithmetic and geometric means differ by only 2% for $\Delta\beta/\beta = 0.4$

From Eqs. (18) and (19) we can write :

$$\frac{d}{d\phi}\left(2\Delta\phi + A\right) \approx \beta^2 \Delta K \tag{21}$$

and in achromatic region, where $\Delta K = 0$, $2\Delta\phi + A$ is invariant in the above approximation.

Some of the above considerations were discussed by Autin and Verdier[4] in a more general context of linear perturbation theory. In a specific application to chromatic effects, Zyngier[1] obtained the same perturbation equations by a somewhat different method, using a complex number representation.

Equations (19) and (20) can be derived from the Hamiltonian :

$$H(B,A,\phi) = A^2 + B^2 + \beta^2(\phi)\ \Delta K(\phi)\ B \tag{22}$$

with A and B the canonical position and momentum coordinates respectively. In an achromatic region $\Delta K = 0$, the Hamiltonian equals $A^2 + B^2$ and is invariant as previously noted. The two first-order equations (19) and (20) yield :

$$\frac{d^2A}{d\phi^2} + 4A = 0 \tag{23}$$

with an identical equation for B. Thus in an achromatic region A and B oscillate in quadrature at twice betatron frequency. They can be considered as representing a chromatic mismatch of the off-momentum betatron envelope function with respect to that of nominal momentum.

The chromatic variables have so far been defined with respect to an arbitrary momentum error δ. It is convenient to normalise these variables to δ by defining formally :

$$B = \lim_{\delta \to 0} \left[\frac{1}{\delta} \cdot \frac{\beta(\delta) - \beta(0)}{\sqrt{\beta(\delta)\ \beta(0)}} \right]$$

and similarly for A, ΔK, $\Delta\phi$. The invariant (for $\Delta K = 0$)

$$W = \frac{1}{2}\left(A^2 + B^2\right)^{1/2} \tag{24}$$

is dimensionless, and in the normalised form is an absolute measure of the strength of a linear chromatic perturbation. The factor $^1/_2$ is a convention (from Ref. 1) arising from the oscillation at twice betatron frequency. It is omitted by some authors. If W^2 is expressed in full, using the definitions (10) and (11) it is seen to be formally analogous to the Courant and Snyder quadratic form[3].

3.1 Simple examples

It is evident that in an idealised machine which is everywhere achromatic W is identically zero. To illustrate the situation in a real machine with low-β insertions it is instructive to consider two simple models in thin lens appoximation.

For a regular FODO lattice we use Eqs. (3), (6) and (7), and by logarithmic differentiation obtain the normalised chromatic variations :

$$\frac{1}{\delta}\frac{\Delta\hat{\beta}}{\hat{\beta}} = 1 - \frac{\tan\frac{\mu}{2}}{\cos\frac{\mu}{2}}$$

$$\frac{1}{\delta}\frac{\Delta\check{\beta}}{\check{\beta}} = 1 + \frac{\tan\frac{\mu}{2}}{\cos\frac{\mu}{2}} \quad . \tag{25}$$

Both α and $\Delta\alpha$ vanish because of symmetry. For $\mu = \pi/3$ we have :

$$\frac{1}{\delta}\frac{\Delta\hat{\beta}}{\hat{\beta}} = \frac{1}{3} \quad \text{and} \quad \frac{1}{\delta}\frac{\Delta\check{\beta}}{\check{\beta}} = \frac{5}{3}$$

with values of $W = {}^{1}/_{6}$, ${}^{5}/_{6}$ respectively. Thus a single lattice period uncorrected for chromaticity generates a value of W of order unity. A simple two-family sextupole arrangement can reduce this to zero so that the lattice contribution to W vanishes to first order.

As a simple model of a low-β insertion we consider a single thin focusing lens distance s from the interaction point, where the envelope function β^* is assumed not to change with momentum. The value of α just before the lens is :

$$\alpha_- = -\frac{s}{\beta^*} \quad .$$

In a typical low-β insertion the first quadrupole changes α by about twice this amount and the value after the lens is :

$$\alpha_+ = \alpha_- - \hat{\beta}k,$$

where $\hat{\beta} = \beta^*(1 + s^2/\beta^{*2})$ is the value at the lens and $k = -K_0\ell_q < 0$ for a focusing lens. Since we assume β^* fixed the change in α_+ due to a momentum error δ is :

$$\Delta\alpha_+ = -\hat{\beta}\Delta k = -\hat{\beta}k\frac{\Delta k}{k} = \hat{\beta}k\delta \ ,$$

whence

$$\Delta\alpha_+ = (\alpha_- - \alpha_+)\delta = -\frac{2s}{\beta^*} \cdot \delta \ .$$

From (17), since β^* and $\hat{\beta}$ do not change, normalising by δ gives :

$$A = \frac{\Delta\alpha_+}{\delta} = -\frac{2s}{\beta^*} \ ,$$

and the contribution to W is :

$$W = \frac{s}{\beta^*} \ .$$

For a LEP physics insertion, $s/\beta^* = 50$ vertically. This illustrates how strong is the local chromatic error introduced by the first low-β quadrupole compared to that of a normal lattice, and justifies considering the latter as relatively achromatic in the subsequent sections of this text. The exact value of W in the vertical plane is somewhat reduced in practice since α_+ is less than $(-\alpha_-)$.

4. TWO DIMENSIONS

The methods of Section 3 can readily be extended to treat the vertical and horizontal planes together, in order to take account of the fact that sextupoles act simultaneously in both planes. For this purpose we introduce the two-dimensional normalised "invariant" :

$$W = \frac{1}{2} \left[B_v^2 + B_h^2 + A_v^2 + A_h^2 \right]^{1/2} \ , \qquad (26)$$

where the normalised B_v, B_h, A_v, A_h, are defined exactly as before, independently in the two planes. The value of W is evidently a measure of the overall chromatic error in both planes.

At any position where the dispersion D_x does not vanish, a sextupole acts as a quadrupole for off-momentum particles and therefore introduces a ΔK. It is evident from Eqs. (14), (15), (18) and (19) that ΔK changes A but not B, and in both planes. The changes in A_h and A_v for $\delta = 1$ due to a short sextupole of strength $K'\ell$ are given by :

$$\begin{aligned} a_h &= -\beta_h K'\ell D_x \\ a_v &= \beta_v K'\ell D_x = -\frac{\beta_v}{\beta_h} \cdot a_h \ . \end{aligned} \qquad (27)$$

Replacing A_h by $(A_h + a_h)$ and A_v by $\left(A_v - \frac{\beta_v}{\beta_h} a_h\right)$ in Eq. (26), we obtain a new value of W which can be minimised with respect to a_h. This "optimum" W :

$$W_{op} = \frac{1}{2}\left[B_v + B_h + \frac{(\beta_v A_h + \beta_h A_v)^2}{(\beta_h^2 + \beta_v^2)}\right]^{1/2} \qquad (28)$$

is the lowest that can be obtained for given B and A with a single sextupole, or with a single family of sextupoles whose members are spaced by integer multiples of π in vertical and horizontal betatron phase. The corresponding sextupole strength required is :

$$(K'\ell)_{op} = \frac{\beta_h A_h - \beta_v A_v}{D_x(\beta_h^2 + \beta_v^2)} \quad . \qquad (29)$$

4.1 Choice of linear lattice parameters

Sextupoles are usually located close to the F and D quadrupoles of a FODO lattice in order to decouple the effects in the two planes as far as possible. We adopt this constraint, which fixes β_v, β_h, and seek conditions on B, A in the two planes to make W_{op} in Eq. (28) a minimum.

The examples of Section 3.1 demonstrate that it is legitimate to consider the major chromatic perturbation as originating in the low-β quadrupoles and propagating into the main lattice almost unchanged except in phase, as shown schematically in Fig. 1. It follows that, in each plane, B and A at any given point in the main lattice are just functions of the phase advance in the respective planes from the low-β quadrupoles to that point.

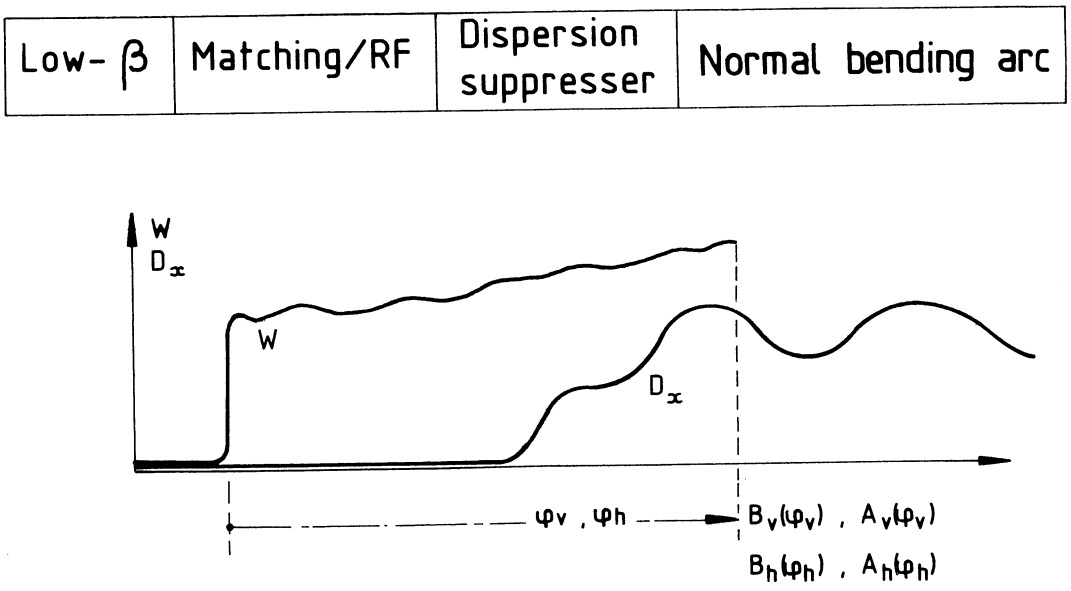


Fig. 1 Variation of chromatic perturbation W and dispersion D_x near the interaction region

We consider W_{op} at the first sextupole in the regular lattice and write Eq. (28) in terms of $B_v(\phi_v)$, $B_h(\phi_h)$, $A_v(\phi_v)$ and $A_h(\phi_v)$. A minimum of W_{op} with respect to both phase advances ϕ_v, ϕ_h occurs for :

$$\frac{\partial}{\partial\phi_v}(W_{op}^2) = \frac{\partial}{\partial\phi_h}(W_{op}^2) = 0 \ ,$$

provided that both $\frac{\partial^2}{\partial\phi_v^2}(W_{op}^2)$ and $\frac{\partial^2}{\partial\phi_h^2}(W^2{}_{op})$ are positive definite.

After some algebra one finds a unique solution corresponding to :

$$B_v = B_v = 0 \quad \text{and} \quad A_v A_h < 0 \ . \tag{30}$$

With these conditions Eq. (28) reduces to :

$$(W_{op})_{min} = \frac{1}{2} \cdot \left[\frac{(\beta_v A_h + \beta_h A_v)^2}{(\beta_v^2 + \beta_h^2)} \right]^{1/2} \tag{31}$$

and the required value of K'ℓ is still given by Eq. (29), either for a single sextupole or for the total strength of a single "coherent" family as before.

The condition given by Eq. (30) is illustrated in the phase-plane diagram of Fig. 2. For simplicity the chromatic mismatches for vertical and horizontal planes are superposed on the same normalised coordinate system (η,η') though the normalisation is necessarily different for the two planes. The vanishing of $B = \Delta\beta/\beta$ in the two planes corresponds to the common vertical tangent and makes $A = \Delta\alpha$. The condition that A_v and A_h have opposite signs follows from the fact that, for an off-momentum particle, a sextupole focuses in one plane and defocuses in the other, as indicated by the arrows.

In Eq. (31) $(W_{op})_{min}$ cannot be made to vanish with a single sextupole family, since it is determined by parameters of the lattice (β_v,β_h) and of the low-β insertion (A_v,A_h) whose values are governed by other considerations. However, the introduction of a second family of sextupoles, suitably phased with respect to the first, enables W to be reduced to zero in first order.

In general, four chromatic variables describe the first-order behaviour of off-momentum particles, and four parameters are therefore required in principle in order to make these variables vanish. Normally this would require four independent sextupole families but, by choosing the linear phase advances in the two planes appropriately, we have introduced a degeneracy, which makes it possible to correct the first-order effects arising from the low-β quadrupoles with only two families. In addition, of course, provision must be made for correcting the lattice chromaticity, either by appropriate adjustments to the K' values or by extra families, which latter may anyway be required for taking care of high-order effects.

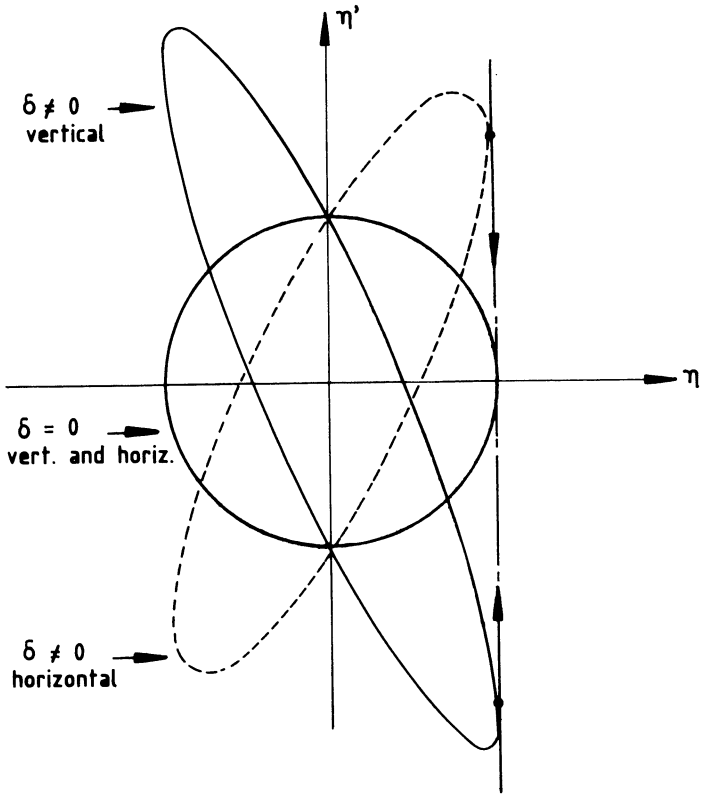


Fig. 2 Normalized phase-plane diagram for the optimum ϕ_v and ϕ_h at first sextupole.

4.2 Concentrated correction

We can illustrate the principles on which the W-matching strategy of chromaticity correction is based by considering a simple model.

At an interaction point we impose the initial condition that the chromatic errors there vanish, i.e. W = 0. The low-β quadrupole doublet introduces strong chromatic perturbations in both planes which propagate through the remainder of the straight section, where the dispersion is zero, into the main arc. The chromatic contribution of this part of the straight section is small compared with that of the low-β quadrupoles. We now adjust the parameters of the straight-section lattice, together with the matching into the low-β insertion and the dispersion suppressor, such that the phase advances in the two planes lead to Eq. (30) being satisfied at the first sextupole in the main arc.

In order to distribute the total required sextupole strength amongst several members of a family acting coherently, the phase advance μ per period of the main lattice must satisfy :

$$n\mu = (2m + 1)\pi;\ m,n \text{ integer.} \qquad (32)$$

For our model we assume a lattice phase advance $\mu = \pi/3$ and choose the simplest configuration n = 3, m = 0, with two sextupole families of two members each. The phase difference of π between the two members of each family helps to cancel nonlinear resonance excitation by the sextupoles[5]. The trajectories of the (B,A) vectors in the vertical and horizontal planes are indicated in Fig. 3(b). The radius of a trajectory is 2W, as seen in Eq. (24), and the polar angle rotates at twice betatron frequency in this representation. Each of the sextupoles, whether F or D, contributes to reducing W (in steps) simultaneously in both planes, and after the fourth sextupole W is brought to zero.

Figure 3(c) illustrates a feature which has so far not been discussed in detail, namely the behaviour of the accumulated chromaticity $\Delta\phi$, which is correlated with A by $2\Delta\phi + A$ = constant (Eq. (21) with $\Delta K = 0$). Like B and A, $\Delta\phi$ oscillates at twice betatron frequency and is negative everywhere in this region, since the large chromaticity produced by the low-β quadrupoles (and the smaller contribution from the straight-section lattice) have not yet been corrected. We now impose a further constraint over and above those of Eq. (30), requiring not only that A_v and A_h have opposite signs, but also specifying which one should be negative at a given sextupole. This choice is made from the following qualitative considerations of a second-order effect.

The local value $\Delta\phi$ of the accumulated chromaticity is a measure of how well the phase advances imposed to satisfy Eq. (30) are maintained as a function of momentum error. Our choice of normalised, first-order variables has removed the linear dependence on δ; consequently the non-vanishing of $\Delta\phi$ implies a variation with δ^2 or higher order. It is therefore desirable to arrange that $\Delta\phi$ be numerically a minimum at the locations of the sextupoles. Because of the dominant, first-order constraint that $A_v A_h < 0$, this is not possible simultaneously in both planes and we therefore impose $|\Delta\phi|$ = minimum in the plane in which a sextupole has the strongest influence, i.e. the plane of β_{max}. This is the more favourable case, depicted in Fig. 3(a), (b) and (c).

Although this simple arrangement illustrates clearly the concepts of W-matching, it unfortunately cannot be used in practice. The basic reason is clear from the examples of Section 3.1; to concentrate the correction of the low-β quadrupoles, chromatically the equivalent of about 50 lattice periods, into two pairs of elements would require exceedingly strong sextupoles. Such strong nonlinear elements greatly enhance the excitation of systematic resonances of third and higher order, leading to orbit instability at quite low betatron amplitudes, where the π phase-separation rule is violated due to the strong amplitude-dependence of phase advance.

4.3 Distributed correction

It is therefore essential to distribute the sextupole correction amongst many members of each family, the aim being to find arrangements which minimise the strength of the strongest sextupoles in the machine. The value of W is then progressively reduced towards the centre of the main bending arc at which point it should be rather small. The symmetry of the machine then ensures that $W \approx 0$ at the next interaction point, consistent with the initial condition imposed.

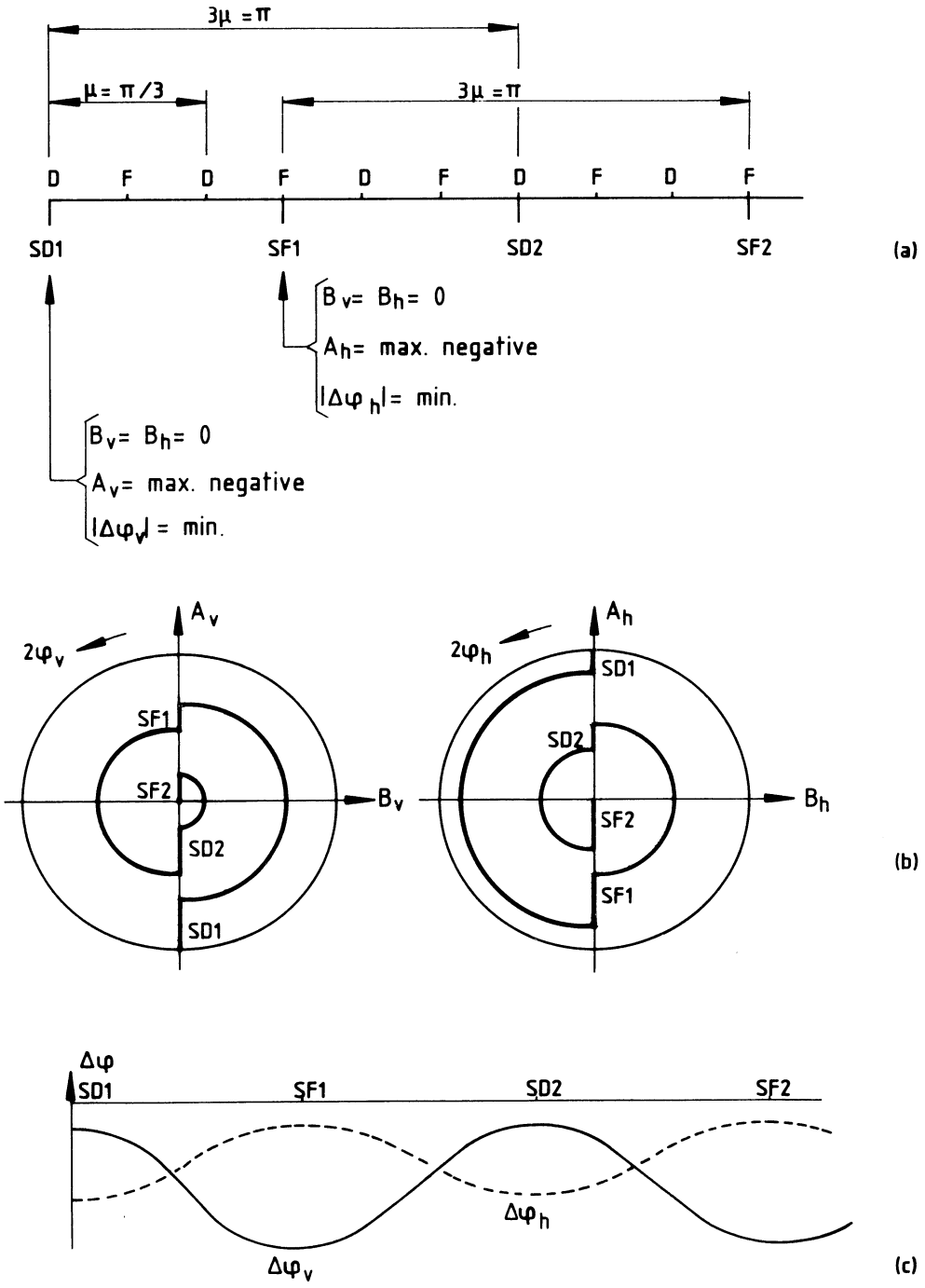


Fig. 3 Example of W-correction with two pairs of sextupoles; a) configuration and constraints, b) loci of the W-vectors, c) accumulated chromaticities $\Delta\phi_v$, $\Delta\phi_h$.

The phase constraint of Eq. (32) shows that members of the same sextupole family occur at intervals of n lattice periods. For m = 0 the phase advance μ is a simple fraction of π, e.g. $\pi/2$, $\pi/3$, etc., and in the design of linear lattices there is normally no reason to choose more exotic values. In fact there are good reasons to avoid configurations with $m > 0$ since this leads to larger separation in betatron phase between members of the same family. The need to correct the low-β chromaticity in the normal lattice necessarily implies local over-correction because of the increased values of K_0', which can lead to excessive accumulation of phase errors for off-momentum particles if adjacent members of a family are too widely spaced.

These higher-order effects require more powerful methods for systematic study and are treated by G. Guignard in this course. However, the first-order theory discussed here has shown itself to be very useful in clarifying the general principles, exposing the importance of basic lattice parameters and indicating the origin of some significant higher-order phenomena.

* * *

REFERENCES

1) H. Zyngier, Report LAL 77/35 (1977).
2) B.W. Montague, LEP Note 165 (1979).
3) E.D. Courant and H.S. Snyder, Annals of Physics 3 (1958) 1.
4) B. Autin and A. Verdier, CERN ISR-LTD/76-14 (1976).
5) K.L. Brown, SLAC-PUB 2257 (1979).

CHROMATICITY: NONLINEAR ABERRATIONS

G. Guignard
CERN, Geneva, Switzerland.

ABSTRACT

For accelerator designers, chromaticity deals with the variations of optics parameters with momentum deviations $= \Delta p/p$. Differential bending gives rise to the dispersion of the closed orbit while differential focusing in the quadrupoles is first seen as a change in the oscillation frequency and amplitude. Sextupoles placed in the lattice where the dispersion is sufficient, are used to correct strong, first-order chromatic effects and this paper on chromaticity discusses this question. Nevertheless, these sextupole fields introduce nonlinear forces and additional tune variations with the amplitude. Furthermore, every quadrupole and sextupole of the structure induces aberrations to all orders in δ. These effects are first described and methods of analysing and possibly minimizing them are then presented in this second lecture on chromaticity. The pertubation theory in the canonical variables as a tool for dealing with the nonlinearities is outlined. In addition, the concept of dynamic aperture is introduced. Recent investigations aiming at the development of analytical or semi-analytical means for estimating dynamic aperture without tracking are mentioned.

1. INTRODUCTION

This paper discusses the nonlinear aberrations associated with the chromaticity and the presence of sextupole fields necessary for its correction. Following a qualitative description of the expected nonlinearities in momentum deviation and in the transverse coordinates (x,z), one defines variables allowing quantitative estimation of the subsequent effects (Section 2). The so-called dynamic aperture, frequently debated and considered as one of the most interesting concepts, is then brought in. Since the most powerful method of studying it remains single-particle tracking, the computer programs able to achieve this are briefly reviewed. However, the main object of the present lecture is to describe the analytical methods which have been investigated in order to deal with nonlinearities.

The pertubation theory using the Hamiltonian formalism is a very important tool in this frame. Therefore, the basic principles are recalled for one-dimensional motion and the use of this theory for minimizing nonlinear aberrations from sextupoles is explained (Section 3). The generalization to two dimensions and the analysis of resonances as well as tune shifts are briefly outlined. The recent efforts, made in parallel with the work on numerical tracking, to develop analytical or semi-analytical tools to find the stability limit in the presence of nonlinearities are summarized (Section 4). The possible use of the Hamiltonian pertubation to calculate the distortions of the invariant tori is explained and the search for approximate solutions by successive linearization of the equation of motion is described.

2. DESCRIPTION OF THE NONLINEAR PERTURBATIONS

2.1 Relevant quantities

Two sorts of nonlinearities are present in an accelerator or storage ring : those in the energy deviation $\delta = \Delta p/p$ and those in the transverse amplitudes x and z. When we are dealing with the minimization of chromatic effects, we are of course directly concerned by the former, but the latter are also important since they may be generated by the sextupoles used for chromaticity correction.

Considering the vector $\vec{W}$ defined in the previous chapter on chromaticity[1] and describing amplitude effects, it is clear that every quadrupole and sextupole (actual sextupole magnets or sextupole-field imperfections) induces local gradient errors ΔK to all orders in δ. Consequently, the betatron functions are also perturbed to all orders, in particular the phase advance ϕ_y. Hence, if the vector $\vec{W}$ changes by steps in its angular component when $\Delta K \neq 0$ and rotates with twice the phase advance ϕ_y in achromatic sections, to first order in δ, nonlinearities in δ are necessarily present in addition ; $\vec{W}$-amplitude nonlinearities which are usually weak and $\vec{W}$-phase nonlinearities, which are large and follow the spread in phase advance mentioned above. This can be visualized by tracking the $\vec{W}$-vector through a superperiod which is achromatic to first order, for instance.

Sextupoles are used for compensating the linear chromatic variations discussed in the previous chapter and field imperfections are present in all magnets. They are responsible for nonlinear forces in x and z, which induce nonlinear kicks in the betatron motion. Direct consequences are the excitation of many resonances, the presence of coupling and possible blow-up of the beam. As is the case for δ nonlinearities, all orders in x and z can be present in general.

There are analytical and numerical means[2,3] to calculate the compensation of the linear chromatic perturbation. The computer program HARMON[3] is able in addition to minimize some δ non-linearities (e.g. second and third-tune derivatives) and the excitation of some resonances.

2.2 Diagnosis means

It is important to have values for the quantities characterizing the nonlinear effects mentioned in Section 2.1, in order to estimate the quality of a compensation scheme and the future performance of the accelerator. Hence, numerical diagnosis means are necessary and the ones which are nowadays most frequently used are listed below :

- The program MAD[4] can track the vector $\vec{W}$ trough the structure and give direct information about its spread with δ.

- The same program is able to calculate the curve giving the tune as a function of δ.

This makes it possible to check the accuracy of the linear compensation (Q' = 0) and to estimate the importance of the quadratic and cubic terms (Fig. 1). Such a curve is also able to show at which value of δ stop-bands or strong resonances are reached.

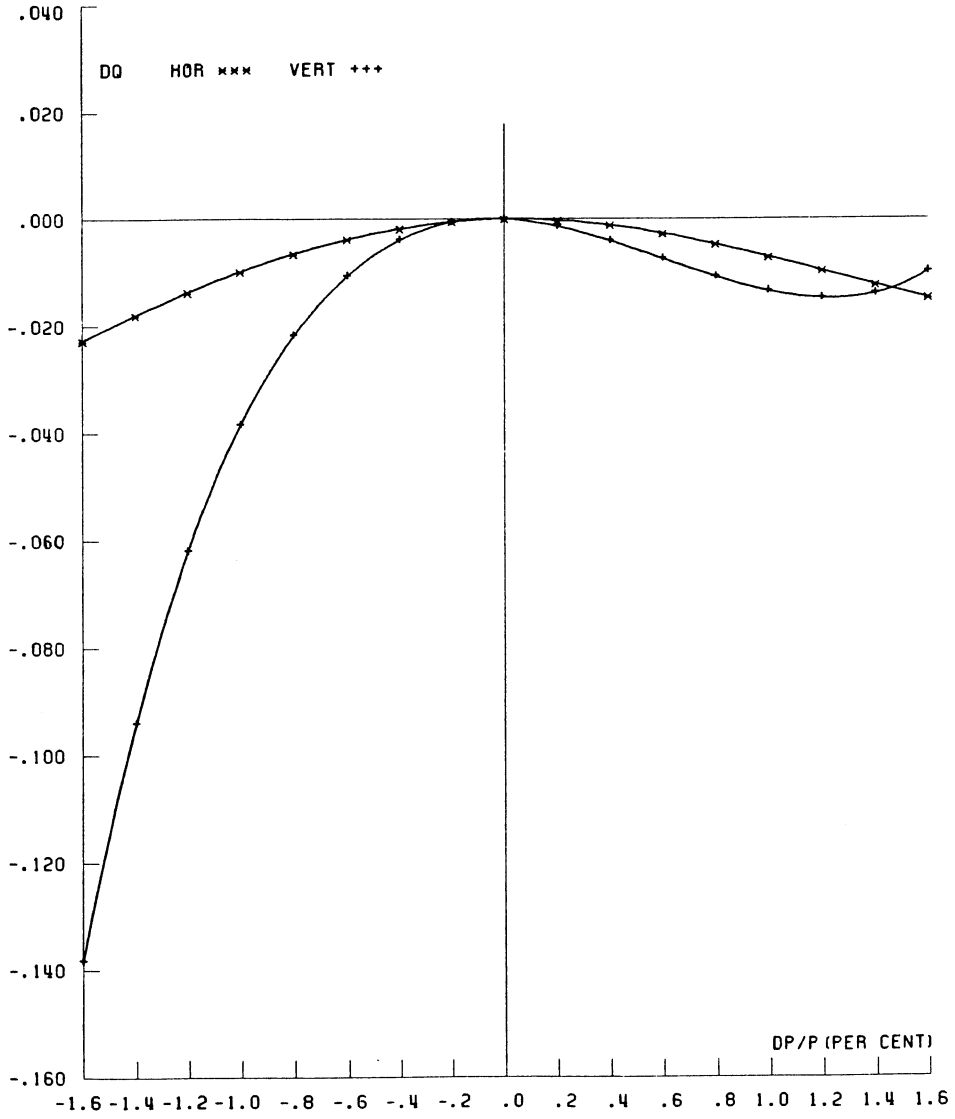


Fig. 1 Example of tune variations with momentum
(LEP with Q_x = 70.35 and Q_z = 78.20, at δ = 0)

- Some information on the quality of the compensation scheme can be derived by tracking with TURTLE[5] through a section of the machine and plotting phase space distortions. The experimental observation was that background conditions in PEP experiments were strongly dependent on the sextupole strengths. For the observed optimum, the tracking over one superperiod[6] showed that tails do not develop in the phase space, in contrast with the other configurations. Tracking is done with 10 000 particles, using a rectangular distribution enhancing the importance of the tails and spreading over 20 and 10 standard deviations of the transverse and

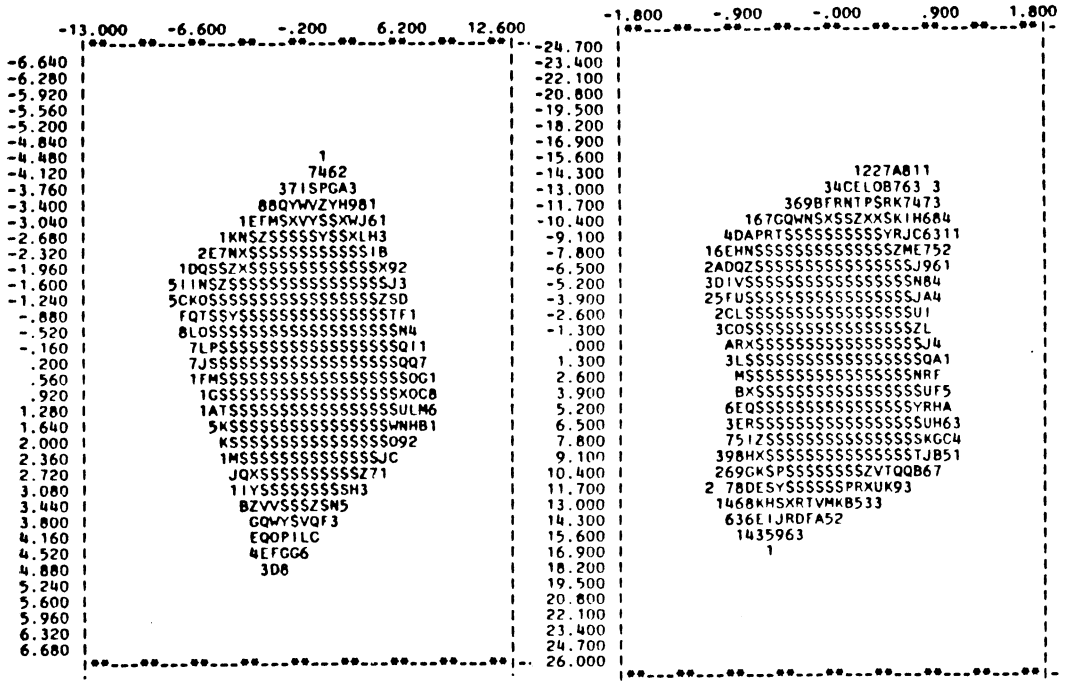
longitudinal beam dimensions, respectively. For each element, second-order transforms are used with no truncation in the cross-terms. An example of LEP results[6] (60^0 configuration) is given in Fig. 2.

- An interesting way to have an overall description of the nonlinear effects and to estimate the performance is to look at the dynamic aperture. By definition, the dynamic aperture d_y and the associated acceptance $A_y = d^2y/\beta y$ describe the betatron amplitudes at which particles can circulate indefinitely as functions of δ. It corresponds to the initial transverse amplitude up to which the betatron motion is stable, meaning it remains bounded. The condition of boundedness must prevail during a time comparable with either one damping time or many synchrotron periods for electrons or protons, respectively.

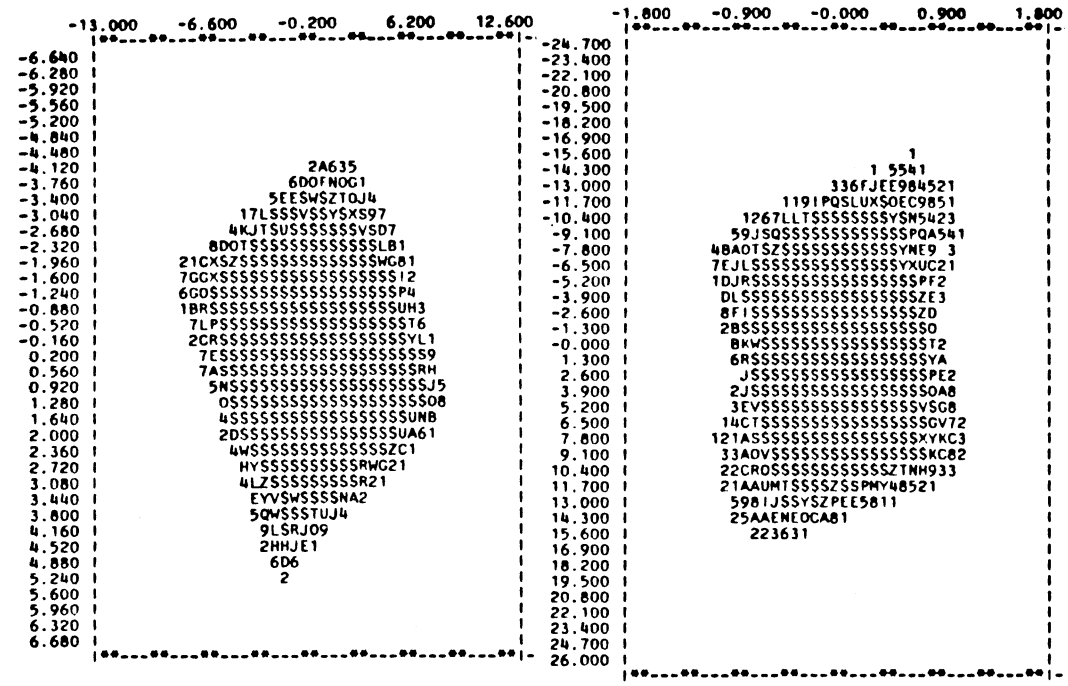
  In principle, the stable region is a volume in (E_x, E_z, δ) space, where E_y are the emittances related to the betatron amplitudes and δ is the relative amplitude of the synchrotron oscillations. To simplify the presentation, cuts of the stable volume are shown graphically. Fig. 3 gives an example of such a cut with $E_z = {}^1/_2E_x$ for the LEP design.

  To calculate the stability limit, designers of accelerators rely above all on particle tracking. Recently, analytical or semi-analytical methods have been investigated with the idea to complement the numerical results and to obtain some useful information about the dynamic aperture. These methods [7-10] are based either on the Hamiltonian formalism and the iterative perturbation in canonical variables or the successive linearization of the equations of motion. These two techniques will therefore be described in this chapter.

  Since particle tracking remains the most powerful method for studying stability, we will briefly mention here some of the computer programs for that purpose, reminding the reader that tracking is treated in another chapter[11]. In practice tracking amounts to calculating orbits of single particles in external fields for a few initial conditions, by approximate integration of Hamilton's equations through individual magnetic elements. Because of the special nature of the problem, one uses particular integration methods. The simplest method is the kick approximation, which treats the nonlinearities as infinitesimally thin lenses. It is used for instance in the programs PATRICIA[12], EVOL[13] and RACETRACK[14]; as a special feature, EVOL includes multiple beam-beam collisions and tune modulation, while RACETRACK deals with orbit distortions. A second method is based on second-order transfer matrices which are used in the programs TRANSPORT[15], MAD[4] and DIMAD[16], the last one allowing for tracking with orbit distortions before and after correction. A third method exploits Lie transforms[17] for constructing sympletic transfer maps to high order (3 or 4) in the canonical variables[17] and has been applied extensively in the program MARYLIE[18], and later in MAD[4]. An example of tracking with PATRICIA is given in Fig. 4.



Horizontal plane, $\delta = 0$ Vertical plane, $\delta = 0$



Horizontal plane, $\delta = 0.82\ ^{0}/_{00}$ Vertical plane, $\delta = 0.82\ ^{0}/_{00}$

Fig. 2 Example of TURTLE tracking over one superperiod (LEP with $\mu_0 = 60^0$ in the lattice)

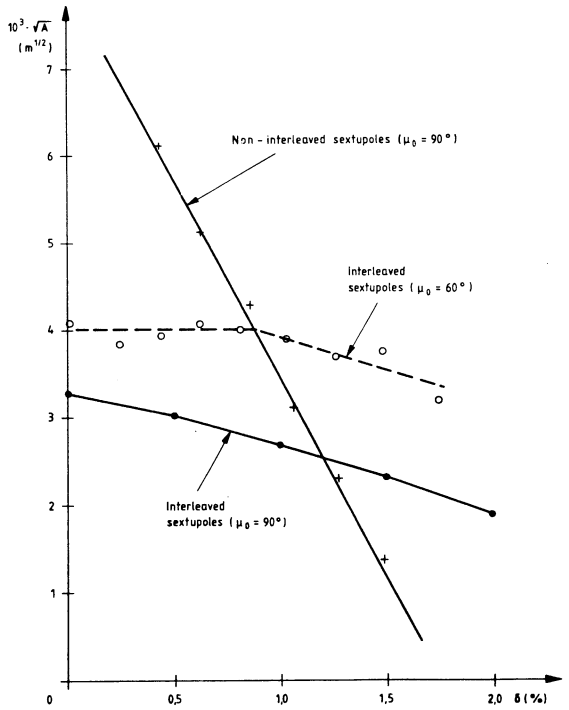


Fig. 3 Examples of cuts $E_z = {}^1/_2\, E_x$ of the stable volume in the phase space (LEP)

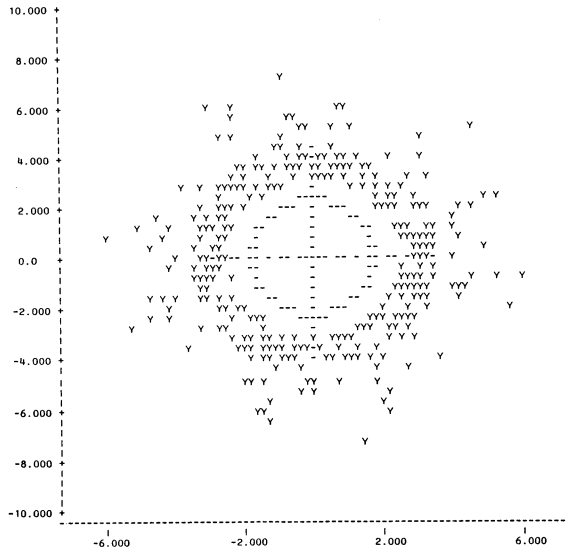


Fig. 4 Example of PATRICIA tracking over 400 turns (LEP) for initial vertical amplitude close to the stability limit

- Distortions of elliptically cross-sectioned emittance tori allow one to visualize the effects of nonlinearities and give indirect information about the proximity of the stability limit or the onset of chaotic motion. Methods for the analytical computation of invariant surfaces in the phase space have been recently studied. Similarly to the dynamic aperture, these methods are based on the successive canonical generating functions[8,10] and the linearization of the equations of motion[9], but also on the first-order distortion functions[19], the solution of the Hamilton-Jacobi equation[20] and the perturbation series exploiting Lie operators[21].

3. PERTURBATION THEORY IN THE CANONICAL VARIABLES

The practical problem to face in the correction of the chromaticy with lumped sextupole magnets is to keep small the nonlinearities in δ and the effects of the transverse nonlinear forces on the betatron motion. This section describes a possible solution to that problem. It consists of analysing the motion with the Hamiltonian formalism, calculating characteristic quantities and minimizing them with the available sextupole strengths (program HARMON[3]); these quantities are mainly the excitation coefficients characterizing broad resonances and their bandwidth, and the stabilizing coefficients controlling the tune variations with amplitude and momentum deviation[22].

3.1 Recap of the basic principles

To illustrate the method, let us consider one-dimensional motion with a nonlinear force, whose equation can be written :

$$y'' + K(s)\, y = a_m(s)\, y^{(m-1)} \quad . \tag{3.1}$$

Using the betatron variables, Eq. (3.1) can be rewritten :

$$u'' + Q^2 u = Q^2\, a_m\, \beta^{(m+2)/2}\, u^{(m-1)} \tag{3.2}$$

and the corresponding Hamiltonian is given by :

$$H = \frac{1}{2}\,(Q^2u^2 + p_u^2) + \frac{1}{m}\, Q^2\, a_m\, \beta^{(m+2)/2}\, u^m \quad . \tag{3.3}$$

In the more convenient action-angle variables I and ϕ, H (3.3) takes the form :

$$H = H_0 + H_1 = QI + \frac{1}{m}\, Q^2\, a_m\, \beta \left(\frac{2I}{Q} \right)^{m/2} \cos^m \phi \tag{3.4}$$

where H_0 is the unperturbed Hamiltonian giving the well-known invariant curves of the linear motion, i.e. circles in the betatron phase space. The perturbation H_1, which is a function of the variable θ used instead of s, describes a nonlinear system which is not always integrable, depending on the function a_m. The problem consists of finding approximate new invariants and the method used in the Hamiltonian formalism can be summarized as follows; one looks for a canonical transformation which makes the new Hamiltonian (called G) independent of θ to first order and almost independent for small amplitudes. This process can be iterated with successive canonical transformations, pushing the θ-dependence to higher orders.

Let us do now the first step explicitely. The canonical transformation from (I,ϕ) to (J,ψ) is defined by the generating function S_{tot} :

$$S_{tot}\,(\phi,J,\theta) = \phi J + S\,(\phi,J,\theta) \quad , \tag{3.5}$$

with the following relations between the old and new variables

$$I = \frac{\partial S_{tot}}{\partial \phi} = J + \frac{\partial S}{\partial \phi}$$

$$\psi = \frac{\partial S_{tot}}{\partial J} = \phi + \frac{\partial S}{\partial J}$$

$$G(\psi,J,\theta) = H(\phi,I,\theta) + \frac{\partial S_{tot}}{\partial \theta}\,(\phi,J,\theta) \quad . \tag{3.6}$$

In order to express both H and G in the last equation (3.6) as functions of the S_{tot}-variables (ϕ,J), we may use Taylor's expansions and apply the first two equations (3.6) :

$$H(\phi,I) = H(\phi,J) + \partial H_I \, \partial S_\phi + \ldots$$

$$G(\psi,J) = G(\phi,J) + \partial G_\psi \, \partial S_J + \ldots \quad , \tag{3.7}$$

where ∂f_I is a notation for the partial derivative of f with respect to I.

Introducing (3.4) and (3.7) into the last equation (3.6) and remembering that $\partial H_{0I} = Q$, one obtains the following relation :

$$\underline{G(\phi,J)} + \partial G_\psi \cdot \partial S_J + \ldots = \underline{H_0(J)} + \underline{Q \cdot \partial S_\phi} + \underline{H_1(\phi,J,\theta)} + \partial H_{1I} \cdot \underline{\partial S_\phi} + \partial S_\theta + \ldots \, , \tag{3.8}$$

where the underlined terms are to first order in the perturbation.

Since the structure of the machine is periodic over the turn, the functions of interest are periodic in θ and ϕ with a period of 2π. It is therefore possible to analyse them in Fourier's series :

$$S = \sum_{n,p} s_{np} \exp\left[i(n\phi - p\theta)\right]$$

$$H_1 = \sum_{n,p} h_{np} \exp\left[i(n\phi - p\theta)\right]$$

$$G = \sum_{n,p} g_{np} \exp\left[i(n\phi - p\theta)\right] \quad , \tag{3.9}$$

where s_{np}, h_{np} and g_{np} are the harmonic coefficients of the generating function, the perturbation Hamiltonian and the transformed Hamiltonian, respectively. Introducing the expressions (3.9) in the first order terms (3.8) and dividing by the phase term which is in factor, gives for every harmonic :

$$g_{np} = inQ \, s_{np} + h_{np} - ip \, s_{np} \quad . \tag{3.10}$$

The generating function (3.5) must be chosen in order to satisfy (3.10). Let us therefore solve (3.10) with respect to the coefficients s_{np} of $S(\phi,J,\theta)$:

$$s_{np} = i \, \frac{h_{np} - g_{np}}{n\,Q - p} \quad . \tag{3.11}$$

The canonical transformation from (I,ϕ) to (J,ψ) must remain finite and the convergence of S implies that s_{np} be bounded.

When Q is rational, i.e. Q = p/n, resonances of the type nQ-p appear because of the small denominator in (3.11). The only way to ensure the boundedness of s_{np} is to make its numerator equal to zero, which means that $g_{np} = h_{np}$ and s_{np} vanishes. In this case, coefficients g_{np} of the new Hamiltonian are different from zero and the first order perturbation theory applies[22,23]; the concepts of isolated resonances, bandwidths, resonance curves and separatrices can be introduced and used for instance in the optimization of the sextupole correction scheme (see Section 3.2).

When, on the contrary, Q is irrational or at least distant from every strong resonance line, the denominator of (3.11) is different from zero and the new coefficient g_{np} can be chosen freely. To make the new Hamiltonian G independent of θ to first order, as desired, the obvious choice is $g_{np} = 0$. Indeed, by virtue of (3.8) and (3.10), G will then contain only higher-order terms in the perturbation :

$$G(\psi,J) = \partial H_{1I}\ \partial S_{\phi} + \ldots \qquad (3.12)$$

The proof of this statement can easily be given. From (3.4), the perturbation H_1 and, as a consequence, the function ∂S_{ϕ} are proportional to $a_m\ I^{m/2}$. Hence, the first term in (3.12) for the new Hamiltonian goes like $a^2_m\ I^{(m-1)}$, i.e. it is to second order in the perturbation a_m and to a higher power in the amplitude.

Successive canonical transformations can then be applied to suppress the dependence on θ of G to second, third and higher orders, if desired, albeit the development becomes tedious.

3.2 Minimization of the perturbation due to sextupoles

As explained in the previous chapter on the chromaticity[1], sextupole magnets are used to compensate the linear perturbation. Section 2.1 of this lecture indicates that in the presence of these elements, nonlinear effects are necessarily observed. In addition, it was mentioned that HARMON[3] can calculate the linear compensation, but also minimize some nonlinearities. Let us describe now how this program achieves this, using the formalism recalled in the preceding section.

Let us note first that the approach of Section 3.1 can be and has been generalized for two dimensions[24,25] (x and z). The expression equivalent to (3.4) is in this case :

$$H = H_0+H_1 = Q_xI_x + Q_zI_z + h^{(N)}_{jklm}\ I_x^{\frac{l+m}{2}}\ I_z^{\frac{l+m}{2}}\ \exp\{i[(j-k)(\phi_x+Q_x\theta)+(l-m)(\phi_z+Q_z\theta)]\}\ , \qquad (3.13)$$

where j+k+l+m = N and N is the "order in amplitude" of the perturbation.

Following exactly Section 3.1, a canonical transformation from (I_x,ϕ_x,I_z,ϕ_z) to (J_x,ψ_x,J_z,ψ_z) has to be defined by a generating function S_{tot} depending on

five variables (generalization of (3.5)). The following steps remain the same : expression of the new Hamiltonian G from H and S_{tot}, Taylor's expansions for H and G, and analysis in Fourier's Series of the three functions H, G and S_{tot}. However, the equivalent equations of (3.7) to (3.9) are more complicated, since we need two phases and two amplitudes, before and after the transformation. Let us write explicitely the coefficients of the Fourier's series of H_1, which are the main characteristic quantities (generalization for two dimensions of h_{np}) :

$$h_{jklmp} = \frac{1}{\pi 2^{(N+3)/2} B\rho\ j!k!l!m!} \int_0^{2\pi} d\theta\ \frac{\partial^{N-1} B_y}{\partial x^{(N-1)}} \left(\frac{\beta_x}{R}\right)^{\frac{j+k}{2}} \left(\frac{\beta_z}{R}\right)^{\frac{l+m}{2}} \times$$

$$\times \exp\{i[(j-k)(\phi_x - Q_x\theta) + (l-m)(\phi_z - Q_z\theta) + p\theta]\} \quad . \tag{3.14}$$

The coefficients h_{Cp} (C standing for jklm) are called[22] excitation coefficients when the indices jklm correspond to a resonance, and stabilizing coefficients when j = k, l = m and p = 0, associated with an attenuation of the resonance effect by detuning.

Eventually, the harmonics of the generating function defined above for a two-dimensional motion are given by a relation similar to (3.11) :

$$s_{Cp} = i\ \frac{h_{Cp} - g_{Cp}}{(j-k)Q_x + (l-m)Q_z - p} \quad , \tag{3.15}$$

in which one recognizes the well-known resonance condition if we stipulate that n_x =j-k and n_z = l-m.

Let us consider the specific problem of the minimization of the perturbation due to the sextupoles used for compensating the linear chromaticy. Exploiting the first-order perturbation theory[22,24,25], isolated resonances can be studied as explained in Section 3.1. The coefficients h_{Cp} of interest are necessarily of order N = j+k+l+m = 3 in amplitude (degree of the sextupole potential).

Hence, the first order perturbation (with $g_{Cp} = h_{Cp}$) evidences the existence of resonances of third order in amplitude. However, second (and higher)-order terms in the pertubation are also present in the development of the new Hamiltonian (generalization of (3.8) for two transverse dimensions). These terms contain products of partial derivatives of H_1 and S (see (3.12)) which have the following forms :

$$h_{Cq} h_{C'q'} J_x^{(j+k+j'+k'-2)/2}\ J_z^{(l+m+l'+m')/2}$$

$$h_{Cq} h_{C'q'} J_x^{(j+k+j'+k')/2}\ \ J_z^{(l+m+l'+m'-2)/2} \quad , \tag{3.16}$$

leaving out factors of proportionality and exponential functions. In (3.16), the condi-

tions $j+k+l+m = j'+k'+l'+m' = 3$ prevail (sextupoles are considered) and the direct consequence is that the order in amplitude of the terms (3.16) is $j+k+j'+k'+l+m+l'+m'-2 = 4$. Hence, sextupoles excite also fourth (and higher)-order resonances. The physical explanation is that trajectory oscillations due to nonlinear kicks propagate around the ring and do affect the kicks of other sextupoles.

The program HARMON[3] includes in its minimization process a set of coefficients h_{Cp} (with $N = 3$), like h_{1011}, h_{1020} and h_{1002} which concern coupled motion. Contributions to second order in the perturbation are indirectly reduced, since their excitation terms are made of products of h's (see (3.16)).

It was said in Section 2 that nonlinearities in δ of the amplitude beating ($\vec{W}$-vectors or betatron function variations), and variations in y (or J_y) and in δ of the tunes Q_y are always present in a ring containing quadrupoles and sextupoles. The Hamiltonian formalism provides means to calculate contributions to these different variations using integrals of the form (3.14). For the amplitude beating, the starting equation is the first in (3.6), which can be written for both transverse coordinates :

$$I_y = J_y + \partial S_{\phi_y} \quad . \tag{3.17}$$

Equation(3.17) gives the variation ΔI_y and hence the fluctuations of the β_y function by virtue of :

$$\frac{\Delta\beta_y}{\beta_y} = \frac{\Delta I_y}{I_y} \quad .$$

The derivatives of the phases ψ_y with respect to θ give the tune variations due to nonlinear potential and are easily related to the Hamiltonian G by the canonical equations of motion :

$$\Delta Q_y = \frac{d\psi_y}{d\theta} = \frac{\partial G}{\partial J_y} \quad . \tag{3.18}$$

Let us stop for a short while at the equation (3.18). The preceding discussion on G can be reused in the present context. In the frame of the first-order perturbation theory, G is restricted to its low-frequency terms (average over the independent variable θ) and contains resonant and stabilizing coefficients. However, in general, the Hamiltonian G contains first, second and higher-order terms in the perturbation, i.e. coefficients h_{Cp} (since $g_{Cp} = h_{Cp}$) and products of h's respectively. This last property of G holds also for its partial derivative $\partial G/\partial J_y$ appearing in (3.18), and consequently for the tune variation ΔQ_y.

As a first example, we can now write the expression of the change of Q_x which is independent of the amplitudes J_y:

$$\Delta Q_x = h_{11000} - 4 \sum_q \left(\frac{h_{2000q}\, h_{0200-q}}{2Q_x - q} + \frac{h_{1000q}\, h_{1200-q}}{Q_x - q} \right) \quad . \tag{3.19}$$

In (3.19), the presence of first and second-order contributions is obvious. From the definition (3.14), the coefficient h_{1000q} contains linearly the dipole component B, while h_{2000q} contains the first derivative B'. In quadrupoles and sextupoles sitting at positions where the dispersion D_x is not vanishing, dipole contributions will be $B'D_x\delta$ and $^1/_2 B''D_x{}^2\delta^2$, respectively. In the sextupoles, quadrupole contributions are $B''D_x\delta$. Hence, the tune dependence on δ and δ^2 is calculable from (3.19).

A second example concerns the variation of Q_x with the amplitude J_x in the presence of right sextupoles :

$$\Delta Q_x = 2J_x \left[h_{22000} - 9 \sum_q \left(\frac{h_{3000q}\, h_{0300-q}}{3Q_x - q} + \frac{1}{3} \frac{h_{2100q}\, h_{1200-q}}{Q_x - q} \right) \right] \quad . \tag{3.20}$$

By definition, all the coefficients h in (3.20) are of order 3 in amplitude and contain the sextupole component B".

Expressions of the type (3.17) to (3.20) allow us to calculate the tune and betatron variations. HARMON[3] includes in its minimization process the coefficients h_{Cq} which control the derivatives Q', Q", Q'" and β' with respect to δ and the variation of Q with the amplitude.

To sum up, perturbation theory in the canonical variables makes it possible to quantify and consequently to minimize some of the nonlinear effects due to sextupoles, necessarily present in the ring.

4. DYNAMIC APERTURE AND ANALYTICAL APPROACH

It would be ideal to have a closed formula for the dynamic aperture defined in Section 2.2, in order to make the design of accelerators easier. Even if this seems very difficult, it is justified to strike a balance between analytical and numerical methods so as to improve the analysis of the effects and "increase the useful information obtained per computational cycle"[21].

4.1 Possible use of the Hamiltonian perturbation

The recent investigations using Hamiltonian formalism aimed at the estimation of the invariant distortion in the presence of nonlinear forces and not directly at the determination of the dynamic aperture, knowing that both are related in some way. It is felt indeed that near or above the stability limit defining the dynamic aperture, either unbounded invariants (going to infinity) do exist or the invariant tori cannot be calculated analytically, because of the onset of stochasticity. As briefly recalled in

Section 3.1, the Hamiltonian perturbation requires iterations to improve the accuracy of the determination of the invariants. Consequently, the critical question is the convergence of this method, mainly close to the stability limit.

Following Section 3.1, the important point is to seek a canonical transformation which makes the new Hamiltonian G dependent only on the action J. The problem then consists of solving the Hamilton-Jacobi equation which has the form (see (3.6)) :

$$G(J) = H_0(J + \partial S_\phi) + H_1(\phi, J+\partial S_\phi, \theta) + \partial S_\theta \quad . \tag{4.1}$$

As explained in Sections 3.1 and 3.2, plain iteration can be done in the perturbation and high order terms are present in the development of the new Hamiltonian. For example, after two iterations the Hamiltonian G is of second order in the perturbation (Eq. (3.16)) and can be written formally as follows[10] for two dimensions :

$$G = Q_x J_x + Q_z J_z + \sum_{m,n} B_{mn} J_x^{m/2} J_z^{n/2} \quad . \tag{4.2}$$

The quantities G, J_x and J_z are constants of the motion and G describes to second order the distortions of the ellipses. The generating function $S(\phi_x, J_x, \phi_z, J_z, \delta)$ makes it possible to write the relations between the coordinates (see (3.6) and (3.17)) before and after the canonical transformation :

$$I_y = J_y + \partial S_{\phi_y} (\phi_y, J_y, \theta)$$

$$\psi_y = \phi_y + \partial S_{J_y} (\phi_y, J_y, \theta) \quad . \tag{4.3}$$

Since J_y are constants of motion, the first equations (4.3) give the distortions of the amplitudes (emittances), provided the following equations for the initial conditions can be solved :

$$I_{y_0} = I_{y_0} (J_x, J_z, \phi_{x_0}, \phi_{z_0}) \quad . \tag{4.4}$$

Some results based on such a second-order perturbation were compared with tracking results[10]. However, it may be difficult to solve the nonlinear algebraic equations (4.4) and the truncation of third and higher-order terms of G implies the perturbation to be small enough. Hence, the second-order approach presented here has difficulties to describe the motion near the stability limit and many plain iterations may be necessary to improve the description, since the convergence becomes slow as mentioned above.

More recently, studies[20] were carried out to introduce iterative methods in which the algebraic complexity at each iteration does not increase as rapidly as with plain iteration. In this work, the Newton method is used for solving the Hamilton-Jacobi equation (4.1). As before (Section 3.1), the linear part of the motion is separated and the Fourier transform of S (Eq. (3.9)) is taken; but no Taylor's expansions are used. The

Hamilton-Jacobi equation can then be stated in terms of the coefficients s_{np}[20] :

$$i(nQ-p)s_{np} = -\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} d\phi\, d\theta\, \exp[-i(n\phi-p\theta)] \times$$

$$\times [H_0(J+\partial S_\phi) - H_0(J) - Q\partial S_\phi + H_1(\phi, J+\partial S_\phi, \theta)], \quad n \neq 0$$

$$ip\, s_{op} = -\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} d\phi\, d\theta\, \exp(ip\theta)\, H(\phi, J+\partial S_\phi, \theta), \quad p \neq 0 \quad , \tag{4.5}$$

choosing arbitrarily $s_{00} = 0$, subtracting $H_0(J)$ and $-Q\partial S_\phi$ on both sides of (4.1), and using the fact that $G(J) - H_0(J)$ makes no contribution to non zero n (angle independent). After dividing by i(nQ-p), Eqs (4.5) may be summarized as :

$$s = A(s) \quad , \tag{4.6}$$

where s is the vector made up of the Fourier amplitudes s_{np} (Eq. (3.9)) and A is an operator defined by a double integral (Eq. (4.5)). If only a finite number of Fourier modes are included in (4.6), this equation can be solved by a Newton iteration under appropriate restrictions on the Hamiltonian (H_1 and its derivative $\partial H_1/J$ not too large) and the phase-space region considered. The interest of this approach is that in the Newton method the first iterate of s already involves terms of all orders in the perturbation H_1, in contrast to the first iteration of the plain iteration described above. This implies faster convergence. In a single-resonance model, the convergence properties were found to be excellent[20] (three Newton iterations suffice near the separatrix). In a two-resonance model, the convergence degenerates when the perturbation increases till the resonance islands overlap, as expected.

Another idea, recently considered[21], is exploiting Lie operators in order to exhibit perturbation series free of small denominators and to generate an averaged Hamiltonian of given order in the perturbation. This method has been applied to three particular problems in one dimension; zeroth-harmonic sextupoles and octupoles, quadrupole field errors and arbitrary harmonic sextupoles and octupoles.

An entirely different way to deal with the question of the dynamic aperture is to go back to the differential equation leaving aside the Hamiltonian formalism and to find good approximations for its solution.

4.2 Successive linearization of equation of motion

Let us rewrite the one-dimensional equation of motion (3.1) in the horizontal plane, with m = 3 since we are concerned with sextupole effects. Hence, introducing a differential operator L :

$$L = \frac{d^2}{ds^2} + K(s) \quad , \tag{4.7}$$

equation (3.1) simply becomes :

$$L[x(s)] - a_3(s)\, x^2(s) = 0 \quad . \tag{4.8}$$

Let us try to solve this differential equation by an iterative process of linearization[9], assuming the nonlinear part a_3 is small (a few percents) as compared to the linear one. The first approximate solution comes from a first linearization of (3.1), i.e. :

$$L[x^{(0)}(s)] = 0 \quad . \tag{4.9}$$

The solution of this purely linear equation is well-known and given in detail in Ref. 26. For the complete solution which must contain the effect of the nonlinear term, we can write by definition :

$$x(s) = x^{(0)}(s) + u(s) \quad , \tag{4.10}$$

and the corresponding equation for u(s) becomes :

$$L(u) - 2a_3(s)\, x^{(0)}(s)u - a_3(s)u^2 = a_3(s)x^{(0)^2}(s) \quad . \tag{4.11}$$

The next approximation consists of dropping the quadratic term in u and this step is called second linearization :

$$L[u^{(0)}] - 2a_3(s)\, x^{(0)}(s)\, u^{(0)} = a_3(s)\, x^{(0)^2}(s) \quad . \tag{4.12}$$

The initial conditions must satisfy :

$$\begin{aligned} x_0 &= x^{(0)}(0) + u^{(0)}(0) \\ x'_0 &= x^{(0)'}(0) + u^{(0)'}(0) \end{aligned} \quad . \tag{4.13}$$

In general, the initial values can be arbitrarily distributed between $x^{(0)}$ and $u^{(0)}$, but the choice :

$$\begin{aligned} x^{(0)}(0) &= x_0 \qquad & u^{(0)}(0) &= 0 \\ x^{(0)'}(0) &= x_0' \qquad & u^{(0)'}(0) &= 0 \end{aligned} \quad , \tag{4.14}$$

proves to be quite adequate in practice. The linearization process can be continued to higher level, giving a more and more precise solution.

As mentioned before, a typical feature of the solutions for the nonlinear betatron motion is to remain bounded up to a certain amplitude and then to become suddenly unbounded. The threshold-amplitude is called the dynamic aperture of the system. The term proportional to $x^{(0)2}$ in (4.11) indicates that, from a certain amplitude onwards, the nonlinear term will contribute dominantly to the behaviour of the solution. At this point, a rapid self-amplification will take place. So far, it seems that the effect of dynamic aperture is a pure large-amplitude effect, whose occurence cannot really be described by perturbative methods.

In real systems, however, the stability limit frequently occurs at small amplitudes, before the nonlinear term becomes dominant, so that perturbation treatment is applicable. This should correspond to another mechanism driving the solution in the self-amplification regime. To look for such a mechanism, let us start from (4.12). The term proportional to $a_3\ x^{(0)}$ on the left hand side of (4.12) represents a focusing force depending on the amplitude, which may induce parametric resonances. The term on the right hand side drives the linear stopband as well as third-order resonances associated with unbounded motion for distinct values of the phase. The consequence is that, in the general case where the phase is not on a resonance, the mechanism driving the solution in the self-amplification regime seems to be the parametric resonances.

Since $u^{(0)}$ is unbounded if the homogeneous part of the solution is unbounded, the homogeneous equation :

$$L[u^{(0)}] - 2a_3(s)\ x^{(0)}(s)\ u^{(0)} = 0 \quad , \tag{4.15}$$

contains the whole information about the dynamic aperture, within the approximation of the second linearization and the assumption that the parametric resonance mechanism prevails.

Let us now discuss the equation (4.15). In general, $x^{(0)}(s)$ is not periodic over one magnetic period described by the functions $K(s)$ and $a_3(s)$. Consequently, Eq. (4.15) can only be reduced to a vector recurrence containing a non-constant transfer matrix :

$$\begin{pmatrix} u^{(0)}_{n+1} \\ u^{(0)'}_{n+1} \end{pmatrix} = M_n \begin{pmatrix} u^{(0)}_{n} \\ u^{(0)'}_{n} \end{pmatrix} \quad . \tag{4.16}$$

However,if the normalized phase advance $\mu/2\pi$ per magnetic period is rational, i.e. $\mu = 2\pi p/q$, p and q being integers, $x^{(0)}$ is periodic with a period equal to q-times the magnetic period and the equation becomes again of Hill's type. The associated matrix over one new period is independent of n :

$$R = \prod_{i=0}^{q-1} M_i \tag{4.17}$$

and its elements are polynoms in x_0 and x_0'. The theory of linear motion[26] tells us that the solution $u^{(0)}$ will be bounded if the following condition is fullfilled :

$$|Tr(R)| < 2 \quad . \tag{4.18}$$

It has to be noted that the condition for boundedness of the motion is an explicit function of x_0, x_0' and can directly be used for an estimation of the stability limit.

In this approach, it was assumed that the second mechanism driving nonlinear instabilities is the occurence of parametric resonances caused by fluctuating transfer matrices in the perturbation equation for $u^{(0)}$, Eq. (4.15). These resonances appear at small amplitudes driving the solution towards higher amplitudes where the self-amplification takes over. It should be recalled that this mechanism does not necessarily prevail for all nonlinear motions and all the values of the parameters, like the strength a_3 and the phase advance μ of the treated example. As already explained, rapid self-amplification can take place suddenly at a certain amplitude and chaotic motion may also arise. This is for instance possible in the two-resonance model, when the resonance islands nearly overlap. In such cases, the linearization process described above also becomes difficult.

As an example,the linearization method was applied to a FODO lattice[9] with two families of sextupoles near the focusing and defocusing quadrupoles, respectively. The normalized phase advance $Q = \mu/2\pi$ is varied from 30/120 to 50/120. Applying the condition (4.18) leads to a curve $x_0 = x_0(Q)$ giving the maximum initial value for which the motion $u^{(0)}$ is still bounded, i.e. the stability limit as a function of the phase (dotted curve in Fig. 5). This curve is compared with the expected limit (full curve in Fig. 5) and with the resonance fixed-point-limit[22,23] calculable near the third integer resonance (the large points in Fig. 5). The expected limit was found by numerical computation of the exact nonlinear transformation over 10^5 cells, while increasing x_0 in steps of 1 mm until unboundedness was reached (overflow). The fixed-point approach breaks down if Q is distant from 1/3 by more than about 1/120.

Comparing the two curves of Fig. 5, the agreement is remarkably good for Q-values between 34/120 and 50/120. Between 34/120 and 30/120, the linearization method (with two iterates) begins to fail, while below 30/120 it does not succeed in giving a stability limit different from zero. Even if the results can possibly be improved by adding more iterates, it is felt that for Q smaller than 30/120 large amplitude effects and chaotic motion set the dynamic aperture. Consequently, for these particular Q-values, the method can not really be applied in the form it has been presented here.

The linearization of the equation of motion and the associated stability condition offer some advantage with respect to tracking of many particles over a large number of periods, since it implies evaluating the product of a reasonable number (about 100) of matrices. This method is being extended to two-dimensional motion and its range of validity explored.

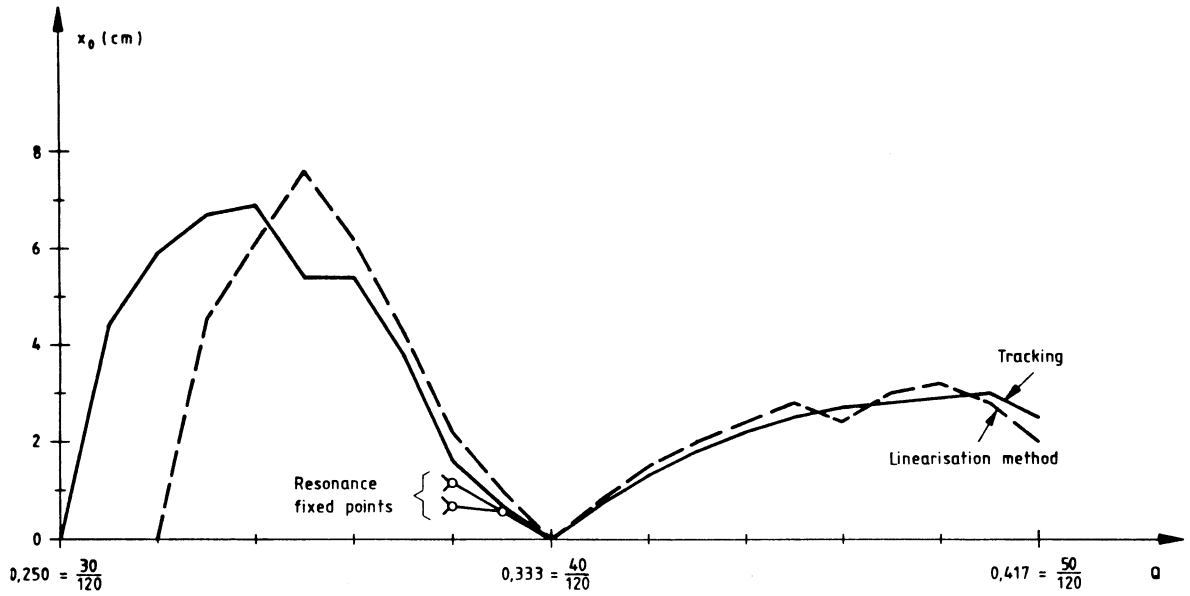


Fig. 5 Horizontal stability limit as a function of the normalised phase advance, for a FODO lattice

* * *

REFERENCES

1) B.W. Montague, Chromatic effects and their first-order corrections, these proceedings.
2) G. Guignard, First order chromatic perturbations and sextupole strength, CERN ISR-TH/82-14 (1982).
3) M. Donald, Chromaticity correction in circular accelerators and storage rings, PEP Note 311, July 1979.
4) F.C. Iselin, the MAD program, Proc. 12th International Conference on High Energy Accelerators, FNAL (1983).
5) D.C. Carey, TURTLE, A computer program for simulating charged particle beam transport systems, Fermilab NAL-64, December 1971.
6) K.L. Brown, M. Donald, R. Servranckx, Optical distortions in electron-positron storage rings, Proc. 12th International Conference on High Energy Accelerators, FNAL, 1983, and K.L. Brown, A. Verdier, A study of beam particle losses in the LEP experimental areas based on PEP experience, LEP Note 557, March 1986.
7) S.Y. Lee and S. Tepikian, the chaotic dynamical aperture, Particle Accelerator Conference, Vancouver, May 1985 (IEEE, New York, 1985).
8) F. Willeke, Analytical study of the Tevatron non-linear dynamics, Fermilab FN-22, July 1985.
9) G. Guignard and J. Hagel, Sextupole correction and dynamic aperture, numerical and analytical tools, CERN LEP-TH/85-3 (1985); Particle Accelerators, Vol.18 (3) 129, (1986).

10) A. Ando, Distortion of beam emittance with non-linear magnetic fields, Particle Accelerators, Vol. 15 (3) 177 (1984).

11) A. Wrulich, Particle tracking, these proceedings.

12) H. Wiedemann, Chromaticity correction in large storage rings, PEP Note 220 (1976) and User's guide for PATRICIA, PEP technical memo, PTM-230 (1981).

13) S. Peggs, Hadron collider behaviour in the non-linear numerical model EVOL, CBN 84-12 (1984).

14) A. Wrulich, RACETRACK, A computer code for the simulation of non-linear particle motion in accelerators, DESY 84-026, March 1984.

15) K.L. Brown, D.C. Carey, Ch. Iselin and F. Rothaker, TRANSPORT, A computer program for designing charged particle beam transport systems, CERN 80-04 (1980).

16) R.V. Servranckx, K.L. Brown, L. Schachinger, D. Douglas, User's Guide to the program DIMAD, SLAC Report 285, UC-28(A), May 1985.

17) A.J. Dragt, Lectures on non-linear orbit dynamics, Physics of High Energy Particle Accelerators, A.I.P. Conf. Proc., 87 (1981).

18) A.J. Dragt, R. Ryne, L.M. Healy, F. Neri, D.R. Douglas and E. Forest, MARYLIE, A program for charged particle beam transport based on Lie Algebraic Methods, to be published.

19) T.L. Collins, Distortion functions, Fermilab FN-84/114 (1984).

20) R.L. Warnock, R.D. Ruth and T. Raubenheimer, Superconvergent tracking and invariant surfaces in phase space, Particle Accelerator Conference, Vancouver, May 1985 (IEEE, New York, 1985).

21) L. Michelotti, Moser-like transformations using the Lie transform, Particle Accelerators, Vol. 16 (4) 233 (1985).

22) G. Guignard, A general treatment of resonances in accelerators, CERN 78-11 (1978).

23) A. Schoch, Theory of linear and non-linear perturbations of betatron oscillations in alternating gradient synchrotrons, CERN 57-23 (1958).

24) R. Hagedorn, Stability and amplitude ranges of two dimensional non-linear oscillations with periodical Hamiltonian, CERN 57-1 (1957).

25) G. Guignard, Effets des champs magnétiques perturbateurs d'un synchrotron sur l'orbite fermée et les oscillations bétatroniques, ainsi que leur compensation, CERN 70-24 (1970).

26) E.D. Courant and H.S. Snyder, Theory of the alternating-gradient synchrotron, Ann. Phys. (U.S.A.) Vol. 3 (1958) 1.

THEORY OF RF ACCELERATION

G. Dôme
CERN, Geneva, Switzerland

ABSTRACT

Formulae for RF acceleration and synchrotron motion are derived from basic principles in the case of an arbitrary RF voltage.

1. ENERGY GAIN AND TRANSIT TIME FACTOR

Particles experience the effect of RF fields when they cross accelerating gaps that basically produce an electric field $\vec{\mathcal{E}}$ parallel to their trajectories. The gap is the space between two electrodes provided with a beam pipe, which for simplicity we take as a circular cylinder of radius a.

Let $\hat{V}(r)$ be the amplitude of the RF voltage impressed across the two electrodes. When a particle with electric charge e (which may be larger than an electron charge) crosses the gap at a distance r from the s-axis (see Fig. 1.1), it gains an energy $\Delta E = e\int \mathcal{E}_s(s,r,t)\,ds$.

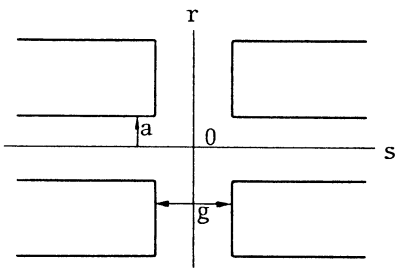


Fig. 1.1 - Longitudinal cross section of an accelerating gap

The time dependence of $\mathcal{E}_s$ is given by

$$\mathcal{E}_s(s,r,t) = \hat{\mathcal{E}}_s(s,r)\ \sin\ (\omega_{RF}t)$$

Traditionally, for circular accelerators the origin of time is taken at the zero crossing of the RF voltage with positive slope. The phase ϕ of the RF voltage when a particle crosses the middle of the accelerating gap (at s = 0) is called the *phase of the particle* with respect to the RF voltage. On the other hand, for circular accelerators in the Russian literature and for linacs, the origin of time is taken at the crest of the RF voltage. The phase φ in that case is such that $\phi = \frac{\pi}{2} + \varphi$. (Strictly speaking, in the previous sentences, the term "RF voltage" should be understood as "RF voltage times the charge e of the particles"). If we neglect the change in velocity of the particle when crossing the gap, the time t when the particle is at position s in the gap reads

$$t = \frac{\phi}{\omega_{RF}} + \frac{s}{v} \quad \text{and} \quad \omega_{RF} t = \phi + \frac{\omega_{RF}}{v} s$$

where v is the particle velocity in the middle of the gap.

For simplicity's sake we assume that the gap is symmetric with respect to the plane s = 0; then

$$\Delta E = e \int \hat{\mathcal{E}}_s (s,r) \sin\left(\phi + \frac{\omega_{RF}}{v} s\right) ds = e \sin\phi \int \hat{\mathcal{E}}_s (s,r) \cos\left(\frac{\omega_{RF}}{v} s\right) ds \qquad (1.1)$$

By representing the fields for $r \leq a$ as Fourier integrals along s, one gets

$$\int_{-\infty}^{+\infty} \hat{\mathcal{E}}_s(s,r) \cos\left(\frac{\omega_{RF}}{v} s\right) ds = \frac{I_0\left(\frac{\omega_{RF}}{v} \cdot \frac{r}{\gamma}\right)}{I_0\left(\frac{\omega_{RF}}{v} \cdot \frac{a}{\gamma}\right)} \int_{-\frac{g}{2}}^{+\frac{g}{2}} \hat{\mathcal{E}}_s(s,a) \cos\left(\frac{\omega_{RF}}{v} s\right) ds$$

With $\hat{V}(r) = \int_{-\infty}^{+\infty} \hat{\mathcal{E}}_s(s,r) ds = \hat{V}(0) \cdot J_0\left(\frac{\omega_{RF}}{c} r\right)$, this may be written as

$$\int_{-\infty}^{+\infty} \hat{\mathcal{E}}_s(s,r) \cos\left(\frac{\omega_{RF}}{v} s\right) ds = \hat{V}(r) \cdot T(r) = \frac{I_0\left(\frac{\omega_{RF}}{v} \cdot \frac{r}{\gamma}\right)}{I_0\left(\frac{\omega_{RF}}{v} \cdot \frac{a}{\gamma}\right)} \cdot \hat{V}(a)\, T(a)$$

where by definition

$$T(r) = \frac{\int_{-\infty}^{+\infty} \hat{\mathcal{E}}_s(s,r) \cos\left(\frac{\omega_{RF}}{v} s\right) ds}{\int_{-\infty}^{+\infty} \hat{\mathcal{E}}_s(s,r)\, ds}$$

is the *transit time factor* at r; it is the ratio of peak energy gained by a particle with velocity v to the same quantity if v were infinite.

T(r) is simplest at r = a, where $\hat{\mathcal{E}}_s$ is zero outside the gap. In many practical cases, a good approximation is obtained when $\hat{\mathcal{E}}_s(s,a)$ is considered to be constant in the gap; then

$$T(a) = \frac{\sin\left(\frac{\omega_{RF}}{v} \cdot \frac{g}{2}\right)}{\left(\frac{\omega_{RF}}{v} \cdot \frac{g}{2}\right)} \qquad (1.2)$$

Finally,

$$\Delta E = eV \sin \phi \tag{1.3}$$

where

$$V = \hat{V}(a) \cdot T(a) \frac{I_0\left(\frac{\omega_{RF}}{v}\frac{r}{\gamma}\right)}{I_0\left(\frac{\omega_{RF}}{v}\frac{a}{\gamma}\right)} = \hat{V}(0) \cdot J_0\left(\frac{\omega_{RF}}{c} a\right) \cdot T(a) \frac{I_0\left(\frac{\omega_{RF}}{v}\frac{r}{\gamma}\right)}{I_0\left(\frac{\omega_{RF}}{v}\frac{a}{\gamma}\right)}$$

with $eV > 0$. Neglecting the second order variation in r due to $I_0\left(\frac{\omega_{RF}}{v}\frac{r}{\gamma}\right)$, we are left with

$$V = \hat{V}(a) \cdot \frac{T(a)}{I_0\left(\frac{\omega_{RF}}{v}\frac{a}{\gamma}\right)} \qquad \text{for all particles} \tag{1.4}$$

It is seen that through the transit time factor and the Bessel function $I_0\left(\frac{\omega_{RF}}{v}\frac{a}{\gamma}\right)$, the effective peak voltage V depends on the particle velocity v. This effect will be neglected in what follows, so that all particles will be considered as experiencing the same peak voltage.

More precise (but more complicated) expressions for ΔE can be found in Ref. 1.

2. HARMONIC NUMBER

For some reference particle (also called synchronous particle), the phase ϕ is kept unchanged (mod 2π) at a value ϕ_s when the particle returns to the same accelerating gap after one revolution along the ring. This requires that $\omega_{RF} = h\,\omega_0$ where $\omega_0 = 2\pi/T_0$ is the angular revolution frequency of the reference particle and h is an integer called *harmonic number*.

Then

$$\omega_{RF}\, T_0 = 2\pi h \tag{2.1}$$

When the ring is large, ω_0 is small and h may be quite a big number.

3. FINITE DIFFERENCE EQUATIONS

For simplicity, let us assume that RF acceleration takes place in N identical cavities evenly spaced along the synchrotron ring. Let n be the number of accelerating cavity traversals by a particle.

Definition of variables (see Fig. 3.1)

p_s, v_s, momentum and velocity of the reference (synchronous) particle

t_n, time of n^{th} cavity traversal by the reference particle

$$\delta p_n = p_n - p_s$$

$$\delta \phi_n = \phi_n - \phi_s$$

In what follows, δ represents a difference taken with respect to the reference particle at a given time; d represents an increment during acceleration.

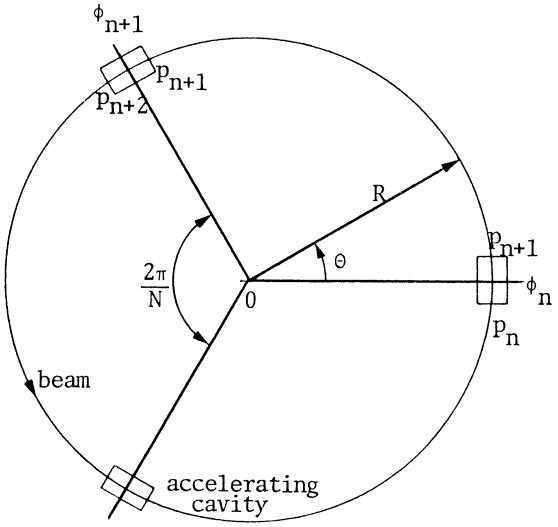


Fig. 3.1 Definition of variables

Besides the general coordinates (R,Θ) whose origin is the accelerator centre, each bending magnet has its own local coordinates (r,θ) whose origin is the centre of the reference particle orbit in the magnet. Any integral with respect to θ is taken in the bending magnets only.

Phase variation between adjacent cavities

In order to keep the phase of the reference particle constant at every cavity traversal, the RF phase must be shifted by $2\pi h/N$ between adjacent cavities. The phase of any particle with respect to the RF voltage is then given by

$$\phi(t) = \int \omega_{RF}\, dt - h\Theta(t) \tag{3.1}$$

where $\Theta(t)$ is the azimuthal position of the particle. With this relation ϕ is not only defined during cavity traversals, but it is defined at any time. In particular, for the reference particle,

$$\frac{d\phi_s}{dt} = \omega_{RF} - h\omega_0 = 0 \tag{3.2}$$

For a particle with an energy deviation $\delta\gamma/\gamma_s$ with respect to the reference particle, the phase is compared to ϕ_s. If T_r is the revolution period, the variation of $\delta\phi_n$ from one cavity to the next is

$$\delta\phi_{n+1} - \delta\phi_n = \phi_{n+1} - \phi_n = \omega_{RF}\frac{T_r}{N} - h\frac{2\pi}{N} = \frac{\omega_{RF}}{N}(T_r - T_0)$$

But $T_r = C/v$ where C is the orbit circumference and v is the particle velocity. For a relative momentum deviation $\delta p/p$,

$$\frac{\delta C}{C} = \alpha \frac{\delta p}{p} \quad \text{by definition of the momentum compaction } \alpha; \quad \alpha = \frac{1}{2\pi R}\int D_x(s)\, d\theta$$

where $D_x(s)$ is the radial dispersion.

$$\frac{\delta v}{v} = \frac{1}{\gamma^2}\frac{\delta p}{p} \quad \text{by relativistic kinematics.}$$

Therefore, if ω_r is the angular revolution frequency of a particle,

$$\frac{\delta\omega_r}{\omega_o} = -\frac{\delta T_r}{T_0} = -\left(\alpha - \frac{1}{\gamma^2}\right)\frac{\delta p}{p} = \eta\frac{\delta p}{p} \qquad \text{where}^{*)} \qquad \eta = \frac{1}{\gamma^2} - \alpha \tag{3.3}$$

If η vanishes for some γ, this particular energy is called the *transition energy* γ_{tr}; when α is independent of γ, α is equal to $1/\gamma_{tr}^2$. Finally,

$$\boxed{\delta\phi_{n+1} - \delta\phi_n = \phi_{n+1} - \phi_n = \frac{\omega_{RF}}{N} T_0 \frac{\delta T_r}{T_0} = -\frac{2\pi h}{N}\eta\frac{\delta p_{n+1}}{p} = -\boxed{\frac{2\pi h}{N}\frac{\eta}{\beta^2}\frac{\delta\gamma_{n+1}}{\gamma}}} \tag{3.4}$$

since $\frac{\delta\gamma}{\gamma} = \beta^2\frac{\delta p}{p}$.

Energy variation between adjacent cavities

Since all accelerating cavities are assumed to be identical, the total RF voltage produced along the ring is NV. With (1.3) we have

$$\Delta E = E_{n+1} - E_n = eV \sin\phi_n - \frac{e}{N}\oint\frac{\partial B_z}{\partial t}\cdot r\, d\theta\, dr \tag{3.5}$$

In the righthand side, the first term represents an energy gain which is lumped in the accelerating cavities, whereas the second term represents an energy gain which is distributed all along the magnets. *Although the second term is usually negligible with respect to the first one, its variation for particles with different energies must not be overlooked.* For the reference particle

$$\Delta E_s = E_{s,n+1} - E_{s,n} = eV \sin\phi_n - \frac{e}{N}\oint\frac{\partial B_{zs}}{\partial t}\cdot r_s\, d\theta\, dr$$

To first order,

$$\frac{N}{2\pi}(\Delta E - \Delta E_s) = \frac{eNV}{2\pi}(\sin\phi_n - \sin\phi_s) - \frac{e}{2\pi}\oint\frac{\partial B_{zs}}{\partial t}\, r_s\, d\theta\, \delta x \tag{3.6}$$

*) This definition of η is the same as the one used in Ref. 2, but another definition of η which differs in sign is also used in the literature.

But

$$\delta x = D_x(s) \cdot \frac{\delta p}{p} \qquad \text{and} \qquad p = -e\, B_z \cdot r = -e \langle B_z \rangle R \tag{3.7}$$

where $< B_z >$ is the average magnetic field along a closed orbit:

$$< B_z > = \frac{1}{2\pi R} \oint B_z \cdot r \, d\theta$$

With (3·7) the last term of (3·6) becomes

$$- \frac{e}{2\pi} \frac{\delta p}{p} \frac{\partial B_{zs}}{\partial t} r_s \int D_x(s) d\theta = - \frac{e}{2\pi} \frac{\delta p}{p} \frac{\partial <B_{zs}>}{\partial t} R_s \, 2\pi \, R_s \, \alpha = -e <B_{zs}> R_s \frac{\delta p}{p} \left[\frac{\dot p}{p} - \frac{1}{\alpha} \frac{\dot R}{R} \right] R_s \, \alpha$$

$$= \delta p \left[\alpha \frac{\dot p}{p} - \frac{\dot R}{R} \right] R_s \tag{3.8}$$

because from (3.7),

$$\frac{\dot p}{p} = \frac{1}{<B_z>} \left[\frac{\partial <B_z>}{\partial t} + \frac{\partial <B_z>}{\partial R} \frac{dR}{dt} \right] + \frac{1}{R} \frac{dR}{dt} = \frac{1}{<B_z>} \cdot \frac{\partial <B_z>}{\partial t} + \underbrace{\left[\frac{R}{<B_z>} \cdot \frac{\partial <B_z>}{\partial R} + 1 \right]}_{\left[\text{average magnetic field index} + 1\right] = \frac{1}{\alpha}} \frac{\dot R}{R} \tag{3.9}$$

Now we must remember that ΔE, ΔE_s are gained in different times T_r/N, T_0/N. As next approximation, E and E_s are considered to be smooth functions of t:

$$\Delta E \approx \frac{T_r}{N} \frac{dE}{dt} \quad , \quad \Delta E_s \approx \frac{T_0}{N} \frac{dE_s}{dt}$$

With (3.3),

$$\Delta E - \Delta E_s = \frac{T_r}{N} \frac{dE}{dt} - \frac{T_0}{N} \frac{dE_s}{dt} = \frac{T_r}{N} \frac{d(E - E_s)}{dt} + \frac{T_r - T_0}{N} \cdot \frac{dE_s}{dt} = \frac{T_r}{N} \frac{d(\delta E)}{dt} - \eta \frac{\delta p}{p} \frac{T_0}{N} \cdot \frac{dE_s}{dt}$$

Using the kinematic relations $dE = v \cdot dp = \omega_r R \, dp$, $\delta E = v \cdot \delta p = \omega_0 R_s \, \delta p$:

$$\frac{N}{2\pi} (\Delta E - \Delta E_s) = \frac{1}{\omega_r} \frac{d(\delta E)}{dt} - \eta \frac{\delta p}{p} \frac{1}{\omega_0} \frac{dE_s}{dt} = \frac{d}{dt} \left(\frac{\delta E}{\omega_0} \right) + \underbrace{\frac{\delta E}{\omega_0} \frac{\dot\omega_0}{\omega_0} - \eta \frac{\delta p}{p} R_s \dot p}_{} + \text{2nd order terms.}$$

$$R_s \delta p \left(\cancel{\frac{\dot v}{v}} - \frac{\dot R}{R} \right) + R_s \delta p \left(\alpha - \cancel{\frac{1}{\gamma^2}} \right) \frac{\dot p}{p} = R_s \delta p \left[\alpha \frac{\dot p}{p} - \frac{\dot R}{R} \right]$$

which is exactly equal to (3·8), i.e. to the last term of (3·6). Finally (3·6) reduces to

$$\boxed{\frac{d}{dt}\left(\frac{E - E_s}{\omega_0}\right) = \frac{eNV}{2\pi}(\sin\phi - \sin\phi_s)} \qquad (3.10)$$

The present derivation of Eq. (3.10) is the same as in Ref. 3 (p.156-163); this equation may also be derived by using the generalized angular momentum $\vec{r} \times (\vec{p} + e\vec{A})$ instead of $\vec{B}$ (Ref. 4). Any correct derivation must take into account the electromotive force induced by a varying magnetic flux; it happened often in the past that Eq. (3.10) was either wrongly derived or wrongly stated.

The corresponding finite difference equation reads

$$\left(\frac{E - E_s}{\omega_0}\right)_{n+1} - \left(\frac{E - E_s}{\omega_0}\right)_n = \frac{eV}{\omega_{0,n}}(\sin\phi_n - \sin\phi_s) \quad \text{where} \quad \frac{E - E_s}{\omega_0} = R_s(p - p_s) \qquad (3.11)$$

Finally, about the set of finite difference equations (3.4), (3.11), one should quote H. Hereward: "These equations are only roughly correct, and it is work to estimate how good they are" (Ref. 4, p. 11).

Betatron electromotive force

The betatron e.m.f. $\mathcal{E}$ along a closed orbit appears in (3-5) as

$$\Delta E = eV \sin\phi + \frac{e\mathcal{E}}{N} \qquad (3.12)$$

where

$$\mathcal{E} = -\frac{\partial}{\partial t}\oint B_z \cdot r\, d\theta\, dr$$

In this integral, the closed orbit should be considered to be fixed with respect to time. Computing the integral first with respect to r, this may be rewritten as

$$\mathcal{E} = -\frac{\partial}{\partial t}\oint b(r,\theta) \cdot B_z\,(x = 0) \cdot r\, d\theta$$

where r is now taken on the closed orbit $x=0$, and $b(r,\theta)$ is an effective magnet width inside the orbit. Using (3-7) it is seen that

$$e\,\mathcal{E} = \frac{\partial}{\partial t}\oint b(r,\theta) \cdot p\, d\theta = \frac{\partial}{\partial t}\left[p \oint b(r,\theta) \cdot d\theta\right]$$

or

$$e\,\mathcal{E} = 2\pi\frac{\partial}{\partial t}\left[p \cdot b\right] \qquad \text{where} \qquad b \equiv \langle b(R) \rangle = \frac{1}{2\pi}\oint b(r,\theta) \cdot d\theta \qquad (3.13)$$

$b = \langle b(R) \rangle$ appears to be an average effective magnet width inside a closed orbit. Since the orbit must be considered as being fixed in time, the dependence of b on time can only be due to possible changes in the configuration of B_z when the magnetic field is increased; and these changes are kept as small as possible. By using (3.7) and (3.9) we obtain

$$\frac{e\mathcal{E}}{2\pi} = b \cdot \frac{\partial p}{\partial t} + p \frac{\partial b}{\partial t} \qquad \text{where} \qquad \frac{1}{p}\frac{\partial p}{\partial t} = \frac{1}{\langle B_z \rangle}\frac{\partial \langle B_z \rangle}{\partial t} = \frac{\dot{p}}{p} - \frac{1}{\alpha}\frac{\dot{R}}{R} \tag{3.14}$$

and where ideally $\frac{\partial b}{\partial t}$ should be negligible.

Since

$$\Delta E \approx \frac{T_r}{N} \cdot \frac{dE}{dt} = \frac{T_r}{N} \cdot \omega_r R \frac{dp}{dt} = \frac{2\pi}{N} R \frac{dp}{dt} \quad ,$$

(3-12) becomes

$$R \frac{dp}{dt} = \frac{eNV}{2\pi} \sin\phi + \frac{e\mathcal{E}}{2\pi} \tag{3.15}$$

Finally, with (3-14),

$$R \frac{dp}{dt} = \frac{eNV}{2\pi} \sin\phi + b \left[\frac{dp}{dt} - \frac{p}{\alpha R}\frac{dR}{dt}\right] + p \frac{\partial b}{\partial t} \tag{3.16}$$

In this relation the $\frac{\partial b}{\partial t}$ - term should be negligible, while the b-term is (at most) of the order $b \frac{dp}{dt}$, which is a factor b/R smaller than the left-hand side. Therefore, when $b \ll R$ (which is the case for all large synchrotrons), acceleration due to the betatron electromotive force is a small fraction of the acceleration produced by the RF voltage.

4. DIFFERENTIAL EQUATIONS FOR AN ARBITRARY RF VOLTAGE

If higher harmonics are added to the fundamental sinusoidal RF field, in Eq. (3.10) and (3.11) $\sin\phi$ must be replaced by a more general function $g(\phi)$ such that

$$g(\phi + 2\pi) = g(\phi) \qquad \text{and} \qquad \int_0^{2\pi} g(\phi)\, d\phi = 0 \tag{4.1}$$

hence

$$g(\phi) = \sum_{n=1}^{\infty} (a_n \sin n\phi + b_n \cos n\phi) \tag{4.2}$$

where one can take $a_1 = 1$, $b_1 = 0$, since the values of a_1, b_1 are defined by normalization and by the choice of the origin of time. Eq. (3.10) then becomes

$$\frac{d}{dt}\left(\frac{\delta E}{\omega_0}\right) = \frac{eNV}{2\pi}\left[g(\phi) - g(\phi_s)\right] \tag{4.3}$$

which has to be combined with the differential form of (3.4):

$$\frac{d}{dt}(\phi - \phi_s) = -h\,\omega_0 \frac{\eta}{\beta^2}\frac{\delta E}{E} \tag{4.4}$$

Instead of ϕ and $\delta E/\omega_0$ as conjugate variables, we shall use ϕ and $\delta E/(h\omega_0)$, so that the elementary phase space area will read

$$\frac{\delta E}{h\omega_0} \wedge \delta\phi = \frac{\delta E}{h\omega_0} \wedge \delta(\omega_{RF} t) = \delta E \wedge \delta t$$

With ϕ and $\delta E/(h\omega_0)$ as conjugate variables, the system (4.3) , (4.4) becomes

$$\boxed{\frac{d}{dt}\left(\frac{\delta E}{h\omega_0}\right) = \frac{eNV}{h\,2\pi}\left[g(\phi) - g(\phi_s)\right]} \tag{4.5}$$

$$\boxed{\frac{d}{dt}(\phi - \phi_s) = -h^2\,\omega_0^2 \cdot \frac{\eta}{\beta^2 E} \cdot \left(\frac{\delta E}{h\omega_0}\right)} \quad \text{where} \quad E = \gamma E_0 = \gamma m_0 c^2 \tag{4.6}$$

With (3.2), the system (4.5) , (4.6) may be derived from the Hamiltonian

$$H = -\frac{1}{2} h^2 \omega_0^2 \frac{\eta}{\beta^2 \gamma E_0}\left(\frac{\delta E}{h\omega_0}\right)^2 + \frac{eNV}{h2\pi}\left[\Gamma\phi + G(\phi)\right] \tag{4.7}$$

where

$$\Gamma \equiv g(\phi_s) \quad \text{and} \quad G(\phi) = -\int g(\phi)\,d\phi = \sum_{n=1}^{\infty}\left(\frac{a_n}{n}\cos n\phi - \frac{b_n}{n}\sin n\phi\right) \tag{4.8}$$

H depends explicitly on time through parameters which vary slowly during acceleration.

5. HAMILTONIAN WITH REDUCED VARIABLES

The study of particle motion can be simplified by using reduced dimensionless variables y and t* instead of $\delta E/(h\omega_0)$ and t. Let

$$\begin{cases} dt = K_1\, dt^* \\ \dfrac{\delta E}{h\omega_0} = K_2\, y \end{cases} \qquad \text{where } K_1,\ K_2 \text{ are slowly varying parameters.} \tag{5.1}$$

With the reduced variables and reduced Hamiltonian H*, the equations of motion read

$$\frac{dy}{dt^*} = -\frac{K_1}{K_2}\frac{\partial H}{\partial \phi} = -\frac{\partial H^*}{\partial \phi} \qquad\qquad \frac{d\phi}{dt^*} = \frac{K_1}{K_2}\frac{\partial H}{\partial y} = \frac{\partial H^*}{\partial y} \tag{5.2}$$

whence

$$H^* = \frac{K_1}{K_2} H = -\frac{1}{2} h^2\omega_0{}^2 \frac{\eta}{\beta^2\gamma E_0} K_1 K_2 y^2 + \frac{eNV}{h2\pi}\frac{K_1}{K_2}\left[\Gamma\phi + G(\phi)\right]$$

By taking

$$-h^2\omega_0{}^2 \frac{\eta}{\beta^2\gamma E_0} K_1 K_2 = 1 \quad \text{and} \quad \frac{eNV}{h2\pi}\frac{K_1}{K_2} = -\operatorname{sgn}(\eta) \tag{5.3}$$

i.e.

$$\frac{1}{K_1} = -\operatorname{sgn}(\eta)\, h\omega_0 \left|\frac{\eta}{\beta^2\gamma}\frac{eNV}{h2\pi E_0}\right|^{\frac{1}{2}}, \qquad K_2 = \frac{E_0}{h\omega_0}\left|\frac{\beta^2\gamma}{\eta}\frac{eNV}{h2\pi E_0}\right|^{\frac{1}{2}} \tag{5.4}$$

H* becomes

$$\boxed{H^* = \frac{Y^2}{2} - \operatorname{sgn}(\eta)\left[\Gamma\phi + G(\phi)\right]} \tag{5.5}$$

Fixed points. From (5.2) they correspond to $\dfrac{\partial H^*}{\partial \phi} = \dfrac{\partial H^*}{\partial y} = 0$

i.e., with (4.8)': $\Gamma + G'(\phi) = 0$ or $\Gamma \equiv g(\phi_s) = g(\phi)$ with $y = 0$.

Because of (4.1), beside ϕ_s there will be in general another value of ϕ satisfying the condition $g(\phi) = \Gamma$. Let ϕ_0 be any one of them; for small $\psi = \phi - \phi_0$,

$$\Gamma\phi + G(\phi) = \left[\Gamma\phi_0 + G(\phi_0)\right] - g'(\phi_0)\frac{\psi^2}{2!} - g''(\phi_0)\frac{\psi^3}{3!} - \cdots$$

$$H^* = -\operatorname{sgn}(\eta)\left[\Gamma\phi_0 + G(\phi_0)\right] + \frac{Y^2}{2} + \operatorname{sgn}(\eta)\cdot g'(\phi_0)\frac{\psi^2}{2} + \operatorname{sgn}(\eta)\cdot g''(\phi_0)\frac{\psi^3}{6} \tag{5.6}$$

If $\operatorname{sgn}(\eta)\cdot g'(\phi_0) > 0$, ϕ_0 is an elliptic fixed point; this is the case of $\phi_0 = \phi_s$ for ϕ_s being a stable fixed point.

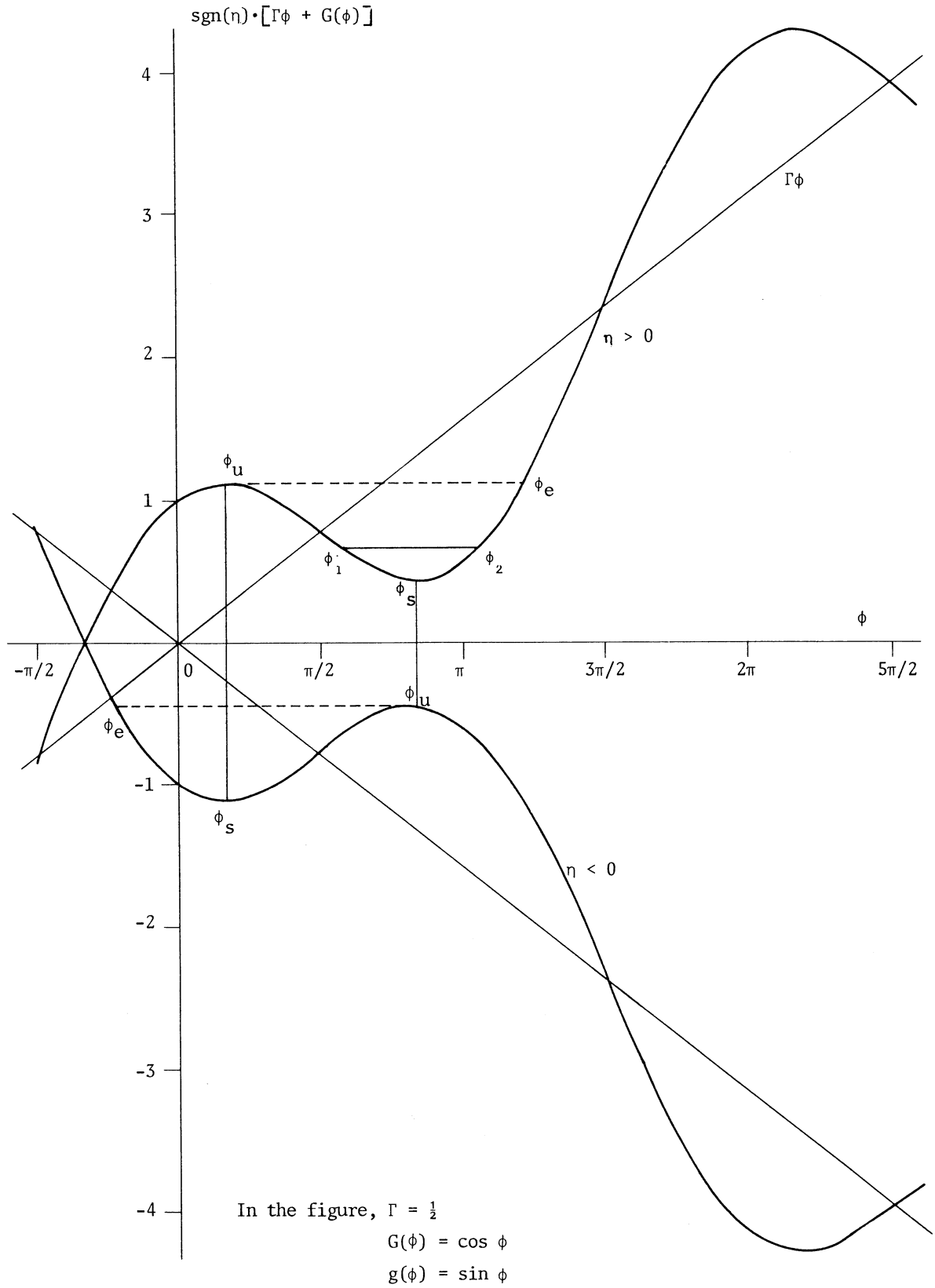
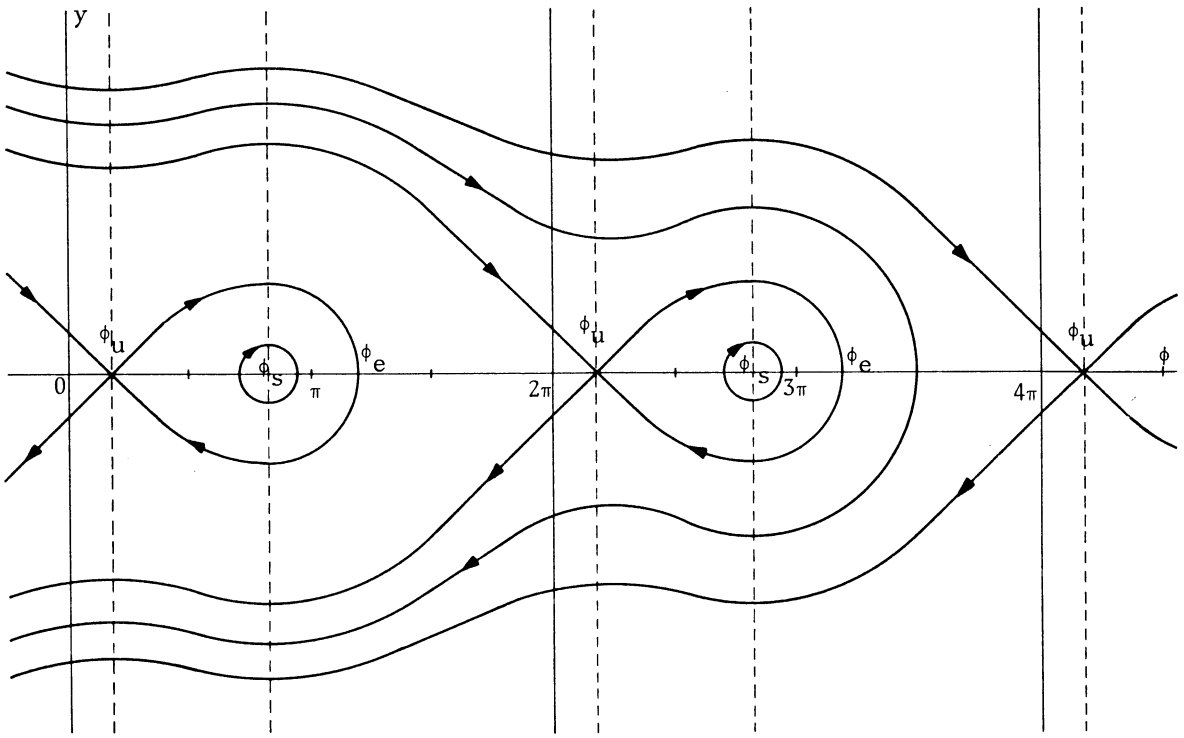
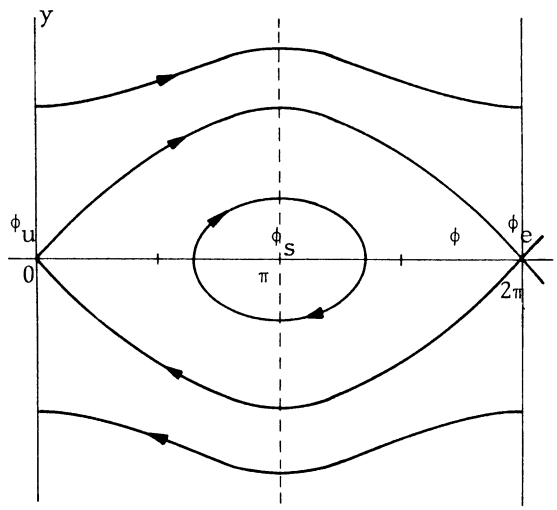


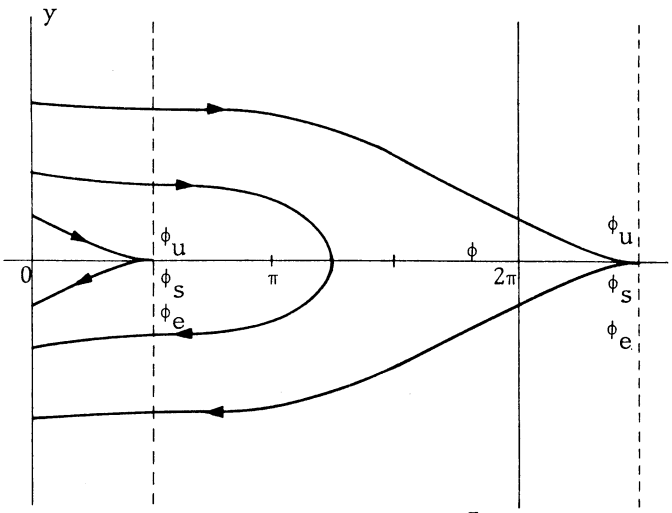
Fig. 5.1 Potential energy as a function of ϕ.



Accelerating bucket ($\frac{\pi}{2} < \phi_s < \pi$)



Stationary bucket ($\phi_s = \pi$)



Vanishing bucket ($\phi_s = \frac{\pi}{2}$)

Fig. 5.2 Trajectories in synchrotron phase space, when $\eta < 0$; when $\eta > 0$, ϕ_s and ϕ_u are interchanged.

The complete phase space is wrapped around a cylinder $0 \leq \phi \leq 2\pi h$.

If $\mathrm{sgn}(\eta) \cdot g'(\phi_0) < 0$, ϕ_0 is a hyperbolic fixed point; this is the case of the other fixed point $\phi_0 = \phi_u$, which is unstable.

ϕ_s is at a minimum of potential energy; ϕ_u is at a maximum. When η changes sign, the two points ϕ_s and ϕ_u are interchanged (see Fig. 5.1). Therefore, when crossing the transition energy, the RF voltage must undergo a phase jump which puts the particles around the new stable fixed point.

Separatrix. The trajectory in phase space passing through the unstable fixed point (ϕ_u, $y = 0$) crosses the ϕ-axis at another point (ϕ_e, $y = 0$). This trajectory is the boundary between trapped and untrapped motion (or between libration and rotation); it is called the *separatrix* (see Fig. 5.2). The phase space domain inside the separatrix is called *bucket;* its area A_s is the longitudinal acceptance of the accelerator.

From (5.5) the equation for the separatrix is

$$\frac{y^2}{2} - \mathrm{sgn}(\eta) \left[\Gamma\phi + G(\phi) - \Gamma\phi_u - G(\phi_u) \right] = 0 \tag{5.7}$$

Taking the derivative with respect to ϕ:

$$y \cdot \frac{dy}{d\phi} - \mathrm{sgn}(\eta) \left[\Gamma - g(\phi) \right] = 0$$

This equation is satisfied with $y = 0$ at $\phi = \phi_u$ and with $\frac{dy}{d\phi} = 0$ at $\phi = \phi_s$. Therefore y is maximum at ϕ_s (this is also the case for any trajectory.)

Bucket width. The bucket width is $(\phi_e - \phi_u)$ where ϕ_e is determined by the equation

$$\Gamma\phi_e + G(\phi_e) = \Gamma\phi_u + G(\phi_u) \tag{5.8}$$

Bucket height. With (5.1),

$$\left(\frac{\delta\hat{E}}{h\omega_0} \right) = \frac{R}{h} \hat{\delta p} = K_2 \hat{y} \quad \text{where} \quad \frac{\hat{y}^2}{2} = -\mathrm{sgn}(\eta) \left[\Gamma\phi_u + G(\phi_u) - \Gamma\phi_s - G(\phi_s) \right] \tag{5.9}$$

Bucket area (per bunch)

$$A_s = K_2 A_s^* \quad \text{where} \quad A_s^* = 2 \int_{\phi_u}^{\phi_e} y \, d\phi = 2\sqrt{2} \int_{\phi_u}^{\phi_e} d\phi \, \left| \Gamma\phi_u + G(\phi_u) - \Gamma\phi - G(\phi) \right|^{\frac{1}{2}} \tag{5.10}$$

$(\eta\Gamma < 0)$

It is not invariant during acceleration.

Period T_s of (large) synchrotron oscillations around the stable fixed point

The phase space trajectories are represented by (5.5) where H* is constant. Let ϕ_1, ϕ_2 be the two phases where y = 0; then (5.5) may be written as

$$\frac{y^2}{2} - \mathrm{sgn}(\eta)\left[\Gamma\phi + G(\phi)\right] = -\mathrm{sgn}(\eta)\left[\Gamma\phi_1 + G(\phi_1)\right] = -\mathrm{sgn}(\eta)\left[\Gamma\phi_2 + G(\phi_2)\right]$$

From (5.2)

$$\frac{d\phi}{dt^*} = \frac{\partial H^*}{\partial y} = y$$

hence

$$dt^* = \frac{d\phi}{y}$$

and with (5.1),

$$T_s = |K_1| T_s^* \quad \text{where} \quad T_s^* = 2\int_{\phi_1}^{\phi_2} \frac{d\phi}{y} = \sqrt{2}\int_{\phi_1}^{\phi_2} d\phi \left|\Gamma\phi_1 + G(\phi_1) - \Gamma\phi - G(\phi)\right|^{-\frac{1}{2}} \tag{5.11}$$

For a general RF voltage this expression involves cumbersome elliptic integrals.

6. SMALL OSCILLATIONS AROUND THE STABLE FIXED POINT

From (5.6), the small amplitude trajectories around ϕ_s are represented by the ellipse equation

$$\frac{y^2}{2} + \left|g'(\phi_s)\right| \cdot \frac{\psi^2}{2} = \frac{C}{2} > 0 \ , \quad \psi = \phi - \phi_s \tag{6.1}$$

It is apparent that all properties of small oscillations around ϕ_s involve the RF voltage only through its slope at ϕ_s.

Period T_{so} of small synchrotron oscillations

The subscript 0 refers here to vanishingly small amplitudes. With (6.1) the general formula (5.11) simplifies to

$$T_{so} = |K_1| T_{so}^* \quad \text{where} \quad T_{so}^* = 2\int_{-\hat{\psi}}^{\hat{\psi}} \frac{d\psi}{y} = 2\int_{-\hat{\psi}}^{\hat{\psi}} \frac{d\psi}{\sqrt{C - |g'(\phi_s)|\psi^2}} = \frac{2\pi}{\sqrt{|g'(\phi_s)|}}$$

This is independent of the amplitude $\hat{\psi}$ as long as the ψ^3 and higher order terms are missing in (5.6), which means as long as $g(\phi)$, i.e. the RF voltage, is a linear function of ϕ.

With (5.4) the synchrotron tune

$$Q_s = \frac{1}{\omega_0}\,\frac{2\pi}{T_s} = \frac{1}{\omega_0 |K_1|}\,\frac{2\pi}{T_s^*}$$

is given by

$$\boxed{Q_{so}^2 = \left| \frac{h\eta}{\beta^2\gamma}\,\frac{eNVg'(\phi_s)}{2\pi E_0} \right| = \frac{1}{2\pi}\,\frac{h\eta}{\beta^2\gamma}\,\frac{eNVg'(\phi_s)}{E_0}} \quad \text{where } Vg'(\phi_s) = \frac{d}{d\phi}\,(\text{RF voltage}) \text{ at } \phi_s \tag{6.2}$$

Height of a trajectory in synchrotron phase space.

For a trajectory of half width $\hat{\psi}$, the height is obtained from (6.1) as $\hat{y}^2 = |g'(\phi_s)|\hat{\psi}^2$ whence, with (5.1) and (5.4),

$$\left(\frac{\hat{\delta E}}{h\omega_0}\right) = K_2\,\hat{y} = \frac{E_0}{h\omega_0} \left| \frac{\beta^2\gamma}{h\eta}\,\frac{eNVg'(\phi_s)}{2\pi E_0} \right|^{\frac{1}{2}} \hat{\psi}$$

Longitudinal emittance. Bunch matching.

The area of a bunch in synchrotron phase space is its longitudinal emittance E_s; it is an invariant by Liouville's theorem. If we call *"emittance of a single particle"* the area $2\pi J$ in phase space which is enclosed by the particle trajectory,

$$2\pi J = \pi\hat{\psi}\left(\frac{\hat{\delta E}}{h\omega_0}\right) = \pi\,\frac{E_0}{h\omega_0} \left| \frac{\beta^2\gamma}{h\eta}\,\frac{eNVg'(\phi_s)}{2\pi E_0} \right|^{\frac{1}{2}} \hat{\psi}^2 = \left| \frac{\pi E_0\,\beta^2\gamma\,eNVg'(\phi_s)}{2\,h^3\,\omega_0^2\,\eta} \right|^{\frac{1}{2}} \hat{\psi}^2 \quad [\text{eV.s}] \tag{6.3}$$

The action J is an *adiabatic invariant*, i.e. it stays constant if the parameters in H are varied infinitely slowly (Ref. 5, p. 154; Ref. 6, p. 110; Ref. 7, p. 234). If at some time a bunch is matched (which means that its border in phase space is just the closed trajectory of the outermost particles) then its emittance E_s is equal to the single particle emittance of its outermost particles. After a change of the parameters in H, the emittance E_s is unchanged but the action of the outermost particles has changed slightly and differently for each particle, which means that the bunch is no longer matched exactly: therefore the matching of a bunch can only be preserved in the adiabatic sense, i.e. if the parameters in H are varied very slowly.

7. MOTION IN THE VICINITY OF THE FIXED POINTS

Take the Hamiltonian (5.5) with reduced variables:

$$H^* = \frac{y^2}{2} - \text{sgn}(\eta)\left[\Gamma\phi + G(\phi)\right]$$

Canonical equations:

$$\frac{dy}{dt^*} = \mathrm{sgn}(\eta)\left[\Gamma - g(\phi)\right] \quad , \qquad \frac{d\phi}{dt^*} = y$$

hence:

$$dt^* = \frac{d\phi}{y} = \frac{d\phi}{\sqrt{2H^* + 2\,\mathrm{sgn}(\eta)\,\left[\Gamma\phi + G(\phi)\right]}} \tag{7.1}$$

In the vicinity of a fixed point ϕ_0 ,

$$y^2 = \underbrace{2H^* + 2\,\mathrm{sgn}(\eta)\left[\Gamma\phi_o + G(\phi_o)\right]}_{C} - \mathrm{sgn}(\eta)\, g'(\phi_o)\,(\phi - \phi_o)^2 + \ldots$$

Motion around the stable fixed point ϕ_s

$$y^2 = \underbrace{C}_{y^2_{max}} - |g'(\phi_s)|\,(\phi - \phi_s)^2 + \ldots$$

$$dt^* = \frac{d\phi}{\sqrt{y^2_{max} - |g'(\phi_s)|\,(\phi - \phi_s)^2}}$$

$$t^* = -\,|g'(\phi_s)|^{-\frac{1}{2}}\ \arccos\left[\frac{\sqrt{|g'(\phi_s)|}}{y_{max}}\,(\phi - \phi_s)\right]$$

or

$$(\phi - \phi_s) = \frac{y_{max}}{\sqrt{|g'(\phi_s)|}}\cos\left[\sqrt{|g'(\phi_s)|}\,t^*\right] \tag{7.2}$$

Motion around the unstable fixed point ϕ_u

$$y^2 = C + |g'(\phi_u)|\,(\phi - \phi_u)^2 + \ldots$$

$$dt^* = \frac{d\phi}{\sqrt{C + |g'(\phi_u)|\,(\phi - \phi_u)^2}}$$

C = 0 : particle on the separatrix

$$t^* = \pm\, |g'(\phi_u)|^{-\frac{1}{2}} \log |\phi - \phi_u| \to \mp\infty \qquad \text{when} \qquad \phi \to \phi_u \tag{7.3}$$

Therefore the motion near ϕ_u is very slow (it is a fixed point!) This relation may be inverted as

$$\phi - \phi_u \sim e^{\pm \sqrt{|g'(\phi_u)|}\, t^*}$$

C > 0 : particle outside the bucket

$$C = y^2_{min}$$

$$t^* = |g'(\phi_u)|^{-\frac{1}{2}} \text{ arg sh } \left[\frac{\sqrt{|g'(\phi_u)|}}{y_{min}} (\phi - \phi_u) \right] \tag{7.4}$$

or

$$\phi - \phi_u = \frac{y_{min}}{\sqrt{|g'(\phi_u)|}} \text{ sh } \left[\sqrt{|g'(\phi_u)|}\, t^* \right]$$

The motion is very slow when $y_{min} \to 0$.

C < 0 : particle inside the bucket

$$C = - |g'(\phi_u)|\, \psi^2_{min} \qquad \text{where} \qquad \psi = \phi - \phi_u$$

$$t^* = |g'(\phi_u)|^{-\frac{1}{2}} \text{ arg ch } \left[\frac{\phi - \phi_u}{\psi_{min}} \right] \tag{7.5}$$

or

$$\phi - \phi_u = \psi_{min} \cdot \text{ch } \left[\sqrt{|g'(\phi_u)|}\, t^* \right]$$

When $\psi_{min} \to 0$, the motion close to ϕ_u becomes very slow, and the synchrotron period T_s becomes infinite. As shown by (7-3) to (7-5), all trajectories are slowed down in the vicinity of the unstable fixed point.

8. STATIONARY BUCKET WITH A HARMONIC CAVITY

When the beam is not accelerated but is simply kept bunched at a fixed energy, $\Gamma = 0$ and *the bucket is called stationary.* In this case, which corresponds to collider operation, the frequency of synchrotron oscillations as a function of phase amplitude $\hat{\psi}$ is given by

$$\omega_s = \frac{\omega_s^*}{|K_1|}$$

where as a first approximation*) (see Appendix A, Eq. (A.8)):

$$\omega_s^{*2} = \operatorname{sgn}(\eta) \cdot \sum_{n=1}^{\infty} a_n \cos n\phi_s \cdot \frac{2}{\hat{\psi}} J_1(n\hat{\psi}) + O(\hat{\psi}^4) \tag{8.1}$$

when

$$g(\phi) = \sum_{n=1}^{\infty} a_n \sin n\phi \ , \quad \Gamma = g(\phi_s) = 0 \ , \quad \operatorname{sgn}(\eta) \cdot g'(\phi_s) = \operatorname{sgn}(\eta) \cdot \sum_{n=1}^{\infty} n\, a_n \cos n\phi_s > 0.$$

Assuming that $a_1 = 1$ is the dominant term,

$$\phi_s = 0 \quad \text{if} \quad \eta > 0$$

$$\phi_s = \pi \quad \text{if} \quad \eta < 0 \ .$$

With (8.1), the synchrotron frequency for vanishingly small amplitudes is

$$\omega_{so}^{*2} = \operatorname{sgn}(\eta) \cdot \sum_{n=1}^{\infty} n a_n \cos n\,\phi_s = \operatorname{sgn}(\eta) \cdot g'(\phi_s) = |g'(\phi_s)| \tag{8.2}$$

in agreement with (6.2).

The relation (8.1) allows shaping the variation of ω_s^2 with $\hat{\psi}$. For example, if ω_s^2 is to be proportional to $\hat{\psi}^2$, it is sufficient to take $a_1 = 1$ and $\operatorname{sgn}(\eta) \cdot n a_n \cos n\phi_s = -1$ for some $n > 1$; then

$$\omega_s^{*2} = \frac{2}{\hat{\psi}} J_1(\hat{\psi}) - \frac{2}{n\hat{\psi}} J_1(n\hat{\psi}) + O(\hat{\psi}^4) = \frac{n^2 - 1}{2}\left(\frac{\hat{\psi}}{2}\right)^2 + O(\hat{\psi}^4) \tag{8.3}$$

If the sign of a_n is reversed so that $\operatorname{sgn}(\eta) \cdot n a_n \cos n\phi_s = 1$, then

$$\omega_s^{*2} = \frac{2}{\hat{\psi}} J_1(\hat{\psi}) + \frac{2}{n\hat{\psi}} J_1(n\hat{\psi}) + O(\hat{\psi}^4) = 2\left[1 - \frac{n^2 + 1}{4}\left(\frac{\hat{\psi}}{2}\right)^2 + O(\hat{\psi}^4)\right] \tag{8.4}$$

---

*) $J_\nu(x)$ is the Bessel function of the first kind of order ν and argument x.

The latter case (8.4) corresponds to the usual operation of a Landau harmonic cavity, which increases the relative spread of synchrotron frequencies as a function of $\hat{\psi}$. In both cases the amplitude of the harmonic voltage is such that

$$\left|\frac{a_n}{a_1}\right| = \frac{1}{n} \quad .$$

9. FORMULAE FOR A SINUSOIDAL RF VOLTAGE

Formulae are simpler when using the phase φ measured from the crest of the RF voltage:

$$\phi = \frac{\pi}{2}\,\mathrm{sgn}(\Gamma) + \varphi$$

$$\phi_s = \frac{\pi}{2}\,\mathrm{sgn}(\Gamma) + \varphi_s \quad , \quad \phi_u = \frac{\pi}{2}\,\mathrm{sgn}(\Gamma) - \varphi_s \quad , \quad \mathrm{sgn}(\varphi_s) = -\,\mathrm{sgn}(\eta)\cdot\mathrm{sgn}(\Gamma) \quad ,$$

$$\sin\phi_s \cdot \mathrm{sgn}(\Gamma) = |\Gamma| = \cos\varphi_s$$

From (4.2) and (4.8),

$$g(\phi) = \sin\phi = \mathrm{sgn}(\Gamma)\cdot\cos\varphi \quad , \qquad \Gamma = g(\phi_s) = \sin\phi_s = \mathrm{sgn}(\Gamma)\cdot\cos\varphi_s \quad , \tag{9.1}$$

$$G(\phi) = \cos\phi = -\,\mathrm{sgn}(\Gamma)\cdot\sin\varphi$$

Bucket width $(\phi_e - \phi_u)$ where

$$\mathrm{tg}(\phi_e - \phi_u - 3\varphi_s) = \frac{1}{10}\,\mathrm{tg}^3\varphi_s\cdot\left[1 + \frac{111}{140}\,\mathrm{tg}^2\varphi_s - 0(\mathrm{tg}^4\varphi_s)\right]^{-1} \tag{9.2}$$

Bucket height From (5.9), $\frac{\hat{\gamma}^2}{2} = 2\left|\sin\varphi_s - \varphi_s\cos\varphi_s\right|$ and, using (5.4):

$$\left(\frac{\delta\hat{E}}{h\omega_0}\right) = \frac{R}{h}\,\delta\hat{p} = \frac{E_0}{h\omega_0}\,2\left|\frac{\beta^2\gamma}{h\eta}\,\frac{eNV\sin\varphi_s}{2\pi E_0}\right|^{\frac{1}{2}}\left[1 - \varphi_s\,\mathrm{cotg}\,\varphi_s\right]^{\frac{1}{2}} = \frac{E_0}{h\omega_0}\,2\left|\frac{\beta^2\gamma}{h\eta}\right|Q_{so}\sqrt{1 - \varphi_s\,\mathrm{cotg}\,\varphi_s} \quad [\mathrm{eV.s}] \tag{9.3}$$

Bucket area From (5.10),

$$A_s^* = 2\sqrt{2}\int_{\phi_u}^{\phi_e} d\phi\,\left|\phi_u\sin\phi_s + \cos\phi_u - \phi\sin\phi_s - \cos\phi\right|^{\frac{1}{2}} \tag{9.4}$$

$$(\eta\Gamma < 0)$$

A_s^* is steadily increasing with $|\varphi_s|$. Its maximum value is obtained for $\varphi_s = \pi/2$ (stationary bucket):

$$A_{s\,max}^* = 2\sqrt{2} \int_0^{2\pi} d\phi \sqrt{1 - \cos\phi} = 16$$

Let

$$\alpha(\Gamma) = \frac{A_s^*}{A_{s\,max}^*} = \frac{A_s}{A_{s\,max}} \quad ; \qquad \alpha(\Gamma) = \frac{3}{10} |\varphi_s|^{5/2} \left[1 - \frac{1}{60}\varphi_s^2 + 0(\varphi_s^4)\right] \tag{9.5}$$

The derivation of the series expansion for $\alpha(\Gamma)$ is given in Appendix B.

With (5.4)

$$A_s = 16 \frac{E_0}{h\omega_0} \left| \frac{\beta^2\gamma}{h\eta} \cdot \frac{eNV}{2\pi E_0} \right|^{\frac{1}{2}} \alpha(\Gamma) \quad \text{per bunch} \qquad [\text{eV.s}] \tag{9.6}$$

Remark: Instead of (9.3), the dimensionless quantity $\hat{\delta p}/m_0c$ is often used. In these coordinates, we have:

Bucket height

$$\frac{\hat{\delta p}}{m_0c} = 2 \left| \frac{\gamma}{h\eta} \cdot \frac{eNV \sin\varphi_s}{2\pi E_0} \right|^{\frac{1}{2}} \cdot \left[1 - \varphi_s \operatorname{cotg} \varphi_s\right]^{\frac{1}{2}} \qquad \left[\frac{\delta p}{m_0c}\right] \tag{9.7}$$

Bucket area

$$A_s \cdot \frac{h}{m_0cR} = 16 \left| \frac{\gamma}{h\eta} \cdot \frac{eNV}{2\pi E_0} \right|^{\frac{1}{2}} \cdot \alpha(\Gamma) \quad \text{per bunch} \qquad \left[\frac{\delta p}{m_0c} \cdot \text{RF rad}\right] \tag{9.8}$$

Period of (large) synchrotron oscillations in a stationary bucket

Besides the trivial case of a linear RF voltage, a sinusoidal RF voltage is the only case where it is possible to compute simply the synchrotron period for any amplitude. With $\Gamma = 0$ and $G(\phi) = \cos\phi$, the general equation (5.11) for the libration period reduces to

$$T_s^* = \sqrt{2} \int_{\phi_1}^{\phi_2} d\phi \, |\cos\phi_1 - \cos\phi|^{-\frac{1}{2}} = \sqrt{2} \int_{-\hat{\psi}}^{\hat{\psi}} d\psi \, |\cos\psi - \cos\hat{\psi}|^{-\frac{1}{2}} = 4K\left(\sin\frac{\hat{\psi}}{2}\right) \tag{9.9}$$

where $\psi = \phi - \phi_s$; $K(k)$ is the complete elliptic integral of the first kind with modulus k.

Therefore

$$\omega_s^{*2} = \left[\frac{2}{\pi} K\left(\sin\frac{\hat{\psi}}{2}\right)\right]^{-2} = 1 - \frac{1}{2}\sin^2\frac{\hat{\psi}}{2} - \frac{3}{32}\sin^4\frac{\hat{\psi}}{2} - \ldots \qquad (9.10)$$

$$= 1 - \frac{1}{2}\left(\frac{\hat{\psi}}{2}\right)^2 + \frac{7}{8.12}\left(\frac{\psi}{2}\right)^4 - \ldots$$

The approximate Eq. (8.1) would yield in this case

$$\omega_s^{*2} \approx \frac{2}{\hat{\psi}} J_1(\hat{\psi}) = 1 - \frac{1}{2}\sin^2\frac{\hat{\psi}}{2} - \frac{1}{12}\sin^4\frac{\hat{\psi}}{2} - \ldots = 1 - \frac{1}{2}\left(\frac{\hat{\psi}}{2}\right)^2 + \frac{1}{12}\left(\frac{\hat{\psi}}{2}\right)^4 - \ldots$$

which shows that the error on the $\hat{\psi}^4$ term in Eq. (8.1) is rather small.

Motion outside a stationary bucket

$$\Gamma = 0 \qquad G(\phi) = \cos\phi$$

$$\text{sgn}(\eta)\cos\phi_u = -1 \qquad G(\phi) = \cos\phi_u \quad \cos(\phi - \phi_u)$$

From (5-5),

$$y^2 = 2H^* - 2\cos(\phi - \phi_u) = (2H^* - 2) + 4\sin^2\left[\frac{\phi - \phi_u}{2}\right] = y^2_{min} + 4\sin^2\left[\frac{\phi - \phi_u}{2}\right]$$

$$dt^* = \frac{d\phi}{2\sqrt{\left[\frac{y_{min}}{2}\right]^2 + \sin^2\left[\frac{\phi - \phi_u}{2}\right]}} = \frac{d\left[\frac{\phi - \phi_u}{2}\right]}{\sqrt{\left[\frac{y_{min}}{2}\right]^2 + 1 - \cos^2\left[\frac{\phi - \phi_u}{2}\right]}}$$

The motion outside a stationary bucket is a "rotation", the period of which is the time needed to increase ϕ by 2π :

$$T^* = \frac{1}{\sqrt{1 + \left[\frac{y_{min}}{2}\right]^2}} \int_0^{2\pi} \frac{d\left[\frac{\phi - \phi_u}{2}\right]}{\sqrt{1 - \frac{\cos^2\left[\frac{\phi - \phi_u}{2}\right]}{1 + \left[\frac{y_{min}}{2}\right]^2}}}$$

or

$$T^* = \frac{2}{\sqrt{1 + \left[\frac{y_{min}}{2}\right]^2}} \cdot K\left[\frac{1}{\sqrt{1 + \left[\frac{y_{min}}{2}\right]^2}}\right] \qquad (9.11)$$

It follows that

$$T^* \to \begin{cases} 2 \log \dfrac{8}{y_{min}} & \text{for} \quad y_{min} \to 0 \\ \dfrac{2\pi}{y_{min}} & \text{for} \quad y_{min} \to \infty \end{cases}$$

When $\frac{y_{min}}{2} \ll 1$, it plays the same role as $\operatorname{cotg} \frac{\hat{\psi}}{2}$ in (9.9); both quantities represent half the minimum distance to ϕ_u , outside and inside the bucket respectively.

10. ADIABATIC DAMPING OF PHASE OSCILLATIONS.

From (6.3):

$$\hat{\psi} = \sqrt{2J} \left| \frac{2\pi\, h^3\, \omega_o^{\,2}\, \eta}{E_o\, \beta^2\, \gamma \cdot eNV\, g'(\phi_s)} \right|^{\frac{1}{4}} \tag{10.1}$$

and

$$\frac{\hat{\delta E}}{h\omega_o} = \frac{2J}{\hat{\psi}} = \sqrt{2J} \left| \frac{2\pi\, h^3\, \omega_o^{\,2}\, \eta}{E_o\, \beta^2\, \gamma \cdot eNV\, g'(\phi_s)} \right|^{-\frac{1}{4}} \tag{10.2}$$

where

$$\frac{\omega_o^{\,2}}{\beta^2} = c^2 \frac{\omega_o^{\,2}}{v_s^{\,2}} = \frac{c^2}{R^2}$$

and

$$2\pi J = \pi \hat{\psi} \left(\frac{\hat{\delta E}}{h\omega_o} \right)$$

is an adiabatic invariant.

If $V g'(\phi_s)$ is kept constant during acceleration, the only quantity in (10.1) which varies (slowly) in a synchrotron is $\frac{\eta}{\gamma}$.

Let us compute

$$\frac{d}{d\gamma} \log \left| \frac{\eta}{\gamma} \right| = \frac{-\frac{2}{\gamma^3}}{\frac{1}{\gamma^2} - \alpha} - \frac{1}{\gamma} = \frac{1}{\gamma} \left[\frac{2}{\alpha\gamma^2 - 1} - 1 \right] = \frac{1}{\gamma} \left[\frac{3 - \alpha\gamma^2}{\alpha\gamma^2 - 1} \right]$$

$$= \frac{1}{\gamma} \left[\frac{3\,\gamma_{tr}^2 - \gamma^2}{\gamma^2 - \gamma_{tr}^2} \right]$$

$$\frac{d}{d\gamma} \log \left| \frac{\eta}{\gamma} \right| > 0 \qquad \text{when} \qquad \gamma_{tr}^2 < \gamma^2 < 3\,\gamma_{tr}^2$$

$$< 0 \qquad \text{in all other cases}$$

Therefore, when $\gamma^2 = 3\,\gamma_{tr}^2$, $\left|\frac{\eta}{\gamma}\right|$ reaches a maximum where

$$\left.\frac{\eta}{\gamma}\right|_{max} = \frac{1}{\gamma}\left[\frac{1}{\gamma^2} - \alpha\right]_{max} = -\frac{2}{\left(3\,\gamma_{tr}^2\right)^{3/2}}$$

We notice that $\quad \frac{\eta}{\gamma} = \frac{2}{\left(3\,\gamma_{tr}^2\right)^{3/2}} \quad$ when $\quad \gamma^2 = \frac{3}{4}\,\gamma_{tr}^2$

The variation of $\left|\frac{\eta}{\gamma}\right|^{\frac{1}{4}}$ i.e. of $\hat{\psi}$ as a function of γ , is shown in Fig. 10.1.

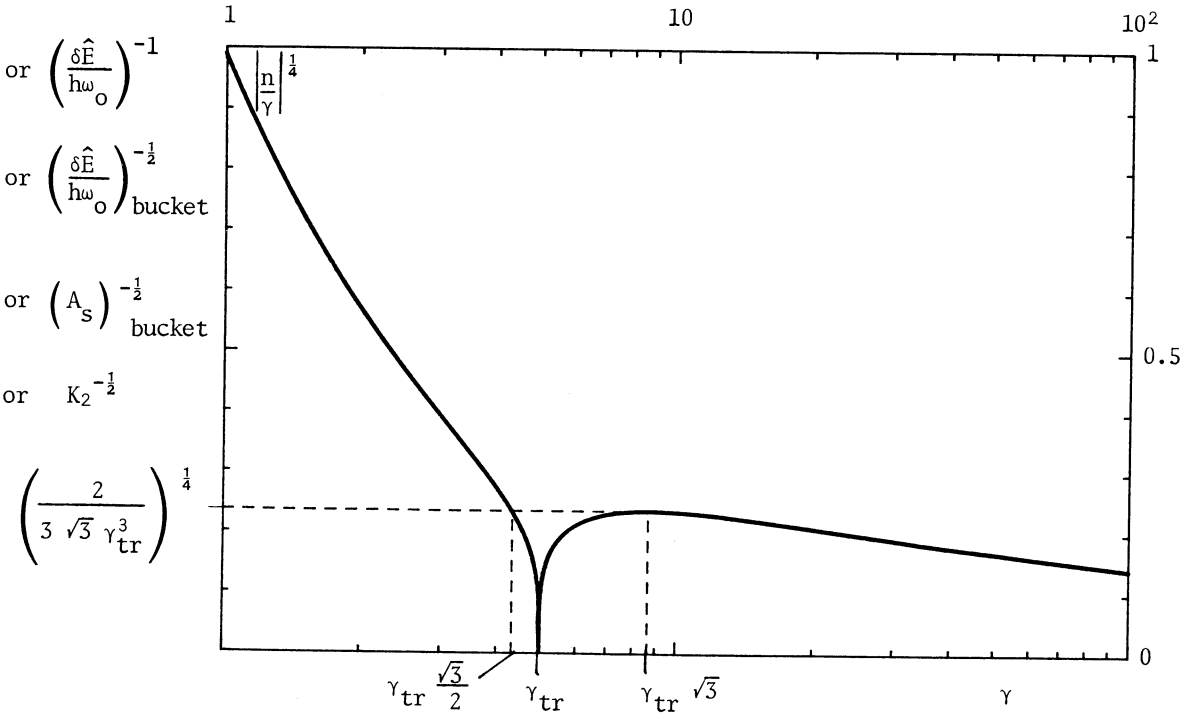


Fig. 10.1 Amplitude of phase oscillation as a function of γ . In the figure, $\gamma_{tr} = 5$.

In case of a constant $|\varphi_s|$ during acceleration, the bucket width $|\phi_e - \phi_u|$ is constant. From (9-3) the bucket height is, for a sinusoidal voltage:

$$\left(\frac{\delta\hat{E}}{h\omega_o}\right)_{\text{bucket}} = 2\left|\frac{2\pi\,h^3\,\omega_o^{\,2}\,\eta}{E_o\,\beta^2\,\gamma\cdot e\,NV\,g'(\phi_s)}\right|^{-\frac{1}{2}}\left[1 - \varphi_s\,\text{cotg}\,\varphi_s\right]^{\frac{1}{2}}$$

where $g'(\phi_s) = \cos \phi_s = \sin \varphi_s$.

The ratio of particle height (10.2) to bucket height is thus:

$$\frac{(\hat{\delta E})_{\text{particle}}}{(\hat{\delta E})_{\text{bucket}}} = \frac{\sqrt{2J}}{2} \left| \frac{2\pi\, h^3\, \omega_o^{\,2}\, \eta}{E_o\, \beta^2\, \gamma \cdot e\, NV\, g'(\phi_s)} \right|^{\frac{1}{4}} \left[1 - \varphi_s \operatorname{cotg} \varphi_s \right]^{-\frac{1}{2}}$$

or, with (10.1),

$$\frac{(\hat{\delta E})_{\text{particle}}}{(\hat{\delta E})_{\text{bucket}}} = \frac{\hat{\psi}}{2\sqrt{1 - \varphi_s \operatorname{cotg} \varphi_s}}$$

Since this ratio behaves like $\hat{\psi}$, it is maximum at injection when the injection energy is below $\gamma_{tr} \frac{\sqrt{3}}{2}$ or above $\gamma_{tr} \sqrt{3}$.

In these cases, if $|\varphi_s|$ and V are kept constant during acceleration, particles captured at injection will stay in the bucket during the whole acceleration process.

Remark: In Fig. 10.1, $\hat{\psi} \to 0$ while $\frac{\delta \hat{E}}{h\omega_o} \to \infty$ when $\gamma \to \gamma_{tr}$;this means that the adiabatic approximation breaks down in the vicinity of γ_{tr} . A more refined treatment (see K. Johnsen, Ref. 2, p.178) shows that $\hat{\psi}$ goes through a minimum while $\frac{\delta \hat{E}}{h\omega_o}$ goes through a maximum at transition.

Separatrix crossing; golf-club

Although the area $2\pi J$ which is enclosed by a particle trajectory in phase space is an adiabatic invariant, it follows from (9.6) that the bucket area A_s behaves like the bucket height (9.3) and is not an adiabatic invariant, as appears clearly in Fig. 10.1. Therefore, one must conclude that the separatrix is not a particle trajectory in phase-space, except when there is no acceleration ($\Gamma=0$) and therefore no time variation of the parameters in the Hamiltonian (4.7). The reason is that, as shown by (7.3), the particle motion along the separatrix is extremely slow in the vicinity of ϕ_u , therefore violating the adiabaticity condition; this condition indeed requires that the variation of the parameters in the Hamiltonian (4.7) be negligible during a synchrotron period.

In order to arrive at a qualitative picture of the actual particle motion, we observe that the reduced Hamiltonian (5.5) does not depend on time. Therefore this is also true for the curves of constant H* in the (y,ϕ) plane (see Fig. 5.2), in particular for the slope of the separatrix at ϕ_u . This would not be true in the $\left(\frac{\delta E}{h\omega_o}, \phi\right)$ phase space because from (5.1)

$$\frac{\delta E}{h\omega_o} = K_2\, y \qquad \text{or} \qquad y = \frac{1}{K_2} \cdot \frac{\delta E}{h\omega_o} \tag{10.3}$$

and K_2 varies (slowly) with time. Since the elementary phase space area $\frac{\delta E}{h\omega_o} \wedge \delta\phi$ is invariant, in the (y,ϕ) plane the elementary area

$$\delta y \wedge \delta\phi = \frac{1}{K_2} \cdot \frac{\delta E}{h\omega_o} \wedge \delta\phi$$

varies as K_2^{-1}. In Fig. 10.1 it is seen that when $\gamma < \gamma_{tr} \frac{\sqrt{3}}{2}$ or $\gamma > \gamma_{tr} \sqrt{3}$, K_2^{-1} decreases with increasing γ (or time); therefore in the (y,ϕ) plane, areas shrink with time and the motion appears to be damped. In particular, the separatrix is no longer a trajectory since it encloses a constant area in (y,ϕ). Instead, one of the trajectories which leaves ϕ_u (or, strictly speaking, which tends to ϕ_u when $t \to -\infty$) will no longer come back to ϕ_u (strictly speaking, tend to ϕ_u when $t \to +\infty$), but it will spiral inwards around ϕ_s; in other words, instead of being a center, ϕ_s has become a focus (see Fig. 10.2). This shows clearly the adiabatic damping of phase oscillations: $\hat{\psi}$ shrinks as $K_2^{-\frac{1}{2}}$, $\frac{\delta\hat{E}}{h\omega_o}$ expands as $K_2^{\frac{1}{2}}$, and from (10.3) $\hat{y}$ also shrinks as $K_2^{-\frac{1}{2}}$. On the other end of the separatrix, the trajectory which tends to ϕ_u must then come from outside the fish-shaped bucket. As a result, in reduced coordinates the longitudinal acceptance of an accelerator is not a fish-shaped bucket, but rather has the shape of a *golf-club*. In the literature, this effect is mostly discussed for low-β linacs (Ref. 8; Ref. 9, p.27), for which the adiabaticity condition is worst fulfilled; since for a linac $\gamma_{tr} = \infty$, the (y,ϕ) plane corresponds to $\gamma < \gamma_{tr}$ in Fig. 10.2.
When $\gamma_{tr} < \gamma < \gamma_{tr}\sqrt{3}$, K_2^{-1} increases with time; in the (y,ϕ) plane areas expand with time and the motion appears to be anti-damped. In particular, one of the trajectories which tends to ϕ_u spirals outwards around ϕ_s, and the handle of the golf-club is directed toward the negative y-axis.

In fact, the motion in the (y,ϕ) plane depicted in Fig. 10.2 as a result of slow variation of parameters in the Hamiltonian, seems to be more generic (i.e. more common) in dynamical systems than the motion depicted in Fig. 5.2. Indeed (Ref. 10, p.29-31) it is exceptional that a trajectory leaving (or arriving at) a saddle point goes to (or comes from) another (or the same, as in Fig. 5.2) saddle point; it rather goes to an attractor or comes from a repellor, which in Fig. 10.2 is the stable fixed point ϕ_s. In practice, the motion depicted in Fig. 10.2 becomes apparent mainly when K_2 varies fast with γ, i.e. (see Fig. 10.1) at low β or near transition.

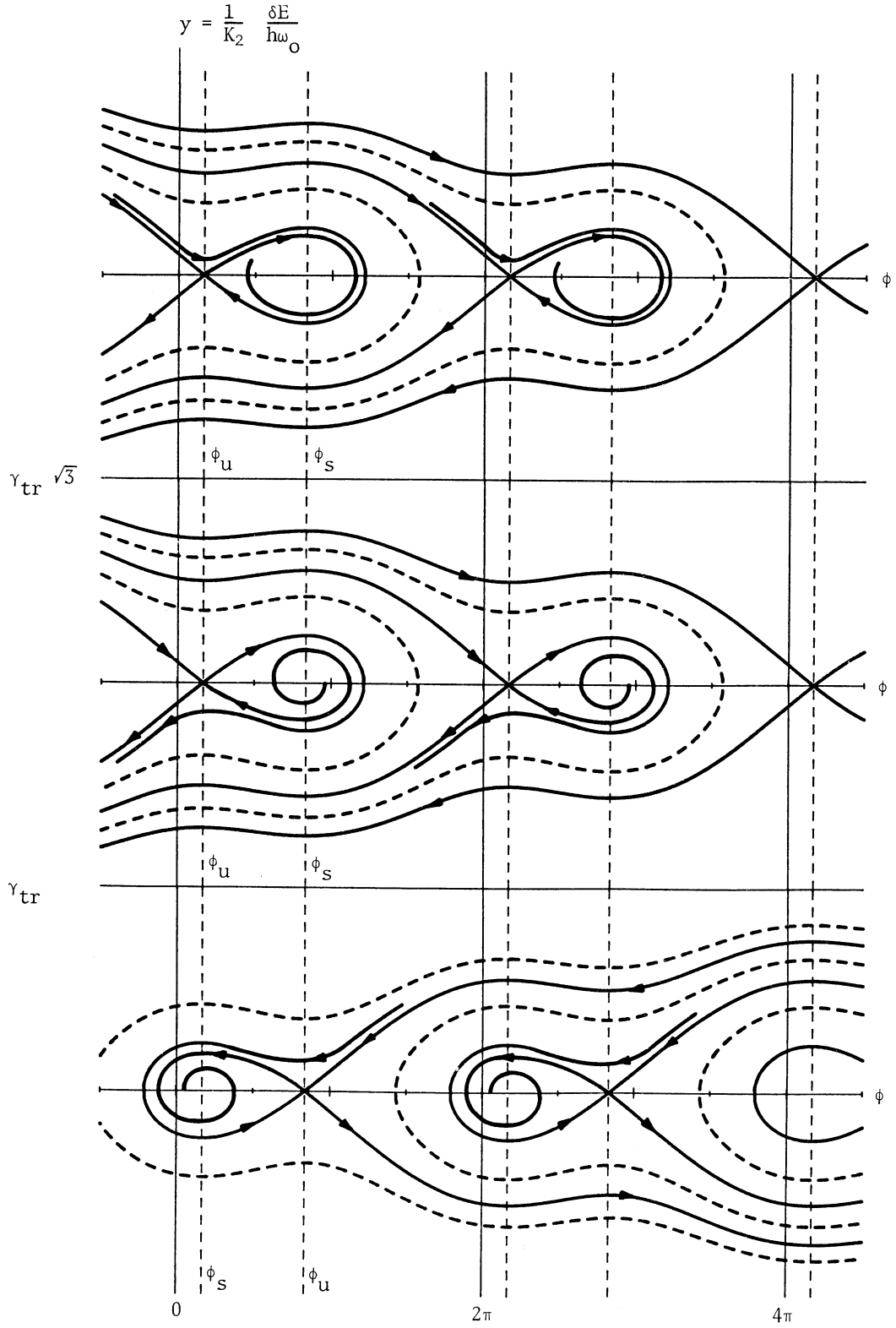


Fig. 10.2 - Actual trajectories in the reduced coordinate y,ϕ plane.

11. BACK TO FINITE DIFFERENCE EQUATIONS. STOCHASTICITY

For an arbitrary RF voltage, the finite difference equations (3.11) and (3.4) read in case of a synchrotron (R_s = constant):

$$\begin{cases} \delta p_{n+1} - \delta p_n = \dfrac{eV}{v_{s,n}}\left[g(\phi_n) - g(\phi_s)\right] & (11.1) \\ \phi_{n+1} - \phi_n = -\dfrac{2\pi h\eta}{N\, p_{s,n+1}} \cdot \delta p_{n+1} & (11.2) \end{cases}$$

where η, v_s, p_s are slowly varying parameters.

This mapping preserves area in the δp_n, ϕ_n plane. In contrast to differential equations, there is no (smooth) constant of motion for the finite difference equations.

Fixed points (mod 2π): If k is any integer,

$$\begin{cases} \phi_n = \phi_s + 2\pi kn, \ \dfrac{h\eta}{Np_s} \cdot \delta p_n = k & \text{is a stable fixed point} \\ \phi_n = \phi_u + 2\pi kn, \ \dfrac{h\eta}{Np_s} \cdot \delta p_n = k & \text{is an unstable fixed point.} \end{cases}$$

The $k \neq 0$ case corresponds to working with the same RF frequency, but with an harmonic number $(h + Nk)$. Indeed, for a given ω_{RF} the synchronous revolution frequencies are such that

$$\omega_0 = \frac{\omega_{RF}}{h}, \quad \frac{\delta\omega_0}{\omega_0} = -\frac{\delta h}{h}$$

to which correspond the synchronous momenta

$$\eta \frac{\delta p}{p_s} = \frac{\delta\omega_0}{\omega_0} = -\frac{\delta h}{h} = -\frac{Nk}{h}$$

This means that the vertical distance between neighbouring buckets is $\dfrac{\delta p}{p_s} = \dfrac{N}{|h\eta|}$. In order to prevent stochastic effects (Ref. 11) from becoming important, the ratio ξ between the full bucket height and the vertical distance between neighbouring buckets should be less than 1 (Chirikov's criterion). From (9.7),

$$\xi = \frac{\dfrac{4}{\beta\gamma}\left|\dfrac{\gamma}{h\eta}\,\dfrac{eNV\sin\varphi_s}{2\pi E_0}\right|^{\frac{1}{2}}\sqrt{1 - \varphi_s \operatorname{cotg}\varphi_s}}{\left|\dfrac{N}{h\eta}\right|} = \frac{4}{N}\left|\frac{h\eta}{\beta^2\gamma}\,\frac{eNV\sin\varphi_s}{2\pi E_0}\right|^{\frac{1}{2}}\sqrt{1 - \varphi_s \operatorname{cotg}\varphi_s}$$

so that Chirikov's criterion reads, with (6.2):

$$\xi = 4 \frac{Q_{so}}{N} \cdot \sqrt{1 - \varphi_s \operatorname{cotg} \varphi_s} < 1$$

Therefore, differential equations are valid only when $Q_{so}/N << 1$. For a finite Q_{so}, the motion near the bucket border becomes chaotic, making the bucket area shrink (Ref. 12); in practice, this effect is still very small for $Q_{so}/N < 0.1$.

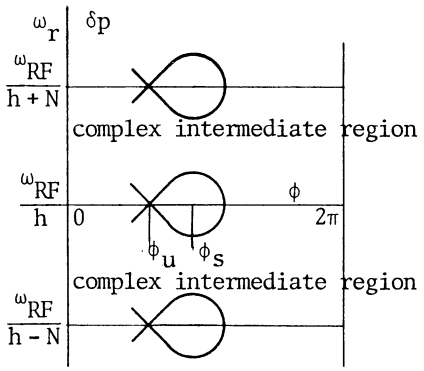


Fig. 11.1 Phase space enlarged to several harmonic numbers at a fixed RF frequency (for $\eta < 0$)

Period of small synchrotron oscillations

In the close vicinity of ϕ_s, the system (11.1), (11.2) reduces to

$$\begin{cases} \left[\frac{2\pi h\eta}{Np_s}\right]_n \left[\frac{2\pi h\eta}{Np_s}\right]^{-1}_{n+1} \cdot P_{n+1} - P_n = - K_0 \, (\phi_n - \phi_s) \\ \phi_{n+1} - \phi_n = P_{n+1} \end{cases} \tag{11.3}$$

where we have put

$$P_n = -\frac{2\pi h\eta}{Np_s} \cdot \delta p_n \tag{11.4}$$

and

$$K_0 = \frac{2\pi h\eta}{N} \frac{eV \, g'(\phi_s)}{p_s v_s} > 0 \tag{11.5}$$

With μ defined by

$$4 \sin^2 \frac{\mu}{2} = K_0 \, , \tag{11.6}$$

the system (11.3) admits of the general solution

$$\phi_n - \phi_s = \mathrm{Re}\left[a \, e^{in\mu}\right] \quad , \qquad P_n = \mathrm{Re}\left[i \, a \, 2 \sin \frac{\mu}{2} \cdot e^{i(n-\frac{1}{2})\mu}\right] \tag{11.7}$$

This represents small synchrotron oscillations with angular frequency

$$\omega_{so} = \frac{N}{T_0}\,\mu = N\,\omega_0\,\frac{\mu}{2\pi}\,.$$

From (11.5) and (11.6) the synchrotron tune for small oscillations is given by

$$\boxed{4\,N^2\,\sin^2\left(\pi\,\frac{Q_{so}}{N}\right) = N^2K_0 = 2\pi h\eta\,\frac{eNVg'(\phi_s)}{p_s v_s} = 2\pi\,\frac{h\eta}{\beta^2\gamma}\,\frac{eNVg'(\phi_s)}{E_0}} \qquad (11.8)$$

The analogous formula (6.2) obtained with differential equations appears to be the limiting case of (11.8) when $N \to \infty$, i.e. when particle acceleration is evenly distributed all around the ring with the total RF voltage NV remaining finite.

Adiabatic damping of phase oscillations

From (11.4) an area element in the P_n, ϕ_n plane reads

$$\delta P_n \wedge \delta\phi_n = -\frac{2\pi h\eta}{N\,p_s}\cdot \underbrace{\delta p_n \wedge \delta\phi_n}_{\text{invariant}} \qquad (11.9)$$

which shows that the mapping (11.3) does not preserve area in the P_n, ϕ_n plane. All successive points of a trajectory described by (11.7) lie on an ellipse with area $\pi\,|a|^2 \sin\mu$; by (11.9) this area is related to the adiabatic invariant action J through

$$\pi\,|a|^2\sin\mu = \left|\frac{2\pi h\eta}{N\,p_s}\right|\,\frac{h}{R_s}\,2\pi J$$

Since from (11.6)

$$\sin^2\mu = K_o\left[1 - \frac{K_o}{4}\right]\quad,$$

we obtain, using (11.5):

$$|a|^2 = 2J\,\frac{h}{R_s}\left|\frac{2\pi h\eta}{N\,p_s}\right| K_o^{-\frac{1}{2}}\left(1 - \frac{K_o}{4}\right)^{-\frac{1}{2}} = 2J\,\frac{h}{R_s}\left|\frac{2\pi h\eta}{N\,p_s}\,\frac{v_s}{eV\,g'(\phi_s)}\right|^{\frac{1}{2}}\left(1 - \frac{K_o}{4}\right)^{-\frac{1}{2}}$$

hence

$$|a| = \sqrt{2J}\left|\frac{2\pi\,h^3\eta}{R_s^2\;m_o\,\gamma\cdot e\,NV\,g'(\phi_s)}\right|^{\frac{1}{4}}\left(1 - \frac{K_o}{4}\right)^{-\frac{1}{4}} \qquad (11.10)$$

This expression generalizes the expression (10.1) of $\hat{\psi}$ to the case of finite difference equations.

12. PHASE DISPLACEMENT ACCELERATION

Empty bucket sweep

For a fixed harmonic number, the RF frequency determines the synchronous revolution frequency or, because

$$\eta \frac{\delta p}{p_s} = \frac{\delta \omega_0}{\omega_0} ,$$

the synchronous momentum of the particles. Let ω_1 and ω_2 be two revolution frequencies located on both sides of the central revolution frequency ω_0 of the stack, well outside the stack, with ω_1 corresponding to a higher momentum than the stack (Fig. 12.1).

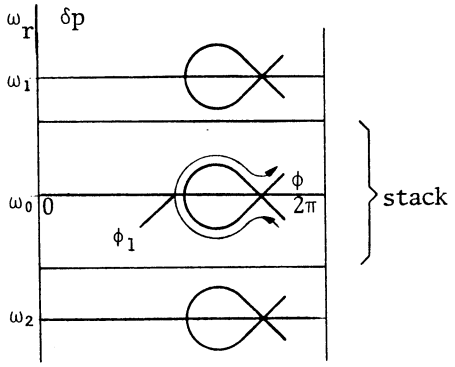


Fig. 12.1 Sweeping an empty bucket through the stack, from ω_1 to ω_2 ($\eta < 0$). With $\Gamma = \sin \phi_s < 0$, particles move upwards around the bucket (when Γ changes sign, all ϕ's change sign).

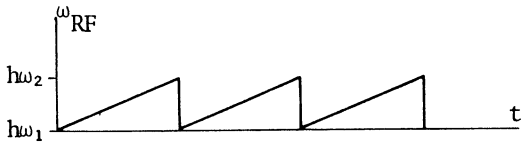


Fig. 12.2 Variation of RF frequency with time ($\eta < 0$).

If the RF frequency is varied from $h\omega_1$ to $h\omega_2$ (Fig. 12.2), an empty bucket is moved completely through the stack in the direction of decreasing p; because phase space is incompressible, the average position of the stack is moved upwards by a quantity equal to (Bucket area/Horizontal axis period). Therefore, according to (9.8), the average momentum of the stack is increased by

$$\left\langle \frac{\Delta p}{m_0 c} \right\rangle = \frac{16}{2\pi} \left| \frac{\gamma}{h\eta} \frac{eNV}{2\pi E_0} \right|^{\frac{1}{2}} \cdot \alpha(\Gamma) \quad \text{per empty bucket sweep} \qquad (12.1)$$

Since this average momentum increase is small, in order to maintain the beam at fixed radius the magnetic field B_z must increase so slightly during an empty bucket sweep, that for all computations B_z may be considered as constant in time. Therefore, the betatron electromotive force may be neglected in (3.15), which for any particle reduces to

$$R \frac{dp}{dt} = \frac{eNV}{2\pi} \sin \phi \tag{12.2}$$

In particular, the stable phase ϕ_s of the empty bucket is determined by

$$R_s \frac{dp_s}{dt} = \frac{eNV}{2\pi} \Gamma \qquad \text{where} \qquad \Gamma = \sin \phi_s < 0 \tag{12.3}$$

i.e. by the slope of the RF frequency with time, since for fixed B_z :

$$\frac{d\omega_{RF}}{dt} = h \frac{d\omega_o}{dt} = h\omega_o \frac{\eta}{p_s} \cdot \frac{dp_s}{dt} = h\eta \frac{v_s}{R_s p_s} \frac{dp_s}{dt}$$

Combined with (12-3) this yields

$$\boxed{\frac{d\omega_{RF}}{dt} = \frac{h\eta}{R_s^2 m_o \gamma_s} \cdot \frac{eNV}{2\pi} \Gamma = \left(\frac{c}{R_s}\right)^2 \frac{h\eta}{\gamma} \frac{eNV}{2\pi E_o} \Gamma} \tag{12.4}$$

For normal acceleration in a synchrotron, where B_z increases noticeably with time, a similar formula applies but with η replaced by $1/\gamma^2$.

The method of phase displacement (Ref. 13) allows acceleration of a stack by an empty bucket with a momentum height which is much smaller than the momentum spread of the stack; this is in contrast with normal acceleration, where the momentum height of the bucket is necessarily larger than the momentum spread of the bunch. Phase displacement acceleration has been successfully used in the ISR to accelerate coasting beams from 26.6 to 31.4 Ge/c (Ref. 14); it necessitated around 200 sweeps of 3 seconds each with a total voltage NV of 12 kV.

Momentum blow up

For any particle, the change in momentum is given by (12.2). Using the reduced time (5.1) and the reduced Hamiltonian (5.5) yields

$$dt = K_1 \, dt^* = K_1 \frac{d\phi}{y} \qquad \text{where} \qquad y = \pm \sqrt{2H^* + 2 \operatorname{sgn}(\eta) \left[\Gamma\phi + \cos \phi\right]}$$

Therefore, during an empty bucket sweep the total change in momentum of a particle which crosses the ϕ-axis at ϕ_1 (see Fig. 12.1) reads

$$\Delta p = 2\int_{\phi_1}^{-\mathrm{sgn}\,(\eta)\cdot\infty} \frac{eNV}{2\pi R} K_1 \frac{\sin\phi \cdot d\phi}{y} \qquad \text{where} \qquad y = \sqrt{2H^* + 2\,\mathrm{sgn}\,(\eta)\left[\Gamma\phi + \cos\phi\right]} > 0 \tag{12.5}$$

$$\text{and} \qquad 0 < \phi_1 < 2\pi \quad \text{excluding the interval } (\phi_e, \phi_u).$$

As a first approximation we shall consider H^* to be a constant; with (5.3) this integral becomes

$$\Delta p = -2\,\mathrm{sgn}\,(\eta)\int_{\phi_1}^{-\mathrm{sgn}\,(\eta)\cdot\infty} \frac{h}{R} K_2 \frac{\sin\phi \cdot d\phi}{y} \qquad \text{where} \qquad y = \sqrt{-2\,\mathrm{sgn}\,(\eta)\left[\Gamma(\phi_1 - \phi) + \cos\phi_1 - \cos\phi\right]}$$

which can be rewritten as

$$\Delta p = -2\,\mathrm{sgn}\,(\eta) \cdot \left\langle \frac{h}{R} K_2 \right\rangle \int_{\phi_1}^{-\mathrm{sgn}\,(\eta)\cdot\infty} \frac{\sin\phi \cdot d\phi}{\sqrt{-2\,\mathrm{sgn}\,(\eta)\left[\Gamma(\phi_1 - \phi) + \cos\phi_1 - \cos\phi\right]}} \tag{12.6}$$

The ensemble average of Δp over the stack is given by (12.1) as

$$\langle \Delta p \rangle = \langle \Delta p_{ST} \rangle \cdot \alpha(\Gamma) \qquad \text{per empty bucket sweep} \tag{12.7}$$

where $\langle \Delta p_{ST} \rangle$ is the average momentum increase due to a stationary bucket sweep:

$$\langle \Delta p_{ST} \rangle = \frac{16}{2\pi} \left| \frac{\gamma}{h\eta} \frac{eNV}{2\pi E_o} \right|^{\frac{1}{2}} m_o c = \frac{16}{2\pi} K_2 \frac{h\omega_o}{\beta E_o} m_o c = \frac{8}{\pi} \left\langle K_2 \frac{h}{R_s} \right\rangle \tag{12.8}$$

when using (5.4) and taking the average of the bucket area throughout the stack. Referred to $\langle \Delta p_{ST} \rangle$, the momentum change (12.6) of an individual particle reads

$$S = \frac{\Delta p}{\langle \Delta p_{ST} \rangle} = -\frac{\pi}{4}\,\mathrm{sgn}\,(\eta) \int_{\phi_1}^{-\mathrm{sgn}\,(\eta)\cdot\infty} \frac{\sin\phi \cdot d\phi}{\sqrt{-2\,\mathrm{sgn}\,(\eta) \cdot \left[\Gamma(\phi_1 - \phi) + \cos\phi_1 - \cos\phi\right]}} \tag{12.9}$$

The integral (12.9) converges when the upper limit of integration tends to ∞; but it must be computed numerically. From (12.7) one must have

$$\langle S \rangle = \alpha(\Gamma) \qquad \text{where} \qquad \alpha(\Gamma) = 0 \qquad \text{for} \qquad |\Gamma| \geq 1 \tag{12.10}$$

which has indeed been verified in the case of a uniform particle distribution in phase space (Ref.15). Detailed numerical computations (Ref. 15), confirmed by measurements, have also shown that:

$$\begin{aligned} &< (S - <S>)^2 > \approx \Gamma^2 && \text{for} \quad |\Gamma| < 1 \\ &\approx \left(\frac{\pi}{4}\right)^3 \cdot \Gamma^{-1} && \text{for} \quad |\Gamma| > 1.5 \end{aligned} \tag{12.11}$$

Therefore, after the n^{th} empty bucket sweep, the mean square momentum spread of the stack is given by

$$\left\langle \left(\frac{\delta p}{m_o c}\right)^2 \right\rangle_n = \left\langle \left(\frac{\delta p}{m_o c}\right)^2 \right\rangle_o + n\,\Gamma^2 \left[\left\langle \frac{\Delta p_{ST}}{m_o c} \right\rangle \right]^2 \tag{12.12}$$

Choice of Γ

The total number n of empty bucket sweeps necessary to increase the momentum of the stack by an amount $< \Delta p >_{wanted}$ reads, with (12.7):

$$n = \frac{< \Delta p >_{wanted}}{< \Delta p >_{sweep}} = \frac{1}{\alpha(\Gamma)} \cdot \frac{< \Delta p >_{wanted}}{< \Delta p_{ST} >} \tag{12.13}$$

whereas the time needed for a single sweep is obtained from (12.3) as

$$T_{sweep} = F \cdot \frac{1}{|\Gamma|} \cdot \frac{< 2\pi R >}{eNV} \cdot 2 \left[(\delta p)_{stack} + M \cdot (\hat{\delta p})_{bucket} \right] \quad \text{per sweep} \tag{12.14}$$

where F and M are safety factors which are taken as (Ref. 16)

$$1 < F \lesssim 2 \qquad \text{and} \qquad 1 < M \lesssim 10$$

With (9.7) and (12.8), the bucket height in (12.14) may be expressed as

$$(\hat{\delta p})_{bucket} = \frac{\pi}{4} < \Delta p_{ST} > \left| \sin \varphi_s - \varphi_s \cos \varphi_s \right|^{\frac{1}{2}} \qquad \text{where} \qquad \cos \varphi_s = |\Gamma|$$

In (12.14) this term is normally much smaller than $(\delta \hat{p})_{stack}$.

From (12.12) and (12.13), in order to keep the momentum blow-up small, one should take $\frac{\Gamma^2}{\alpha(\Gamma)}$ as small as possible. On the other hand, from (12.13) and (12.14), in order to keep

the total time $n \cdot T_{sweep}$ short, one should take $|\Gamma| \cdot \alpha(\Gamma)$ as large as possible. Since $|\Gamma| \cdot \alpha(\Gamma)$ reaches a maximum for $|\Gamma| \approx 0.4$, the best choice is certainly $|\Gamma| < 0.4$. These and other considerations (E. Ciapala: Ref. 2, p.217-220), based on the effects of RF phase noise on particle diffusion and on the effects of a change in the empty bucket area across the stack (due to a variation of γ_{tr} across the machine aperture) (Ref. 17) have, in the case of the ISR led to the choice

$$-0.3 < \Gamma < -0.1 \qquad \text{with mostly} \qquad \Gamma = -0.2$$

13. LINEAR ACCELERATORS

A linear accelerator may be considered as the limiting case of a synchrotron where $R \to \infty$. The momentum compaction is zero:

$$\alpha = 0 \qquad \text{hence} \qquad \gamma_{tr} = \infty$$

and, from (3-3),

$$\eta = \frac{1}{\gamma^2} > 0$$

Therefore a linac is always below transition.

In a linac, the distance L between accelerating cavities becomes the cell length of the accelerating structure:

$$L = \frac{C}{N} = \frac{2\pi R}{N} \qquad \text{is fixed while} \qquad R \to \infty \,, \quad N \to \infty \tag{13.1}$$

$$\frac{h}{R} = \frac{\omega_{RF}}{\omega_o R} = \frac{\omega_{RF}}{v_s} \qquad \text{remains finite while} \qquad R \to \infty \,, \quad h \to \infty \tag{13.2}$$

$$\boxed{\theta = \frac{2\pi h}{N} = \frac{2\pi h}{2\pi R} \cdot L = \frac{\omega_{RF}}{v_s} \; L = k_{RF} \, L \cdot \frac{1}{\beta_s}} \qquad \text{where} \qquad k_{RF} = \frac{\omega_{RF}}{c} \tag{13.3}$$

is the RF phase-shift that must be provided between adjacent cells in order to keep the the phase of the reference particle constant along the accelerating structure (This phase shift should hopefully not be confused with the azimuthal variable θ defined in section 3).

In contrast to synchrotrons, ω_{RF} in linacs does not vary with time, and there are no harmonics on the RF frequency, i.e.

$$g(\phi) = \sin \phi$$

Moreover, linacs are usually designed so as to maintain the phase shift per cell θ constant along the accelerating structure (which is then said to operate in the θ-mode); in that case, as shown by (13.3), L/v_s is constant along the structure: the cell length increases with particle velocity.

Non-relativistic linacs

Finite difference equations

Dividing (3.11) by h, and using (13.3) in (3.4) we obtain

$$\begin{cases} \delta E_{n+1} - \delta E_n = eV \left\{ g(\phi_n) - g(\phi_s) \right\} & (13.4) \\ \delta\phi_{n+1} - \delta\phi_n = -\dfrac{\omega_{RF}}{v_s} \cdot L \cdot \dfrac{\eta}{\beta^2\gamma}\, \delta\gamma_{n+1} = -k_{RF} L \cdot \dfrac{\delta\gamma_{n+1}}{\beta^3\gamma^3} = k_{RF} L \cdot \delta \left(\dfrac{1}{\beta_{n+1}} \right) & (13.5) \end{cases}$$

where V is the peak RF voltage in a single cell.

This mapping preserves area in the $(\delta E_n , \delta\phi_n)$ plane; in fact, the variable conjugate to $\delta\phi_n$ is $\dfrac{\delta E_n}{\omega_{RF}} = \dfrac{E_o}{\omega_{RF}}\, \delta\gamma_n$. The equations above can be written down at once when we observe that for any particle

$$\begin{cases} E_{n+1} - E_n = eV \cdot g(\phi_n) & (13.6) \\ \phi_{n+1} - \phi_n = \omega_{RF} \cdot \dfrac{L}{v_{n+1}} - \theta = k_{RF} L \left(\dfrac{1}{\beta_{n+1}} - \dfrac{1}{\beta_s} \right) & (13.7) \end{cases}$$

Frequency of small phase oscillations

It is given by (11.8), where

$$2\pi \frac{Q_{so}}{N} = \frac{\omega_{so}}{\omega_o} \cdot \frac{2\pi}{N} = \frac{\omega_{so}}{v_s} \frac{2\pi R}{N} = \frac{\omega_{so}}{v_s} L$$

Using (13.3) in (11.8) then yields

$$4 \sin^2 \left(\frac{\omega_{so}}{v_s} \cdot \frac{L}{2} \right) = \theta \frac{\eta}{\beta^2\gamma} \frac{eV\, g'(\phi_s)}{E_o}$$

or

$$\sin^2 \left(\frac{\omega_{so}}{\omega_{RF}} \cdot \frac{\theta}{2} \right) = \frac{\theta}{4} \frac{1}{\beta_s^2\gamma_s^3} \cdot \frac{eV \cos\phi_s}{E_o} = \left(\frac{\theta}{2} \right)^2 \cdot \frac{1}{\beta_s\gamma_s^3} \cdot \frac{eV \cos\phi_s}{k_{RF} L E_o} \qquad (13.8)$$

where V/L is the average accelerating electric field on the linac axis.

Differential equations. Hamiltonian

Taking the distance s along the linac axis as the independent variable, one immediately writes (13.4), (13.5) in differential form as

$$\begin{cases} \dfrac{d}{ds}(\delta\gamma) = \dfrac{eV}{E_o L}\left[g(\phi) - g(\phi_s)\right] = -\dfrac{\partial H}{\partial \phi} & (13.9) \\ \qquad\qquad \text{where} \qquad \delta\gamma = \gamma - \gamma_s \\ \dfrac{d}{ds}(\phi - \phi_s) = -k_{RF} \cdot \dfrac{\delta\gamma}{\beta_s^3 \gamma_s^3} = \dfrac{\partial H}{\partial(\delta\gamma)} & (13.10) \end{cases}$$

Using (4.8) this system may be derived from the Hamiltonian

$$H = -\frac{1}{2}\frac{k_{RF}}{\beta_s^3\gamma_s^3}(\delta\gamma)^2 + \frac{eV}{E_o L}\left[\Gamma\phi + G(\phi)\right] \qquad (13.11)$$

where the potential energy reads, for a sinusoidal RF voltage:

$$\Gamma\phi + G(\phi) = \phi \sin\phi_s + \cos\phi$$

$$= \left(\frac{\pi}{2} + \varphi\right)\cos\varphi_s - \sin\varphi = -(\sin\varphi - \varphi\cos\varphi_s) + \frac{\pi}{2}\cos\varphi_s$$

Instead of the phase ϕ used in synchrotron theory, it is common in linac theory to use the phase φ measured from the crest of the RF voltage. Since $\mathrm{sgn}(\varphi_s) = -\mathrm{sgn}(\eta)\cdot\mathrm{sgn}(\Gamma)$, it follows that φ_s is negative in a linear accelerator.
The variable conjugate to ϕ is $\delta\gamma$ or, in the physical phase space, $\dfrac{\delta E}{\omega_{RF}} = \dfrac{E_o}{\omega_{RF}} \cdot \delta\gamma$;

therefore area is preserved in the $(\delta\gamma, \phi)$ plane.

Adiabatic damping of phase oscillations

The Hamiltonian (13.11) depends on s through $\beta^3\gamma^3$, even if V/L and ϕ_s are assumed to be constant along the linac. Therefore H is not exactly a constant of motion, and the curves H = constant with $\beta^3\gamma^3$ and V/L taken as constants are only approximate trajectories.
In order to have a more accurate picture of the actual trajectories, it is better, as in Fig. 10.2, to use a representation in the reduced coordinate (y,ϕ) plane. From (5.1) and (5.4),

$$y = \frac{1}{K_2}\cdot\frac{\delta E}{\omega_{RF}} = \left|\frac{\beta^2\gamma}{\eta}\,\frac{eV}{E_o\theta}\right|^{-\frac{1}{2}}\cdot\delta\gamma = \left|\beta^3\gamma^3\,\frac{eV}{E_o\, k_{RF}\, L}\right|^{-\frac{1}{2}}\cdot\delta\gamma \qquad (13.12)$$

where again we have used (13.3). In the (γ,ϕ) plane, the trajectories look like the bottom part $(\gamma < \gamma_{tr})$ of Fig. 10.2, where we recognize the familiar golf-club of low-β linacs (Ref. 8; Ref. 9, p.27).

Relativistic linacs

They are built for $\beta_s = 1$, i.e. $\gamma_s = \infty$. By (13.3) this entails that for θ constant, L is constant along the linac: therefore the accelerating structure is exactly periodic, with geometrical period L. Since v_s is constant, the synchronous particle is not accelerated, and

$$\Gamma = g(\phi_s) = \sin \phi_s = 0$$

Since $\eta > 0$, $\phi_s = 0$ at the stable fixed point.

Therefore the particles are accelerated inside a stationary bucket, but the reference energy is at infinity.

Finite difference equations

Since $\gamma_s = \infty$, we can no longer use the system (13.4), (13.5); instead we must use the system (13.6), (13.7) which reads, for any particle:

$$\begin{cases} \gamma_{n+1} - \gamma_n = \dfrac{eV}{E_o}\, g(\phi_n) & (13.13) \\[2ex] \phi_{n+1} - \phi_n = k_{RF}\, L \left(\dfrac{1}{\beta_{n+1}} - 1\right) \quad \text{where} \quad \dfrac{1}{\beta_{n+1}} = \left(1 - \dfrac{1}{\gamma_{n+1}^2}\right)^{-\frac{1}{2}} & (13.14) \end{cases}$$

This mapping preserves area in the (γ_n, ϕ_n) plane; but like the mapping (11.1), (11.2), it does not admit of any (smooth) constant of motion. In order to obtain an (approximate) constant of motion, we must again go over to differential equations.

Differential equations. Hamiltonian

In differential form, the system (13.13), (13.14) becomes

$$\begin{cases} \dfrac{d\gamma}{ds} = \dfrac{eV}{E_o L}\, g(\phi) = -\dfrac{\partial H}{\partial \phi} & (13.15) \\[2ex] \dfrac{d\phi}{ds} = k_{RF} \left(\dfrac{1}{\beta} - 1\right) = \dfrac{\partial H}{\partial \gamma} \quad \text{where} \quad \dfrac{1}{\beta} = \left(1 - \dfrac{1}{\gamma^2}\right)^{-\frac{1}{2}} & (13.16) \end{cases}$$

Using (4.8) this system may be derived from the Hamiltonian

$$H = k_{RF}\,(\beta - 1)\,\gamma \quad + \quad \frac{eV}{E_o L} \cdot G(\phi) \tag{13.17}$$

where, for a sinusoidal voltage: $G(\phi) = \cos\phi = -\sin\varphi$

If, as is often the case, the average accelerating field V/L is maintained constant along the accelerator, H does not depend explicitly on s : it is therefore a constant of motion, and the particle trajectories in the (γ,ϕ) plane are the curves H = constant. By putting with (13.3),

$$K = \frac{eV}{k_{RF}\,L \cdot E_o} = \frac{1}{\theta} \cdot \frac{eV}{E_o} \tag{13.18}$$

it is seen that (13.17) is proportional to the reduced Hamiltonian

$$H^* = -\,(1 - \beta)\,\gamma \quad + \quad K \cdot \cos\phi = -\,\frac{1}{(1+\beta)\,\gamma} \quad + \quad K \cdot \cos\phi \tag{13.19}$$

which depends on a single dimensionless parameter K.

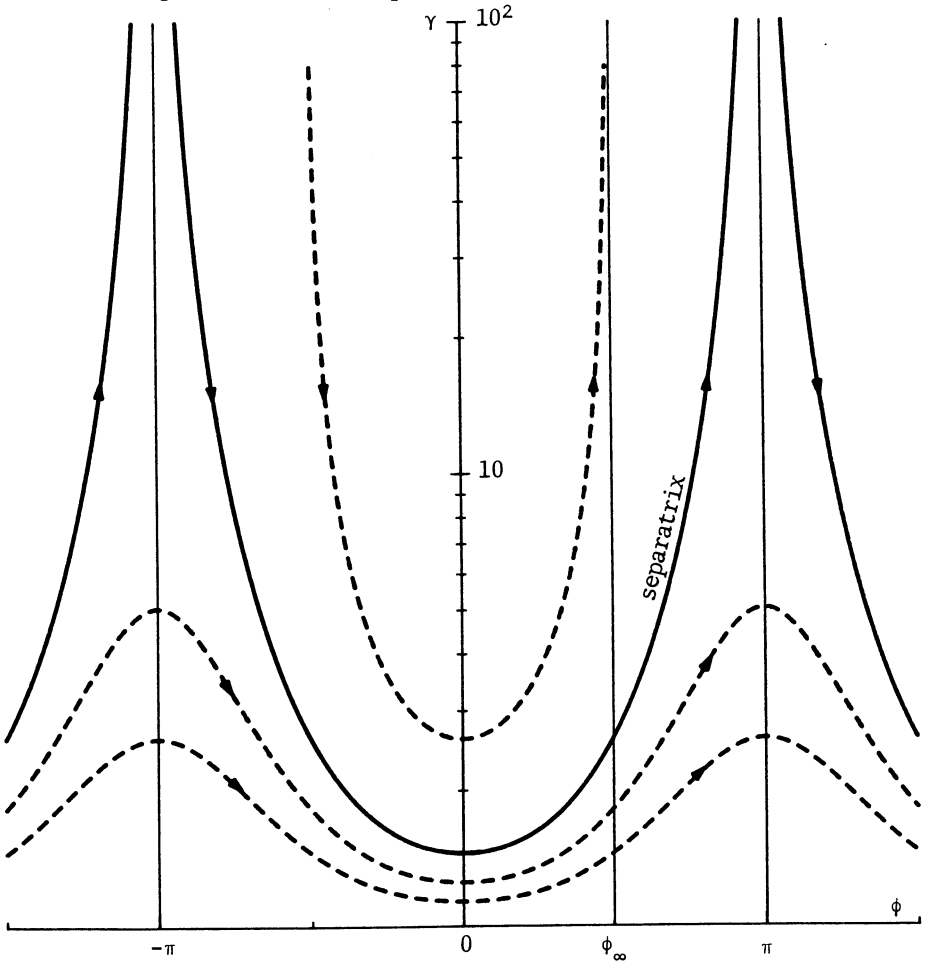


Fig. 13.1 - Longitudinal phase space of a relativistic linac ($\phi_s = 0$, $\gamma_s = \infty$). In the figure, K = 0.2 ; this value corresponds to the SLAC 3.2 km electron linac operated at 20 GeV (with $\theta = 2\pi/3$).

Particle trajectories $H^* = \text{constant}$ are shown in Fig. 13.1. When $\gamma \to \infty$, ϕ approaches an asymptotic value ϕ_∞ given by

$$H^* = K \cdot \cos \phi_\infty \tag{13.20}$$

If $H^* > K$, there is no corresponding real trajectory in the γ, ϕ plane.

If $-K < H^* < K$, the trajectories are asymptotic to a line parallel to the γ-axis at ϕ_∞, with $0 < \phi_\infty < \pi$.

If $H^* < -K$, γ reaches a finite maximum at $\phi = \pi$.

All trajectories are symmetrical with respect to $\phi = 0$ and have period 2π in ϕ. The particular curve which corresponds to $\phi_\infty = \pi$ separates the bounded and unbounded motions: it is thus the separatrix; at the same time it is an actual trajectory, because the bucket is stationary and K is assumed to be constant.

In (13.15), (13.16), the variable conjugate to ϕ is γ or, in the physical phase space, $\frac{E}{\omega_{RF}} = \frac{E_o}{\omega_{RF}} \cdot \gamma$; therefore the particle motion preserves area in the (γ,ϕ) plane.

In order to reach high energies, particles are injected with initial conditions such that $\phi_\infty \approx \frac{\pi}{2}$.

* * *

REFERENCES

(1) B. Schnizer, A simple derivation of proton linac beam dynamics equations, Particle Accelerators, Vol. 2 (1971), p. 141-147.

(2) CERN Accelerator School: General Accelerator Physics, Paris, September 1984, CERN 85-19, Vol. I.

(3) A.A. Kolomensky and A.N. Lebedev, Theory of cyclic accelerators, North Holland, Amsterdam, 1966.

(4) H.G. Hereward, What are the equations for the phase oscillations in a synchroton? CERN 66-6, 1966.

(5) L.D. Landau and E.M. Lifschitz, Mechanics, Pergamon Press, Oxford, 1960.

(6) W. Yourgrau and S. Mandelstam, Variational principles in dynamics and quantum theory, 3rd edition, Pitman, London, 1968.

(7) E.A. Hylleraas, Mathematical and theoretical phyiscs, Wiley, New York, 1970, Vol. 2.

(8) R.L. Gluckstern: Phase acceptance in linacs, 6th International Conference on High Energy Accelerators, CEAL - 2000, Cambridge Electron Accelerator, 1967, p. 153-156.

(9) H.G. Hereward: The general theory of linear accelerators, in "Linear Accelerators", edited by P.M. Lapostolle and A.L. Septier, North-Holland, Amsterdam, 1970.

(10) R. Abraham and C. Shaw: Dynamics, Part 3: Global behaviour, Aerial Press, Santa Cruz, California, 1983.

(11) L. Jackson Laslett, Stochasticity, 9th International Conference on High Energy Accelerators, Stanford, 1974, p. 394-401.

(12) G. Dôme, Frontière du domaine de stabilité autour de l'origine pour la transformation "standard" de Chirikov, CERN-SPS/82-20 (ARF); or Théorie de l'itération et ses applications, Editions du CNRS, Paris, 1982, p. 195-204.

(13) K.R. Symon and A.M. Sessler, Methods of radio frequency acceleration in fixed field accelerators, CERN Symposium 1956, p. 44-58.

(14) E. Ciapala, S. Myers, C. Wyss, Phase displacement acceleration of high intensity stacks in the CERN ISR, CERN-ISR-OP/77-13; or 1977 Particle Accelerator Conference, IEEE Transactions on Nuclear Science, Vol. NS-24, 1977, p. 1431-1433.

(15) E.W. Messerschmid: Scattering of particles by phase displacement acceleration in storage rings, CERN-ISR-TH/ 73-31.

(16) S. Myers: private communication, February 1987.

(17) S. Myers and E. Peschardt: The efficiency of acceleration by phase displacement in the presence of RF noise, 1979 Particle Accelerator Conference, IEEE Transactions on Nuclear Science, Vol. NS-26, 1979, p. 3529-3531.

APPENDIX A Synchrotron frequency in a stationary bucket with a harmonic cavity

$$\phi_s = 0 \quad \text{if} \quad \eta > 0$$

$$\phi_s = \pi \quad \text{if} \quad \eta < 0$$

$$g(\phi) = \sum_{n=1}^{\infty} a_n \sin(n\phi) \qquad G(\phi) = \sum_{n=1}^{\infty} \frac{a_n}{n} \cos(n\phi) \qquad \text{with } a_1 = 1$$

Canonical equations:

$$\begin{cases} \dfrac{dy}{dt^*} = \operatorname{sgn}(\eta) \left(\Gamma - g(\phi) \right) \\ \dfrac{d\phi}{dt^*} = y \end{cases}$$

Differential equation for ϕ:

$$\frac{d^2\phi}{dt^{*2}} - \operatorname{sgn}(\eta) \left(\Gamma - g(\phi) \right) = 0$$

For a stationary bucket, $\Gamma = 0$ and this equation reduces to

$$\frac{d^2\phi}{dt^{*2}} + \operatorname{sgn}(\eta) \sum_{n=1}^{\infty} a_n \sin(n\phi) = 0$$

Let $\psi = \phi - \phi_s$; then

$$\frac{d^2\psi}{dt^{*2}} + \operatorname{sgn}(\eta) \sum_{n=1}^{\infty} a_n \cos(n\phi_s) \sin(n\psi) = 0$$

Putting

$$c_n = \operatorname{sgn}(\eta) \cdot a_n \cos(n\phi_s) \tag{A.1}$$

we finally have

$$\boxed{\frac{d^2\psi}{dt^{*2}} + \sum_{n=1}^{\infty} c_n \sin(n\psi) = 0 \qquad \text{with} \qquad c_1 > 0} \tag{A.2}$$

Depending on the origin of time, the solution of (A.2) may in specific cases appear as an even or an odd function of time. Let us assume that it is an even function of time, then we may write the solution of (A.2) in the form

$$\psi = \sum_{m=0}^{\infty} \alpha_m \cos(m\tau) \qquad \text{where} \qquad \tau = \omega^* t^* \tag{A.3}$$

hence

$$\frac{d^2\psi}{dt^{*2}} = -\omega^{*2} \cdot \sum_{m=0}^{\infty} m^2 \alpha_m \cos(m\tau) \tag{A.4}$$

Using the expansions

$$e^{jz \sin\theta} = \sum_{p=-\infty}^{+\infty} J_p(z)\, e^{jp\theta}$$

$$e^{jz \cos\theta} = \sum_{p=-\infty}^{+\infty} j^p\, J_p(z)\, e^{jp\theta}$$

we obtain

$$e^{jn\psi} = \prod_{m=0}^{\infty} e^{jn\,\alpha_m \cos m\tau}$$

$$= \prod_{m=0}^{\infty} \sum_{p=-\infty}^{+\infty} j^p J_p(n\,\alpha_m)\, e^{jpm\tau}$$

$$= \sum_{\sum_m m\,p_m = -\infty}^{+\infty} \prod_m j^{p_m} J_{p_m}(n\,\alpha_m) \cdot e^{j\tau \sum_m m\,p_m}$$

where p_m is a full set of integers (from $-\infty$ to $+\infty$) corresponding to m.
Combining (A.4) with (A.2) yields

$$\omega^{*2} \cdot \sum_{r=0}^{\infty} r^2\, \alpha_r \cos(r\tau) = \sum_{n=1}^{\infty} c_n \sin(n\psi)$$

$$= \sum_{n=1}^{\infty} c_n\, \mathrm{Im} \sum_{\sum_m m\,p_m = -\infty}^{+\infty} \prod_m j^{p_m} J_{p_m}(n\,\alpha_m) \cdot \left[\cos\left(\tau \sum_m m\,p_m\right) + j \sin\left(\tau \sum_m m\,p_m\right) \right]$$

Let

$$\sum_{m=0}^{\infty} m\,p_m = \pm\, r \qquad \text{and} \qquad \sum_{m=0}^{\infty} p_m = q \tag{A.5}$$

By changing p_m into $-p_m$ we see that all sin-terms disappear; we are thus left with

$$\omega^{*2}\, r^2\, \alpha_r = \sum_{q\ \text{odd} = -\infty}^{+\infty} \sum_{n=1}^{\infty} c_n\, j^{q-1} J_{p_0}(n\alpha_0)\, J_{p_1}(n\alpha_1)\, J_{p_2}(n\alpha_2) \ldots$$

It is easy to see that in the $(\psi,\dot{\psi})$ plane, a trajectory must be symmetrical with respect to both axes; this entails that in (A.3), m = 1, 3,...

and (A.5) reduces to

$$\begin{aligned} p_1 + 3p_3 + 5p_5 + \ldots &= \pm r \\ p_1 + p_3 + p_5 + \ldots &= q \end{aligned} \qquad \text{(A.6)}$$

hence $\pm r - q = 2p_3 + 4p_5 + \ldots$ is even.

Since q must be odd, r must also be odd and we are left with

$$\sum_{q \text{ odd}=-\infty}^{+\infty} \sum_{n=1}^{+\infty} c_n \, j^{q-1} \, J_{p_1}(n\alpha_1) \, J_{p_3}(n\alpha_3) \, J_{p_5}(n\alpha_5) \ldots = \omega^{*2} \, r^2 \, \alpha_r \qquad \text{(A.7)}$$

$r = 1, 3, \ldots$

In the following, we neglect terms of order α_5 and higher; this amounts to taking $p_5 = p_7 = \ldots = 0$ and to putting $J_0(n\alpha_5) = J_0(n\alpha_7) = \ldots = 1$. Then (A.6) reduces to

$$p_1 + 3p_3 = \pm r$$

$$q = p_1 + p_3 = \pm r - 2p_3$$

Changing p_m into $-p_m$ changes $+r$ into $-r$ but leaves the summation in (A.7) unchanged. Therefore, it is sufficient to keep $+r$ only if one introduces a factor 2 in the left-hand side of (A.7).

<u>r = 1</u>

$$p_1 + 3p_3 = 1 \qquad \text{with} \qquad q = 1 - 2p_3$$

Then (A.7) becomes successively

$$\sum_{n=1}^{\infty} c_n \; 2 \sum_{p_3=-\infty}^{+\infty} (-1)^{p_3} \, J_{1-3p_3}(n\alpha_1) \, J_{p_3}(n\alpha_3) = \omega^{*2} \, \alpha_1$$

$$\sum_{n=1}^{\infty} c_n \; 2 \Big[J_1(n\alpha_1) \, J_0(n\alpha_3) - J_2(n\alpha_1) \, J_1(n\alpha_3) + J_4(n\alpha_1) \, J_1(n\alpha_3) - J_5(n\alpha_1) \, J_2(n\alpha_3) + J_7(n\alpha_1) \, J_2(n\alpha_3) - \ldots \Big] = \omega^{*2} \, \alpha_1$$

$$\boxed{\sum_{n=1}^{\infty} n \, c_n \left\{ 2 \frac{J_1(n\alpha_1)}{n\alpha_1} J_0(n\alpha_3) - 2 \frac{J_1(n\alpha_3)}{n\alpha_1} \Big[J_2(n\alpha_1) - J_4(n\alpha_1) \Big] - 0 \left(\alpha_1^4 \, , \, \alpha_3^2\right) \right\} = \omega^{*2}} \qquad \text{(A.8)}$$

r = 3

$$p_1 + 3p_3 = 3 \qquad \text{with} \qquad q = 3 - 2p_3$$

Then (A.7) becomes successively

$$\sum_{n=1}^{\infty} c_n \, 2 \sum_{p_3=-\infty}^{+\infty} (-1)^{1-p_3} \, J_{3-3p_3}(n\alpha_1) \, J_{p_3}(n\alpha_3) = \omega^{*2} \, 3^2 \, \alpha_3$$

$$\sum_{n=1}^{\infty} c_n \, 2 \sum_{p=-\infty}^{+\infty} J_{3p}(n\alpha_1) \, J_{p+1}(n\alpha_3)$$

$$= \sum_{n=1}^{\infty} c_n \, 2 \Big[J_o(n\alpha_1) \, J_1(n\alpha_3) - J_3(n\alpha_1) \, J_o(n\alpha_3) + J_3(n\alpha_1) \, J_2(n\alpha_3) - J_6(n\alpha_1) \, J_1(n\alpha_3)$$

$$+ J_6(n\alpha_1) \, J_3(n\alpha_3) - \ldots \Big] = 3^2 \, \omega^{*2} \cdot \alpha_3$$

$$\sum_{n=1}^{\infty} n \, c_n \left\{ \frac{2J_1(n\alpha_3)}{n\alpha_3} \left[J_o(n\alpha_1) - J_6(n\alpha_1) \right] - 2 \frac{J_3(n\alpha_1)}{n\alpha_3} \left[J_o(n\alpha_3) - J_2(n\alpha_3) \right] \right.$$

$$\left. + 0 \, (\alpha_1{}^6 \, . \, \alpha_3{}^2) \right\} = 3^2 \, \omega^{*2} \qquad \text{(A.9)}$$

Keeping only the first terms in (A.9) and (A.8) yields

$$\begin{cases} \displaystyle\sum_{n=1}^{\infty} n \, c_n \left\{ J_o(n\alpha_1) - 2 \frac{J_3(n\alpha_1)}{n\alpha_3} + 0(\alpha_1{}^6) + 0(\alpha_3{}^2) \right\} = 3^2 \, \omega^{*2} & \text{(A.10)} \\ \displaystyle\sum_{n=1}^{\infty} n \, c_n \left\{ 2 \frac{J_1(n\alpha_1)}{n\alpha_1} - \frac{n\alpha_3}{n\alpha_1} J_2(n\alpha_1) + 0(\alpha_1{}^6) + 0(\alpha_1{}^3\alpha_3) + 0(\alpha_3{}^2) \right\} = \omega^{*2} & \text{(A.11)} \end{cases}$$

In particular, by letting $\alpha_1 \to 0$ in (A.11) we obtain

$$\omega_o^{*2} = \sum_{n=1}^{\infty} n\, c_n \tag{A.12}$$

Combining (A.10) and (A.11) yields

$$\sum_{n=1}^{\infty} n\, c_n \left\{ 3^2 \frac{2J_1(n\alpha_1)}{n\alpha_1} - 3^2 \frac{\alpha_3}{\alpha_1} J_2(n\alpha_1) - J_o(n\alpha_1) + 2 \frac{J_3(n\alpha_1)}{n\alpha_3} \right.$$

$$\left. + O(\alpha_1{}^6) + O(\alpha_1{}^3\alpha_3) + O(\alpha_3{}^2) \right\} = 0$$

$$\sum_{n=1}^{\infty} n\, c_n \left\{ 3^2\, 2 \frac{J_1(n\alpha_1)}{n\alpha_1} - J_o(n\alpha_1) - \frac{\alpha_3}{\alpha_1} 3^2 J_2(n\alpha_1) + \frac{\alpha_1}{\alpha_3} 2 \frac{J_3(n\alpha_1)}{n\alpha_1} \right.$$

$$\left. + O(\alpha_1{}^6) + O(\alpha_1{}^3\alpha_3) + O(\alpha_3{}^2) \right\} = 0$$

$$\underbrace{-\frac{\alpha_3}{\alpha_1} \sum_{n=1}^{\infty} n\, c_n \left[3^2 \cdot 2 \frac{J_1(n\alpha_1)}{n\alpha_1} - J_o(n\alpha_1) \right]}_{B} + \underbrace{\left(\frac{\alpha_3}{\alpha_1}\right)^2 3^2 \cdot \sum_{n=1}^{\infty} n\, c_n\, J_2(n\alpha_1)}_{A}$$

$$\underbrace{-\sum_{n=1}^{\infty} n\, c_n\, 2 \frac{J_3(n\alpha_1)}{n\alpha_1}}_{C} = \left[O(\alpha_1{}^6) + O(\alpha_1{}^3\alpha_3) + O(\alpha_3{}^2) \right] \frac{\alpha_3}{\alpha_1} \tag{A.13}$$

Because

$$2\nu \frac{J_\nu(x)}{x} = J_{\nu-1}(x) + J_{\nu+1}(x) \quad ,$$

it is seen that:

$$A = 3^3 \cdot C + O(\alpha_1{}^4) = O(\alpha_1{}^2) \qquad \text{while} \qquad B = O(1)$$

Solve (A.13) for α_3 / α_1 :

$$\frac{\alpha_3}{\alpha_1} = \frac{-2C}{B + \sqrt{B^2 + 4AC}} \approx \frac{-C}{B + \frac{AC}{B}} = - \frac{BC}{B^2 + AC} \left[1 + 0(\alpha_1{}^6) \right] = - \frac{C}{B} \left[1 + 0(\alpha_1{}^4) \right]$$

$$= - \frac{BC}{B^2 + 3^3C^2} \left[1 + 0(\alpha_1{}^6) \right] = - \frac{C}{B} \left[1 + 3\left(\frac{3C}{B}\right)^2 \right]^{-1} \cdot \left[1 + 0(\alpha_1{}^6) \right].$$

Finally

$$\alpha_3 = - \frac{\sum_{n=1}^{\infty} c_n \, 2 \, J_3(n\alpha_1)}{\sum_{n=1}^{\infty} n \, c_n \left[3^2 \cdot 2 \frac{J_1(n\alpha_1)}{n\alpha_1} - J_o(n\alpha_1) \right]} \left[1 + 3\left(\frac{3C}{B}\right)^2 \right]^{-1} \left[1 + 0(\alpha_1{}^6) \right] \qquad \text{(A.14)}$$

The synchrotron frequency is then obtained as a function of

$$\hat{\psi} = \psi(0) = \sum_{m=1,3,\ldots}^{\infty} \alpha_m$$

by using (A.14) in (A.8) and remembering that $\alpha_3 = 0(-\alpha_1{}^3)$.

Remark: The series

$$\psi(\tau) = \sum_{m=1,3,\ldots}^{\infty} \alpha_m \cos(m\tau)$$

is equivalent to

$$\psi\left(\tau - \frac{\pi}{2}\right) = \sum_{m=1,3,\ldots}^{\infty} \alpha_m \sin\left(m \frac{\pi}{2}\right) \sin(m\tau)$$

APPENDIX B Bucket Area $\alpha(\Gamma)$

From (9-1), $\Gamma = \sin \phi_s = \sin \phi_u$ $\quad |\Gamma| = \cos \varphi_s$

Put $\psi = - \operatorname{sgn}(\eta\Gamma) \cdot (\phi - \phi_u) > 0$

From (9-4) and (9-5),

$$\alpha = - \frac{2}{16} \operatorname{sgn}(\eta\Gamma) \int_{\phi_u}^{\phi_e} \sqrt{2} \sqrt{\underbrace{- \operatorname{sgn}(\eta) \left[\Gamma\phi_u + \cos \phi_u \right] + \operatorname{sgn}(\eta) \left[\Gamma\phi + \cos \phi \right]}} \cdot d\phi$$

$$- \operatorname{sgn}(\eta) \cos \phi_u \left[1 - \cos \psi \right] - \left[\Gamma\psi - \sin \phi_u \sin \psi \right]$$

$$= |\sin \varphi_s| \, (1 - \cos \psi) - \cos \varphi_s \, (\psi - \sin \psi)$$

Therefore

$$\alpha = \frac{2}{16} \int_0^{|\phi_e - \phi_u|} \sqrt{2} \sqrt{|\sin \varphi_s| \, (1 - \cos \psi) - \cos \varphi_s \, (\psi - \sin \psi)} \cdot d\psi$$

or

$$\alpha = \frac{1}{2} \sqrt{\cos \varphi_s} \int_{\psi=0}^{\psi=|\phi_e - \phi_u|} d\frac{\psi}{2} \cdot \sin \frac{\psi}{2} \sqrt{|\operatorname{tg} \varphi_s| - t} \qquad \text{(B.1)}$$

where

$$t = \frac{\psi - \sin \psi}{1 - \cos \psi}$$

As a series in ψ,

$$t = \frac{\psi}{3} + \frac{\psi^3}{2.5.9} + \frac{\psi^5}{2.4.5.7.9} + \frac{\psi^7}{3^3.4^2.5^2.7} + \frac{\psi^9}{3^3.4^2.8.7.9.11} + \frac{691\,\psi^{11}}{3^5.4^4.5^3.7^2.11.13} + \ldots \qquad \text{(B.2)}$$

Inverting the series (B.2) one obtains

$$\frac{\psi}{2} = \frac{3}{2} t \left[1 - \frac{2}{3.5} \left(\frac{3}{2} t \right)^2 + \frac{2.3}{5^2.7} \left(\frac{3}{2} t \right)^4 - \frac{2^2}{3.5^3} \left(\frac{3}{2} t \right)^6 \right.$$

$$\left. + \frac{2.1103}{3^2.5^3.7^2.11} \left(\frac{3}{2} t \right)^8 - \frac{2^2.7171}{5^5.7^2.11.13} \left(\frac{3}{2} t \right)^{10} + \ldots \right] \qquad \text{(B.3)}$$

For $t = \operatorname{tg} \varphi_s$, this relation yields $\psi = \phi_e - \phi_u$ as a series in $\operatorname{tg} \varphi_s$; converted into a series in φ_s it reads

$$\phi_e - \phi_u = 3\,\varphi_s + \frac{1}{2.5}\,\varphi_s^3 + \frac{29}{2^3.5^2.7}\,\varphi_s^5 + \frac{13.17}{2^4.3.5^3.7}\,\varphi_s^7 + \frac{346943}{2^7.3^3.5^3.7^2.11}\,\varphi_s^9 + \cdots \quad (B.4)$$

From (B.4) one can derive the faster converging expression (9.2).

From (B.3) one can write

$$1 - \cos\frac{\psi}{2} = \frac{1}{2!}\left(\frac{3}{2}t\right)^2 - \frac{3.7}{5!}\left(\frac{3}{2}t\right)^4 + \frac{3^2.11.17}{7!\,5}\left(\frac{3}{2}t\right)^6 - \frac{3.1783}{8!\,5}\left(\frac{3}{2}t\right)^8 + \frac{3^3.(2791423)}{11!\,5^2.7}\left(\frac{3}{2}t\right)^{10} - \cdots$$

$$= \sum_{n=2,4,..}^{\infty} a_n\, t^n \quad (B.5)$$

by definition of the a_n's . Using this expansion transforms (B.1) into

$$\alpha = \frac{1}{2}\sqrt{\cos\varphi_s}\int_{t=0}^{t=|\operatorname{tg}\varphi_s|} d\left(1 - \cos\tfrac{\psi}{2}\right)\sqrt{|\operatorname{tg}\varphi_s| - t}$$

$$\alpha = \frac{1}{2}\sqrt{\cos\varphi_s}\left\{\left[\left(1 - \cos\tfrac{\psi}{2}\right)\sqrt{|\operatorname{tg}\varphi_s| - t}\right]_{t=0}^{t=|\operatorname{tg}\varphi_s|} + \int_0^{|\operatorname{tg}\varphi_s|}\left(1 - \cos\tfrac{\psi}{2}\right)\frac{dt}{2\sqrt{|\operatorname{tg}\varphi_s| - t}}\right\}$$

$$\alpha = \frac{1}{2}\sqrt{\cos\varphi_s}\sum_{n=2,4,..}^{\infty} a_n \underbrace{\int_0^{|\operatorname{tg}\varphi_s|} t^n \frac{dt}{2\sqrt{|\operatorname{tg}\varphi_s| - t}}}_{1}$$

$$|\operatorname{tg}\varphi_s|^{n+\frac{1}{2}}\int_0^1 \frac{x^n}{2}(1-x)^{-\frac{1}{2}}\,dx = |\operatorname{tg}\varphi_s|^{n+\frac{1}{2}}\,\frac{\Gamma(n+1)\,\Gamma(\frac{1}{2})}{2\Gamma(n+\frac{3}{2})}$$

$$= |\operatorname{tg}\varphi_s|^{n+\frac{1}{2}}\,\frac{(1)_n}{(\frac{3}{2})_n}$$

hence

$$\alpha = \frac{1}{2}\sqrt{|\sin\varphi_s|}\sum_{n=2,4,..}^{\infty} a_n \frac{(1)_n}{\left(\frac{3}{2}\right)_n}(\mathrm{tg}\,\varphi_s)^n \qquad (B.6)$$

where the a_n's are given by (B.5).

With $\mathrm{tg}^2\varphi_s = \dfrac{\sin^2\varphi_s}{1-\sin^2\varphi_s}$, (B.6) can be converted into a series in $\sin^2\varphi_s$:

$$\alpha = \frac{3}{10}\left|\sin\varphi_s\right|^{\frac{5}{2}}\left[1 + \frac{2}{5}\sin^2\varphi_s + \frac{2^2.37}{7^2.13}\sin^4\varphi_s + \frac{2^3.13.347}{5^2.7^2.11.17}\sin^6\varphi_s + \frac{2^4.(490487)}{5^3.7^3.11^2.13}\sin^8\varphi_s + \frac{2^5.(100403329369)}{5^6.7^3.11^2.13^2.17.19}\sin^{10}\varphi_s + \ldots\right] \qquad (B.7)$$

and then to a series in φ_s^2:

$$\alpha(\varphi_s) = \frac{3}{10}\left|\varphi_s\right|^{\frac{5}{2}}\left[1 - \frac{1}{3.5}\left(\frac{\varphi_s}{2}\right)^2 + \frac{1607}{2.3.5.7^2.13}\left(\frac{\varphi_s}{2}\right)^4 + \frac{939031}{2.3^3.5^2.7^2.11.17}\left(\frac{\varphi_s}{2}\right)^6 + \frac{397918669}{2^3.3^2.5^3.7^3.11^2.13}\left(\frac{\varphi_s}{2}\right)^8 + \frac{240180071548099}{2^3.3^2.5^6.7^3.11^2.13^2.17.19}\left(\frac{\varphi_s}{2}\right)^{10} + \ldots\right] \qquad (B.8)$$

Cut after $\left(\dfrac{\varphi_s}{2}\right)^4$, this expression yields at the worst point for convergence $\left(\varphi_s = \frac{\pi}{2}\right)$:

$\alpha\left(\frac{\pi}{2}\right) = 0.919262$ instead of the correct value 1 .

BEAM-BEAM EFFECTS

L.R. Evans and J. Gareyte
CERN, Geneva, Switzerland

ABSTRACT

The beam-beam interaction imposes severe limitations on the performance of colliding beam storage rings. In linear colliders the self-pinching effect of the beam-beam force can enhance their perfomance.

1. INTRODUCTION

Particles circulating in a colliding beam storage device experience localised periodic kicks when crossing the opposing beam. As the intensity increases, this beam-beam interaction has a profound effect on the beam dynamics, ultimately limiting the performance of all existing lepton storage rings as well as the SPS hadron collider. It is therefore not surprising that a great deal of experimental, theoretical and computational effort has gone into trying to understand the underlying physics.

For lepton machines, computer simulations have been particularly productive in understanding and predicting machine performance limitations due to this effect. In contrast, in hadron colliders computer simulation is hampered paradoxically by the conservative nature of the beam dynamics and one has to rely more heavily on approximate analytical models based on the theory of nonlinear resonance effects[1-5].

In the near future the first linear collider will be commissioned[6]. Although the beams only pass once through one another, a very strong beam-beam interaction, commonly known as beam disruption, will occur. In contrast with storage rings this disruption has a potentially beneficial effect, strongly focussing both beams and increasing the luminosity by a substantial factor.

2. THE BEAM-BEAM FORCE

We consider first the simplest case of head-on collisions between two round Gaussian bunches of length L with n particles per unit length and with a transverse density distribution

$$\rho(r) = \frac{ne}{2\pi\sigma^2} e^{-r^2/2\sigma^2} \quad . \tag{2.1}$$

The Lorentz force on a test particle at a radius r is

$$\underline{F} = e(\underline{E} + \underline{v} \times \underline{B}) = e(E_r \pm \beta c B_\phi) \cdot \underline{r} \qquad (2.2)$$

where the negative sign corresponds to a particle in the same bunch and the positive sign to a particle in the other beam, $\underline{r}$ is the unit vector.

The radial electric field E_r and the poloidal magnetic induction B_ϕ can be obtained from Gauss' theorem and Ampère's law respectively.

$$2\pi r E_r = \frac{1}{\varepsilon_o} \int_o^r 2\pi r' \rho(r') \, dr'$$

so

$$E_r = \frac{ne}{2\pi r \varepsilon_o} (1 - e^{-r^2/2\sigma^2}) , \qquad (2.3)$$

$$2\pi r B_\phi = \mu_o \int_o^r 2\pi r' \beta c \rho(r') \, dr'$$

and

$$B_\phi = \frac{ne\mu_o \beta c}{2\pi r} (1 - e^{-r^2/2\sigma^2}) . \qquad (2.4)$$

Then

$$F_r = \frac{ne^2}{2\pi r \varepsilon_o} (1 \pm \beta^2) (1 - e^{-r^2/2\sigma^2}) . \qquad (2.5)$$

The angular kick $\Delta x'$ due to the beam-beam interaction is then given by

$$\Delta x' = \frac{Ne}{2\pi \varepsilon_o \beta c B\rho} \cdot \frac{1}{x} (1 - e^{-x^2/2\sigma^2}) \qquad (2.6)$$

where N is the number of particles per bunch and we consider only the plane y = 0.

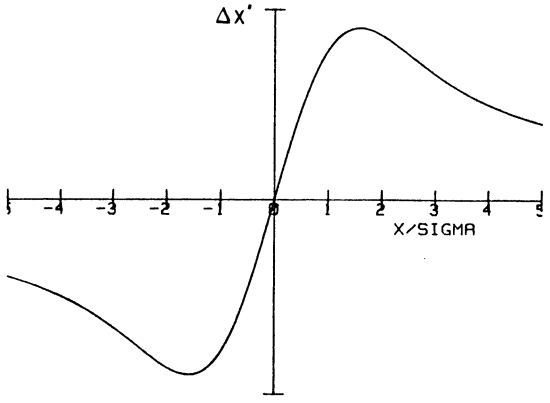


Fig. 1 Beam-beam kick for a round Gaussian beam

For the general case ($\sigma_x \neq \sigma_y$) the beam-beam kick can be obtained by solving Poisson's equation for the generalised electromagnetic potential of an elliptical bunch[7]

$$V(x,y) = \frac{ne}{4\pi\varepsilon_o} \int_o^\infty dt \, \frac{1 - e^{-\left(\frac{x^2}{2\sigma_x^2 + t} + \frac{y^2}{2\sigma_y^2 + t}\right)}}{\sqrt{(2\sigma_x^2 + t)(2\sigma_y^2 + t)}} \quad . \tag{2.7}$$

The kicks $\Delta x'$ and $\Delta y'$ due to the beam-beam interaction are then

$$\Delta x' = -\frac{\partial V}{\partial x} \quad \text{and} \quad \Delta y' = -\frac{\partial V}{\partial y} \quad .$$

2.1 The Linear Tune Shift

For small amplitude particles, the beam-beam kick is identical to that given by a linear lens of focal length f given by

$$\begin{aligned} 1/f &= \Delta x'/x \\ &= \frac{Ne}{4\pi\varepsilon_o \beta c B \rho \sigma^2} \\ &= \frac{Nr_o}{\gamma\sigma^2} \end{aligned} \tag{2.8}$$

where r_o is the classical particle radius $r_o = e^2/(4\pi\varepsilon_o m_o c^2)$.

To investigate the perturbation of the lattice functions by this lens we compute the perturbed one-turn transfer matrix

$$\begin{bmatrix} \cos(\mu + \Delta\mu) & \beta^* \sin(\mu + \Delta\mu) \\ -\frac{1}{\beta^*}\sin(\mu + \Delta\mu) & \cos(\mu + \Delta\mu) \end{bmatrix} \tag{2.9}$$

$$= \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f} & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\mu & \beta_o^* \sin\mu \\ -\frac{1}{\beta_0^*}\sin\mu & \cos\mu \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f} & 1 \end{bmatrix}$$

where $\mu + \Delta\mu$ and β^* are the perturbed lattice functions.

Then

$$\begin{aligned} \cos(\mu + \Delta\mu) &= \cos\mu - \frac{\beta_0^*}{2f}\sin\mu \\ &= \cos\mu - 2\pi\xi \sin\mu \end{aligned} \tag{2.10}$$

where
$$\xi = \frac{N r_o \beta^*}{4\pi\sigma^2\gamma} \quad . \tag{2.11}$$

For small $\Delta\mu$,
$$\xi \simeq \frac{\Delta\mu}{2\pi} = \Delta Q_L \tag{2.12}$$

where ΔQ_L is the linear tune shift due to the beam-beam perturbation.

The corresponding perturbation of the beta function is given by
$$\frac{\Delta\beta^*}{\beta_o^*} = -2\pi\xi \cot \mu \quad . \tag{2.13}$$

2.2 Stability of Linear Incoherent Motion

In the linear approximation, the motion of a test particle in the presence of the other beam is stable if the absolute value of the trace of the one-turn transfer matrix is less than 2

$$\cos \mu - 2\pi\xi \sin \mu \leq 1$$

$$\xi \leq \frac{1}{2\pi} \cot (\mu/2). \tag{2.14}$$

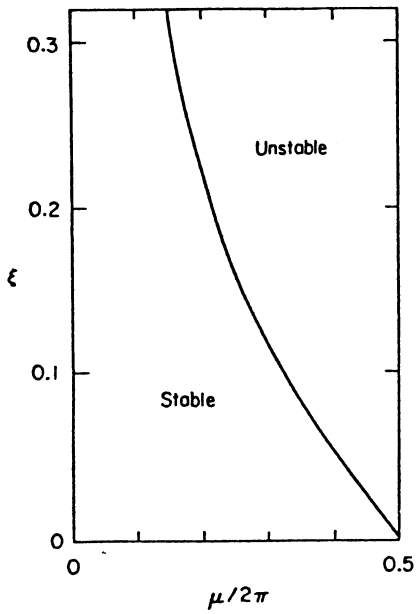


Fig. 2 Stability region for a weak beam executing small oscillations. μ is the betatron phase advance between collision points

The experimentally measured limiting beam-beam parameter in lepton machines is of the order of 0.03 to 0.05 and in the SPS hadron collider it is an order of magnitude lower. The linear model clearly predicts a threshold which is much too high.

2.3 Stability of Coherent Motion

For two beams of similar intensity, if one beam is slightly displaced with respect to the other, coherent oscillations are induced which under certain conditions can lead to instability [8,9]. With one bunch per beam, two modes are possible, the 0-mode where both beams move up and down together, and the π -mode where the two beams move in opposite directions. With m bunches per beam, 2 m modes of oscillation are possible. The stability of the system can still be computed by linear matrix theory. For the case of one and three bunches per beam the stability boundary is shown in Fig. 3[9]. Clearly the threshold is substantially reduced compared with the incoherent case, although for the appropriate choice of working point it is still substantially higher than experimentally observed thresholds.

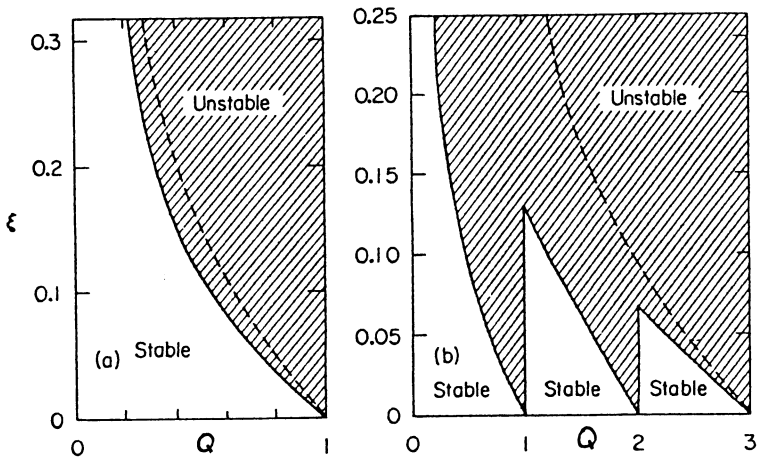


Fig. 3 Stability region for two strong rigid beams executing small center-of-mass oscillations for (a) two colliding bunches and (b) six colliding bunches. The figures are periodic in Q, the total tune of the storage ring; the periods are 1 in (a) and 3 in (b). The dashed lines show the strong-weak stability limit and are reproduced from Fig. 2.

As can be seen from Fig. 1 the beam beam force is intrinsically nonlinear above about 1.5 σ. The nonlinearity has important consequences on the beam dynamics. The effect will be discussed in more detail in Section 5.

2.4 Scaling of the Linear Tune Shift

For the general case of an elliptical beam with Gaussian distribution, the tune shift parameter ξ is given by

$$\xi_{x,y} = \frac{Nr_o e}{2\pi\gamma\ (\sigma_x + \sigma_y)} \cdot \frac{\beta^*}{\sigma_{x,y}} \tag{2.15}$$

where β^* is the value of the beta function at the interaction point. This parameter scales differently with energy for hadron and lepton machines.

In a hadron machine the normalised emittance ($\varepsilon\beta\gamma$) is a conserved quantity. Then putting

$$\sigma^2 = \frac{(\varepsilon\beta\gamma)\ \beta^*}{4\gamma}$$

and assuming a round beam ($\sigma_x = \sigma_y$), then

$$\xi = \frac{Nr_p}{\pi(\varepsilon\beta\gamma)} \tag{2.16}$$

which is independent of both γ and of the value of β^* at the crossing point.

For lepton machines the situation is different because the equilibrium beam size scales proportional to γ. The tune shift parameter then has a strong γ^{-3} dependence. Consequently, in a machine like LEP where the beams will be injected at low energy it is vital to separate the beams completely during injection and throughout acceleration. In addition, the beams are normally flat at the collision points so the tune shift parameter also depends on the beta functions at the interaction point.

2.5 Nonlinear Beam-Beam Interaction

In the preceding sections, only the linear part of the beam-beam force was considered. In fact, the beam-beam interaction is an intrinsically nonlinear phenomenon and this nonlinearity gives rise to two effects. Firstly, it introduces a dispersion of the tune with amplitude, the beam-beam tune spread. Secondly, the nonlinear kick together with the localised nature of the interaction drives nonlinear resonances wherever the machine tunes satisfy a relationship of the form

$$mQ_x + nQ_y = \text{Integer}$$

where m and n are even integers for head-on collisions between bunches.

These nonlinear resonances can profoundly influence the topology of the phase space. There is a great deal of evidence, both experimental and from computer simulations, that shows that these resonances play an important rôle in determining the nature of the beam-beam interaction.

In the next two sections, some experimental data from lepton and hadron machines is discussed. The physical manifestation of the phenomenon turns out to be quite different in the two types of machine.

Lepton machines are in some ways both simpler and more complex to understand than hadron machines. The tune spread is normally at least an order of magnitude larger than in hadron machines, so the beams straddle many nonlinear resonances and their synchrotron satellites. However, there is a strong damping mechanism through synchrotron radiation emission to counteract the beam-beam interaction, giving rise to an equilibrium situation. This equilibrium is generally achieved after a few damping times and renders the problem particularly suitable to computer simulation. In the next section some results of computer simulations are discussed and compared with real machine data.

In hadron machines the tune spread is sufficiently small that the beams can be kept clear of low-order resonances. However, as the experimental data will show, resonances of order 10 or even under some conditions of order 16, have been shown to have a catastrophic effect on beam lifetime at quite modest values of the beam-beam tune shift. There is no, or negligible, radiation damping so an equilibrium situation cannot exist as it does in lepton machines. This makes quantitative predictions using computer simulation difficult, although considerable insight can still be obtained from such simulations. However, for a more complete understanding they must be supplemented by a more detailed analysis of the nature of nonlinear beam-beam resonances.

3. EXPERIMENTAL AND NUMERICAL DATA FROM e^+e^- MACHINES

The most comprehensive compilation known to the authors of data from the world's lepton colliders can be found in Ref. 10. One of the most striking features of this data is shown in Fig. 3 of this reference, reproduced below for convenience.

Figure 4 shows the luminosity versus current observed in seven lepton colliders. The luminosity L is given by

$$L = \frac{I^2}{4\pi M e^2 f\ \sigma_x \sigma_y} \qquad (3.1)$$

where f is the revolution frequency, M the number of bunches per beam, I the current per beam (assumed equal in the two beams) and $\sigma_{x,y}$ are the standard deviations of the beam sizes at the crossing point.

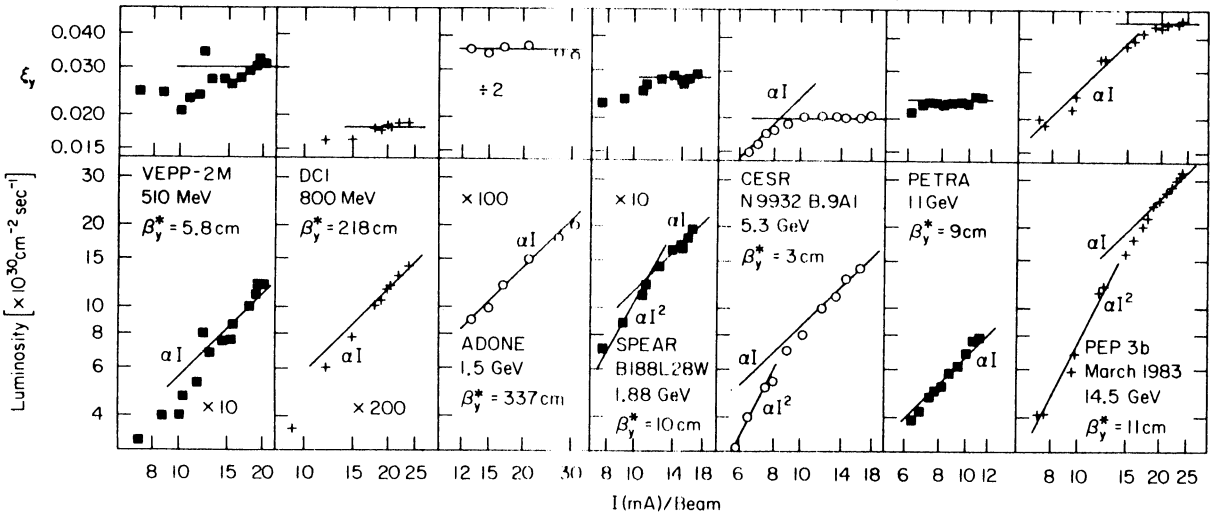


Fig. 4 Luminosity and vertical tune shift parameter versus current for seven electron-positron colliders[10]. Note that the tune shift saturates at some current value above which the luminosity grows linearly.

The behaviour of all seven machines is remarkably similar. At low current, the luminosity increases approximately as the current squared, in agreement with equation (3.1) whereas at high current the luminosity is more proportional to I rather than I^2. Even more striking, the upper plots show the calculated vertical linear beam-beam tune shift parameter ξ_y as a function of current. For a flat beam ($\sigma_y << \sigma_x$) ξ_y is given by

$$\xi_y = \frac{Ir_o\beta^*}{2\pi Mef\gamma\sigma_x\sigma_y} \qquad (3.2)$$

It can be seen that at high current, instead of the linear dependence on current predicted from equation 3.2, the linear tune shift saturates at some value between 0.02 and 0.05. This limiting tune shift is commonly called the 'beam-beam limit'.

In order for σ_y to be constant, the product $\sigma_x\sigma_y$ must increase linearly with current. In practice, the horizontal beam size is observed to change very little and therefore, at the beam-beam limit the vertical beam size must grow linearly with current. This is an important difference between lepton and hadron machines. In e^+e^- machines, at a given current an equilibrium distribution exists which is a balance

between the heating of the beam due to quantum fluctuations and the beam-beam interaction and the cooling due to synchrotron radiation damping. The beam size has been observed to blow up by as much as a factor of five before the lifetime is affected [11].

The fact that an equilibrium distribution is established in a few damping times (10^3-10^4 turns) makes the beam-beam problem in e^+e^- machines particularly amenable to computer simulation.

3.1 Computer Models of the Beam-Beam Interaction

Computer simulations of the beam-beam interaction have been made in practically all laboratories in which e^+e^- machines exist. The types of simulation can be divided into two main classes, strong-weak and strong-strong.

In strong-weak simulations test particles in the 'weak' beam are tracked through a linear lattice followed by nonlinear beam-beam kicks due to the 'strong' beam which itself is not perturbed by its interaction with the weak beam. Therefore the beam-beam kicks can generally be precalculated and stored in a look-up table with appropriate interpolation. This method is economic in computer time but lacks quantitative predictive power. However, this kind of simulation is useful for studying the beam-beam interaction in, for example, the SPS collider.

More sophisticated simulations treat the strong-strong case, where the changes in beam size and distribution are periodically used to recompute the beam-beam kicks as the calculation proceeds. This kind of simulation can be used to compute the final equilibrium beam size and can quantitatively predict the machine performances (luminosity, beam-beam limit etc.) when the relevant physics is introduced into the problem.

Many different effects can be included in such simulations. In general, as well as the transverse dynamics the longitudinal motion must be taken into account. The synchrotron motion results in a modulation of the arrival time of a particle at the interaction point producing a modulation of the strength of the beam-beam force. This results in the generation of beam-beam synchro-betatron resonances (Section 5) and can sometimes strongly affect the performance[12].

Quantum fluctuations and synchrotron radiation damping are also taken into account. In addition, small errors such as the variation in phase advance between intersection points, small offsets between the beams and spurious dispersion can all have non-negligible effects.

Figure 5 shows one result from a strong-strong simulation at CESR [13], where contours of constant relative luminosity are plotted in the tune space. Areas marked with crosses indicate regions of bad lifetime. These regions are closely correlated to the location of nonlinear resonance lines.

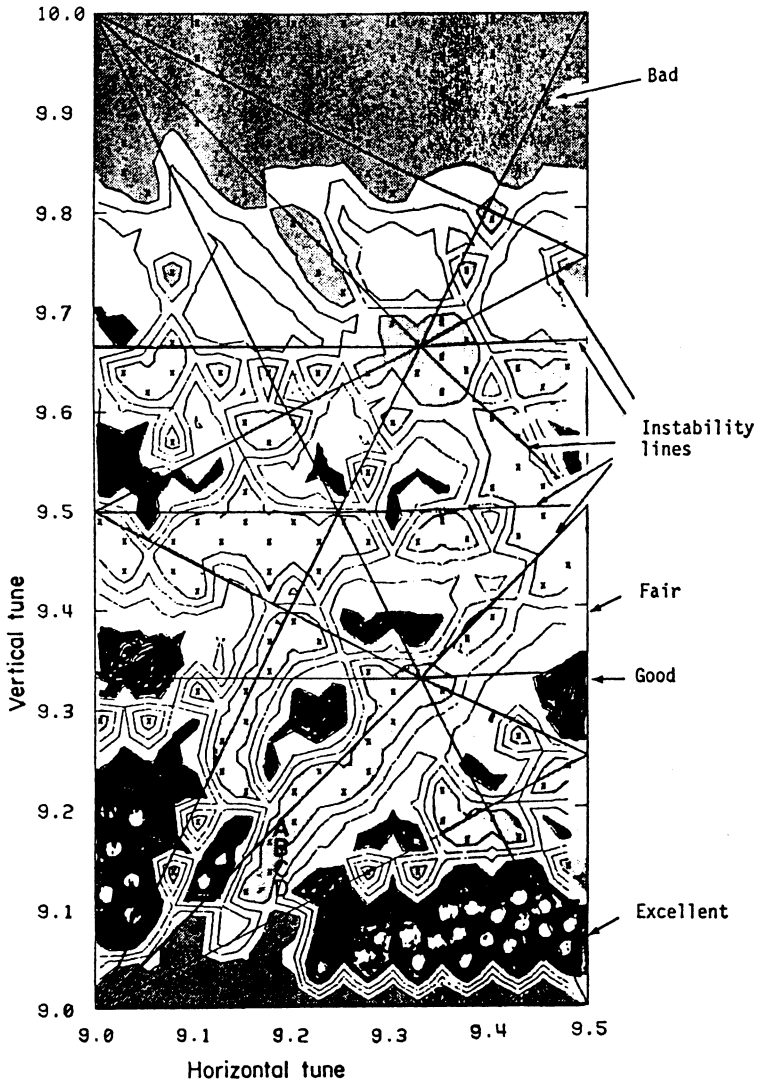


Fig. 5 Beam-beam simulation results for CESR[13]. The contours are at equally spaced relative levels of luminosity. Crosses indicate bad lifetime. The straight lines define the positions of strong nonlinear resonances.

Similar results have been obtained in simulations of LEP [12]. Figure 6 shows the computed luminosity as a function of the vertical tune over a wide range. The regions of low luminosity are again strongly correlated to large azimuthal Fourier harmonics of nonlinear resonances. Note that these simulations have real quantitative predictive value.

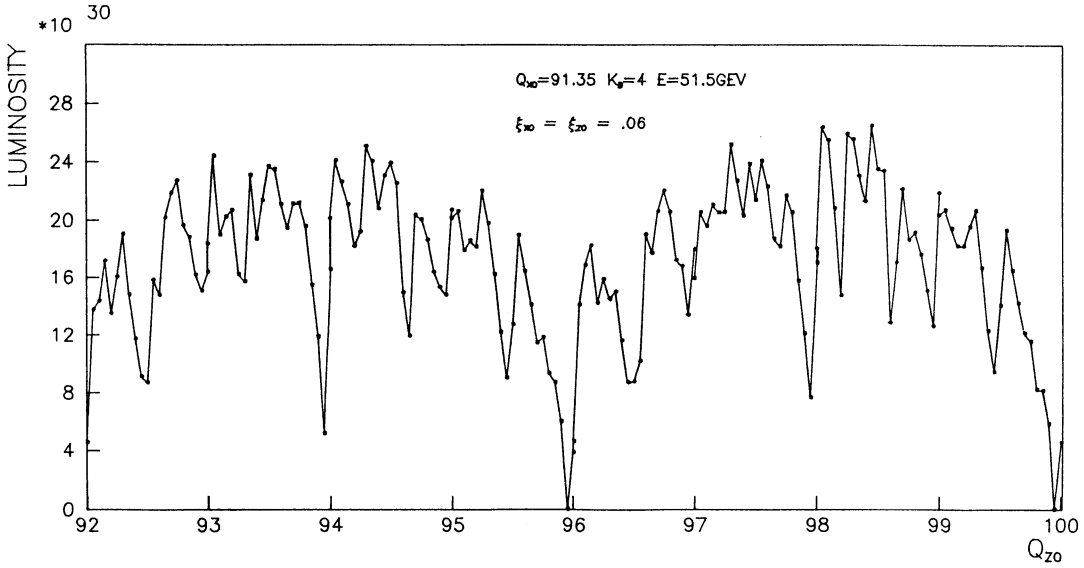


Fig.6 Luminosity ($cm^{-2}.s^{-1}$) as a function of the vertical tune over a wide range computed for LEP[12]. The regions of low luminosity coincide with strong nonlinear resonances.

Figure 7 shows a similar simulation for Petra[14], where the equilibrium vertical beam height is plotted as a function of both radial and vertical tunes. The curves on the left correspond to a perfect machine. Again, at tune values corresponding to nonlinear resonances the beam height increases. The situation is much worse when small machine imperfections are added, in this case small variations in vertical tune between interaction points and spurious vertical dispersion. The result of these imperfections is to excite azimuthal Fourier components of the nonlinear beam-beam force driving resonances which would not normally be present due to the symmetry of the system. The curves on the right show how the number and strength of the resonances increases when imperfections are added.

The predictive value of beam-beam simulations for lepton machines is illustrated in Fig. 8. These curves were generated using the LEP simulation code[12] modified for PEP[15]. The top curve corresponds to the normal PEP working point Q_H = 20.175, Q_V = 25.275. The experimentally measured luminosity shows good agreement with the simulation. The program was then used to probe the tune space in search of a better working point, which was found to be Q_H = 21.275, Q_V = 18.175. When the machine was tuned to this new working point (fig. 8b) the luminosity increased as predicted by the simulation, by about 40%.

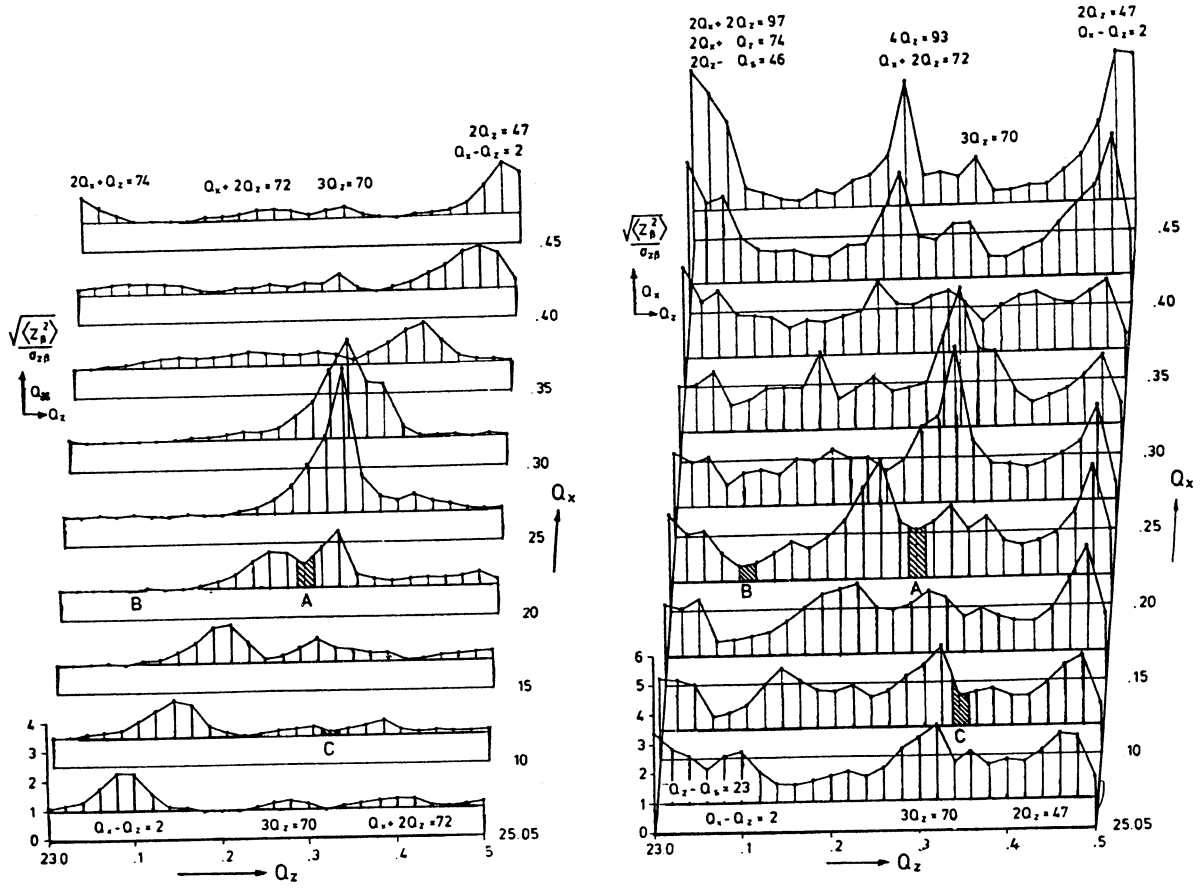
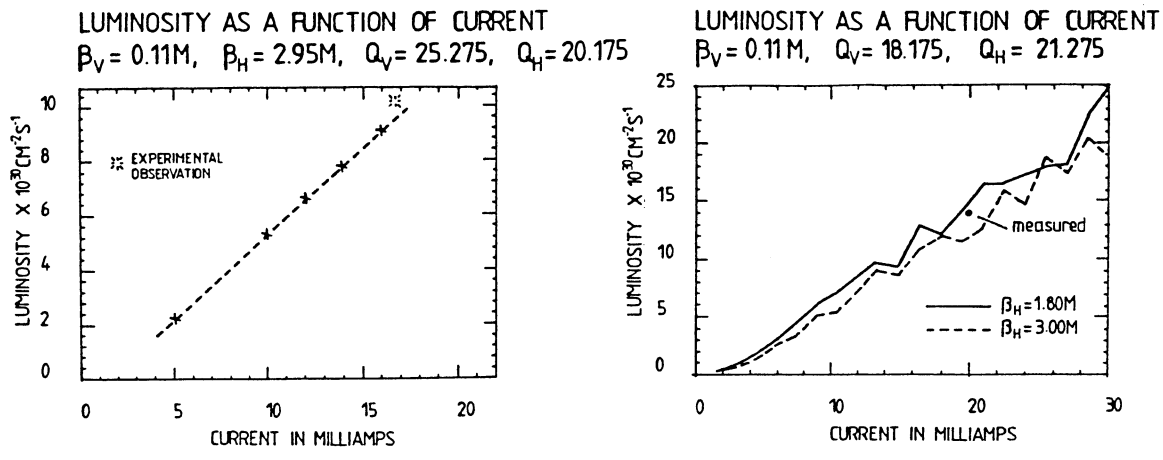


Fig. 7 Simulated vertical beam height in PETRA[14] as a function of vertical and horizontal tunes a) without machine imperfections and b) with small imperfections.



a) b)

Fig. 8 Luminosity as a function of current predicted for PEP[15] with the LEP simulation code[12] and compared with experimental observations. The better working point b) was predicted by the simulation.

4. EXPERIMENTAL DATA FROM HADRON MACHINES

At the time of writing, only two hadron colliders have operated, the ISR and the SPS proton-antiproton colliders at CERN, soon to be followed by the Tevatron proton-antiproton collider at Fermilab.

In the ISR the beams were debunched and crossed horizontally at an angle so that in the horizontal plane there was almost no tune shift as a particle was kicked first one way and then the other as it crossed the opposing beam. The only substantial tune shift was in the vertical plane, and this was much smaller ($\sim 4\times10^{-4}$) than that obtained in the SPS ($\sim 4\times10^{-3}$). The ISR stacked beams on a working line which straddled 7th , 8th and 9th-order resonances (Fig. 9) with very little effect due to the beam-beam interaction.

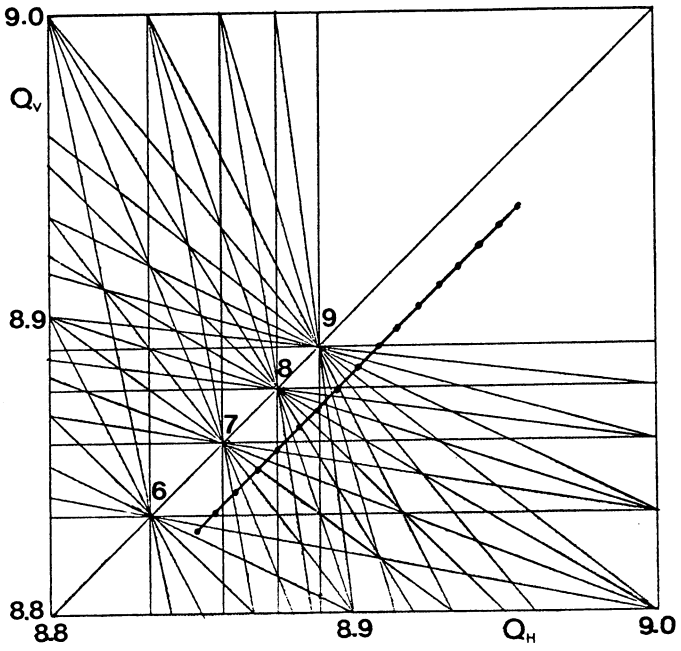


Fig. 9 A typical ISR working line crossing 7th, 8th and 9th-order resonances.

However, the presence of nonlinear resonances could be detected by exciting the beam and measuring the response (beam transfer function)[16]. Figure 10 shows such a measurement on a stack, first with no beam in the other ring and then with successively increasing current. Depletion of the density distribution can be observed at resonant tune values. The amplitude dependence of the tune spread has a stabilizing influence. As the amplitude of a particle increases due to the resonances the tune changes to push the particle off resonance.

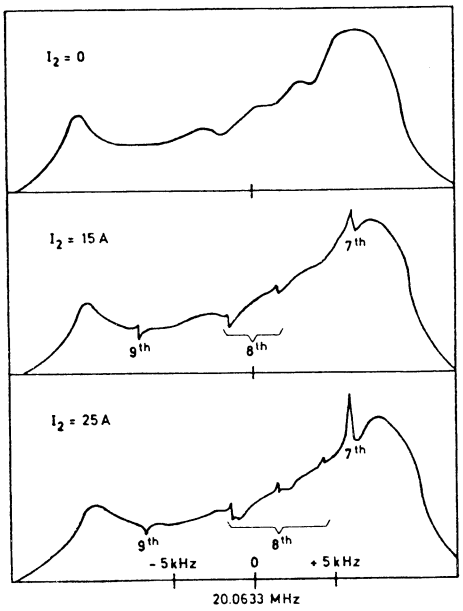


Fig. 10 Vertical beam transfer function measurement of the amplitude distribution of beam 1 as a function of the current in beam 2, showing the beam-beam resonances of order 7 to 9.

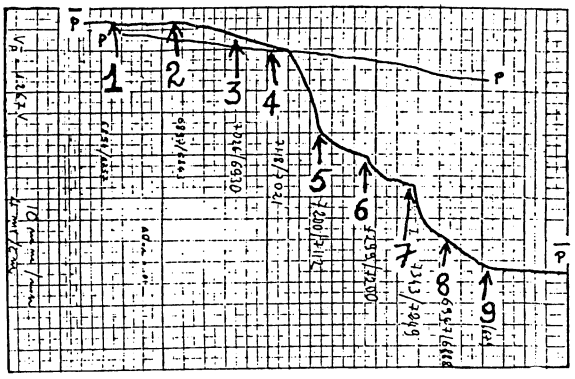
The SPS proton-antiproton collider is more similar to an electron-positron storage ring. The beams are bunched and collisions are head-on, giving approximately the same beam-beam tune shifts in the two planes.

Very strong beam-beam effects are observed. Figure 11 shows a scan of the tune diagram with three proton bunches and a single weak antiproton bunch (6 crossings per revolution) and with a beam-beam tune shift on the antiprotons of 3×10^{-3} per crossing[3]. Figure 11a) shows a chart recorder output of the intensity of one of the proton bunches together with the antiproton bunch on a very sensitive scale. Figure 11b) shows the tune diagram between 3rd and 4th-order resonances, where 7th, 10th and 11th-order resonances are indicated.

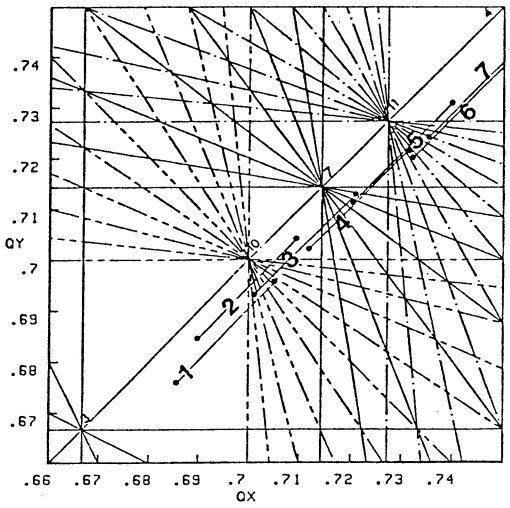
The intensity decay rate was measured at different positions in the working diagram, indicated by the lines marked 1,2 etc. The meaning of these lines is the following. The proton bunch (with which the tune is measured) has negligible spread and can be considered to occupy a point in the working diagram indicated by the lower point on each line. The small-amplitude antiprotons which experience the linear part of the beam-beam force, are shifted upwards in tune by the total beam-beam spread and occupy a point at the other end of each line. Large amplitude antiprotons occupy much of the space between.

The decay rate of the antiprotons is extremely sensitive to the tune, increasing rapidly as the antiprotons touch the 10th-order resonances. In contrast, the proton decay rate under these conditions where proton and antiproton emittances were comparable, was quite insensitive to the tune. In order to have reasonable operating conditions for physics data taking, the tune must be restricted to a very small region of the working diagram corresponding to that of point 1 in the figure. Therefore the beam-beam interaction imposes severe constraints on machine performance.

One way to reduce the tune spread in the beams is to separate the protons and antiprotons at the unwanted collision points (the SPS contains only two experimental areas at adjacent long straight sections Such a separation is possible in the SPS by making global orbit deformations in the opposite sense for protons and antiprotons using a set of four



a)



b)

Fig. 11 Scan of the SPS tune space between 3rd and 4th order resonances with a single weak antiproton bunch and three strong proton bunches.

electrostatic deflectors[17]. When the separation is switched on the beam lifetime improves considerably (Fig. 12).

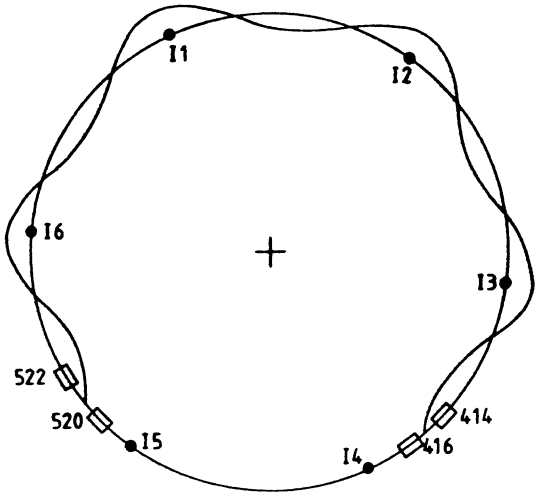


Fig. 12a Schematic diagram of SPS separation scheme. The proton and antiproton orbits are deformed in opposite directions.

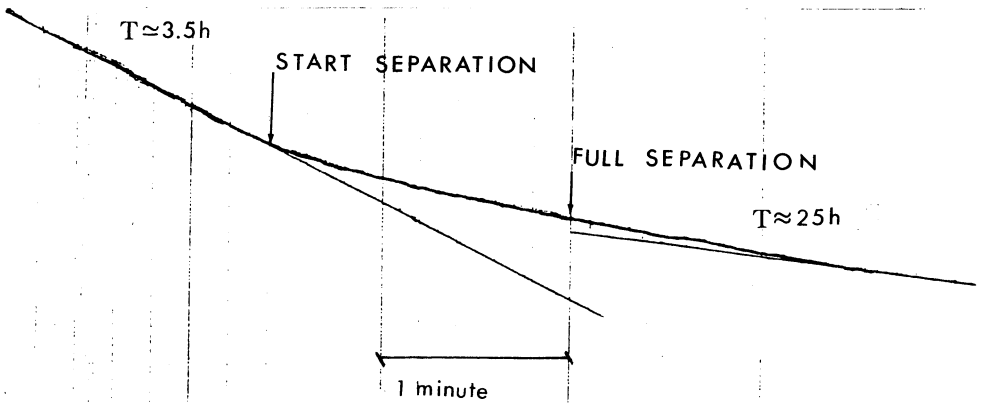


Fig. 12b Intensity decay of an antiproton bunch as the separation is brought up.

Another important effect observed at the SPS is the self-scraping of large emittance particles when the two beams have unequal emittances. Figure 13 shows the decay rates of three antiproton bunches which were injected with successively bigger emittance. The effect on the decay rates can be easily seen. Antiprotons whose amplitude exceeds the average dimensions of the proton beam are rapidly peeled off and the decay rate is

initially high. As a result the antiproton emittance shrinks during the early phase of storage. The 'dynamic acceptance' of the machine in the presence of the beam-beam interaction is therefore not much more than the strong beam emittance.

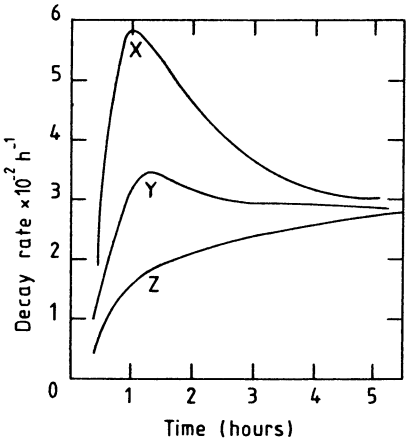


Fig. 13 Decay rates of three antiproton bunches with different emittances. The normalized emittances $E = \varepsilon\beta\gamma/\pi$ where $E_x = 17$, $E_y = 15$, $E_z = 12$. The proton emittance $E_p = 16$ and $\xi = .004$.

5. NONLINEAR BEAM-BEAM RESONANCES

The theory of nonlinear resonances in circular accelerators has been treated elsewhere in these proceedings[4, 5]. Before applying the theory to the beam-beam interaction, some of its general features will be briefly reviewed. Detailed derivations are not given here. For these, the reader is referred to Refs. 1-3 and 18-20.

The motion of a test particle in the presence of a nonlinear perturbation is governed by the perturbed Hamiltonian

$$H = H_o + H_1$$

$$= \tfrac{1}{2}(p_x^2 + p_y^2 + K_x(s)x^2 + Ky(s)y^2) + V(x,y)\delta(s-s_o) \qquad (5.1)$$

where the perturbing potential $V(x,y)$ due to some nonlinear element at azimuthal location $s = s_o$ is considered to be small. The "time" dependence of the nonlinear kick is introduced through the periodic δ-function having period 2π. Note that instead of the time coordinate the azimuthal position s is used.

The general method of solving such a problem[2,18] is first to make two successive canonical transformations, a Courant and Snyder transformation to remove the time dependence of the unperturbed Hamiltonian H_o followed by an action-angle transformation. The periodic δ function is also replaced by a Fourier series expansion.

The next step, for the case of a single multipole is to isolate the 'slowly varying' part of the Hamiltonian when the tune is close to a 'resonant' value $nQ - p \approx 0$. One finally arrives at a 'resonant invariant' (see E.J.N. Wilson's chapter for a detailed derivation) of the form

$$K = (Q - p/n)J + A_n J^{n/2} + B_n J^{n/2} \cos \psi \qquad (5.2)$$

where J is the action variable (Q/2 times the emittance of the orbit) and ψ is the 'slow' phase. The first term in the above expression is the distance of the tune from the resonant tune. The second term corresponds to a variation of tune with amplitude, the nonlinear detuning, and the third term is the 'resonance excitation' term.

The same procedure can be carried out for the case of the beam-beam interaction. However, in this case care must be taken to isolate all the slowly varying terms in the resonant Hamiltonian. These terms come about because the beam-beam force can be decomposed into an infinite series of multipoles. The resonant invariant is of the form

$$K = (Q-p/n)\,\alpha + \xi\, u(\alpha) + \xi\, V_n(\alpha) \cos n\psi \qquad (5.3)$$

where the action variable α has been normalised such that $\alpha = \varepsilon\beta/2\sigma^2$.

The tune shift due to the resonances is given by

$$\frac{\partial\psi}{\partial\theta} = \frac{\partial K}{\partial\alpha}$$

$$= (Q-p/n) + \xi\, u'(\alpha) + \xi\, V_n'(\alpha) \cos n\psi \ . \qquad (5.4)$$

Here, $\xi u'(\alpha)$ is the amplitude dependent tune shift, the 'nonlinear detuning' and $\xi V_n'(\alpha)$ is the 'resonance width'. Note that the linear tune shift ξ appears as a scaling parameter.

The functions can be expressed as infinite series[3]:

$$u'(\alpha) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}\,(2m)!\,\alpha^{m-1}}{2^{2m-1}\,(m!)^3} \qquad (5.5)$$

$$V_n'(\alpha) = \sum_{m = n/2}^{\infty} \frac{(-1)^{m-1}\, \alpha^{m-1}\, (2m)!}{2^{2m-2}\, m!\, (m + \frac{n}{2})!(m - \frac{n}{2})!} \quad . \qquad (5.6)$$

Alternatively they can be expressed in terms of modified Bessel functions[21]:

$$u'(\alpha) = \frac{2}{\alpha} [1 - e^{-\alpha/2} I_o(\alpha/2)]$$

$$V_n'(\alpha) = (-1)^{\frac{n}{2}+1} \cdot \frac{4}{\alpha} \cdot e^{-\alpha/2}\, I_{\frac{n}{2}}(\alpha/2) \quad .$$

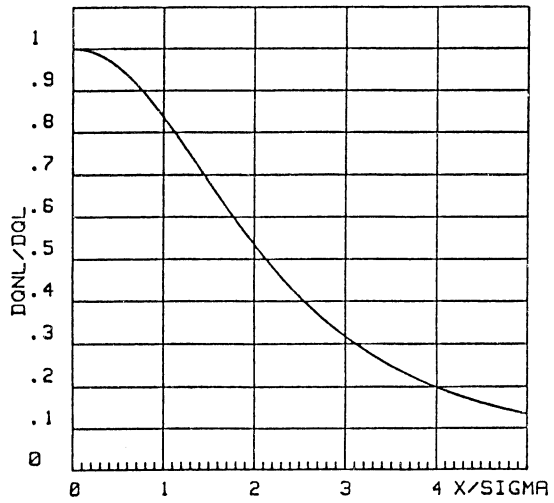


Fig. 14 The nonlinear detuning function u', with $\alpha = x^2/2\sigma^2$

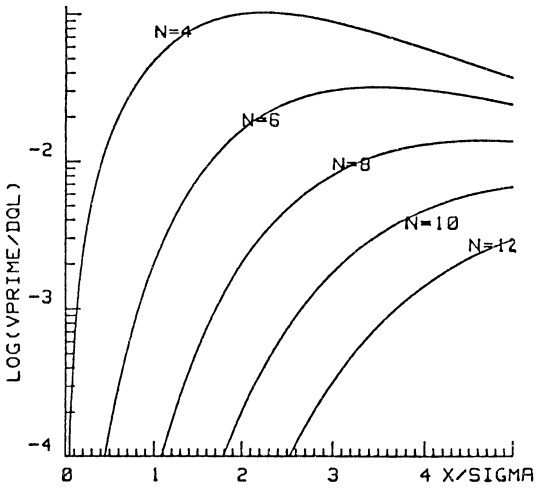


Fig. 15 The resonance width function for resonances of order 4 to 12

In two dimensions the tune shift and resonances width depends on both horizontal and vertical coordinates. The beam occupies a 'footprint' in the tune space which is shown below for the case of a round beam.

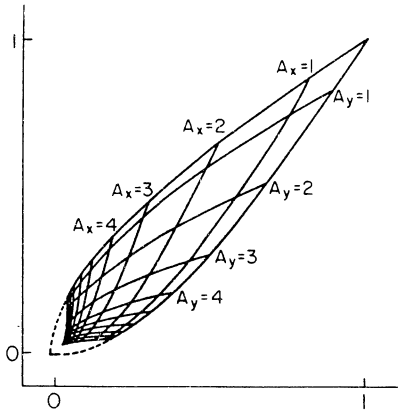


Fig. 16 Nonlinear detuning as a function of x and y amplitudes $A = x/\sigma$ for a round beam

Given the resonant invariant 5.3, particle trajectories can be constructed in the 'slow' phase space (α,ψ). When this is done, it is found that the phase space trajectories are stable to a very high value of the linear tune shift. As an example, Fig. 17 shows the trajectories in the vicinity of the 4th-order resonances for a linear tune shift ξ = 0.04. A characteristic island structure can be observed at an amplitude at which the perturbed tune is approximately equal to the resonant tune.

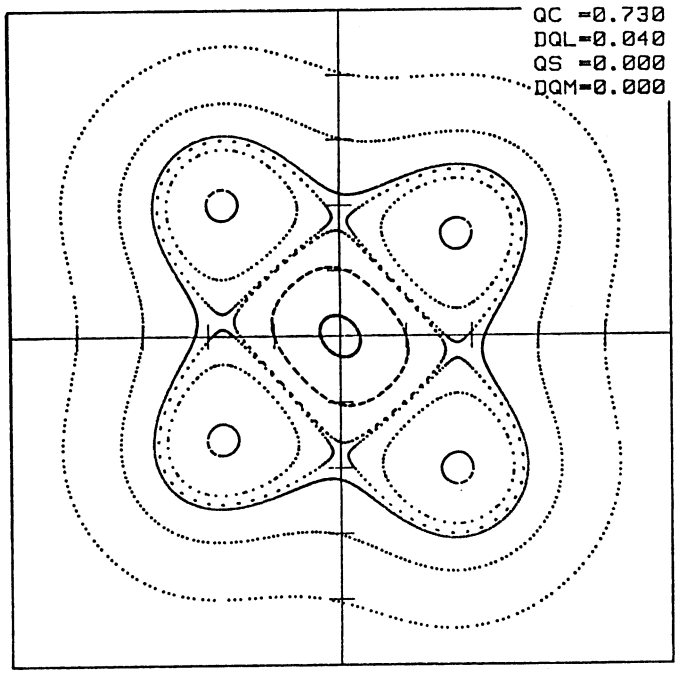


Fig. 17 Phase-space trajectories in the vicinity of a 4th-order beam-beam resonances for ξ = 0.04.

5.1. Resonance Overlap

The single-resonance model is clearly inadequate for explaining the beam-beam interaction due to the fact that increasing the beam-beam tune shift increases the destabilizing effect of the resonance excitation and stabilizing effect of the nonlinear detuning proportionately. However, it was first pointed out by Chirikov[22] that the working area is covered by an infinite number of resonance lines for which the tune value is a rational fraction. Although the high-order resonances have narrow width, the fact that there is an infinite number of them may have a significant effect. If they have sufficient width they overlap in the tune space and particle motion will be unstable even if the working area is free of low-order resonances. In Fig. 18 the particular case of the trajectory in the vicinity of a 6th-order beam-beam resonance is shown at the very

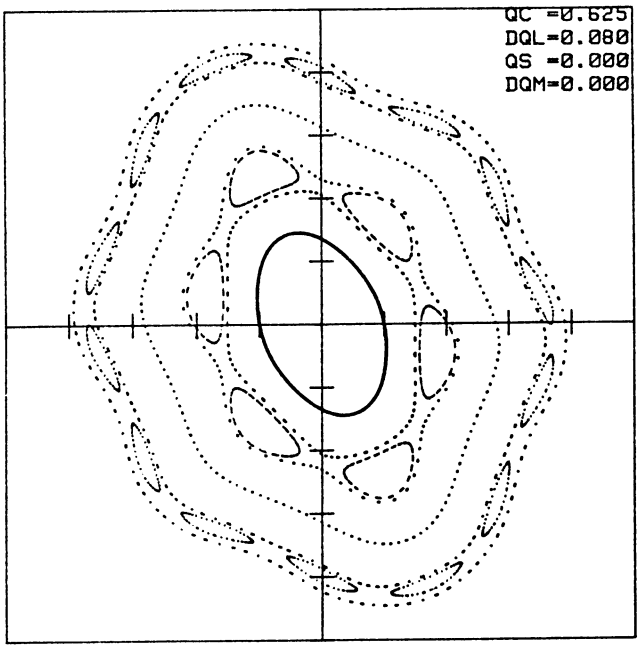


Fig.18 Particle trajectories in the vicinity of a 6th-order beam beam resonance for ξ = 0.08. At large amplitude the characteristic structure of a 14th order resonance can be observed.

high value of the linear tune shift ξm= 0.08. At large amplitude another resonance of order 14 can be observed. If the tune shift increases even further, the two sets of islands will approach each other, and other high-order resonances will appear, finally leading to chaotic motion. However, the linear tune shift required for this phenomenon to occur is still much higher than observed experimentally.

5.2 Synchrobetatron Resonances

The fact that the beam is bunched can have a profound influence on the topology of the transverse phase space. Particles performing energy oscillations can experience a modulation of the transverse tunes due to a number of mechanisms. In lepton machines with short bunches this can arise due to the modulation of the arrival time at the interaction point, resulting in a variation of the strength of the beam-beam kick. In hadron machines a more important source is due to the small residual chromaticity or a small ripple on the quadrupole power supply. Synchrobetatron resonances are also excited by two beams crossing at an angle [23, 24].

Tune modulation at frequency $Q_m = f_m/f_r$ results in the splitting of a nonlinear resonance $n_1Q_H+n_2Q_V=p$ into an infinite number of sidebands

$$n_1Q_H + n_2Q_V = p + k\,Q_m.$$

For a one-dimensional resonance of order n, the sidebands are separated by Qm/n and reduced in strength by the factor $J_k(n\hat{Q}/Qm)$, where $\hat{Q}$ is the amplitude of the modulation. For small Q_m, these sidebands are very close together and can give rise to resonance overlap and stochastic behaviour [20, 25] at a much lower threshold than in the static case, as we shall see below.

5.3 Computer Simulation

Although the resonance invariant (5.3) can be modified to take into account the synchrotron motion, a simple and powerful method of obtaining phase space trajectories is through particle tracking. For a round Gaussian beam the nonlinear beam-beam kick is given by

$$\Delta x' = \frac{8\pi\xi\sigma^2 x}{\beta^* r^2}\left(1 - e^{-r^2/2\sigma^2}\right)$$
$$\Delta y' = \frac{8\pi\xi\sigma^2 y}{\beta^* r^2}\left(1 - e^{-r^2/2\sigma^2}\right) \qquad (5.7)$$

with $r^2 = x^2 + y^2$.

Transforming to new variables $\bar{x} = x/\sigma$, $\bar{x}' = \beta^* x'/\sigma$ we get the position and angle of a particle on turn n + 1 from its coordinates on turn n

$$\bar{X}_{n+1} = \bar{X}_n \cos 2\pi Q_{xn} + \bar{X}'_n \sin 2\pi Q_{xn}$$
$$\bar{X}'_{n+1} = \bar{X}_n \sin 2\pi Q_{xn} + \bar{X}'_n \cos 2\pi Q_{xn} - \Delta\bar{X}_{n+1}$$
$$\bar{Y}_{n+1} = \bar{Y}_n \cos 2\pi Q_{yn} + \bar{Y}'_n \sin 2\pi Q_{yn}$$
$$\bar{Y}'_{n+1} = \bar{Y}_n \sin 2\pi Q_{yn} + \bar{Y}'_n \cos 2\pi Q_{yn} - \Delta\bar{Y}_{n+1}$$
$$Q_{n+1} = Q_o + \hat{Q} \sin (2\pi Q_m n) .$$

This kind of transformation is called nonlinear mapping. The problem of the stability of such maps is one of considerable current interest in a wide range of disciplines outside the field of particle accelerators.

Now in order to observe the sidebands due to tune modulation the correct timescale for the problem must be chosen. Figure 19 shows a plot of the phase space for an unperturbed tune of 0.7 with a linear beam-beam tune shift ξ=0.01, and a tune modulation frequency Q_m=0.004. Here the phase-space coordinates of a particle have been plotted once per

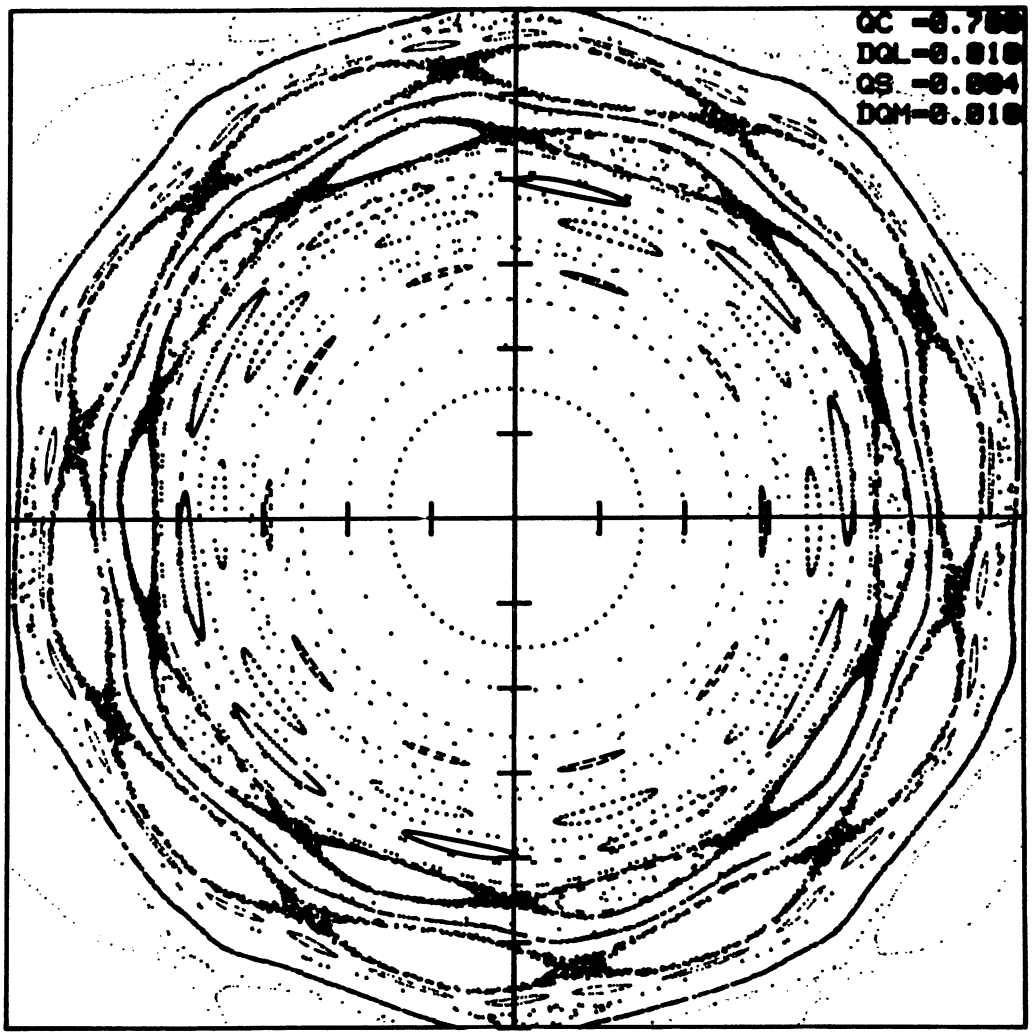


Fig. 19 Phase-space trajectories for an unperturbed tune of 0.7 with ξ = 0.01, modulation tune Q_m = 0.004 and amplitude Q = 0.01. Synchrotron sidebands of a 10th-order resonances of order 2 (outer) to 8 are visible at amplitudes corresponding to tune values $0.7 + nQ_m/10$. The 7th sideband is not visible because the Bessel function goes through zero.

synchrotron period, revealing many sidebands of the 10th-order resonance. In this plot, sidebands 2,3,4,5,6 and 8 can be identified. The 7th sideband is not present because the Bessel function $J_7(10\hat{Q}/Q_m)$ goes through zero for the parameters chosen. In this example the islands are well separated and the phase space is stable.

5.4 Stochastic Threshold

The presence of synchrotron satellites enormously increases the density of resonances thereby reducing the Chirikov threshold for

resonance overlap and stochastic behaviour. This threshold can be computed from the resonant invariant by equating the island widths to their separation. The threshold linear tune shift for stochastic behaviour is given by[3, 26]:

$$\xi = \frac{Q_m}{4n} \sqrt{\frac{1}{U''(\alpha)\, V_n(\alpha)\, J_k(n\hat{Q}/Q_m)}} \quad . \tag{5.7}$$

Figure 20 shows the ratio Q_m/ξ as a function of amplitude for the case of a 10th-order resonance.

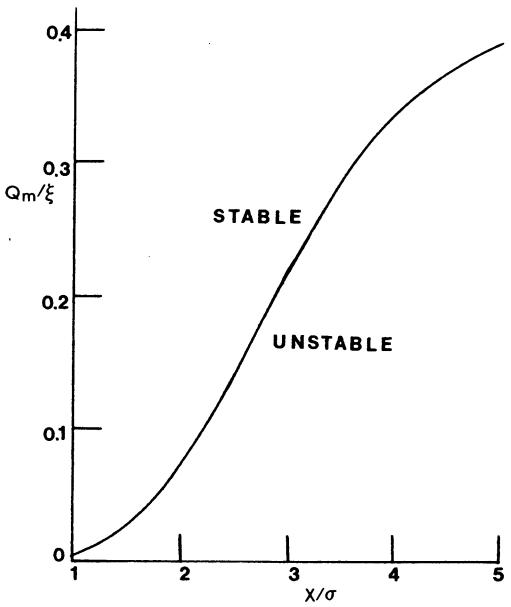


Fig. 20 Stochastic threshold as a function of amplitude computed for a 10th - order beam-beam resonance.

It is of some interest to compare this analytically computed threshold with the results of a computer simulation. Figure 21 shows the phase space trajectory of a particle with initial amplitude of 4σ in the vicinity of a 10th order resonance and with the very modest beam-beam tune shift parameter $\xi = 1.5 \times 10^{-3}$ and a tune modulation amplitude $\hat{Q}$ of only 4×10^{-4}. The theory predicts that the phase space should be stable when the tune modulation frequency is higher than 5.1×10^{-4}. This agrees quite well with the computer experiment.

The most dominant sources of tune modulation in the SPS collider are due to the unavoidable small residual chromaticity combined with the synchrotron motion and current ripple on the main power supplies. This second source is particularly dangerous because it is at low frequency and a great deal of effort has gone into eliminating it up to a point where

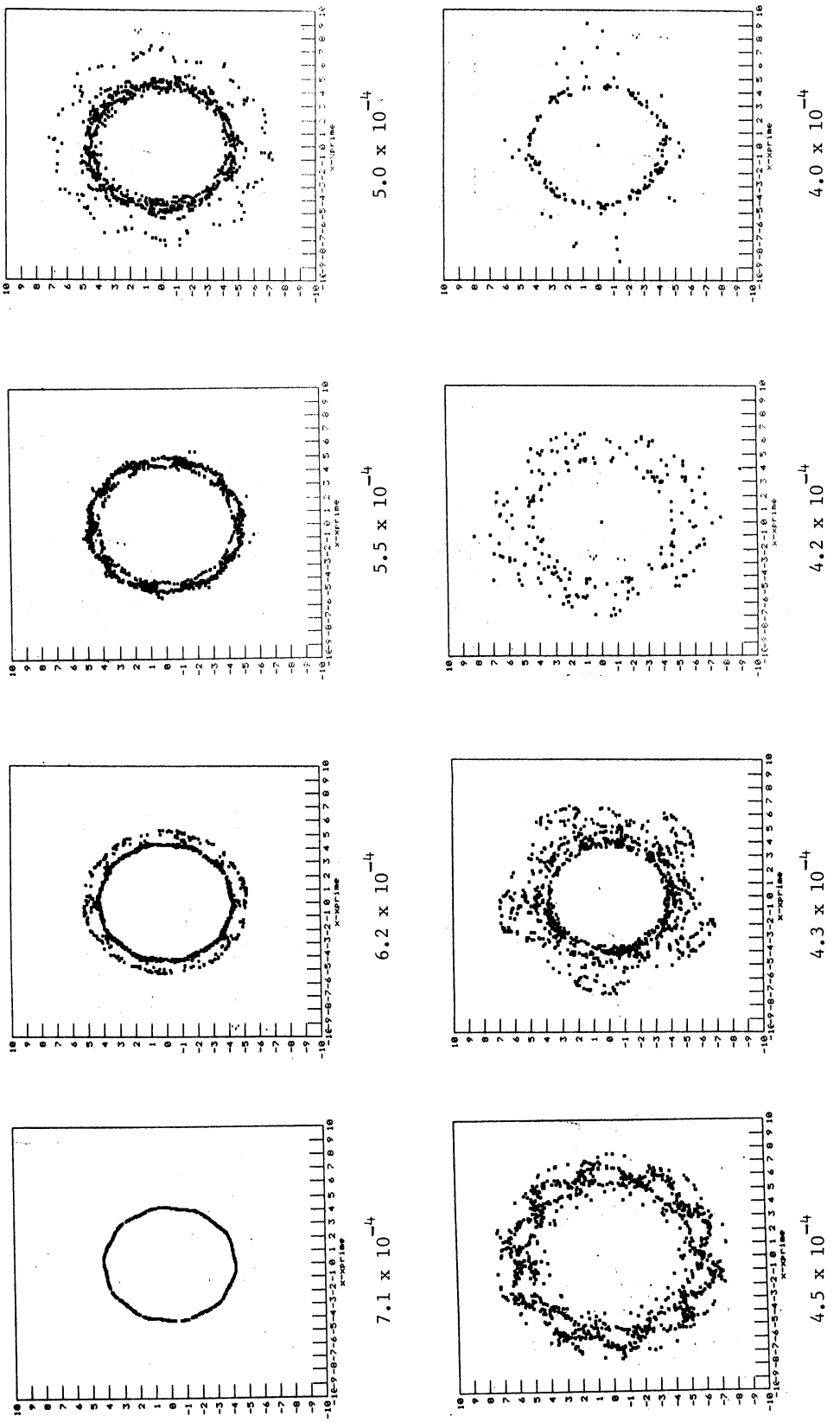


Fig. 21 Phase-space trajectories for $\xi = 1.5 \times 10^{-3}$ and $x/\sigma = 4$ as the tune modulation frequency is varied. The stochastic threshold predicted by the analytic theory is at $Q_m = 5.1 \times 10^{-4}$.

the residual modulation is practically unmeasurable on the sensitive Schottky system used to monitor the machine tunes ($\hat{Q} \approx 4 \times 10^{-4}$). This frequency dependence of the stochastic threshold may also have important consequences for the very big hadron colliders under consideration at the present time, where the synchrotron frequency is low.

6. BEAM DISRUPTION

This is an extreme form of the beam-beam interaction which will be of considerable importance in single pass linear colliders like the SLC[6]. For two beams of different sign (e^+e^-) the electromagnetic fields due to the beam-beam interaction produce a 'pinch' effect, where both beams are focused. Figure 22 [27] shows a computer simulation of this effect.

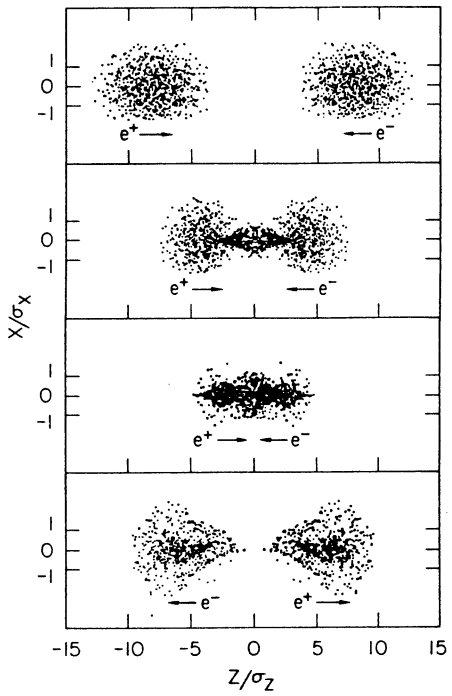


Fig. 22 Pinch effect due to colliding bunches of electrons and positions[26].

The magnitude of the effect is normally parameterized in terms of a disruption parameter D, defined as the ratio of the r.m.s. bunch length to the focal length.

Now
$$\frac{1}{f} = \frac{nr_e}{\sigma_x^2 \gamma} \tag{6.1}$$

so
$$D = \frac{nr_e\sigma_z}{\gamma\sigma_x^2} \quad . \tag{6.2}$$

Or, in terms of the beam-beam tune shift ξ

$$D = 4\pi\xi \frac{\sigma_z}{\beta^*} \quad . \tag{6.3}$$

Lepton storage rings generally operate with $\sigma_z/\beta^* \simeq 1$ and the $\xi \simeq 0.05$, giving an effective maximum disruption parameter for a storage ring of the order of 0.6.

In linear colliders it is of interest to operate at a substantially larger value of the disruption parameter because the pinching of the beams can result in a substantially higher luminosity. Figure 23 shows the luminosity gain as a function of D computed for the case of the SLC. For $D \sim 5$ this simulation predicts a factor of 6 improvement in luminosity. For values of D above about 10 the luminosity gain starts to drop off as the beams pinch each other so strongly as to start to defocus each other within the length of a bunch.

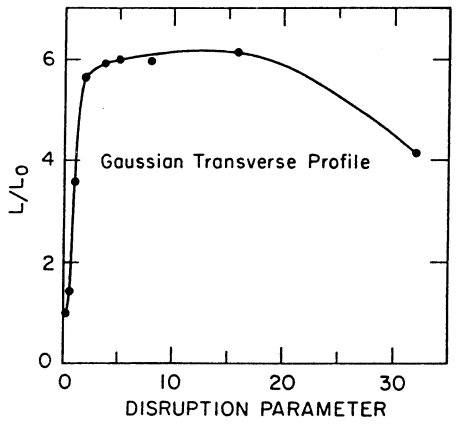


Fig. 23 Luminosity gain in an e^+e^- linear collider as a function of the disruption parameter D[27]. For large D the beams are so strongly disrupted that the luminosity falls off.

7. CONCLUSIONS

Over the last ten years or so, a great deal of effort has gone into trying to understand the details of the beam-beam interaction. For lepton machines, computer simulation has proved to be a powerful tool. Simulation codes have now been developed to the point where they have real predictive value.

For hadron machines the situation is less satisfactory. Although it has not been possible to produce a quantitative predictive model,

analytical calculations supported by computer simulations have shown that synchrobetatron resonances can reduce the threshold for stochastic beam behaviour to a level where the beam-beam interaction has been shown experimentally to play an important role. In future hadron colliders the effect of the low synchrotron tune and the requirement of a non-zero crossing angle at the collision points will have to be given serious consideration.

In linear colliders a new beam-beam effect should manifest itself. Hopefully the SLC will manage to get into a range where the physics of this effect can be investigated experimentally.

* * *

REFERENCES

1] A. Schoch, CERN 57-21, (1958).
2] G. Guignard, CERN 78-11, (1978).
3] L.R. Evans, CERN 84-15, p. 319, (1984).
4] J.S. Bell, These Proceedings.
5] E.J.N. Wilson, ibid.
6] SLAC Linear Collider Conceptual Design Report, SLAC-229, (1980).
7] B. Montague, CERN 68-38, (1968).
8] A. Piwinski, Proc. 8th Int. Conf. on High-Energy Accelerators, Geneva, 1971 (CERN, Geneva, 1971), p. 357.
9] A. Chao, AIP Conf. Proceedings, 127, p. 202, (1983).
10] J. Seeman, SLAC-PUB-3825, (1985).
11] A. Piwinski, DESY 83-028, (1983).
12] S. Myers, CERN-ISR/RF/82-06, (1982).
13] S. Peggs and R. Talman, Proc. 11th Int. Conf. on High-Energy Accelerators, Geneva, 1980 (Birkhäuser Basle, 1980), p. 754.
14] A. Piwinski, IEEE Trans. Nucl. Sci., NS-30, p. 2378, (1983).
15] A. Hutton, PEP-Note-375, (1982).
16] G. Guignard, AIP Conf. Proceedings No. 57, p. 69, (1979).
17] L.R. Evans, A. Faugier, R. Schmidt, IEEE Trans. Nucl. Sci. NS-32, p. 2209, (1985).
18] E.D. Courant, R.D. Ruth, W.T. Weng, AIP Conf. Proceedings, 127, p. 294, (1983).
19] A. Jejcic, J. Le Duff, Proc. 8th Int. Conf. on High-Energy Accelerators, p. 354, CERN (1971).
20] J. Tennyson, AIP Conf. Proceedings, 87, p. 345, (1981).
21] M. Month, BNL 19533, (1975).
22] B. Chirikov, Physics Reports, 52, p. 263, (1979).
23] A. Piwinski, IEEE Trans. Nucl. Sci. NS-24, p. 1408, (1977).
24] A. Piwinski, IEEE Trans. Nucl. Sci, NS-32, p. 2240, (1985).
25] E.D. Courant, ISABELLE Tech. Note 163, (1980).
26] S. Peggs, Particle Accelerators, 17, p.11, (1985).
27] R. Hollebeek, Proc. Beam-Beam Interaction Seminar, SLAC-PUB-2624, p 165, (1980).

SYNCHRO-BETATRON RESONANCES

A. Piwinski*)
CERN, Geneva, Switzerland

ABSTRACT

The three most important mechanisms for exciting a coupling between the transverse betatron oscillations and the longitudinal synchrotron oscillations in a synchrotron are described. The first two are single-beam effects arising from dispersion in accelerating cavities and transverse fields which vary over the bunch length. The third mechanism is a beam-beam effect seen in colliders with a finite crossing angle. The general resonance condition requires the sum of integer multiples of the transverse and longitudinal tunes to be an integer and the resonances are called synchro-betatron or satellite resonances. The later name arises since the synchrotron tune is much smaller than the betatron tunes and hence the resonances appear as satellites close to the simpler transverse resonances.

1. INTRODUCTION

Synchro-betatron resonances or satellite resonances are excited when the synchrotron and betatron frequencies satisfy the relation

$$kQ_x + \ell Q_z + mQ_s = n \qquad (1)$$

where k, ℓ, m and n are integers and Q_x, Q_z and Q_s are the betatron and synchrotron frequencies in units of the revolution frequency.

We will discuss here the three most important effects which excite satellite resonances. The first one occurs if there is a dispersion in an accelerating cavity. This effect was first observed in NINA[1] and later on in SPEAR[2,3] and PETRA[4]. The second effect is given by transverse fields which vary in the longitudinal direction over the bunch. The transverse fields can be produced in the cavities by the accelerating voltage[5] or by the bunch itself[6]. The fields produced by the bunch and by the dispersion limit the currents in PETRA[4].

Finally a third effect arises from the beam-beam interaction when there is a crossing angle. This effect has limited the currents in DORIS I [7,8], and it played a role during the design of HERA[9], the ep storage ring now under construction.

---

*) On leave from DESY, Hamburg, Fed. Rep. Germany.

2. DISPERSION IN A CAVITY

For each particle with an arbitrary energy deviation there exists a closed orbit that is given by the product of the dispersion D times the relative momentum deviation or approximately, the relative energy deviation:

$$x_{c.o.}(\ell) = D_x(\ell) \frac{\Delta p}{p} \approx D_x(\ell) \frac{\Delta E}{E} . \tag{2}$$

Around this closed orbit the particles perform betatron oscillations. The dispersion can be horizontal or vertical and we can consider horizontal or vertical betatron oscillations. In the cavity the energy is changed and therefore the closed orbit is shifted (see Fig. 1).

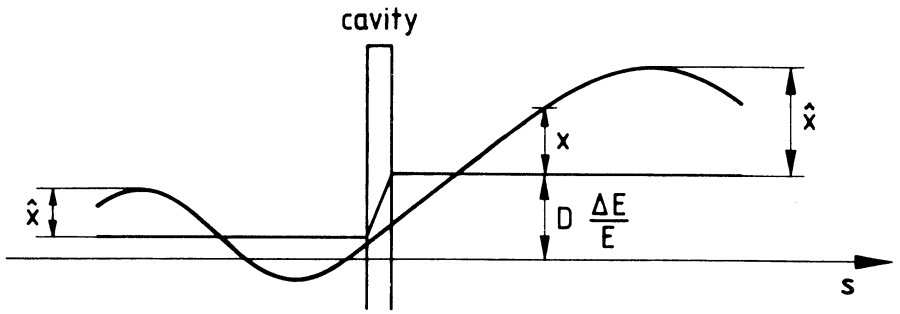


Fig. 1 Change of betatron coordinates in a cavity

Since the total coordinate x cannot be changed in such a short interval, the betatron coordinate is changed:

$$\delta_x = - D_x \frac{\delta E}{E} = - D_x \frac{eU}{E} \left[\sin \left(\phi + \frac{2\pi}{\lambda} s\right) - \sin \phi\right] \tag{3}$$

with e = elementary charge, U = cavity voltage, ϕ = synchronous phase, λ = wavelength, s = longitudinal position.

In a similar way one obtains

$$\delta x' = - D'_x \frac{\delta E}{E} = - D'_x \frac{eU}{E} \left[\sin \left(\phi + \frac{2\pi}{\lambda} s\right) - \sin \phi\right] . \tag{4}$$

We take into account only that part of the energy change which varies with the synchrotron frequency. The other constant part which replaces the radiation losses does not play a role for a resonance.

Equations (3) and (4) show that there is a coupling from the synchrotron oscillation to the betatron oscillation. But the synchrotron oscillation is also influenced by the betatron motion. The synchrotron oscillation can be described in terms of the relative energy deviation $\Delta E/E$ and the longitudinal position s. The longitudinal position is changed in the curved sections of the machine. The change of s per revolution is given by

$$\delta s = \int \frac{1}{\rho} \left(x + D_x \frac{\Delta E}{E}\right) d\ell \tag{5}$$

where ρ is the radius of curvature and $d\ell$ is an element of the length on the design orbit. Evaluation of the integral gives

$$\delta s = A_1 x + A_2 x' - \alpha_M C \frac{\Delta E}{E} \tag{6}$$

with

$$A_1 = -\frac{1}{\beta_x} \left(D_x \sin \mu_x - F_x (1 - \cos \mu_x) + A_2 \beta_x'/2\right)$$

$$A_2 = -D_x (1 - \cos \mu_x) - F_x \sin \mu_x$$

$$F_x = D_x' \beta_x - D_x \beta_x'/2$$

where μ_x = betatron phase advance, β_x = amplitude function, α_M = momentum compaction factor, C = circumference.

x and x' refer to the beginning, D and F to the beginning or the end of the revolution.

Equation (6) shows a coupling from the betatron oscillation to the synchrotron oscillation which is caused by the path lengthening due to the betatron oscillation. If we assume linear betatron and synchrotron oscillations we can find an exact solution for the coupled motion. The phases μ of the eigenvalues of the revolution matrix, on the resonance $Q_x \pm Q_s = n$, are then given by [10]

$$\mu = \mu_{x,s} \pm (1 - \cos \mu_x) \sqrt{\frac{-(D_x^2 + F_x^2)}{\sin \mu_x \, \alpha_M C \beta_x}} \, . \tag{7}$$

Equation (7) shows that an instability occurs, i.e. the amplitudes increase exponentially, when $\sin \mu_x = 0$. In this case we have a difference resonance, since $Q_s = \mu_x/2\pi - n$ is always smaller than 0.5.

The nonlinear satellite resonances, which occur for large synchrotron amplitudes, can be investigated with the help of approximation methods[10]. Only one result is given here. One can derive an invariant of motion which is determined by

$$\hat{x}^2 + \frac{\beta_x \alpha_M C}{2\pi m Q_s} \left(\frac{\hat{\Delta E}}{E}\right)^2 = \text{const.} \tag{8}$$

where $\hat{x}$ and $\hat{\Delta E}$ are oscillation amplitudes.

Since m can be positive or negative [Eq.(1)], the betatron and synchrotron amplitude can increase or decrease at the same time or they can exchange their oscillation energy periodically. So an instability occurs again for a difference resonance, whereas in the case of a coupling resonance of horizontal and vertical betatron oscillation the difference resonance is stable and the sum resonance is unstable. The reason for this different behaviour is the assumption that the particle energy is above the transition energy. Below the transition energy one has to replace the momentum compaction factor α_M by $(\alpha_M\gamma^2 - 1)/(\gamma^2 - 1)$. So below the transition energy ($\alpha_M\gamma^2 < 1$) the satellite resonances show a behaviour similar to that of the coupling resonances of horizontal and vertical betatron oscillations.

For most cases $\hat{x}^2$ is smaller than the second term in Eq. (8), and the change of the betatron amplitude is larger than the change of the synchrotron amplitude.

Computer simulations were done with the following equations.

At the cavity:

$$\begin{aligned} x_{n+1} &= x_n - DR \\ x'_{n+1} &= x'_n - D'R \\ \frac{\Delta E_{n+1}}{E} &= \frac{\Delta E_n}{E} + R \\ s_{n+1} &= s_n \end{aligned} \tag{9}$$

with

$$R = \frac{eU}{E} \left[\sin\left(\phi + \frac{2\pi}{\lambda} s\right) - \sin\phi\right] .$$

Between cavities:

$$x_{n+2} = \cos \mu \, x_{n+1} + \beta \sin \mu \, x'_{n+1}$$

$$x'_{n+2} = - \frac{\sin \mu}{\beta} x_{n+1} + \cos \mu \, x'_{n+1}$$

$$\frac{\Delta E_{n+2}}{E} = \frac{\Delta E_{n+1}}{E} \qquad (10)$$

$$s_{n+2} = s_{n+1} - \alpha_M C \frac{\Delta E_{n+1}}{E} + A_1 x_{n+1} + A_2 x'_{n+1} \quad .$$

The result of the simulations is shown in Fig. 2. With the larger synchrotron frequency (U = 2.5 MV) the first satellite is stronger. With the smaller synchrotron frequency (U = 0.5 MV) the nonlinearities of the synchrotron oscillation are stronger and the satellites of higher order are stronger[10].

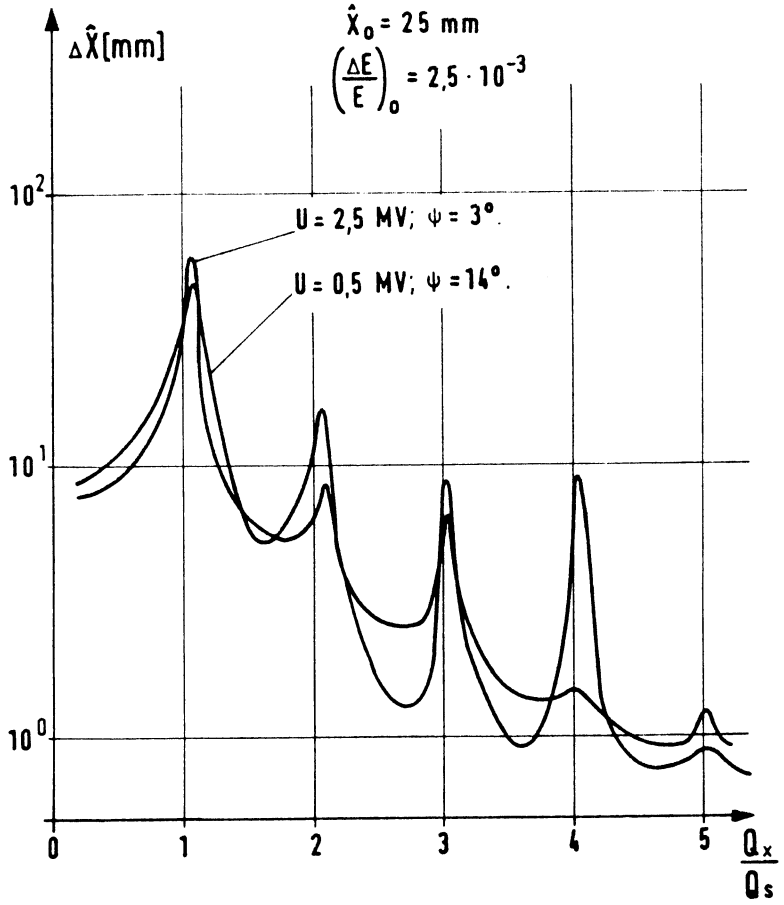


Fig. 2 Variation of the maximum betatron amplitude with betatron frequency

3. TRANSVERSE FIELDS WITH LONGITUDINAL VARIATION

Transverse fields which vary in longitudinal direction inside the bunch can be produced by the accelerating voltage[5] or by the bunch itself[6]. In both cases they are caused by asymmetries of the cavities or by an asymmetric passage of the bunch. The most important case occurs when the transverse fields are excited by an off-centre passage of the bunch (see Fig. 3). Since very small displacements of the bunch in the cavities can produce strong fields the excitation of satellite resonances is hard to suppress especially during energy ramping.

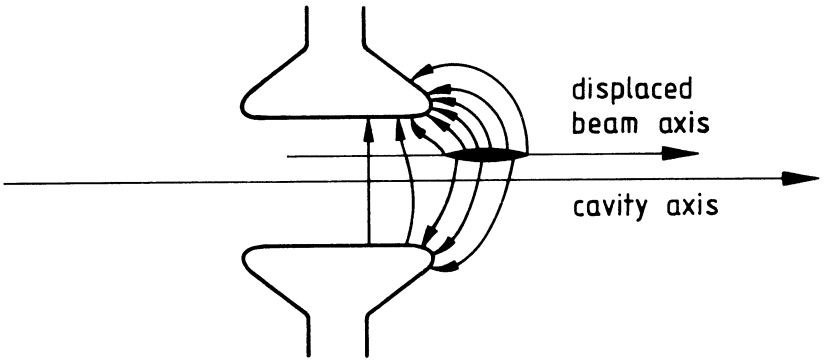


Fig. 3 Off-axis traversal of an r.f. gap

Figure 3 shows that the longitudinal distribution of the transverse kick is strongly nonlinear so that satellites of high order can be excited.

The excitation of the satellite resonances can be calculated exactly only for the first satellite of the integer, i.e. for $Q_\beta \pm Q_s = n$. The transverse kick

$$\delta z' = \frac{e}{p} \int (E_z + vB_x)\, dt \tag{11}$$

with E_z, B_x = electric and magnetic field and v = particle velocity, can then be linearized and one obtains

$$\delta z' = s \frac{e}{p} \int \left(\frac{\partial E_z}{\partial s} + v \frac{\partial B_x}{\partial s}\right) dt \ . \tag{12}$$

The integral is taken along the path of the particle between two limits where the fields vanish. With Maxwell's equation

$$\frac{\partial E_z}{\partial s} - \frac{\partial E_s}{\partial z} = \frac{\partial B_x}{\partial t}$$

one obtains

$$\delta z' = s \frac{e}{p} \left(\int \frac{\partial E_s}{\partial z} dt + B_x \Big|_{t_1}^{t_2} \right) = As \tag{13}$$

with

$$A = \frac{e}{p} \int \frac{\partial E_s}{\partial z} dt \ .$$

If the transverse fields have a longitudinal gradient the longitudinal field must have a transverse gradient, and that means that the betatron oscillation influences the synchrotron oscillation. That part of the energy change which varies linearly with the vertical position of a particle in the bunch is given by

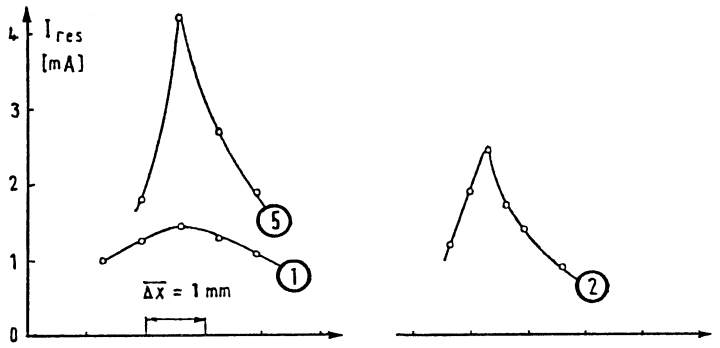
$$\frac{\delta E}{E} = z \frac{e}{E} \int \frac{\partial E_s}{\partial z} v \, dt = \frac{v^2}{c^2} Az = Az \ . \tag{14}$$

Equations (13) and (14) describe a linear coupling which can be investigated again using the matrix formalism. On the linear resonance $Q_z \pm Q_s = n$ the phases of the eigenvalues of the revolution matrix can be written as

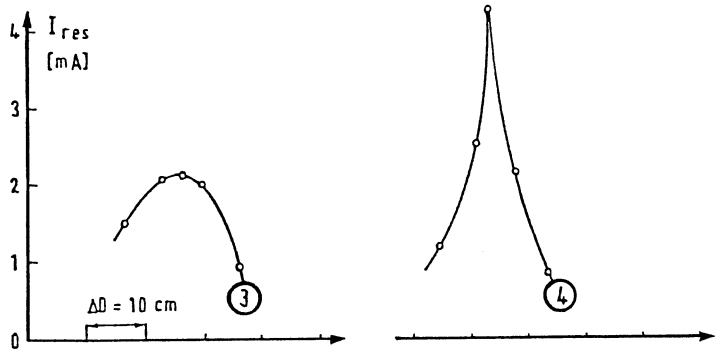
$$\mu = \mu_{z,s} \pm \frac{A}{2} \sqrt{- \frac{\beta_{zc} \alpha_M C}{\sin \mu_z}} \tag{15}$$

where β_{zc} is the amplitude function at the cavity. Equation (15) shows again that the oscillation amplitudes increase exponentially on a difference resonance. Also the invariant for the amplitudes given by Eq. (8) can be derived for the case of coupling due to transverse fields.

Since the excitation by a dispersion and the excitation by transverse fields are additive [Eqs. (3), (4), (11), (14)] both mechanisms can amplify or compensate each other depending on the phases of the dispersion and the orbit displacements in the cavities. Figure 4 shows the compensation of the 3rd horizontal satellite in PETRA[4] with both a horizontal dispersion and horizontal orbit bumps in the cavities. The criterion for the compensation was the residual current which was left when the horizontal betatron frequency was exactly on the resonance frequency. Since most of the satellites are current dependent there is a residual current which is obtained asymptotically after some minutes. At first an orbit bump in a cavity section was varied which gave a weak maximum for the residual current. This bump was left in the maximum, and another orbit bump was varied which was orthogonal to the first one. It was also left in the maximum, and then two bumps in the large quadrupoles near the interaction points were varied which produced two orthogonal dispersions in the whole ring, i.e. also in the cavities. After optimising these two bumps a sharp maximum was obtained which now also occurred when the first bump was varied again.



Residual current as a function of two orthogonal bumps in the cavity section



Residual current as a function of two bumps which produce two orthogonal dispersions

Fig. 4 Compensation of the horizontal satellite $Q_x = 25 + 3Q_s$ with $Q_s = 0.058$

Since the compensation of the satellites cannot be maintained during energy ramping the synchrotron and betatron frequencies must be controlled and kept constant. Figure 5 shows the working point during injection and energy ramping, and with beam-beam interaction at high energy. The maximum currents could be injected and ramped when the horizontal betatron frequency was exactly in the middle between the first and the second satellites and the vertical betatron frequency was between the third and the fourth satellites. The distance from the resonances $Q_x - Q_z = 2$ and $Q_x + 2Q_z = 72$ must be large in order to avoid dangerous coupling. In luminosity runs the frequency spreads ΔQ_x and ΔQ_z due to the beam-beam interaction could overlap the fourth vertical satellite since the satellites are weaker at high energy and also because they are weakened by the strong non-linearities of the beam-beam forces.

4. BEAM-BEAM INTERACTION WITH A CROSSING ANGLE

Figure 6 shows two bunches crossing at an angle of 2ϕ. A particle with a distance s from the centre of its own bunch but without transverse displacement passes the centre of the opposing bunch at a distance of $s \cdot \tan \phi \approx s\phi$. The particle therefore gets a transverse kick, and this means a coupling from the synchrotron oscillation to the betatron oscillation since the kick, i.e. the change of the betatron angle, depends on the longitudinal synchrotron coordinate s.

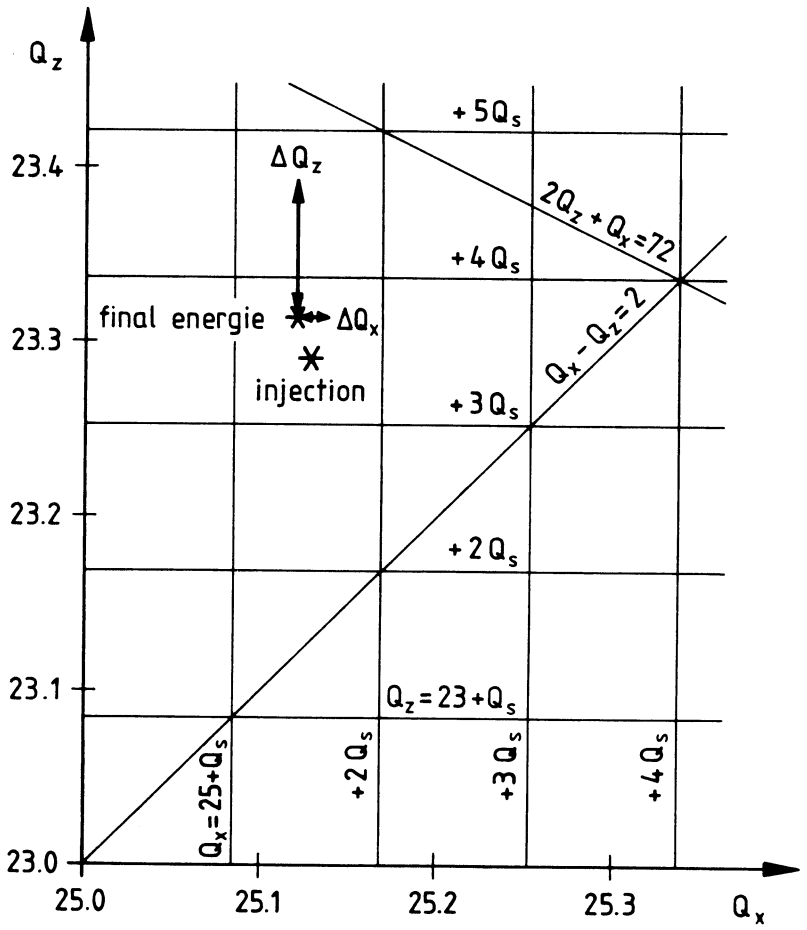


Fig. 5 Working point in PETRA during injection and ramping

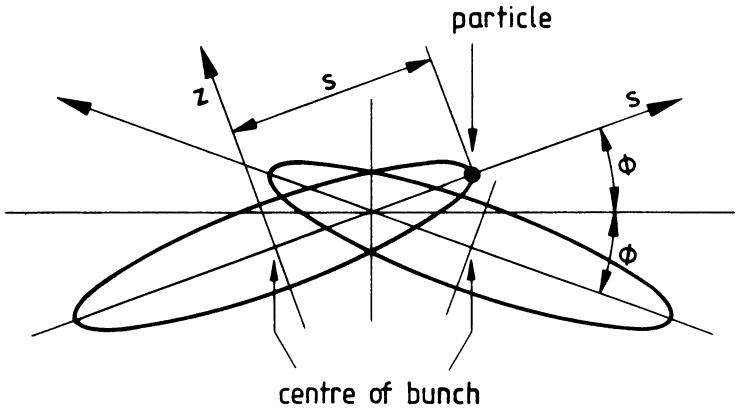


Fig. 6 Beam-beam interaction at a crossing angle

If the particle has also a transverse displacement z the kick is given by

$$\delta z' = f(z + s\phi) \ . \qquad (16)$$

The function f describes the space charge forces and is shown in Fig. 7.

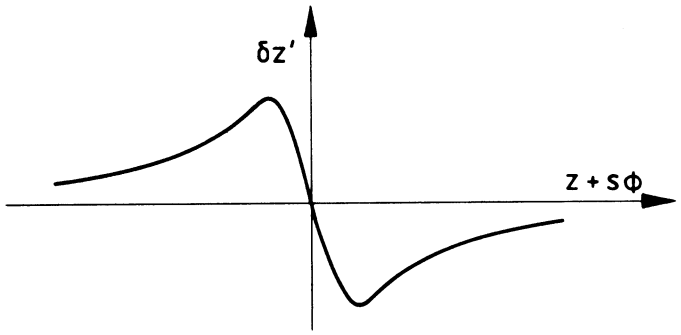


Fig. 7 Dependence of the transverse kick on $z + s\phi$

The synchrotron oscillation, however, is also influenced by the betatron oscillation. The change of the momentum is, in the case of a vertical crossing angle, always vertical. This can be seen by transforming the two bunches into the reference system where they collide head-on.

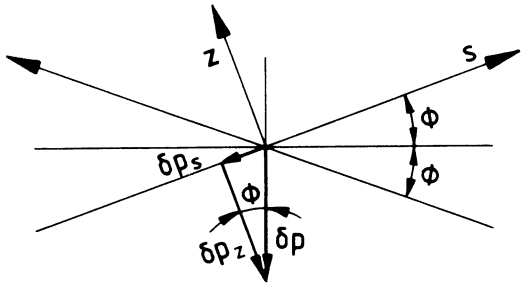


Fig. 8 Resolution of the momentum change

The resolution of the momentum change gives

$$\delta p_s = \phi \delta p_z \tag{17}$$

and the energy change is

$$\frac{\delta E}{E} \approx \frac{\delta p}{p} \approx \frac{\delta p_s}{p} = \phi \frac{\delta p_z}{p} = \phi \delta z'$$

and

$$\frac{\delta E}{E} = \phi f(z + s\phi) \ . \tag{18}$$

The complete coupling is now described by Eqs. (16) and (18). For small oscillation amplitudes the function f may be linearized and one obtains a linear coupling:

$$f(z + s\phi) = - \frac{4\pi\xi}{\beta_z^*} (z + s\phi) \tag{19}$$

with

$$\xi = \frac{r_e N_b \beta_z^*}{2\pi\gamma\sigma_{zel}(\sigma_x + \sigma_{zel})}$$

$$\sigma_{zef} = \sqrt{\sigma_z^2 + \phi^2\sigma_s^2}$$

r_e = electron radius, N_b = number of particles per bunch, β^*_z = amplitude function at the interaction point, γ = particle energy divided by its rest energy, σ_s, σ_x, σ_z = standard deviations for longitudinal, horizontal and vertical bunch dimensions, respectively.

In the case of linear coupling one can apply again the matrix formalism and obtain for the phases μ of the eigenvalues[7,8]

$$\mu = \mu_{z,s} + 2\pi\xi \pm 2\pi\xi\phi \sqrt{\frac{-\alpha_M c}{\beta_z^* \sin \mu_z}} \quad . \tag{20}$$

Equation (20) shows again that an instability occurs on a difference resonance (sin $\mu_z > 0$), and, in this case, also the invariant for the amplitudes Eq. (8) can be derived.

To get an idea of the strength of this instability the rise times for the storage ring DORIS I are given. The parameters are[7]: 2ϕ = 24 mrad, ξ_z = 0.01, α_μ = 0.018, c = 288 m, β^* = 1 m, Q_s = 0.034, f_0 = 1.04 MHz. With these parameters one gets a rise time of 0.3 msec.

For large oscillation amplitudes high order satellites of non-linear betatron resonances can be excited. The most effective method for analysing these satellites is a computer simulation. Figure 9 shows the result of a simulation for DORIS I. Between the interaction points the betatron coordinates are transformed linearly and the synchrotron coordinates are transformed taking into account the non-linear synchrotron potential [Eqs. (9) and (10) with D = 0, D' = 0]. At the interaction points the transformation of z' and $\Delta E/E$ is given by Eqs. (16) and (18) whereas z and s are constant. The function f(x) is:

$$f(x) = 8\pi\xi \frac{\sigma^2}{x\beta^*} \left(e^{-\frac{x^2}{2\sigma^2}} - 1\right) \tag{21}$$

which is exact for a round beam. σ is the standard deviation for the transverse particle distribution (σ = 0.23 mm for DORIS I).

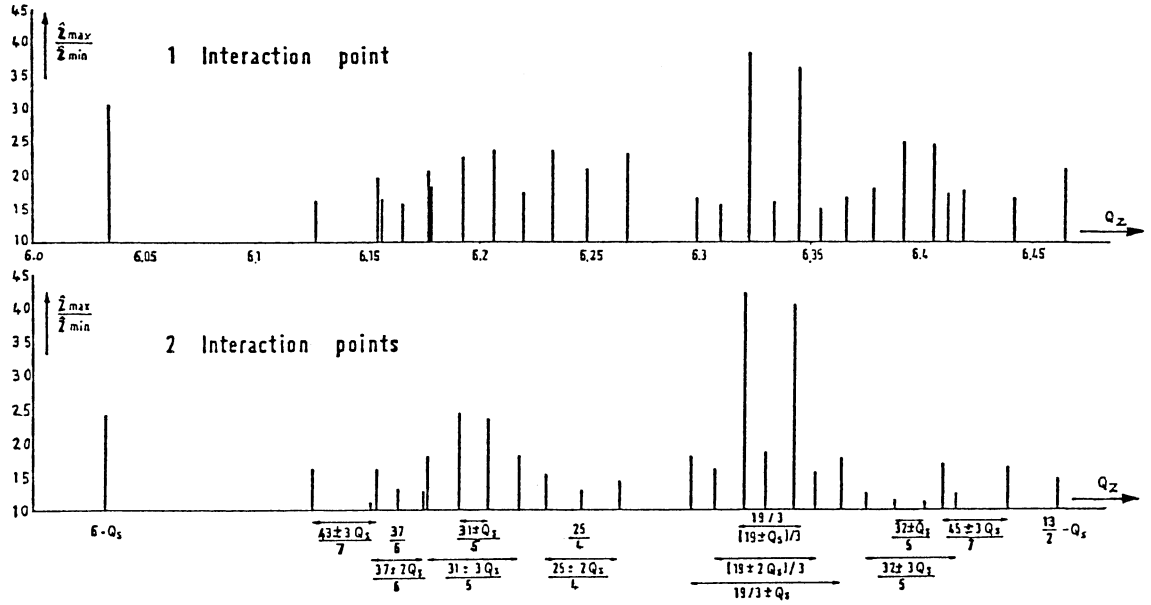


Fig. 9 Maximum betatron amplitude as a function of the betatron frequency

Many measurements were done in DORIS I with weak positron bunches colliding with strong electron bunches. The vertical betatron frequency of the positrons was varied continuously by a computer whereas the synchrotron frequency was kept constant. Simultaneously the current of the positron bunches was plotted. An example of such a measurement is shown in Fig. 10 for one interaction point.

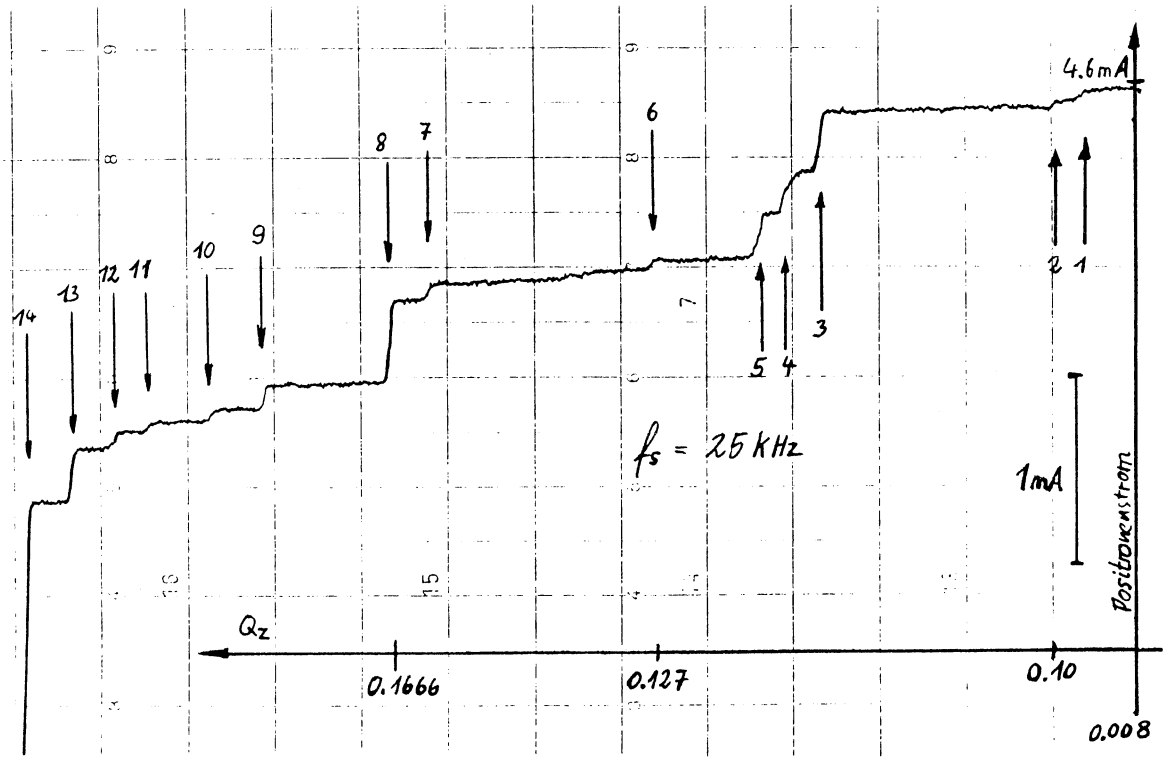


Fig. 10 Variation of the positron current with the betatron frequency

The sharp losses yield exactly the frequencies of the satellite resonances. Because of the very small width of the resonances and because of the speed of the variation, the only resonances which could be observed were those which led to a lifetime of less than 15 min. In several measurements with one interaction point the following 25 vertical resonances were seen.

$6 + 3Q_s$, $6 + 4Q_s$,

$49/8$, $((49+Q_s)/8$, $(49-2Q_s)/8$,

$43/7$, $(43\pm Q_s)/7$, $(43-2Q_s)/7$

$37/6$, $(37-Q_s)/6$, $(37-2Q_s)/6$

$31/5$, $(31\pm Q_s)/5$, $(31-2Q_s)/5$, $(31\pm 3Q_s)/5$, $(31\pm 4Q_s)/5$

$25/4$, $(25-Q_s)/4$, $(25-2Q_s)/4$

$28-3Q_x$, Q_x-1-Q_s,

The width of the satellite resonances was measured for several cases. It is only about 0.0005. This small width is caused by the strong non-linear tune shift of the beam-beam interaction. If the oscillation amplitude increases the particle comes out of resonance and the amplitude growth comes to an end. The frequency variation is small only for large amplitudes and hence the amplitude can increase further.

The luminosity in DORIS I was limited by these satellite resonances. This ocurred because in order to suppress multibunch instabilities an rf-quadrupole and a decoupling transmitter were used to give different bunches different betatron and synchrotron frequencies and this always resulted in a loss of some of the 2×120 bunches.

In the first proposal for the e-p storage rings HERA, which is now under construction, a horizontal crossing angle was foreseen since it has some advantages over head-on collision, e.g. less synchrotron radiation hits the detectors and the optics for the two rings are completely independent. It was assumed that with a smaller space charge parameter ($\xi = 0.0006$ instead of 0.01 as in DORIS I) and without spreads in betatron and synchrotron frequencies a good working point for the protons could be found. Computer simulations, however, have shown that in this case even more satellite resonances would be excited[9] and the crossing angle was abandoned.

One reason for the larger number of satellite resonances was the fact that the proton bunch length was larger than the electron bunch length, and the proton bunch length times half the crossing angle ϕ was larger than the electron bunch width ($\phi\sigma^+_s > \sigma^-_x$). The ratio $\phi\sigma^+_s/\sigma^-_x$ is 4 in the case of HERA whereas the equivalent ratio $\phi\sigma_s/\sigma_{z,eff}$ is only 0.5 for DORIS I. This means that the synchrotron motion of the protons covers more non-linearities (see Fig. 7) and more satellite resonances are excited. Another reason is that even weak satellites with a rise time of several hours become dangerous since there is no damping for protons.

The computer simulations for HERA were similar to the simulations for DORIS I, but a flat electron bunch was assumed (σ^-_x/σ^-_z = 15) and the function f(x + ϕs) (Eq. 16) was calculated by using a table and by interpolating in a two-dimensional grid[9]). The number of simulated revolutions was 50,000 instead of 2000 for DORIS I. A result of such a simulation is shown in Fig. 11. The lower curve shows the maximum of the beam width, i.e. the maximum of the average of the coordinates of 64 particles, and the upper curve shows the maximum amplitude which occurred during the 50,000 revolutions for one of the 64 particles. The maximum initial amplitude was 2.2 σ_{xo}.

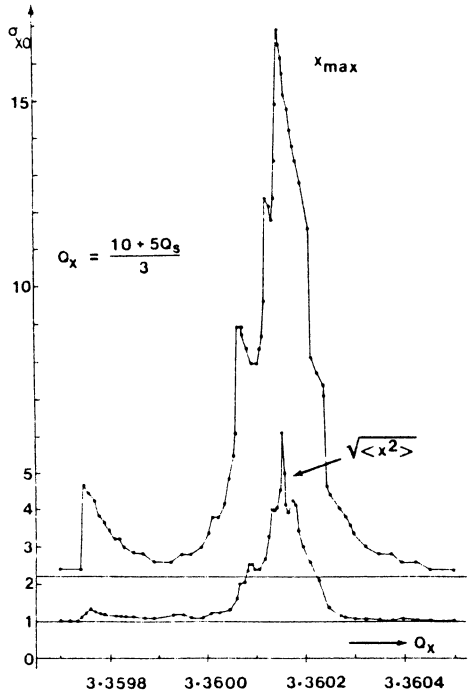


Fig. 11 Dependence of the maximum betatron amplitude and the beam width on the betatron frequency

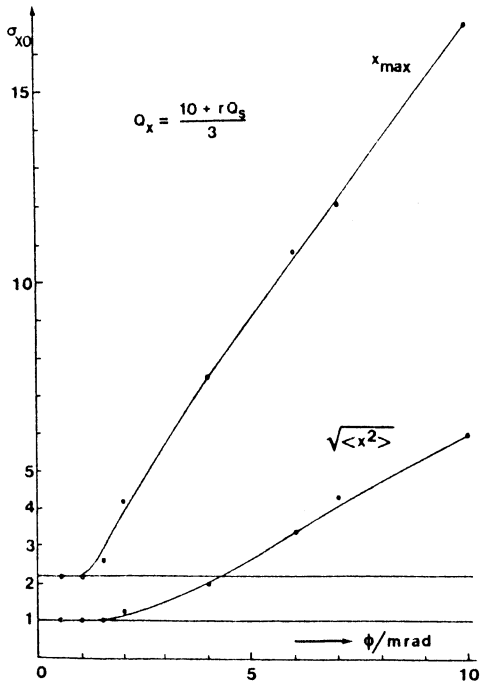


Fig. 12 Dependence of the maximum amplitude and the maximum beam width on the crossing angle

Figure 12 shows the maxima of the two curves in Fig. 11 as a function of the crossing angle. A strong increase of the amplitudes begins with crossing angles of the order of $2\phi \approx 2\sigma^-_x/\sigma^+_s$, i.e. in this case with $\phi \approx 2$ mrad.

The simulations have shown satellites of nonlinear resonances up to the order 11. Figure 13 shows all satellites of the resonance $Q_x = 4n + 10/3$ where n is an integer. Only odd integers are excited. This follows from the antisymmetric distribution of the space charge forces (see Fig. 7). If the resonance condition for the betatron frequency is written as

$$Q_x = \frac{p + rQ_s}{q} \qquad (22)$$

then only those satellites can be excited for which the relation

$$r + q = \text{even}$$

is satisfied. This holds, of course, only if the centres of the two bunches collide. Many satellites are also suppressed if the interaction points are equidistant. If one takes into account that the crossing angles can have equal signs or alternating signs one gets the condition for the excitation of satellites

$$\frac{p}{N_i} = \text{integer} \quad \text{(equal signs)}$$

$$\frac{2p}{N_i} + r = \text{even} \quad \text{(alternating signs)}$$

where N_i is the number of interaction points. If, however, the betatron phase advances between interaction points are different, i.e. due to machine errors, then more resonances can be excited.

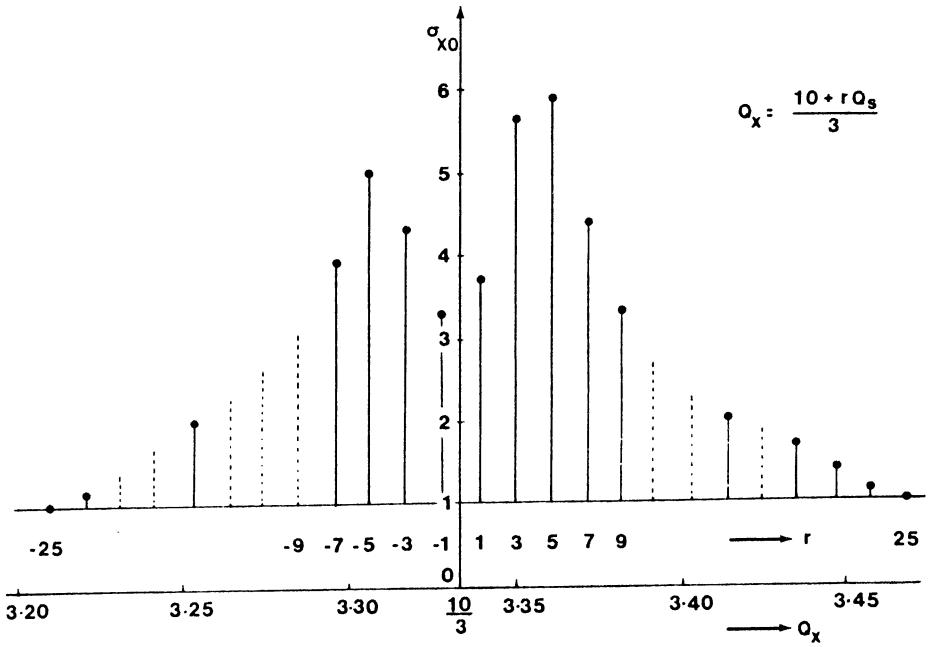


Fig. 13 Maximum beam width on all satellites of the resonance $Q_x = 10/3$ (——— simulated, ------ estimated)

REFERENCES

1) M.C. Crowley-Milling, I.I. Rabinowitz, 1977 Particle Accelerator Conference, IEEE Trans. on Nuclear Science, NS-18, No. 3, 1052 (1977).

2) SPEAR Group, 1975 Particle Accelerator Conf., IEEE Trans. on Nuclear Science, NS-22, 1366 (1975).

3) A.W. Chao, E. Keil, A.S. King, M.J. Lee, P.L. Morton and J.M. Peterson, SPEAR-187, August 1975.

4) A. Piwinski, 11th Intern. Conf. on High Energy Accelerators, (CERN 1980), p. 754 (1980).

5) N.A. Vinokurov et al., 10th Intern. Conf. on High Energy Accelerators (Protvino, July 1977), p. 272.

6) R.M. Sundelin, 1979 Particle Accelerator Conf., IEEE Trans. on Nuclear Science, NS-26, No. 3, 3604 (1979).

7) A. Piwinski, DESY 77/18, March 1977.

8) A. Piwinski, 1977 Particle Accelerator Conf., IEEE Trans. on Nuclear Science, NS-24, No. 3, 1408 (1977).

9) A. Piwinski, 1985 Particle Accelerator Conf., Vancouver, (1985) .

10) A. Piwinski and A. Wrulich, DESY 76/07, February 1976.

BETATRON COUPLING WITH RADIATION

G. Guignard
CERN, Geneva, Switzerland.

ABSTRACT

This lecture deals with the analysis of the coupling of vertical and horizontal oscillations, in the presence of electromagnetic radiations associated with the acceleration of relativistic electrons in a macroscopic force field. It explains how the linear theory of resonances for proton beams circulating in three-dimensional magnetic fields has been extended to radiating particles. This analytical treatment gives the possibility to predict the equilibrium mean-square amplitudes and emittances when coupling and vertical momentum dispersion are included. The application to electron rings with a separated function structure and a large radius of curvature is presented.

1. INTRODUCTION

The problem of betatron coupling with radiation is solved in this paper by applying twice over the method of the perturbation of the constants. At first, the betatron coupling is treated in the formalism of the classical Hamiltonian, the perturbation consisting of the linear forces which couple both the transverse motions[1,2]. In contrast with previous work [3], longitudinal-field components are included as well as the transverse ones. Using a complete set of canonical variables, the mapping of the phase space is always symplectic, and knowing the solution of the unperturbed motion, the equations of the perturbed one can be solved explicitely (Section 2). After this, the effect of radiation and of quantum fluctuations are treated as time variations of the amplitudes of the coupled motions[4]. Averages of these variations over a finite time interval can be written in a general form (Section 3). Then, looking for stationary conditions, it is possible to give expressions for the equilibrium amplitudes and associated beam emittances, combining the two perturbation theories mentioned above (Section 4). This analytical description of linear coupling in the presence of radiation makes it possible to investigate the question of emittance control under particular conditions (Section 5).

2. PERTURBATION TREATMENT OF LINEAR BETATRON COUPLING

2.1 Recap of the principles

The basic principles of the Hamiltonian formalism used in this paper are described for instance in Ref. 1. If we call $\{q_\rho, p_\rho\}$, with $\rho = 1,N$, the canonical variables and θ the independent variable (angle at the machine centre), the Hamiltonians of the unperturbed and perturbed motions can be written as $H_0(p,q,\theta)$ and $H_1(p,q,\theta)$, respectively.

The equations of the unperturbed motion are well known :

$$\dot{q}_\rho = \frac{\partial H_0}{\partial p_\rho} \quad ; \qquad \dot{p}_\rho = - \frac{\partial H_0}{\partial q_\rho} \quad , \tag{2.1}$$

and the solution can be written formally as follows :

$$q_\rho = q_\rho \, (a_j, \theta) \quad : \quad p_\rho = p_\rho \, (a_j, \theta) \, , \tag{2.2}$$

where a_j are constants of the unperturbed motion.

In the presence of the perturbation H_1, we look for the dependence of the quantities a_j on the variable θ, using the equation :

$$\frac{da_j}{d\theta} = [a_j, H_1] = \sum_\rho \left(\frac{\partial a_j}{\partial q_\rho} \frac{\partial H_1}{\partial p_\rho} - \frac{\partial H_1}{\partial q_\rho} \frac{\partial a_j}{\partial p_\rho} \right) \, , \tag{2.3}$$

the second equality defining the Poisson's brackets. Using Eq.(2.2), it is possible to rewrite the Hamiltonian of the perturbation as a function of a_j and θ only,

$$H_1 \, (p,q,\theta) = U \, (a_j, \theta) \, . \tag{2.4}$$

Hence, the equations of the perturbed motion are :

$$\frac{da_j}{d\theta} = \sum_m [a_j, a_m] \frac{\partial U}{\partial a_m} \quad . \tag{2.5}$$

2.2 Application to betatron motion

To describe the betatron motion, the canonical variables q and p are now replaced by the transverse coordinates x,z and their conjugate momenta p_x, p_z. The unperturbed Hamiltonian is simply :

$$H_0 = \frac{1}{2} \left[K_1(\theta) x^2 + K_2(\theta) z^2 + p_x{}^2 + p_z{}^2 \right] \, , \tag{2.6}$$

where K_1 and K_2 are the horizontal and vertical focusing functions.

Since the equations of motion associated with Eq.(2.6) are of Hill's form, Floquet's theorem applies and gives the explicit solution :

$$\begin{aligned} x &= a_1 \, u && \exp(iQ_x\theta) + \text{c.c.} \\ p_x &= a_1 (u' + iQ_x u) && \exp(iQ_x\theta) + \text{c.c.} \\ z &= a_2 \, v && \exp(iQ_z\theta) + \text{c.c.} \\ p_z &= a_2 (v' + iQ_z v) && \exp(iQ_z\theta) + \text{c.c.} \end{aligned} \quad , \tag{2.7}$$

where c.c. means complex conjugate and ' denotes the derivative with respect to θ. Quantities u and v are the Floquet's functions :

$$u(\theta) = \sqrt{\frac{\beta_x(\theta)}{2R}} \exp[i(\phi_x - Q_x\theta)]$$

$$v(\theta) = \sqrt{\frac{\beta_z(\theta)}{2R}} \exp[i(\phi_z - Q_z\theta)] \quad . \tag{2.8}$$

Let us now consider the betatron motion perturbed by linear coupling associated with a 3-dimensional magnetic field. In agreement with the equations for the coupled motion[1], the Hamiltonian of the perturbation is in this case :

$$H_1 = \frac{1}{2} (2Kxz - Mzp_x + Mxp_z + \frac{1}{4} M^2z^2 + \frac{1}{4} M^2x^2) \quad , \tag{2.9}$$

where K stands for the tilted-quadrupole field :

$$K(\theta) = \frac{R^2}{2B\rho} \left(\frac{\partial B_x}{\partial x} - \frac{\partial B_z}{\partial z} \right) \quad , \tag{2.9a}$$

and M for the longitudinal (solenoidal) field :

$$M(\theta) = \frac{R}{B\rho} B_\theta \quad . \tag{2.9b}$$

Following the recipe of Section 2.1, Eq.(2.7) must be used to express H_1 (Eq.(2.9)) as a function U of the a_j's (Eq.(2.4)) and then solved for these quantities; we get from Eq.(2.7) :

$$a_1 = \frac{1}{W(u)} [(u^{*\prime} - iQ_x u^*) x - u^* p_x] \exp(-iQ_x\theta)$$

$$a_1^* = - \frac{1}{W(u)} [(u' + iQ_x u) x - u p_x] \exp(iQ_x\theta) \quad . \tag{2.10}$$

Similar expressions for a_2, a_2^* (* denoting the complex conjugate) are obtained by replacing u and x in Eq.(2.10) by v and z, respectively.

The function W, sometimes called the Wronskian, is defined in the following equation :

$$W(u) = u (u^{*\prime} - iQ_x u^*) - (u' + iQ_x u) u^* \quad . \tag{2.11}$$

Given the form of the Floquet's functions (Eq.(2.8)), we have simply W(u) = W(v) = -i, and Eqs.(2.10) simplify themselves accordingly. Introducing then explicitely Eq.(2.8) into

the complete expressions (2.10) for the a_j's makes it possible to rewrite Eq.(2.5) for our particular application to betatron motion :

$$\frac{da_1}{d\theta} = i\,\frac{\partial U}{\partial a_1^*} \qquad \frac{da_2}{d\theta} = i\,\frac{\partial U}{\partial a_2^*}$$

$$\frac{da_1^*}{d\theta} = -\,i\,\frac{\partial U}{\partial a_1} \qquad \frac{da_2^*}{d\theta} = -\,i\,\frac{\partial U}{\partial a_2} \quad . \tag{2.12}$$

2.3 Solution in the presence of linear coupling

The perturbation Hamiltonian H_1 (Eq.(2.9)) is a quadratic function in y and p_y, while these canonical variables depend linearly on the quantities a_j (Eq.(2.7)) and contain oscillating terms of frequency Q_y (Eq.(2.8)). It results directly from these facts that the function U associated with H_1 (Eq.(2.4)), has necessarily the following form for linearly-coupled betatron motions[1] :

$$U(a_1,a_1^*,a_2,a_2^*,\theta) = \sum_{j,k,l,m=0}^{2} \sum_{q=-\infty}^{+\infty} h_{jklmq}\, a_1^j a_1^{*k} a_2^l a_2^{*m} \times$$

$$\times \exp\{i[(j-k)Q_x + (l-m)Q_z + q]\theta\} \quad . \tag{2.13}$$

An additional sum over the integers q appears in Eq.(2.13). It corresponds to a development in a Fourier series of the Hamiltonian, which is related to the periodicity of the synchrotron or storage ring over the circumference.

The harmonic coefficients h contain by definition the magnetic characteristics K (2.9a) and M (2.9b) as well as the Floquet's functions (2.8) and their derivatives with respect to θ.

So far, everything is rigorous. Now we do the usual assumption that the low frequency part of the function U (2.13) gives the important variations of the a_j's[2]. The associated conditions, directly obtainable from (2.13), are :

a) Zero frequency terms (or stabilizing terms)
   $q = 0$ with either $j = k = 0$ and $l = m = 1$
   or $j = k = 1$ and $l = m = 0$

b) Resonant terms ($n_1Q_x + n_2Q_z - p = 0$)
   $q = -p$ with $j - k = n_1$ and $l - m = n_2$
   and $n_1 = 0,1,2$ and $n_2 = 2, \pm 1, 0$.

With these conditions, the function U can be reduced to :

$$
\begin{aligned}
U = {} & h_{11000}\, a_1 a_1^* + h_{00110}\, a_2 a_2^* && \text{frequency shifts} \\
& + h_{1010\text{-}p}\, a_1 a_2 \exp[i(Q_x + Q_z - p)\theta] + \text{c.c.} && \text{sum resonance} \\
& + h_{1001\text{-}p}\, a_1 a_2^* \exp[i(Q_x - Q_z - p)\theta] + \text{c.c.} && \text{difference resonance} \\
& + h_{2000\text{-}p}\, a_1^2 \exp[i(2Q_x - p)\theta] + \text{c.c.} && \\
& + h_{0020\text{-}p}\, a_2^2 \exp[i(2Q_z - p)\theta] + \text{c.c.} && \text{one-dimensional resonances .}
\end{aligned}
\qquad (2.14)
$$

We have derived[1] explicit solutions for each of these resonances, including simultaneously the frequency shifts. However, in order to simplify the present lecture, we introduce now the two following assumptions :

a) The frequency shifts and the one-dimensional resonances can be neglected, since the corresponding coefficients h are integrals containing the square of M (2.9b), i.e. second order terms in the magnetic field[1];

b) The sum resonance can also be disregarded, provided that the working point is sufficiently close to the difference resonance as in most usual cases.

In this context, the function U (Eq.(2.14)) restricts itself to the third term only and the complete definition of the coefficient h, renamed κ for the special case of the difference resonances, becomes :

$$
\kappa = h_{1001\text{-}p} = \frac{1}{4\pi R} \int_0^{2\pi} \sqrt{\beta_x \beta_z}\ \left[K + \frac{MR}{2}\left(\frac{\alpha_x}{\beta_x} - \frac{\alpha_z}{\beta_z}\right) - i\,\frac{MR}{2}\left(\frac{1}{\beta_x} + \frac{1}{\beta_z}\right)\right] \times
$$

$$
\times \exp\left[i(\phi_x - \phi_z - \theta\Delta)\right] d\theta \quad , \qquad (2.15)
$$

with $\Delta = Q_x - Q_z - p$.

With all the assumptions made in this section, the effect of the perturbation due to linear coupling is described by two equations derived from (2.12) by using (2.14) and (2.15) :

$$
\begin{aligned}
\frac{da_1}{d\theta} &= i\kappa^* a_2 \exp(-i\theta\Delta) \\
\frac{da_2}{d\theta} &= i\kappa\, a_1 \exp(i\theta\Delta) \quad .
\end{aligned}
\qquad (2.16)
$$

Transforming these two equations of first order into one equation of second order, the general solution for the a_j's can be obtained and written as follows :

$$a_1 = \kappa^* \left[\frac{A_1}{\omega_+} \exp\,(i\omega_+\theta) + \frac{A_3}{\omega_-} \exp(i\omega_-\theta)\right]$$

$$a_2 = \left[A_1 \exp\,(i\omega_+\theta) + A_3 \exp(i\omega_-\theta)\right] \exp(i\theta\Delta) \qquad (2.17)$$

with

$$\omega_\pm = -\frac{\Delta}{2} \pm \sqrt{\left(\frac{\Delta}{2}\right)^2 + |\kappa|^2}\ .$$

The complete solution is finally derived by introducing Eqs.(2.17) and (2.8) into the solution of the unperturbed motion (Eq.(2.7)). At this point, let us define the vector $\vec{Y}$ whose components are the canonical variables (Eq.(2.7)), with $p_x = x' + Mz/2$ and $p_z = z' - Mx/2$. This definition allows us to write the complete solution mentioned above as follows :

$$Y_j = \sum_{k=1}^{4} w_{jk}(\theta)\, A_k \quad , \qquad j = 1 \text{ to } 4 \quad , \qquad (2.18)$$

where j is numbering the $\vec{Y}$-components (2.7),

A_1 and A_3 are the complex constants appearing in Eq.(2.17),

$A_2 = A_1^*$ and $A_4 = A_3^*$, and

$$w_{\substack{11\\13}} = \frac{\kappa^*}{\omega_\pm}\sqrt{\frac{\beta_x}{2R}}\exp[i(\phi_x + \omega_\pm\theta)],$$

$$w_{\substack{21\\23}} = \frac{\kappa^*}{\omega_\pm}\sqrt{\frac{R}{2\beta_x}}\,(i-\alpha_x)\exp[i(\phi_x + \omega_\pm\theta)],$$

$$w_{\substack{31\\33}} = \sqrt{\frac{\beta_z}{2R}}\exp[i(\phi_z - \omega_\mp\theta)],$$

$$w_{\substack{41\\43}} = \sqrt{\frac{R}{2\beta_z}}\,(i-\alpha_z)\exp[i(\phi_z - \omega_\mp\theta)], \qquad (2.19)$$

$$w_{j2} = w_{j1}^* \ , \quad w_{j4} = w_{j3}^* \ .$$

The two different subscripts of the w's are directly associated with the two indices of ω. The particular equalities between complex conjugate quantities are an obvious consequence of the presence of complex conjugate functions in Eq.(2.7). The form of the w-functions (Eq.(2.19)) implies that the following relations are satisfied :

$$w_{jk}(\theta+2\pi) = w_{jk}(\theta)\exp(i2\pi\lambda_{jk}) \quad , \qquad (2.20)$$

with $\lambda_{j2} = -\lambda_{j1}$, $\lambda_{j4} = -\lambda_{j3}$.

The four groups of four w-functions defined by Eq.(2.19) can be regarded as the eigenfunctions defining the four eigenvectors associated with the complete solution $\vec{Y}$ of betatron motion in the presence of coupling. It is obvious from the eigenfunctions (Eq.(2.19)) that the single-particle motions (Eq.(2.18)) contain fast and slow oscillations associated with the phases ϕ_y and $\omega_{\pm\theta}$, respectively. As a result from this, the coherent oscillations following a kick in one transverse direction have the pattern represented in Fig.1 [5,6].

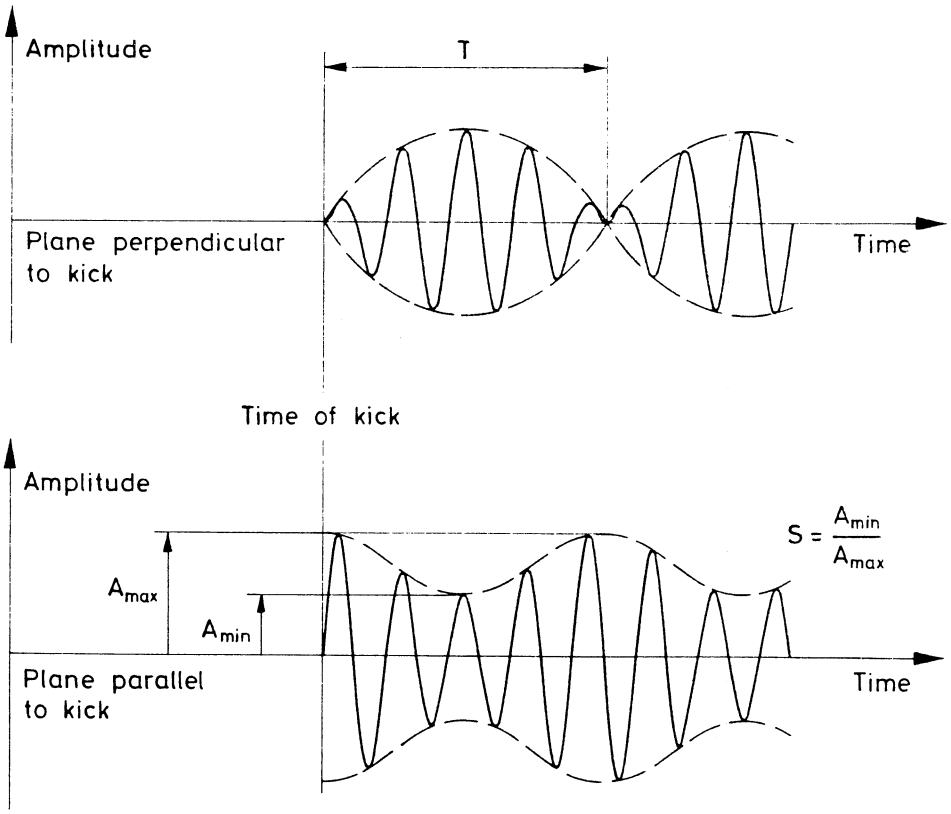


Fig. 1 Coherent oscillations following a kick

3. AMPLITUDE VARIATION DUE TO RADIATION AND ACCELERATION

3.1 General expression of the amplitude variation

Coupled betatron oscillations, treated in the previous section by a perturbation method, are in turn enhanced by photon emission and damped by the longitudinal acceleration as well as by the average energy loss in the presence of the focusing component of the magnetic field.

Starting from the constants A_k of the coupled motion, we now look for their variations due to these effects[4]. Assuming that these constants change slowly with

respect to the quantum fluctuations and the period of the coupled betatron oscillations, they will reach an equilibrium between excitation and damping after a few damping times. It is precisely these equilibrium values of A_k which we want to derive in the following sections.

It follows from Eq.(2.19) that the quantity F defined below is an invariant of the motion :

$$F(w_{j1},w_{jk}) = w_{11}\, w_{2k} - w_{21}\, w_{1k} + w_{31}\, w_{4k} - w_{41}\, w_{3k} \quad . \tag{3.1}$$

As a consequence of the invariance of F and of the properties (2.20), we can write :

$$F[w_{j1}(\theta),\, w_{jk}(\theta)] = F\,[w_{j1}(\theta+2\pi),\, w_{jk}(\theta+2\pi)] =$$

$$= \exp[i2\pi\,(\lambda_{j1} + \lambda_{jk})]\; F\,[w_{j1}(\theta),\, w_{jk}(\theta)] \quad , \tag{3.2}$$

and the equality between the first and last term of Eq.(3.2) (which can only be satisfied if the exponential is equal to 1) induces the following orthonormality :

$$F(w_{j1},w_{jk}) = 1 \quad \text{if} \quad w_{j1} = w_{jk}^*$$

$$F(w_{j1},w_{jk}) = 0 \quad \text{if} \quad w_{j1} \neq w_{jk}^* \quad . \tag{3.3}$$

This property of orthonormality can now be used to solve Eq.(2.18) for the A_k's. Let us write indeed :

$$F(Y_j,w_{jk}) = \sum_{l=1}^{4} A_l\, F(w_{jl},w_{jk}) = A_k^*\, F(w_{jk}^*,w_{jk}) \quad , \tag{3.4}$$

making use of Eqs.(2.18) and (3.3). Then we derive from the last equality :

$$A_k = \frac{F(Y_j,w_{jk}^*)}{F(w_{jk},w_{jk}^*)} \quad . \tag{3.5}$$

Since we want to look for the variations of the A_k's due to changes in the canonical variables (vector $\vec{Y}$), Eq.(3.5) is the key equation. In general, any variation of $|A_k|$ due to trajectory changes δY_j can be expressed as follows :

$$\delta|A_k|^2 = A_k^*\,\delta A_k + A_k \delta A_k^* + |\delta A_k|^2 = \frac{1}{F(w_{jk},w_{jk}^*)}\,\{A_k^*\, F(\delta Y_j,w_{jk}^*) -$$

$$- A_k\, F(\delta Y_j,w_{jk}) - \frac{1}{F(w_{jk},w_{jk}^*)}\; F(\delta Y_j,w_{jk}^*)\, F(\delta Y_j,w_{jk})\} \quad . \tag{3.6}$$

3.2 Application to photon emission and acceleration

Considering the photon emission, it is well known[7] that the equilibrium orbit and the betatron variables are changed by a quantity which is proportional to the photon energy ε and to the dispersion $\vec{D}$, whose components are (D_x, D_x', D_z, D_z') :

$$\delta\vec{Y} = \frac{\varepsilon}{E_0}\ \vec{D} \quad . \tag{3.7}$$

In the presence of longitudinal acceleration δE in a cavity, only the transverse momenta are changed while the transverse coordinates remain constant :

$$\delta\vec{Y} = (\delta x,\ \delta p_x,\ \delta z,\ \delta p_z) = (0,\ -p_x,\ 0,\ -p_z)\ \frac{\delta E}{E_0} \quad , \tag{3.8}$$

E_0 being the nominal energy.

Let us first deal with the orbit change due to photon emission. Since $\delta\vec{Y}$ is proportional to ε (Eq.(3.7)) and since δY_j appears linearly as well as quadratically in Eq.(3.6), it is necessary to evaluate the average $\langle\varepsilon\rangle$ and the mean square $\langle\varepsilon^2\rangle$ of the quantum emission, over a time interval Δt, and to multiply both by the mean emission rate N. We will now consider successively these two terms :

a) If P_γ stands for the rate of loss of energy by radiation, we have :

$$N\ \Delta t\ \langle\varepsilon\rangle = \frac{1}{c}\ P_\gamma\ \Delta\ell \quad , \tag{3.9}$$

where $\Delta\ell$ is the path length interval and

$$P_\gamma = \frac{2}{3}\ r_e c\ \gamma^3\ E\ \frac{1}{\rho^2}$$

$$\Delta\ell = \Delta s\ \left(1 + \frac{x}{\rho_x} + \frac{z}{\rho_z}\right) \quad .$$

The quantity $1/\rho^2$ in P_γ is proportional to the square of the field B^2. Taking into account the possible presence of field gradients, we must develop B^2 in a series for x and z. Keeping only first order terms,

$$\frac{B^2}{B_0^2} = 1 + G_x x + G_z z \quad , \tag{3.10}$$

with $\quad G_x = \frac{2}{B_0^2}\left(B_z \frac{\partial B_z}{\partial x} - B_x \frac{\partial B_z}{\partial z}\right)$ and $G_z = \frac{2}{B_0^2}\left(B_x \frac{\partial B_z}{\partial x} + B_z \frac{\partial B_z}{\partial z}\right)$.

Regrouping all the first order terms in x and z which appear in the product $P_\gamma \Delta\ell$, we obtain :

$$\frac{1}{c} P_\gamma \Delta\ell = \frac{1}{c} P_{\gamma_0} \Delta s \, (1 + C_x x + C_z z) \quad , \tag{3.11}$$

with $C_x = 1/\rho_x + G_x$ and $C_z = 1/\rho_z + G_z$. On the right hand side, P_{γ_0} is calculated on the central trajectory $x = z = 0$ with nominal field B_0.

Putting all these results together and replacing x and z by their development in eigenfunctions (Eq.(2.18)), the expression we are looking for is :

$$N\Delta t \, \langle \varepsilon \rangle = \Delta t \, P_{\gamma_0} \left[1 + \sum_{k=1}^{4} A_k \, (C_x w_{1k} + C_z w_{3k}) \right] \quad . \tag{3.12}$$

b) If Q_ε stands for the mean value of the product $N\langle\varepsilon^2\rangle$, we have simply :

$$N\Delta t \, \langle \varepsilon^2 \rangle = \Delta t Q_\varepsilon \tag{3.13}$$

with

$$Q_\varepsilon = \frac{55}{24\sqrt{3}} r_e \hbar c^2 \gamma^6 \, E \, \frac{1}{\rho^3} \quad .$$

In the expression (3.6) for the amplitude variation, we still need to evaluate the function $F(\delta Y_j, w_{jk}) = \varepsilon F(D_j/E_0, w_{jk})$, and the similar one $F(\delta Y_j, w_{jk}^*)$. This is simple if we make use of the definition (3.1) :

$$F(D_j/E_0, w_{jk}) = \frac{1}{E_0} \left(D_x w_{2k} - D_x' w_{1k} + D_z w_{4k} - D_z' w_{3k} \right) \quad . \tag{3.14}$$

Let us now turn to the question of longitudinal acceleration δE (Eq.(3.8)). For the same reasons previously evoked, linear and quadratic terms in δE will be present in Eq.(3.6). If the quadratic term δE^2 becomes negligible towards the limit $\Delta t \to 0$, the linear term averages to (using eigenfunctions again) :

$$\delta p_y = - \frac{\langle \delta E \rangle}{E_0} p_y = - \frac{\langle \delta E \rangle}{E_0} \sum_{k=1}^{4} w_{\substack{2k \\ 4k}} (\theta) \, A_k \quad . \tag{3.15}$$

Introducing Eq.(3.15) in the linear terms of Eq.(3.6) and assuming logically that the average $\langle \delta E \rangle$ must exactly compensate for the radiation loss, gives :

$$\delta \left| A_k \right|^2_{\text{accel.}} = - \frac{P_{\gamma_0} \Delta t}{E_0} \left| A_k^2 \right| \quad . \tag{3.16}$$

Putting together Eqs.(3.12), (3.13) and (3.14) into Eq.(3.6) for the photon emission effect and adding the contribution (3.16) of the acceleration, we can derive the following expression for the amplitude variation :

$$\langle \delta |A_k|^2 \rangle = - \frac{P_\gamma \Delta t}{E_0} |A_k^2|$$

$$- \frac{P_\gamma \Delta t}{E_0} \frac{|A_k^2|}{F(w_{jk}, w_{jk}^*)} 2i \langle Im[(C_x w_{1k}^* + C_z w_{3k}^*)(D_x w_{2k} - D_x' w_{1k} + D_z w_{4k} - D_z' w_{3k})] \rangle$$

$$- \frac{Q_\epsilon \Delta t}{E_0^2 F^2(w_{jk}, w_{jk}^*)} \langle |D_x w_{2k} - D_x' w_{1k} + D_z w_{4k} - D_z' w_{3k}|^2 \rangle \quad . \tag{3.17}$$

The equality $A^*F(\delta Y, w^*) - AF(\delta Y, w) = -2i\ Im[A^*F(\delta Y, w^*)]$ has been applied and the subscript $_0$ of P_γ abandonned for simplicity. The three terms of (3.17) give respectively the amplitude variation associated with the acceleration, radiation damping and quantum excitation.

4. EQUILIBRIUM IN THE CASE OF BETATRON COUPLING WITH RADIATION

4.1 Equilibrium amplitudes

The finite amplitude variations with radiation for a finite time interval Δt are written explicitely in Eq.(3.17). On the limit of infinitesimal interval ($\Delta t \to dt$), Eq.(3.17) becomes with the usual notation[7] :

$$\frac{d|A_k|^2}{dt} = - 2\alpha_k |A_k^2| + Q_k \quad , \tag{4.1}$$

where α_k are the damping coefficients

$$\alpha_k = \langle \frac{P_\gamma}{2E_0} \rangle J_k \quad , \tag{4.2}$$

which are proportional to the damping partition numbers J_k

$$J_k = 1 + \langle \frac{Im[(C_x w_{1k}^* + C_z w_{3k}^*)(D_x w_{2k} - D_x' w_{1k} + D_z w_{4k} - D_z' w_{3k})]}{Im(w_{1k} w_{2k}^* + w_{3k} w_{4k}^*)} \rangle \quad , \tag{4.3}$$

and where Q_k are the transverse beam amplitude coefficients

$$Q_k = \langle \frac{1}{4} \frac{Q_\epsilon}{E_0^2} \rangle \langle \frac{|D_x w_{2k} - D_x' w_{1k} + D_z w_{4k} - D_z' w_{3k}|^2}{Im^2(w_{1k} w_{2k}^* + w_{3k} w_{4k}^*)} \rangle \quad . \tag{4.4}$$

These relations use the identity $F(w_{jk}, w_{jk}^*) = 2i Im(w_{1k} w_{2k}^* + w_{3k} w_{4k}^*)$.

Within the assumption made in Section 3.1, a stationary state will occur after a few damping times and it corresponds to the condition $d\,|A_k|^2/dt = 0$. Hence, the equilibrium amplitudes are (Eq.(4.1)) :

$$|A_k^2| = \frac{Q_k}{2\alpha_k} \quad , \tag{4.5}$$

and this is the important result to be used in Eq.(2.18). Betatron coupling is present through the eigenfunctions (Eq.(2.19)), the fact that both vertical and horizontal dispersions have an influence and the need of four coefficients (k = 1 to 4) in order to describe the whole motion.

4.2 Equilibrium emittances

Betatron oscillations are characterized by the transverse invariants of the motion, which define the commonly used emittances E_y. If by definition E_y represents the invariant mean-square amplitudes of the transverse oscillations, we must have :

$$E_y = \frac{\langle y^2 \rangle}{\beta_y} \quad . \tag{4.6a}$$

Starting from the solution (2.18) and using the eigenfunctions (2.19) as well as the equalities between complex conjugates, we can rewrite Eq.(4.6a) in the following manner :

$$E_y = \frac{\langle y^2 \rangle}{\beta_y} = \frac{2}{\beta_y} \left(|A_1^2| |w_{\substack{1\,1\\3\,1}}|^2 + |A_3^2| |w_{\substack{1\,3\\3\,3}}|^2 \right) \quad , \tag{4.6b}$$

the two subscripts of the w's being associated with the horizontal and vertical coordinates, respectively.

In order to simplify the next calculations, let us now assume that the accelerator or storage ring of interest is large and has separated functions. This means that the radius of curvature ρ is large and there is no gradient in the dipoles ($C_x \simeq C_z \simeq 0$).

Consequently, $\langle D/\rho \rangle$ is small with respect to 1 and Eq.(4.2) becomes simply :

$$\alpha_k = \left\langle \frac{P_\gamma}{2E_0} \right\rangle \qquad (\text{with } J_k = 1) \ . \tag{4.7}$$

Hence, putting together Eqs.(4.4) to (4.7) and introducing the explicit forms (2.19) of the eigenfunctions make it possible to write the emittances as follows :

$$E_x = \left\langle \frac{Q_\varepsilon}{RE_0P_\gamma} \right\rangle \left[\frac{\omega_+^2 |\kappa|^2}{(\omega_+^2 + |\kappa|^2)^2} \langle H_1 \rangle + \frac{\omega_-^2 |\kappa|^2}{(\omega_-^2 + |\kappa|^2)^2} \langle H_3 \rangle \right]$$

$$E_z = \left\langle \frac{Q_\varepsilon}{RE_0P_\gamma} \right\rangle \left[\frac{\omega_+^4}{(\omega_+^2 + |\kappa|^2)^2} \langle H_1 \rangle + \frac{\omega_-^4}{(\omega_-^2 + |\kappa|^2)^2} \langle H_3 \rangle \right] \quad , \tag{4.8}$$

where P_γ and Q_ε are given below Eqs.(3.9) and (3.13), respectively. It remains to define the functions H_1 and H_3 which are simply the numerator of the second bracket in Eq.(4.4) of Q_1 and Q_3 :

$$H_{\frac{1}{3}} = \frac{|\kappa|^2}{\omega_\pm^2} \frac{R}{2\beta_x} \left[D_x^2 + \frac{1}{R^2}(\beta_x D_x' + R\alpha_x D_x)^2\right] + \frac{R}{2\beta_z}\left[D_z^2 + \frac{1}{R^2}(\beta_z D_z' + R\alpha_z D_z)^2\right]$$

$$+ \frac{R}{\omega_\pm \sqrt{\beta_x \beta_z}} \mathrm{Re}\left[\kappa^*(\alpha_x - i)(\alpha_z + i)\exp(i\phi)\right] D_x D_z + \frac{\sqrt{\beta_x \beta_z}}{\omega_\pm R} \mathrm{Re}\left[\kappa^* \exp(i\phi)\right] D_x' D_z'$$

$$+ \frac{1}{\omega_\pm}\sqrt{\frac{\beta_z}{\beta_x}} \mathrm{Re}\left[\kappa^*(\alpha_x - i)\exp(i\phi)\right] D_x D_z' + \frac{1}{\omega_\pm}\sqrt{\frac{\beta_x}{\beta_z}} \mathrm{Re}\left[\kappa^*(\alpha_z - i)\exp(i\phi)\right] D_x' D_z \quad . \quad (4.9)$$

The subscripts of H are associated with the two indices of ω (see Eq.(2.17)), and the phase ϕ is written for :

$$\phi = \phi_x - \phi_z - \theta\Delta \quad . \quad (4.10)$$

The two first terms of (4.9) are directly proportional to the dispersion invariants (named I_x and I_z), which appear naturally in E_x and E_z respectively[7]), in the absence of coupling. The other terms are obviously coupled terms for the dispersion.

Turning back to the expressions of the equilibrium emittances (4.8) and (4.9), which appear to be fairly complicated, let us look at two borderline cases :

a) For vanishing linear coupling ($\kappa \to 0$), the four terms in the square brackets of Eq.(4.8) have finite limits. Two of them are equal to zero, while the two remaining ones become equivalent to :

$$\frac{\omega_+^2 |\kappa|^2}{(\omega_+^2 + |\kappa|^2)^2} \langle H_1 \rangle = \frac{R}{2} \langle I_x \rangle \; ; \quad \frac{\omega_-^4}{(\omega_-^2 + |\kappa|^2)^2} \langle H_3 \rangle = \frac{R}{2} \langle I_z \rangle \quad . \quad (4.11)$$

In this case, the transverse equilibrium emittances and their ratio $g = E_z/E_x$ are simply given by :

$$E_y = \left\langle \frac{Q_\varepsilon}{2E_0 P_\gamma} \right\rangle \langle I_y \rangle$$

$$g = \frac{\langle I_z \rangle}{\langle I_x \rangle} \quad . \quad (4.12)$$

As expected, the vertical emittance is nul, if the vertical dispersion vanishes in addition to κ.

b) For vanishing vertical dispersion (but $\kappa \neq 0$), the functions (4.9) can be written as follows :

$$H_{\frac{1}{3}} = \frac{|\kappa|^2}{\omega_{\pm}^2} \frac{R}{2} \langle I_x \rangle \quad . \tag{4.13}$$

Introducing Eq.(4.13) into Eq.(4.8) gives for the emittances and their ratio g :

$$E_x = \left\langle \frac{Q_\varepsilon I_x}{2E_0 P_\gamma} \right\rangle \frac{2\left(\frac{|\kappa|}{\Delta}\right)^2 + 1}{4\left(\frac{|\kappa|}{\Delta}\right)^2 + 1}$$

$$E_z = \left\langle \frac{Q_\varepsilon I_x}{2E_0 P_\gamma} \right\rangle \frac{2\left(\frac{|\kappa|}{\Delta}\right)^2}{4\left(\frac{|\kappa|}{\Delta}\right)^2 + 1} \tag{4.14}$$

$$g = \frac{(|\kappa|/\Delta)^2}{(|\kappa|/\Delta)^2 + 1/2} \quad .$$

The corresponding curves for E_x and g are plotted as functions of the ratio $|\kappa|/\Delta$ in Fig. 2. Both Eq.(4.14) and Fig. 2 show that on the limit $|\kappa| \gg \Delta$ (sometimes called full coupling) the transverse emittances are equal and take half the value of the horizontal emittance at $\kappa = 0$.

In general, coupling and vertical dispersion are not vanishing, so that not only I_x and I_z are contributing but also the products $D_x D_z$, $D_x' D_z'$, $D_x D_z'$ and $D_x' D_z$ in agreement with expression (4.9).

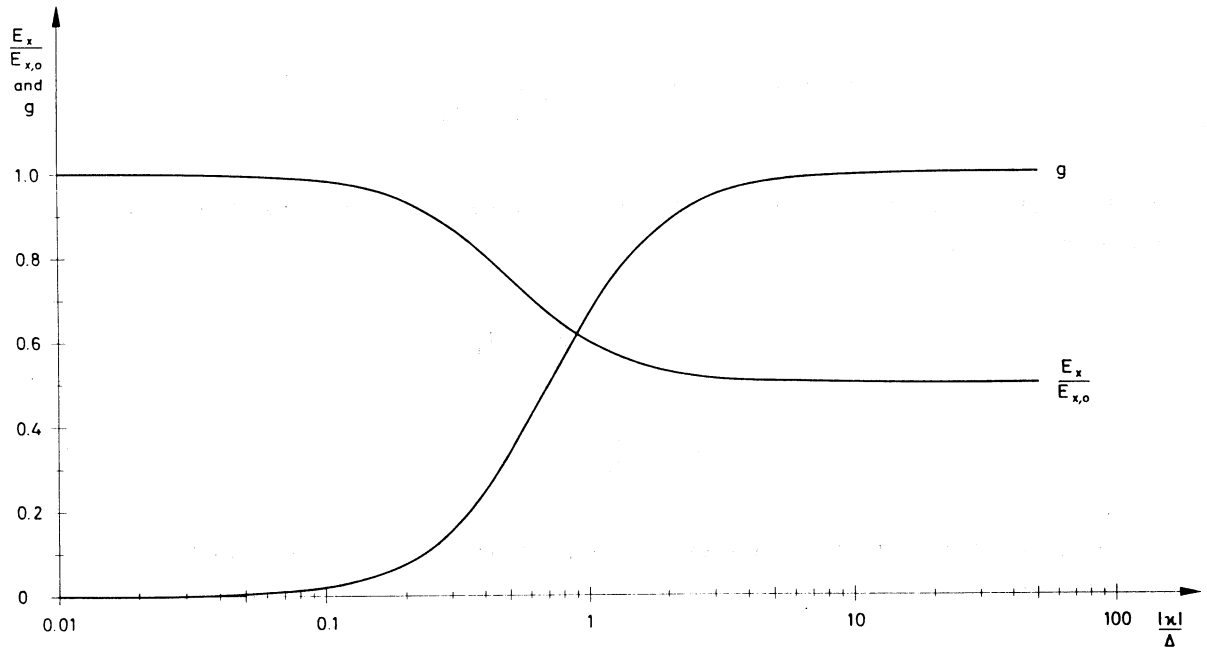


Fig. 2 Horizontal emittance and emittance ratio as function of κ/Δ, with vanishing vertical dispersion

5. APPLICATION TO EMITTANCE CONTROL

The expressions for the emittances in the presence of betatron coupling and radiation allow an estimation of the range in which the emittance ratio g can be controlled[8].

Let us consider the particular example of the LEP design[9] for which the assumptions made in the previous section are valid. Indeed, the tune values are $Q_x = 70.35$ and $Q_z = 78.20$, the radius of curvature is large and there are separated functions. The values of the dispersion invariants have been estimated to be $\langle I_x \rangle = 1.75\ 10^{-2}$ m and $\langle I_z \rangle \simeq 1.8\ 10^{-4}$ m, the vertical dispersion being associated with the expected machine imperfections and closed orbit after correction[9]. It is also important to mention that, in the LEP design, there exist four tilted-quadrupole schemes located in sections where $D_x = 0$ (for the perfect ring) in order to minimize the effect on D_z of a coupling compensation. These schemes are made of four tilted quadrupoles on either side of every experiment (outside the experimental free space) for compensating the solenoidal fields present at these points and controlling the residual machine coupling.

Let us estimate firstly what is the minimum emittance ratio which can be reached after an orbit correction achieving a root-mean-square deviation of about 1 mm. The best we can do consists of compensating exactly the linear coupling such as $\kappa = 0$. Equations (4.12) then apply directly and give for our particular LEP example : $g_{min} \simeq 1\%$.

Secondly, let us try to determine the upper limit of g which can be reached either by trimming the tilted quadrupoles (changing the coefficient κ) or by modifying the parameter Δ (i.e. the distance from the coupling resonance). To increase g above its minimum of 1%, one has to inflate $|\kappa|$. When g reaches a value of the order of a few percent, the κ-contributions dominate the D_z-contributions in Eq.(4.9) in such a manner that Eq.(4.14) and the associated curve of Fig. 2 give a good approximation for the change of g with $|\kappa|/\Delta$. It follows from these results that the strengths of the tilted quadrupoles are sufficient at the energy of 55 GeV to reach $g_{max} \simeq 12.5\%$ for $\Delta = 0.15$. Changing then the tune values until $\Delta = 0.012$, it is even possible to obtain $g_{max} \simeq 96\%$. Remembering that the nominal value is g = 4% in LEP, the present analysis allowed us to show that it is possible to cover a sufficiently large range of values around the nominal one and also to get full coupling if necessary, while compensating simultaneously for the linear effects of the experimental solenoids.

* * *

REFERENCES

1. G. Guignard, The general theory of all sum and difference resonances in a three-dimensional magnetic field in a synchrotron, CERN 76-06 (1976).
2. G. Guignard, A general treatment of resonances in accelerators, CERN 78-11 (1978).
3. G. Leleux, Influence of linear coupling on radiation damping and equilibrium beam sizes in electron storage ring, Proc. 5th Int.Conf. on High-Energy Accelerators, Frascati, 1965 (CNEN, Rome, 1966) p. 286.
4. A. Piwinski, private communication.
5. P.J. Bryant and G. Guignard, Methods for measuring the complex coupling coefficient for the second-order difference resonance, CERN ISR-MA/75-42 (1975).
6. P.J. Bryant, P. Galbraith, J.P. Gourber, G. Guignard, K. Takikawa, Measurement of the excitation of the coupling resonance, Part. Acc. Conf., Chicago, 1977 (IEEE, New York, 1977).
7. M. Sands, The physics of electron storage rings, An introduction, SLAC-121 (1970).
8. G. Guignard, Adjustment of emittance ratio by coupling control in electron-positron storage rings, Proc. 11th Int. Conf. on High-Energy Accelerators, CERN, Geneva, 1980 (Birkhäuser, Basle, 1980), p. 682.
9. LEP Design Report, Vol. II, The LEP main ring, CERN-LEP/84-01 (1984).

KINETIC THEORY AND THE VLASOV EQUATION

C N Lashmore-Davies

Culham Laboratory, Abingdon, Oxon. OX14 3DB, UK

(Euratom/UKAEA Fusion Association)

ABSTRACT

An outline of the kinetic theory of plasmas is given starting from the N-particle distribution. The derivation of the BBGKY system of equations is discussed emphasising the physical arguments rather than the mathematical details. It is shown how this system of equations leads to the Vlasov equation. The derivation of this equation introduces the important idea of the self-consistent field. A brief discussion of some of the properties of the Vlasov equation is then given. Binary collision processes are then considered beginning with a description of various relaxation models. The more accurate Fokker-Planck equation is then discussed which is appropriate to small angle scatterings, the emphasis again being given to the physical justification and content of the equation. Finally, a brief account is given of how the kinetic theory can be used to derive the various fluid models of plasmas.

1. INTRODUCTION

The kinetic theory is the most fundamental aspect of the mathematical description of plasmas. The aims of the theory are to start from an assembly of N charged particles moving under the influence of their own charges and whatever external fields might be present and to derive the dynamical properties of plasmas. There is a vast body of research on the foundations of the kinetic theory of plasmas and an even vaster body of work which rests on these results. In this lecture the aim will be to concentrate on the physical content of the kinetic theory, indicating the assumptions and approximations that have gone into the derivations without going into too much mathematical detail.

The kinetic theory of plasmas is a truly N-body problem since due to the long range nature of the Coulomb interaction between the charged particles making up the plasma every particle interacts with every other particle. Thus, the main task for a satisfactory kinetic theory of plasma is to find a convincing way of dealing with this situation. The chapter will be divided into three parts. In the first part the derivation of the Vlasov equation will be outlined and a brief description of the physical content and properties of this equation will be given. In the second part, the way in which binary collision processes are incorporated into the kinetic theory will be discussed and in the last part we shall show how the fluid treatment of plasmas can be derived from the kinetic description.

2. THE VLASOV EQUATION

In order to appreciate the physical content of Vlasov's equation we need to start from the general position of an assembly of N charged particles. Of course, N is so large (for typical laboratory plasmas $N \sim 10^{13} - 10^{14}$ cm^{-3}) that even if we had perfect diagnostics of the positions and velocities of all the particles at some instant of time we would be unable to deal with such a wealth of information. We shall confine the present discussion to a one-component plasma in which the ions can be thought of as immobile serving only to neutralize the electron charge in equilibrium. We shall also ignore magnetic interactions and consider only the Coulomb interaction. We begin by writing down the Hamiltonian for the N-electron system where the i^{th} electron has six phase-space coordinates $\underset{\sim}{x} \equiv (\underset{\sim}{q}_i, \underset{\sim}{p}_i)$, where the $\underset{\sim}{q}_i$ and $\underset{\sim}{p}_i$ are canonical Hamiltonian variables.

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i<j=1}^{N} \frac{e^2}{4\pi\varepsilon_0 \left| \underset{\sim}{q}_i - \underset{\sim}{q}_j \right|} \quad . \tag{2.1}$$

Since we are not interested in the detailed motion of all N-particles we need some averaging procedure to enable us to reduce the complexity of the system. We now introduce the N-electron probability distribution $D_N(\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_N, t)$ normalized to unity which describes the state of the N-electron system at any instant of time. We are interested in the evolution of the system in time and the evolution of D_N is given by Liouville's theorem

$$\frac{\partial D_N}{\partial t} + \{D_N, H_N\} = 0 \tag{2.2}$$

where $\{\ \}$ denotes the usual Poisson bracket. Starting from this equation there is a systematic technique for reducing equation (2.2) for the N-electron probability to a one- or two-electron distribution function. This procedure leads to the BBGKY hierarchy of equations (after the names of its originators), details of which can be found elsewhere[1,2]. Here we shall content ourselves with a brief description of the steps in the argument.

The first step towards simplification is the definition of a reduced probability distribution f_s/V^s where V is the configuration space volume of the system and

$$\frac{f_s}{V^s} = \int D_N \, d\underset{\sim}{x}_{s+1} \, d\underset{\sim}{x}_{s+2} \cdots d\underset{\sim}{x}_{s+N} \tag{2.3}$$

where the function f_s is now a distribution involving only s-particles. With the aid of equation (2.2) we can obtain the equation for f_s which is

$$\frac{\partial f_s}{\partial t} + \sum_{i=1}^{s} \frac{\underset{\sim}{p}_i}{m} \cdot \frac{\partial f_s}{\partial \underset{\sim}{q}_i} - \sum_{i=1}^{s} \sum_{j=1}^{s} \frac{\partial \phi_{ij}}{\partial \underset{\sim}{q}_i} \cdot \frac{\partial f_s}{\partial \underset{\sim}{p}_i} = n_0 \sum_{i=1}^{s} \int \frac{\partial \phi_{i\,s+1}}{\partial \underset{\sim}{q}_i} \cdot \frac{\partial f_{s+1}}{\partial \underset{\sim}{p}_i} \, d\underset{\sim}{x}_{s+1} \tag{2.4}$$

where $\phi_{ij} \equiv e^2 / \left| \underset{\sim}{q}_i - \underset{\sim}{q}_j \right|$ $\quad j \neq i$ and n_0 is the particle density. No approximation has been made to obtain this result and we see that the equation determining f_s contains f_{s+1}

and thus the equation for f_{s+1} contains f_{s+2} and so on back to D_N. We do not, therefore, appear to be any better off. However, we now appeal to a physical argument to justify the use of simplifying approximations. In order to deal with this problem of closure the Mayer cluster expansion is often used:

$$\begin{aligned} f_1(\underset{\sim}{x}_1) &= f_1(\underset{\sim}{x}) \\ f_2(\underset{\sim}{x}_1, \underset{\sim}{x}_2) &= f_1(\underset{\sim}{x}_1)f_1(\underset{\sim}{x}_2) + P(\underset{\sim}{x}_1, \underset{\sim}{x}_2) \end{aligned} \tag{2.5}$$

and so on. In view of the long range nature of the Coulomb interaction one might expect that $P(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ would be comparable in magnitude to $f_1(\underset{\sim}{x}_1)f_1(\underset{\sim}{x}_2)$ i.e. pairs of particles would be strongly correlated. However, the Coulomb interaction results in particles moving so as to preserve quasi-neutrality. Each particle attracts a screening charge cloud of the opposite sign so as to neutralize its own charge. The effect of this screening is that over almost the whole of phase space P/f_1f_1 is a small quantity of the order $g \equiv (n_0\lambda_D^3)^{-1}$ where $\lambda_D^2 \equiv \varepsilon_0 kT/n_0e^2$ and ε_0, k, T and e are the dielectric constant of vacuum, Boltzmann's constant, the electron temperature and electron charge. [$n_0\lambda_D^3 >> 1$ is often taken as the definition of a plasma - a typical value for a hot laboratory plasma is $n_0 \sim 10^{14}$ cm^{-3}, $T_e \sim 10^6$°K giving $n_0 \lambda_D^3 \sim 10^4$.]

We can now use this as a basis for simplifying equation (2.4). Assuming that three particle correlations are even weaker than two particle correlations we can take $s = 2$ and write equation (2.4) as

$$\begin{aligned} \frac{\partial f_1}{\partial t} + \underset{\sim}{p}_1 \cdot \frac{\partial f_1}{\partial \underset{\sim}{q}_1} &= n_0 \int \frac{\partial \phi_{12}}{\partial \underset{\sim}{q}_1} \cdot \frac{\partial}{\partial \underset{\sim}{p}_1} f_1(\underset{\sim}{x}_1)f_1(\underset{\sim}{x}_2)d\underset{\sim}{x}_2 \\ &+ n_0 \int \frac{\partial \phi_{12}}{\partial \underset{\sim}{q}_1} \cdot \frac{\partial}{\partial \underset{\sim}{p}_1} P(\underset{\sim}{x}_1, \underset{\sim}{x}_2)\, d\underset{\sim}{x}_2 \quad . \end{aligned} \tag{2.6}$$

We do not write down the equation for $P(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ but since we neglect three-particle correlations equation (2.6) together with the equation for $P(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ now form a closed system. However, even this system is too complicated for general use and we are forced to approximate further. First note the form of equation (2.6). The first term on the right hand side of the equation describes the uncorrelated motions and the second is due to two particle correlations. We shall see later that the second term describes the effect of binary collision processes.

The effect of the screening is such that two-particle correlations are very weak except over a very small region of phase space of volume $\sim g^2/n_0$. In view of this we now aim to describe the system in terms of a 'smoothed-out' distribution which should be a good approximation over almost the whole of phase space. This smoothing procedure can be appreciated by returning to equation (2.4). If the equation is non-dimensionalized in terms of the natural time and length scales for the system of ω_p^{-1} and λ_D where $\omega_p^2 = n_0e^2/\varepsilon_0 m$ it will be found that the third term on the left hand side of the equation is of

order g . Neglect of this term then allows the equation to be solved with the uncorrelated function

$$f_s(\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_s) = f_1(\underset{\sim}{x}_1) f_1(\underset{\sim}{x}_2) \ldots f_1(\underset{\sim}{x}_s) \quad . \tag{2.7}$$

We also note that f_s is the distribution function of s-test particles. If we assume that the separation of these s-particles is large then $\left|\underset{\sim}{q}_i - \underset{\sim}{q}_j\right|^{-1} \to 0$ for $i, j = 1, 2 \ldots s$, $i \neq j$ which again allows us to obtain the uncorrelated solution, equation (2.7).

The smoothing procedure is equivalent to allowing e , m and $1/n_0$ to become very small ($\sim g$) while at the same time keeping e/m , kT/m and $n_0 e^2/m$ constant. In the continuum limit $g = 0$ the charges and masses are completely smoothed out and we have a continuous, six-dimensional phase fluid. Instead of a singular distribution function consisting of a set of δ-functions we now have a smooth, continuously differentiable distribution function. In the limit $g = 0$, equation (2.6) reduces to

$$\frac{\partial f}{\partial t} + \underset{\sim}{v}_1 \cdot \frac{\partial f}{\partial \underset{\sim}{x}_1} - \frac{n_0}{m} \left[\int \frac{\partial \phi_{12}}{\partial \underset{\sim}{x}_1} f(\underset{\sim}{x}_2, \underset{\sim}{v}_2) \, d\underset{\sim}{x}_2 \, d\underset{\sim}{v}_2 \right] \cdot \frac{\partial f}{\partial \underset{\sim}{v}_1} = 0 \tag{2.8}$$

where we have changed from q , p variables to $\underset{\sim}{x}$, $\underset{\sim}{v}$ where $\underset{\sim}{x}$ is now the position coordinate and $\underset{\sim}{v}$ the velocity. In contrast to a normal fluid the six-dimensional phase fluid described by equation (2.8) contains the thermal information relating to the distribution of velocities. The final step in this procedure is to interpret the third term in equation (2.8).

In order to do this we note that the electric field acting on the electron at position $\underset{\sim}{x}_1$ at time t , can be written

$$e \, \underset{\sim}{E}(\underset{\sim}{x}_1, t) = n_0 \int \frac{\partial \phi_{12}}{\partial \underset{\sim}{x}_2} f(\underset{\sim}{x}_2, \underset{\sim}{v}_2, t) \, d\underset{\sim}{x}_2 \, d\underset{\sim}{v}_2 \quad . \tag{2.9}$$

Making explicit use of the fact that ϕ_{12} is a Coulomb interaction and taking the divergence of equation (2.9) we obtain from (2.8) and (2.9) the final result

$$\frac{\partial f}{\partial t} + \underset{\sim}{v} \cdot \frac{\partial f}{\partial \underset{\sim}{x}} - \frac{e}{m} \underset{\sim}{E}(\underset{\sim}{x}, t) \cdot \frac{\partial f}{\partial \underset{\sim}{v}} (\underset{\sim}{x}, \underset{\sim}{v}, t) = 0 \tag{2.10}$$

$$\underset{\sim}{\nabla} \cdot \underset{\sim}{E} = \frac{n_0 e}{\varepsilon_0} \int f(\underset{\sim}{x}, \underset{\sim}{v}, t) \, d\underset{\sim}{v} \tag{2.11}$$

where we have now dropped the subscript 1 from $(\underset{\sim}{x}, \underset{\sim}{v})$ since this is now redundant. Equation (2.10) is the required Vlasov equation which forms the basis of most of plasma physics. The Vlasov equation is coupled to Poisson's equation through the electric field $\underset{\sim}{E}$. The electric field appearing in Vlasov's equation is the SELF-CONSISTENT field due to all the other electrons in the system moving in an uncorrelated way, i.e. uncorrelated to any individual particle but correlated to all particles by the self-consistent field. The existence of the self-consistent field is the reason why a plasma is so rich in its

collective effects and is the solution to the problem posed at the outset - namely, how to take account of the fact that every particle could interact with every other. The concept of the self-consistent field is the same as that employed in the Hartree model of atomic structure. It may be viewed as an iterative solution to the Vlasov-Poisson system, i.e. assume a distribution f, calculate the resulting electric field $\underset{\sim}{E}$ and then determine the correction to f and so on until the electric field resulting from f produces the same distribution function.

Before discussing the properties of the Vlasov equation we note that the above derivation used the concept of Debye screening applied to a neutral plasma. For a non-neutral plasma it is not obvious that a similar concept would still hold. A discussion of electrostatic shielding of a test electron embedded in a magnetically confined pure electron gas column has been given by Davidson.[3] He concluded that for rigid rotor equilibria with nearly uniform density the shielding length is the local Debye length for the case that the test electron is at rest relative to the mean motion of the column. A further condition is that there should be many Debye lengths across the electron gas column. The properties of non-neutral plasmas are also discussed by Malmberg and deGrassie[4].

2.1 Properties of the Vlasov Equation

The discussion which leads to the Vlasov-Poisson system of equations given by (2.10) and (2.11) is easily generalized to include magnetic interactions. The complete Vlasov-Maxwell system of equations is then

$$\frac{\partial f}{\partial t} + \underset{\sim}{v} \cdot \frac{\partial f}{\partial \underset{\sim}{x}} - \frac{e}{m} \left(\underset{\sim}{E}(\underset{\sim}{x}, t) + \underset{\sim}{v} \times \underset{\sim}{B}(\underset{\sim}{x}, t) \right) \cdot \frac{\partial f}{\partial \underset{\sim}{v}} = 0 \tag{2.12}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{H} = \underset{\sim}{J} + \varepsilon_0 \frac{\partial \underset{\sim}{E}}{\partial t} \tag{2.13}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{H} = -\mu_0 \frac{\partial \underset{\sim}{H}}{\partial t} \tag{2.14}$$

$$\underset{\sim}{J}(\underset{\sim}{x}, t) = -e \int f(\underset{\sim}{x}, \underset{\sim}{v}, t)\, \underset{\sim}{v}\, d\underset{\sim}{v} \quad . \tag{2.15}$$

Noting that the "particle orbit" corresponding to equation (2.12) is

$$\frac{d}{dt} \underset{\sim}{x}(t) = \underset{\sim}{v}(\underset{\sim}{x}, t) \tag{2.16}$$

$$\frac{d}{dt} \underset{\sim}{v}(\underset{\sim}{x}, t) = -\frac{e}{m} \left(\underset{\sim}{E}(\underset{\sim}{x}, t) + \underset{\sim}{v}(\underset{\sim}{x}, t) \times \underset{\sim}{B}(\underset{\sim}{x}, t) \right) \tag{2.17}$$

we can write the equation

$$\frac{d}{dt} f(\underset{\sim}{x}, \underset{\sim}{v}, t) = 0 \tag{2.18}$$

thus enabling the equation to be integrated once we know the particle orbit (method of

characteristics). The Vlasov equation thus appears as the analogue of Liouville's equation in the phase space of a single particle. As a general method of solution equation (2.18) is not very helpful since the orbit is in the self-consistent electromagnetic field which is not known until f is determined. However, the method is very useful in small amplitude solutions when the equation can be linearized. The integration is then carried out routinely along the unperturbed or equilibrium orbit. Equilibrium solutions to equation (2.12) are also obtained as functions of the constants of the motion.

A fundamental property of Vlasov's equation is that it is time reversible, i.e. if all velocities were reversed and time was made to run backwards, all previous motions would be reproduced. In other words the equation is invariant under the transformation

$$\underset{\sim}{v} \to -\underset{\sim}{v} \quad , \quad \underset{\sim}{x} \to \underset{\sim}{x} \quad , \quad t \to -t \quad .$$

This is because the random binary collisions are not included in the equation. It is these collisions which introduce physical dissipation and irreversibility into the system. The interactions which are responsible for the self-consistent field cannot be interpreted as collisions in the usual sense. This is because the effect is produced by a large number of particles and is therefore a macroscopic effect. Since an average over a large number of particles is involved the process becomes certain rather than random and cannot cause an increase in the entropy of the system. This can be demonstrated explicitly by means of an H-theorem where

$$H(t) = \int f \ln f \, d\underset{\sim}{v} \, d\underset{\sim}{x} \tag{2.19}$$

and it can be shown that Vlasov's equation then satisfies

$$\frac{dH(t)}{dt} = 0 \quad . \tag{2.20}$$

One final property of the Vlasov equation worth mentioning is the following. Referring to the case of electrostatic interactions any $f_{0j}(v)$ will be an equilibrium solution of equation (2.10) provided only

$$\sum_j q_j \int f_{0j}(\underset{\sim}{v}) \, d\underset{\sim}{v} \, d\underset{\sim}{x} = 0 \quad . \tag{2.21}$$

In particular the Vlasov equation does not single out the Maxwell-Boltzmann distribution. The effect of binary collisions must be included to ensure that the distribution will relax to a Maxwell distribution. The above observations should be borne in mind when discussing the phenomenon of collisionless damping or Landau damping as it is usually known.

3. THE EFFECT OF BINARY COLLISIONS

The Vlasov equation is often called the collisionless Boltzmann equation. In the

justification of this equation with the aid of equation (2.6) we neglected the two particle correlation function $P(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ on the grounds of the effect of screening. However, as mentioned in the previous section it is the binary collisions which introduce dissipation into the system which in turn is reponsible for classical transport phenomena. Let us now consider how the effect of binary collisions is incorporated into the kinetic theory. The discussion will be divided into two parts. In the first part we shall describe the simplest method of allowing for collisions through the introduction of a relaxation time. In the second part we shall discuss the more accurate Fokker-Planck model. This model applies specifically to a system, such as a plasma, where most of the scatterings are through small angles.

3.1 Relaxation Models

The simplest model of all arises through the introduction of a relaxation time τ which in general will be velocity dependent. If we consider only changes to the distribution function due to binary collisions, which we write $(\partial f/\partial t)_{coll}$, then the relaxation model is

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = \frac{(f_0 - f)}{\tau(v)} \tag{3.1}$$

where f_0 is the equilibrium distribution. Clearly, any small departure from the equilibrium distribution will decay exponentially on a timescale proportional to τ. The effect of collisional relaxation can be incorporated into the Vlasov equation simply by replacing the zero on the right hand side of the equation with the above relaxation term.

This simple relaxation model, however, suffers from the defect that it does not conserve particle number as may be seen by integrating equation (3.1) over velocity space. For an oscillatory process number would be conserved on average but not instantaneously. This defect can be remedied by the Krook[5] model which is

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\frac{f}{\tau} + \frac{n(x, t)}{n_0}\frac{f_0}{\tau} \tag{3.2}$$

and n_0 is the equilibrium density. Clearly, equation (3.2) does now conserve number density instantaneously. The interpretation of equation (3.2) is that particles are absorbed at a rate proportional to f and re-emitted at a rate proportional to the local density with a Maxwell distribution. The particle distribution is therefore brought to rest as a whole with the result that although conserving particle number, equation (3.2) does not conserve momentum or energy. This does not mean that this model is therefore of no value. It applies, for example, to electron neutral collisions where the momentum and energy would be lost from the electron fluid while the electron number would be conserved. Such a model is therefore appropriate to a weakly ionized plasma. Indeed, since an atom is so much more massive than an electron the electron would effectively be brought to rest.

The above collision model is linearly proportional to f which is appropriate to electron-neutral collisions. However, for a fully ionized plasma where electron-electron or ion-ion collisions will be important such collisions will be proportional to f^2. In addition, for these collisions momentum and energy must be conserved. Krook et al[5] have developed a simplified model to take account of such collisions as follows

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\frac{n(x,\ t)}{n_0}\frac{f}{\tau} + \frac{n(x,\ t)^2}{n_0}\frac{\Phi}{\tau} \tag{3.4}$$

where $$\Phi = \left[\frac{m}{2\pi kT(\underset{\sim}{x},\ t)}\right]^{3/2} \exp\left\{-\frac{m}{2kT(\underset{\sim}{x},\ t)}\left[\underset{\sim}{v} - \underset{\sim}{q}(\underset{\sim}{x},\ t)\right]^2\right\} \tag{3.5}$$

$$n(\underset{\sim}{x},\ t) = \int f\ d\underset{\sim}{v} \tag{3.6}$$

$$\underset{\sim}{q}(\underset{\sim}{x},\ t) = \frac{1}{n(\underset{\sim}{x},\ t)}\int \underset{\sim}{v}\ f\ d\underset{\sim}{v} \tag{3.7}$$

$$\frac{3kT}{m}(\underset{\sim}{x},\ t) = \frac{1}{n(\underset{\sim}{x},\ t)}\int (\underset{\sim}{v} - \underset{\sim}{q})^2\ f\ d\underset{\sim}{v} \tag{3.8}$$

and $$\int \Phi\ d\underset{\sim}{v} = 1 \tag{3.9}$$

where $\underset{\sim}{q}(\underset{\sim}{x},\ t)$ and $T(\underset{\sim}{x},\ t)$ define the flow velocity and temperature at $\underset{\sim}{x}$ and t. The interpretation now is that the re-emitted particles (after the collision) at $(\underset{\sim}{x},\ t)$ emerge with a Maxwellian distribution centred on the mean flow velocity $\underset{\sim}{q}(\underset{\sim}{x},\ t)$ corresponding to the temperature $T(\underset{\sim}{x},\ t)$. It is easily verified that this model conserves particle number, momentum and energy and is therefore appropriate to describe like-particle collisions.

Another commonly used relaxation model is that for a Lorentzian plasma. This refers to a plasma in which the ions are in a highly ionized state. Since the collision cross sections for electron-ion and electron-electron collisions are in the ratio Z^2 where Z is the ionic charge the electron-ion collision frequency can be used but its velocity dependence must be included. Thus

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\nu(v)(f - f_0) \tag{3.10}$$

and $$\nu(v) = \nu_0 \frac{v_T^3}{v^3} \tag{3.11}$$

and the equation can be put into the particle conserving form if required.

3.2 Fokker-Planck Model

The kinetic theory of gases is described by the Boltzmann equation. The collision integral in this equation describes large angle scattering due to close encounters. For

Coulomb interactions in a plasma the slow fall off of the interaction causes the Boltzmann collision integral to diverge. This is an indication that distant encounters (small-angle scattering) are important. However, we know that the Coulomb interaction is screened so that the integral is cut off and does not diverge. Nevertheless, many small-angle collisions can accumulate to produce large-angle deflections and will therefore be important.

In Section 2 we noted that Vlasov's equation could be obtained as an expansion in the plasma parameter $g \equiv (n_0 \lambda_D^3)^{-1}$. Vlasov's equation is the zero-order result of this procedure. The equation obtained to first order in g is the Fokker-Planck equation which takes account of these small angle scatterings. The derivation of the Fokker-Planck equation by such an expansion can be found elsewhere[6,7]. Here, we shall confine ourselves to a simple physical argument to justify the form of the Fokker-Planck equation.

In this approach we make direct use of the fact that changes in the particle velocity distribution are caused by many small angle scatterings. Let $F(\underset{\sim}{v}, \Delta\underset{\sim}{v})$ be the probability that the particle velocity changes from $\underset{\sim}{v}$ to $\underset{\sim}{v} + \Delta\underset{\sim}{v}$ in a time interval Δt due to collisions. The distribution at time t can therefore be expressed in terms of the distribution at the slightly earlier time $t - \Delta t$ as follows

$$f(\underset{\sim}{v}, t) = \int f(\underset{\sim}{v} - \Delta\underset{\sim}{v}, t - \Delta t)\, F(\underset{\sim}{v} - \Delta\underset{\sim}{v}, \Delta\underset{\sim}{v})\, d(\Delta\underset{\sim}{v}) \quad . \tag{3.12}$$

Now due to the long range nature of the Coulomb force most of the collisions will be small angle and we may therefore expand the integrand

$$\begin{aligned} f(\underset{\sim}{v}, t) = \int \Big\{ & f(\underset{\sim}{v}, t)\, F(\underset{\sim}{v}, \Delta\underset{\sim}{v}) - \Delta t\, F(\underset{\sim}{v}, \Delta\underset{\sim}{v}) \Big(\frac{\partial f}{\partial t}\Big)_{\text{coll}} \\ & - \Delta\underset{\sim}{v} \cdot \frac{\partial}{\partial \underset{\sim}{v}} \big[f(\underset{\sim}{v}, t)\, F(\underset{\sim}{v}, \Delta\underset{\sim}{v}) \big] \\ & + \frac{1}{2} \Delta v_i \Delta v_k \frac{\partial^2}{\partial v_i \partial v_k} \big[f(\underset{\sim}{v}, t)\, F(\underset{\sim}{v}, \Delta\underset{\sim}{v}) \big] \Big\} \end{aligned} \tag{3.13}$$

where a repeated index denotes a summation. Using the fact that the probability is normalized

$$\int F(\underset{\sim}{v}, \Delta\underset{\sim}{v})\, d(\Delta\underset{\sim}{v}) = 1$$

we obtain

$$\Big(\frac{\partial f}{\partial t}\Big)_{\text{coll}} = \frac{1}{\Delta t} \int \Big\{ -\Delta\underset{\sim}{v} \cdot \frac{\partial}{\partial \underset{\sim}{v}} (fF) + \frac{1}{2} \Delta v_i \Delta v_k \frac{\partial^2}{\partial v_i \partial v_k} (fF) \Big\} d(\Delta\underset{\sim}{v}) \quad . \tag{3.14}$$

We now write this as

$$\Big(\frac{\partial f}{\partial t}\Big)_{\text{coll}} = - \frac{\partial}{\partial v_i} \Big\{ \frac{\langle \Delta v_i \rangle}{\Delta t} f \Big\} + \frac{1}{2} \frac{\partial^2}{\partial v_i \partial v_k} \Big\{ \frac{\langle \Delta v_i \Delta v_k \rangle}{\Delta t} f \Big\} \tag{3.15}$$

where $\langle \Delta v_i \rangle$ and $\langle \Delta v_i \Delta v_k \rangle$ are the average values of these quantities over the time interval Δt and are defined as

$$\langle \Delta \underset{\sim}{v} \rangle = \int F(\underset{\sim}{v}, \Delta \underset{\sim}{v}) \Delta \underset{\sim}{v} \, d(\Delta \underset{\sim}{v}) \tag{3.16}$$

$$\langle \Delta v_i \Delta v_k \rangle = \int F(\underset{\sim}{v}, \Delta \underset{\sim}{v}) \Delta v_i \Delta v_k \, d(\Delta \underset{\sim}{v}) \quad . \tag{3.17}$$

Finally, we obtain

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = - \frac{\partial}{\partial v_i} \left(A_i f\right) + \frac{1}{2} \frac{\partial}{\partial v_i} \left(D_{ik} \frac{\partial f}{\partial v_k}\right) \tag{3.18}$$

where

$$A_i = \frac{\langle \Delta v_i \rangle}{\Delta t} - \frac{1}{2} \frac{\partial}{\partial v_k} \frac{\langle \Delta v_i \Delta v_k \rangle}{\Delta t} \tag{3.19}$$

and

$$D_{ik} = \frac{\langle \Delta v_i \Delta v_k \rangle}{\Delta t} \quad .$$

Equation (3.18) is the Fokker-Planck equation. The first term has the character of a dynamical friction which decreases or increases the population of a group of particles depending on whether they are moving faster or slower than the corresponding equilibrium particles. The second term acts like a diffusion in velocity space and has the effect of spreading a velocity perturbation. The equilibrium Maxwell distribution results from a balance between these two terms.

We also note that the right hand side of the Fokker-Planck equation is the divergence in velocity space $- \frac{\partial s_i}{\partial v_i}$ of the vector

$$s_i = A_i f - \frac{1}{2} D_{ik} \frac{\partial f}{\partial v_k} \tag{3.20}$$

where s_i is the particle flux in velocity space. The Fokker-Planck equation is therefore an equation of continuity thus guaranteeing the conservation of the number of particles. In addition the quantities A_i and D_{ik} are connected by the equilibrium condition that the particle flux in velocity space must be zero.

In order to make practical calculations with the Fokker-Planck equation we need explicit forms for the quantities A_i and D_{ik}. To complete the discussion of the Fokker-Planck equation we will quote the result for these terms. A fuller account can be found in references 5 and 6.

The coefficients in the Fokker-Planck equation are obtained by considering the details of binary collisions of particles of species s with all other particles of the plasma[7]. The result is

$$\left(\frac{\partial f_s}{\partial t}\right)_{coll} = -\sum_{s'} \frac{q_s^2 q_s'^2, \log \Lambda}{8\pi\varepsilon_0^2 m_s} \frac{\partial}{\partial v_i} \int \left(\frac{\delta_{ik}}{u} - \frac{u_i u_k}{u^3}\right) \left\{\frac{f_s}{m_{s'}}(\underset{\sim}{v}) \frac{\partial f_{s'}}{\partial v_k'}(\underset{\sim}{v}') - \frac{f_{s'}}{m_s}(\underset{\sim}{v}') \frac{\partial f_s}{\partial v_k}(\underset{\sim}{v})\right\} d\underset{\sim}{v} \tag{3.21}$$

where $\underset{\sim}{u} = \underset{\sim}{v} - \underset{\sim}{v}'$, $\left(\frac{\delta_{ik}}{u} - \frac{u_i u_k}{u^3}\right)$ comes from an average over scattering angles and $\log \Lambda$ is the factor which arises from cutting off the Coulomb interaction at small angles due to the effect of Debye screening.

Note that the friction and diffusion coefficients depend on the distribution function with the result that the equation is nonlinear. The quadratic dependence on f for like particle collisions is shown explicitly.

The Fokker-Planck equation is often written in another form

$$\frac{\partial f_s}{\partial t} = -\sum_{s'} \frac{q_s^2 q_s'^2, \log \Lambda}{4\pi\varepsilon_0^2 m_s^2} \frac{\partial}{\partial v_i} \left[f_s(\underset{\sim}{v}) \frac{\partial}{\partial v_k} h(\underset{\sim}{v}) - \frac{1}{2} \frac{\partial}{\partial v_k} \left(f_s(\underset{\sim}{v}) \frac{\partial^2 g}{\partial v_i \partial v_k}\right)\right] \tag{3.22}$$

where $h(\underset{\sim}{v})$ and $g(\underset{\sim}{v})$ are the Rosenbluth potentials

$$g(\underset{\sim}{v}) = \int f_{s'}(\underset{\sim}{v}') \, u \, d\underset{\sim}{v}' \tag{3.23}$$

$$h(\underset{\sim}{v}) = \left(1 + \frac{m_s}{m_{s'}}\right) \int f_{s'}(\underset{\sim}{v}') \, u^{-1} \, d\underset{\sim}{v}' \quad . \tag{3.24}$$

For many problems the distribution can be assumed to be close to thermal equilibrium so that the collision term can be linearized about a Maxwellian. Finally, when all scattering processes are included the Fokker-Planck equation satisfies the important constraints of conservation of number density, momentum and energy.

4. FLUID MODELS

In many situations a simpler description than the full kinetic model is desirable, either because the kinetic treatment would be too complicated or because the simpler model contains all the essential physics. In the final part of this lecture we will show how these simpler, fluid models, may be obtained from the kinetic equation.

If the distribution function $f_j(\underset{\sim}{x}, \underset{\sim}{v}, t)$ is known, then the various physically significant quantities can be obtained as follows:

$$n_j(\underset{\sim}{x}, t) = \int f_j(\underset{\sim}{x}, \underset{\sim}{v}, t) \, d\underset{\sim}{v} \tag{4.1}$$

$$\frac{\partial}{\partial x_m} m_j \int v_n v_m f_j \, dv = \frac{\partial}{\partial x_m} (P_j)_{nm} + m_j u_n \frac{\partial}{\partial x_m} (n_j u_m) + n_j u_m \frac{\partial}{\partial x_m} (m_j u_n) \quad .$$

We may therefore write

$$m_j \frac{\partial}{\partial \underset{\sim}{x}} \int \underset{\sim}{v}\underset{\sim}{v} f_j \, dv = \nabla \cdot \underset{\approx}{P}_j + m_j \underset{\sim}{u}_j \nabla \cdot (n_j \underset{\sim}{u}_j) + n_j \underset{\sim}{u}_j \cdot \nabla (m_j \underset{\sim}{u}_j) \quad . \tag{4.11}$$

Finally,

$$\int \underset{\sim}{v}(\underset{\sim}{E} + \underset{\sim}{v} \times \underset{\sim}{B}) \cdot \frac{\partial f_j}{\partial \underset{\sim}{v}} \, dv = -\underset{\sim}{E} n_j - (\underset{\sim}{u}_j \times \underset{\sim}{B}) n_j \tag{4.12}$$

where (4.12) was obtained by considering the (nm) element, integrating by parts, and using $(\partial/\partial v_m)(\underset{\sim}{v} \times \underset{\sim}{B})_m = 0$ again. Combining equations (4.10), (4.11) and (4.12) and using the continuity equation we finally obtain the first moment (or momentum) equation

$$m_j n_j \frac{\partial}{\partial t} \underset{\sim}{u}_j + n_j m_j (\underset{\sim}{u}_j \cdot \nabla) \underset{\sim}{u}_j + \nabla \cdot \underset{\approx}{P}_j = q_j n_j \underset{\sim}{E} + q_j n_j (\underset{\sim}{u}_j \times \underset{\sim}{B}) \quad . \tag{4.13}$$

This is usually written in the form

$$\frac{\partial \underset{\sim}{u}_j}{\partial t} + (\underset{\sim}{u}_j \cdot \nabla) \underset{\sim}{u}_j + \frac{\nabla \cdot \underset{\approx}{P}_j}{n_j m_j} = \frac{q_j}{m_j} (\underset{\sim}{E} + \underset{\sim}{u}_j \times \underset{\sim}{B}) \quad . \tag{4.14}$$

Notice that the equation for each moment contains a term involving the next higher moment - the closure problem. Thus, the equation for n_j contains $\underset{\sim}{u}_j$ and the equation for $\underset{\sim}{u}_j$ contained $\underset{\approx}{P}_j$. As already mentioned there is no rigorous method of closure for this set of equations. Instead some physical justification is sought for terminating the hierarchy.

The next equation in the series is the energy equation obtained by taking the moment of $m_j \underset{\sim}{v} \cdot \underset{\sim}{v}/2$. This is usually as far as one goes with the moment method and we shall not write down the general energy equation. Instead we shall briefly consider the various physical approximations which are used in a fluid description.

The simplest model is the COLD-plasma approximation in which the term $\nabla \cdot \underset{\approx}{P}_j$ is simply neglected thus closing the hierarchy of equations after the zeroth and first moments. The plasma is thus described as two cold, interpenetrating fluids which may be expected to be a good approximation in a variety of wave propagation problems provided the phase velocity of the wave is much greater than the thermal velocities of the particles. The two fluid approximation is often extended to the case of two warm fluids in which the pressure tensor is now included but assumed to be diagonal, i.e. the pressure is isotropic. Thus

$$\underset{\approx}{P}_j = P_j \underset{\approx}{I} \tag{4.16}$$

where T_j is assumed to be constant. The warm two fluid model gives rise to new effects

and also exhibits the slow thermalization behaviour, characteristic of a high temperature plasma, which allows a plasma to persist in a state in which the two species may have widely different temperatures.

The final fluid approximation we would like to mention is the ONE-FLUID or magneto-hydrodynamic (MHD) model. This is appropriate under conditions where the electrons and ions move together maintaining charge neutrality. This occurs at low frequencies, i.e. $\omega << \Omega_i$, and long wavelengths $\lambda >> v_{Ti}/\Omega_i$ where Ω_i is the ion cyclotron frequency. Adding together the continuity equations for the electrons and ions we obtain the equation of continuity for mass flow:

$$\frac{\partial \rho}{\partial t} + \underset{\sim}{\nabla}\cdot(\rho \underset{\sim}{u}_0) = 0 \tag{4.17}$$

where $\rho \equiv \sum_j n_j m_j$

and $\underset{\sim}{u}_0 \equiv \sum_j n_j \underset{\sim}{u}_j m_j \Big/ \sum_j n_j m_j$.

Adding the momentum equations gives the equation for the centre of mass motion

$$\frac{\partial \underset{\sim}{u}_0}{\partial t} + (\underset{\sim}{u}_0 \cdot \underset{\sim}{\nabla})\underset{\sim}{u}_0 = - \underset{\sim}{\nabla}\cdot \underset{\approx}{P} + \underset{\sim}{J} \times \underset{\sim}{B} \tag{4.18}$$

where $\underset{\approx}{P} = \underset{\approx}{P}_i + \underset{\approx}{P}_e$ refers to the centre-of-mass velocity as is appropriate to a one-fluid description. The pressure term is usually assumed to be diagonal and $\underset{\sim}{J}$ is the electric current given by

$$\underset{\sim}{J} = \sum_j n_j q_j \underset{\sim}{u}_j$$

Equation (4.18) is reminiscent of the Navier-Stokes equation of fluid mechanics without the viscous force but containing the body force $\underset{\sim}{J} \times \underset{\sim}{B}$ due to the fluid being magnetized. The one-fluid model is usually closed by making the adiabatic approximation which assumes a local Maxwellian so that

$$\left(\frac{\partial}{\partial t} + \underset{\sim}{u}_0 \cdot \underset{\sim}{\nabla}\right)\left(\frac{P}{\rho^\gamma}\right) = 0 \tag{4.19}$$

where γ is the ratio of specific heats. Equation (4.19) results from the energy equation by neglecting the heat flow and assuming perfect conductivity, i.e.

$$\underset{\sim}{E} + \underset{\sim}{u}_0 \times \underset{\sim}{B} = 0 \quad .$$

* * *

REFERENCES

1) D. C. Montgomery and D. A. Tidman, Plasma Kinetic Theory (McGraw-Hill, New York, 1964) Chap. 4 and references therein.

2) C. N. Lashmore-Davies, Plasma Physics and Instabilities, CERN 81-13(1981) Chap. I.

3) R.C.Davidson, J. Plasma Phys. 6, 229 (1971)

4) J.H.Malmberg and J.S.deGrassie, Phys. Rev. Letters 35, 577 (1975).

5) P. L. Bhatnagar, E. P. Gross and M. Krook, Physical Review 94, 511 (1954).

6) N. Rostoker and M. N. Rosenbluth, Physics of Fluids 3, 1 (1960).

7) N. A. Krall and A. W. Trivelpiece, Principles of Plasma Physics (McGraw-Hill, New York, 1973) Chap. 6.

8) R. A. Cairns, Plasma Physics (Blackie, Glasgow, 1985) pp. 49-54.

WAVES IN PLASMAS

C N Lashmore-Davies
Culham Laboratory, Abingdon, Oxon. OX14 3DB, GB
(Euratom/UKAEA Fusion Association)

ABSTRACT

A brief discussion of the relevance of linear waves in infinite uniform plasmas is given. The linear waves of a field-free plasma are then discussed in some detail in order to illustrate the methods which are used in more complicated problems. High-frequency waves in a magnetized plasma are then described. The main emphasis is to show how the magnetic field affects the wave phenomena. The two specific cases of propagation parallel and perpendicular to the magnetic field are examined. Next, low-frequency waves in a magnetized plasma are discussed from the point of view of ideal MHD. Finally, the problem of Raman scattering is discussed in some detail. This phenomenon illustrates how the nonlinear terms previously neglected can introduce additional effects. Finally, the preceding analysis is used to introduce the subject of the beat wave accelerator.

1. INTRODUCTION

A magnetized plasma is one of the richest wave media yet studied. It owes this richness to the long-range nature of the Coulomb interaction, and it is this long-range interaction which makes the many-body or collective effects in a plasma so subtle. One of the most important concepts arising from the collective effects is that of the self-consistent electromagnetic field, which is the field which a given plasma particle experiences due to the presence of all the other particles. This self-consistent field is crucial to the theory of waves in the plasma medium.

The idea is as follows. Suppose there is a small electromagnetic field present in the plasma. This produces forces on the plasma particles, resulting in currents and charge perturbations which act as source terms for further electromagnetic fields, which will then produce further plasma motions, and so on. This system of field perturbations and particle motions is iterated until the assumed electromagnetic field is itself produced by the resulting plasma motion. For the theory of linear waves in a plasma, the self-consistency condition gives the dispersion relation which contains all the information concerning the wave motion.

In this lecture we shall be mainly concerned with linear waves although an example of a nonlinear wave phenomenon will be discussed at the end. In order to discuss linear waves we must first of all define the equilibrium state. We shall confine ourselves to the idealization of an infinite, uniform plasma. Having specified the equilibrium, the linear waves describe the possible modes of oscillation when this equilibrium is subject to a small perturbation such that products of perturbations in the equations of motion can be

ignored. Physically, this means that the wave amplitude is small enough that effects such as the beating of two waves to produce sum and difference frequencies or modification of the equilibrium are negligible. Many of the most interesting problems in fact are concerned with these effects. However, we must first understand the linear modes of oscillation of the plasma medium. The theory of linear waves does not contain any criterion on how small the wave amplitude should be for the theory to be valid. To obtain such a criterion we must appeal to a nonlinear analysis.

Let us briefly consider the relevance of our idealized model, bearing in mind that most plasmas are either finite, non-uniform, varying in time or any combination of these. The primary justification is on the grounds of simplicity. A study of the simplest of all models serves to identify many of the basic phenomena and then forms a framework for more complicated situations. If the wavelength is less than the plasma radius ($\lambda << a$) or the characteristic scale length L of some non-uniformity ($\lambda << L$), or if the frequency of the wave is much greater than the inverse of the plasma lifetime $\tau(\omega >> \tau^{-1})$, then these complications will produce only small corrections to the simple theory.

Another noteworthy feature of the simple theory is that no modes are lost by the simplification - the more realistic models only modify the properties of existing modes of the simple model. Even though many of the wave motions observed in plasmas relate to a nonlinear saturated state it is usually the case that the frequencies and wavelengths are determined quite closely by the linear theory. A study of the properties of linear waves does therefore have relevance to physical problems.

Finally, it is worth mentioning that the plasma instabilities which can occur when there is a source of free energy result from one or two of these linear waves which can draw on this free energy. A knowledge of the linear waves in a plasma is therefore essential to the study of stability.

Let us now consider the simple example of linear waves in an infinite field-free, uniform plasma in some detail.

2. WAVES IN A FIELD-FREE PLASMA

In any problem in plasma physics there is always a choice of plasma model. For our present purposes the two-fluid (hydrodynamic) model is adequate to illustrate the properties of linear waves in a field-free plasma. By field free we mean that there is no equilibrium magnetic field. A steady electric field would not change the nature of the waves but might cause some of them to become unstable. Here, we are only concerned with stable wave motions. As always the plasma is coupled to the electromagnetic field through Maxwell's equations. We therefore require solutions to the following system of coupled partial-differential equations

$$\frac{\partial n_e}{\partial t} + \underset{\sim}{\nabla}\cdot\left(n_e \underset{\sim}{v}_e\right) = 0 \qquad (2.1)$$

$$\frac{\partial \underset{\sim}{v}_e}{\partial t} + (\underset{\sim}{v}_e \cdot \underset{\sim}{\nabla})\underset{\sim}{v}_e + \gamma_e \frac{\kappa T_e}{n_e m_e} \underset{\sim}{\nabla} n_e = -\frac{e}{m_e} \underset{\sim}{E} - \frac{e}{m_e} \underset{\sim}{v}_e \times \underset{\sim}{B} \qquad (2.2)$$

$$\frac{\partial n_i}{\partial t} + \underset{\sim}{\nabla} \cdot (n_i \underset{\sim}{v}_i) = 0 \qquad (2.3)$$

$$\frac{\partial \underset{\sim}{v}_i}{\partial t} + (\underset{\sim}{v}_i \cdot \underset{\sim}{\nabla})\underset{\sim}{v}_i = \frac{e}{m_i} \underset{\sim}{E} + \frac{e}{m_i} \underset{\sim}{v}_i \times \underset{\sim}{B} \qquad (2.4)$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{H} = \underset{\sim}{J} + \varepsilon_0 \frac{\partial \underset{\sim}{E}}{\partial t} \qquad (2.5)$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{E} = -\mu_0 \frac{\partial \underset{\sim}{H}}{\partial t} \quad . \qquad (2.6)$$

The above equations are written in MKS units where n_e , n_i , $\underset{\sim}{v}_e$ and $\underset{\sim}{v}_i$ are the electron and ion number densities and the electron and ion fluid velocities. We have assumed a zero pressure (cold) ion fluid and a finite pressure (temperature T_e) electron fluid where γ_e is the ratio of specific heats of the electron fluid. The current density in equation (2.5) is

$$\underset{\sim}{J} = n_i e \underset{\sim}{v}_i - n_e e \underset{\sim}{v}_e \quad . \qquad (2.7)$$

The magnetic field appearing in the above equations is, of course, an oscillating field due to the presence of a wave whose properties are still to be determined. Equations (2.1) - (2.7) are highly nonlinear and in order to describe the linear wave motions we must linearize this system of equations. To do this we must first define the equilibrium which, in this case, is particularly simple. The equilibrium configuration consists of a uniform density plasma with no average electron or ion flows. There is therefore no equilibrium electric current and since the electron and ion fluid densities are uniform, there is no equilibrium space charge and hence no electric field. All variables are now separated into an equilibrium part and a part which represents a small perturbation, varying both in space and time. The linearization is carried out by substituting the variables expressed in this way into equations (2.1) - (2.7) and neglecting all quantities which involve products of perturbed variables. Equations (2.1) - (2.7) are thus transformed into a set of linear equations in which the variables are the perturbed-field variables and the coefficients of the equations are constants due to the uniform equilibrium. The equations can therefore be Fourier analyzed in space and time. Since there is no preferred direction it is sufficient to assume that all the perturbed fields vary as

$$\exp i(kz - \omega t)$$

The set of simultaneous, linear, partial differential equations are therefore reduced to the following set of algebraic equations

$$-i\omega n_{e1} + ikn_0 v_{e1z} = 0 \qquad (2.8)$$

$$-i\omega \underset{\sim}{v}_{e1} + \gamma_e \frac{\kappa T_e}{n_0 m_e} ikn_{e1} = -\frac{e}{m_e} \underset{\sim}{E}_1 \qquad (2.9)$$

$$-i\omega n_{i1} + ikn_0 v_{i1z} = 0 \qquad (2.10)$$

$$-i\omega \underset{\sim}{v}_{i1} = \frac{e}{m_i} \underset{\sim}{E}_1 \tag{2.11}$$

$$i\underset{\sim}{k} \times \underset{\sim}{H}_1 = n_0 e(\underset{\sim}{v}_{i1} - \underset{\sim}{v}_{e1}) - i\omega\varepsilon_0 \underset{\sim}{E}_1 \tag{2.12}$$

$$i\underset{\sim}{k} \times \underset{\sim}{E}_1 = i\omega\mu_0 \underset{\sim}{H}_1 \tag{2.13}$$

where perturbed variables carry a subscript 1.

The self-consistency condition on the fields and particle motions is now obtained by requiring that the above set of equations be simultaneously satisfied. This condition is given by the vanishing of the determinant of the coefficients and results in the dispersion relation for the system which is usually written as

$$D(\omega, \underset{\sim}{k}) = 0 \quad . \tag{2.14}$$

The solutions $\omega(\underset{\sim}{k})$ of this equation then give all the properties of the waves of the system.

Let us now obtain the dispersion relation from equations (2.8)-(2.13). First, solving for the fluid velocities in terms of the components of the electric field $\underset{\sim}{E}_1$ and substituting into equation (2.7) we obtain the conductivity tensor defined by

$$\underset{\sim}{J}_1 = \underset{\approx}{\sigma}(\omega, \underset{\sim}{k}) \underset{\sim}{E}_1 \tag{2.15}$$

where

$$\underset{\approx}{\sigma}(\omega, \underset{\sim}{k}) = \begin{vmatrix} \frac{in_0e^2}{\omega m_i} + \frac{in_0e^2}{\omega m_e} & 0 & 0 \\ 0 & \frac{in_0e^2}{\omega m_i} + \frac{in_0e^2}{\omega m_e} & 0 \\ 0 & 0 & \frac{in_0e^2}{\omega m_i} + \frac{in_0e^2}{\omega m_e}\left[1 - \gamma_e \frac{k^2 v_{Te}^2}{\omega^2}\right]^{-1} \end{vmatrix}$$

and $v_{Te}^2 \equiv \kappa T_e / m_e$. With the aid of equations (2.12), (2.13) and (2.15) we obtain

$$\underset{\sim}{k} \times (\underset{\sim}{k} \times \underset{\sim}{E}_1) = -\frac{\omega^2}{c^2} \underset{\approx}{\varepsilon}(\omega, \underset{\sim}{k}) \underset{\sim}{E}_1 \tag{2.16}$$

where

$$\underset{\approx}{\varepsilon}(\omega, \underset{\sim}{k}) = \underset{\approx}{I} + \frac{i}{\omega\varepsilon_0} \underset{\approx}{\sigma}(\omega, \underset{\sim}{k}) \tag{2.17}$$

is the dielectric tensor of the plasma. The dispersion relation follows immediately from equation (2.16) and is

$$\begin{vmatrix} -k^2 + \frac{\omega^2}{c^2}\varepsilon_{xx} & 0 & 0 \\ 0 & -k^2 + \frac{\omega^2}{c^2}\varepsilon_{yy} & 0 \\ 0 & 0 & \frac{\omega^2}{c^2}\varepsilon_{zz} \end{vmatrix} = 0 \quad . \tag{2.18}$$

It is clear that the transverse ($\underset{\sim}{E} \perp \underset{\sim}{k}$) modes are decoupled from the longitudinal ($\underset{\sim}{E} \parallel \underset{\sim}{k}$) modes.

2.1 Longitudinal modes

The dispersion relation for these modes is

$$\varepsilon_{zz} = 0 \tag{2.19}$$

ignoring the $\omega^2 = 0$ solutions. There are two solutions of equation (2.19), a high frequency branch

$$\omega^2 = \omega_{pe}^2 + \gamma_e k^2 v_{Te}^2 \tag{2.20}$$

involving only the electrons, where $\omega_{pe}^2 \equiv n_0 e^2/\varepsilon_0 m_e$, known as the LANGMUIR wave and a low frequency branch involving both electrons and ions

$$\omega^2 = \frac{k^2 c_s^2}{\left(1 + \gamma_e k^2 \lambda_{de}^2\right)} \tag{2.21}$$

known as the ION ACOUSTIC wave. For long wavelengths the ion acoustic wave is non-dispersive and travels with phase and group velocity both equal to $c_s \equiv \left(\gamma_e \kappa\, T_e/m_i\right)^{1/2}$. At shorter wavelengths $k\lambda_{de} \gtrsim 1$, where $\lambda_{de} \equiv v_{Te}/\omega_{pe}$, the upper limit of frequency of the ion acoustic wave is the ion plasma frequency ω_{pi} where $\omega_{pi}^2 \equiv n_0 e^2/\varepsilon_0 m_i$.

2.2 Transverse modes

The dispersion relation, equation (2.18), shows that there are two independent transverse modes, both linearly polarized. The dispersion relations of these modes are

$$-k^2 + \frac{\omega^2}{c^2}\varepsilon_{xx} = 0 \tag{2.22}$$

$$-k^2 + \frac{\omega^2}{c^2}\varepsilon_{yy} = 0 \quad . \tag{2.23}$$

Since $\varepsilon_{xx} = \varepsilon_{yy}$ the two transverse waves have the same propagation properties, one being polarized with its electric field in the x-direction and the other with its electric field in the y-direction. With the aid of equation (2.17) and the expression for the

conductivity tensor we obtain

$$\varepsilon_{xx} = 1 - \frac{\omega_{pe}^2}{\omega^2} \quad . \tag{2.24}$$

Using (2.24) in equations (2.22) and (2.23) we finally obtain the dispersion relation for the two transverse modes

$$\omega^2 = \omega_{pe}^2 + c^2k^2 \quad . \tag{2.25}$$

This is also a high-frequency wave involving only the electrons. Transverse electromagnetic waves do not propagate in a plasma for frequencies below the plasma frequency, i.e. for $\omega < \omega_{pe}$, $k^2 < 0$ corresponding to an evanescent wave. The longitudinal Langmuir wave also has this property being cut off at $\omega = \omega_{pe}$. The properties of the linear waves in a field free plasma are conveniently summarized with the aid of an ω-k diagram shown in figure 1.

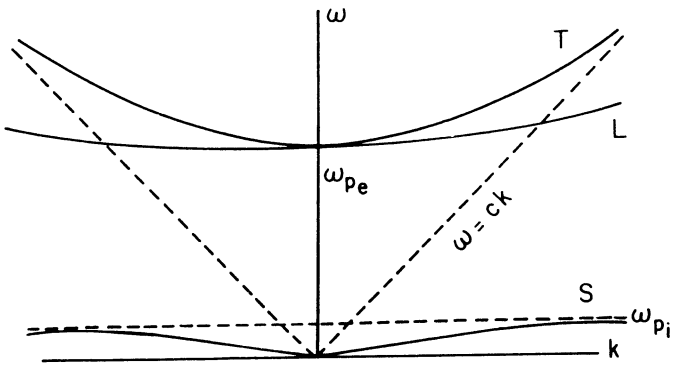


Fig. 1 Dispersion relation of ion acoustic (S), Langmuir (L) and transverse waves in a field free plasma

One final point worth making which is brought out by figure 1 concerns the Langmuir wave. It can be seen that the phase velocity (ω/k) can vary widely ranging from values much larger than the velocity of light c down to values of the order of the electron thermal speed.

3. WAVES IN A MAGNETIZED PLASMA

Let us now consider the effect of an equilibrium magnetic field on the linear waves in a plasma. We shall not go into as much detail as in the previous section but will point out the differences introduced by an equilibrium magnetic field and how the wave properties are altered. We shall keep things as simple as possible by considering only high-frequency waves which do not involve the ions. The only change to the system of equations given by (2.1), (2.2), (2.5) and (2.6) is that the linearized momentum equation for the electron fluid now becomes

$$\frac{\partial \underset{\sim}{v}_{e1}}{\partial t} + \gamma_e \frac{\kappa T_e}{n_0 m_e} \underset{\sim}{\nabla} n_{e1} = - \frac{e}{m_e} \left(\underset{\sim}{E}_1 + \underset{\sim}{v}_{e1} \times \underset{\sim}{B}_0 \right) \qquad (3.1)$$

where the additional term due to the presence of the equilibrium magnetic field $\underset{\sim}{B}_0$ appears on the right hand side of equation (3.1). We assume that the magnetic field is uniform in space, constant in time and points in the z-direction, i.e. $\underset{\sim}{B}_0 = (0,0,B_0)$. Since the magnetic field now defines a preferred direction in the plasma we must allow for perturbations to the equilibrium which vary as

$$\exp i(\underset{\sim}{k} \cdot \underset{\sim}{x} - \omega t)$$

where, without loss of generality we may assume $\underset{\sim}{k} = (0,k_y,k_z)$. We shall not go through the derivation of the dielectric tensor in this case. Even with the neglect of ion motion the algebra is still rather tedious. Instead, we simply note the important fact that due to the presence of the equilibrium magnetic field in equation (3.1) the x- and y-components of the fluid velocity are now coupled together, in contrast to the field free case. This results in some of the off-diagonal elements of the dielectric tensor being non-zero. Once we have obtained the dielectric tensor $\underset{\approx}{\varepsilon}(\omega,\underset{\sim}{k})$ we can again write down the dispersion relation from equation (2.16). This time however all the elements of $\underset{\approx}{\varepsilon}(\omega,\underset{\sim}{k})$ are non-zero for the most general case. The dispersion relation can be written formally as

$$\begin{vmatrix} -k^2 + \frac{\omega^2}{c^2}\varepsilon_{xx} & \frac{\omega^2}{c^2}\varepsilon_{xy} & \frac{\omega^2}{c^2}\varepsilon_{xz} \\ \frac{\omega^2}{c^2}\varepsilon_{yx} & -k_z^2 + \frac{\omega^2}{c^2}\varepsilon_{yy} & k_y k_z + \frac{\omega^2}{c^2}\varepsilon_{zy} \\ \frac{\omega^2}{c^2}\varepsilon_{zx} & k_y k_z + \frac{\omega^2}{c^2}\varepsilon_{yz} & -k_y^2 + \frac{\omega^2}{c^2}\varepsilon_{zz} \end{vmatrix} = 0 \quad . \qquad (3.2)$$

In order to see the consequences of the presence of the equilibrium magnetic field we do not need to analyze the most general case. Instead we shall consider two special cases.

3.1 Propagation parallel to the magnetic field

For propagation parallel to the magnetic field $k_y = 0$. It is then the case that

$$\varepsilon_{xz} = 0 = \varepsilon_{zx}$$

$$\varepsilon_{yz} = 0 = \varepsilon_{zy}$$

with the result that the dispersion relation, equation (3.2) reduces to

$$\left\{ \left(k_z^2 - \frac{\omega^2}{c^2}\varepsilon_{xx}\right)\left(k_z^2 - \frac{\omega^2}{c^2}\varepsilon_{yy}\right) - \frac{\omega^4}{c^4}\varepsilon_{xy}\varepsilon_{yx} \right\} \frac{\omega^2}{c^2}\varepsilon_{zz} = 0 \quad . \qquad (3.3)$$

We therefore find that, for parallel propagation the longitudinal modes are again decoupled from the transverse modes and in fact are unaffected by the magnetic field. This was to be expected since $\underset{\sim}{E}_1 \parallel \underset{\sim}{B}_0$. The longitudinal dispersion relation is again

$$\varepsilon_{zz} = 0 \tag{3.4}$$

which yields the Langmuir waves previously discussed.

The first change brought about by the magnetic field is shown by the transverse dispersion relation. Since ε_{xy} and ε_{yx} are non-zero, E_{1x} and E_{1y} are now coupled. This was also to be expected from our observation concerning the electron momentum equation. The symmetry of the present problem is such that

$$\varepsilon_{yx} = -\varepsilon_{xy}$$

and

$$\varepsilon_{xx} = \varepsilon_{yy} \quad .$$

The transverse dispersion relation can then be written

$$\left(k_z^2 - \frac{\omega^2}{c^2}\,\varepsilon_{xx}\right)^2 = -\frac{\omega^4}{c^4}\,\varepsilon_{xy}^2$$

so that the two transverse modes are given by

$$k_z^2 - \frac{\omega^2}{c^2}\,\varepsilon_{xx} = \pm\, i\,\frac{\omega^2}{c^2}\,\varepsilon_{xy} \quad . \tag{3.5}$$

Equation (3.5) can now be used in equation (2.16) to obtain the wave polarization giving

$$E_{1x} = \pm\, i\, E_{1y} \quad . \tag{3.6}$$

We therefore obtain the result that transverse electromagnetic waves propagating parallel to the magnetic field are right or left hand circularly polarized. The dispersion relation for these transverse modes can now be obtained by substituting the following explicit expressions for ε_{xx} and ε_{xy} ,

$$\varepsilon_{xx} = 1 - \frac{\omega_{pe}^2}{\left(\omega^2 - \Omega_e^2\right)} \tag{3.7}$$

$$\varepsilon_{xy} = \frac{i\,\omega_{pe}^2\,\Omega_e}{\omega\left(\omega^2 - \Omega_e^2\right)} \tag{3.8}$$

into equation (3.5). The result is

$$\frac{c^2 k_z^2}{\omega^2} = 1 - \frac{\omega_{pe}^2}{\omega\left(\omega \pm \Omega_e\right)} \tag{3.9}$$

where $\Omega_e \equiv e\,B_0/m_e > 0$. The cut-off frequencies, $\omega_{1,2}$, of the right and left hand

modes are no longer given by ω_{pe} but now depend on the magnetic field as well as the density

$$\omega_{1,2} = \frac{\Omega_e}{2}\left[\mp 1 + \left(1 + 4\frac{\omega_{pe}^2}{\Omega_e^2}\right)^{1/2}\right] \quad . \tag{3.10}$$

However, the most striking effect of the magnetic field is the existence of a propagating branch of the right hand mode for $\omega << \Omega_e$ and $\omega_{pe}^2 >> \omega\Omega_e$ giving

$$\frac{c^2k_z^2}{\omega^2} \approx \mp\frac{\omega_{pe}^2}{\omega\Omega_e} \quad . \tag{3.11}$$

The left hand wave is evanescent but the right hand wave propagates. This branch is known as the WHISTLER due to the dependence of the group velocity on $\omega^{1/2}$. The received signal has therefore a falling tone. Whistlers are electromagnetic waves which can propagate with phase velocities, $v_{ph} << c$ and have been observed in the ionosphere, laboratory discharges, and in a rod of indium at liquid-helium temperatures. The properties of the waves propagating along the magnetic field can again be displayed in the ω-k diagram shown in figure 2.

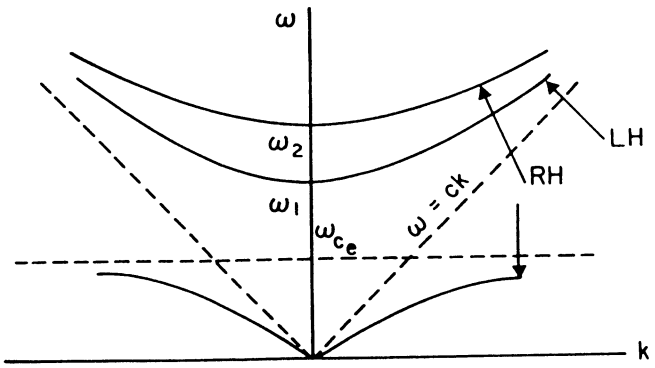


Fig. 2 Dispersion relation for the electromagnetic waves propagating along the equilibrium magnetic field (RH - right hand circular polarization, LH - left hand circular polarization) where $\omega_{1,2}$ are defined by equation (3.10)

3.2 Propagation perpendicular to the magnetic field

For this case $k_z = 0$ and again we find

$$\varepsilon_{xz} = \varepsilon_{zx} = \varepsilon_{yz} = \varepsilon_{zy} = 0 \quad .$$

Under these conditions the general dispersion relation, equation (3.2) reduces to

$$\left\{\left(k_y^2 - \frac{\omega^2}{c^2}\varepsilon_{xx}\right)\frac{\omega^2}{c^2}\varepsilon_{yy} + \frac{\omega^4}{c^4}\varepsilon_{xy}\varepsilon_{yx}\right\}\left(k_y^2 - \frac{\omega^2}{c^2}\varepsilon_{zz}\right) = 0 \quad . \tag{3.12}$$

Equation (3.12) contains two factors. The second one gives

$$k_y^2 = \frac{\omega^2}{c^2}\,\varepsilon_{zz} \quad . \tag{3.13}$$

For a cold plasma $\varepsilon_{zz} = 1 - \dfrac{\omega_{pe}^2}{\omega^2}$ so that equation (3.13) becomes

$$\frac{c^2k_y^2}{\omega^2} = 1 - \frac{\omega_{pe}^2}{\omega^2} \tag{3.14}$$

which is the same result as for transverse waves in a field-free plasma. This is due to the fact that this wave is linearly polarized with its electric field aligned along the equilibrium magnetic field. This wave is therefore called the ordinary mode.

Now consider the other factor of equation (3.12). This illustrates another effect of the magnetic field which is to couple the longitudinal field E_y to the transverse field E_x. For a cold plasma the dielectric tensor elements are again given by equations (3.7) and (3.8). Substituting these equations into the first factor of equation (3.12) we obtain

$$\frac{c^2k_y^2}{\omega^2} = \frac{\{(\omega^2 - \omega_{pe}^2)^2 - \omega^2\Omega_e^2\}}{(\omega^2 - \omega_{pe}^2 - \Omega_e^2)} \quad . \tag{3.15}$$

The cut-off frequencies are the same as for the previous case but now the wave has a resonance at $\omega_{UH}^2 = \omega_{pe}^2 + \Omega_e^2$ where ω_{UH} is called the upper hybrid frequency. Equation (2.16) can again be used to obtain the wave polarization resulting in

$$\left|\frac{E_{1y}}{E_{1x}}\right| = \left|\frac{\varepsilon_{yx}}{\varepsilon_{yy}}\right| = \frac{\Omega_e}{\omega}\,\frac{\omega_{pe}^2}{(\omega^2 - \omega_{UH}^2)} \quad . \tag{3.16}$$

When $\omega >> \Omega_e$, $|E_{1y}| << |E_{1x}|$ and the wave is almost transverse whereas for $\omega \to \omega_{UH}$, $|E_{1y}| >> |E_{1y}|$ and the wave is almost longitudinal. The ω-k diagram for the case of propagation perpendicular to $\underset{\sim}{B}_0$ is shown in figure 3.

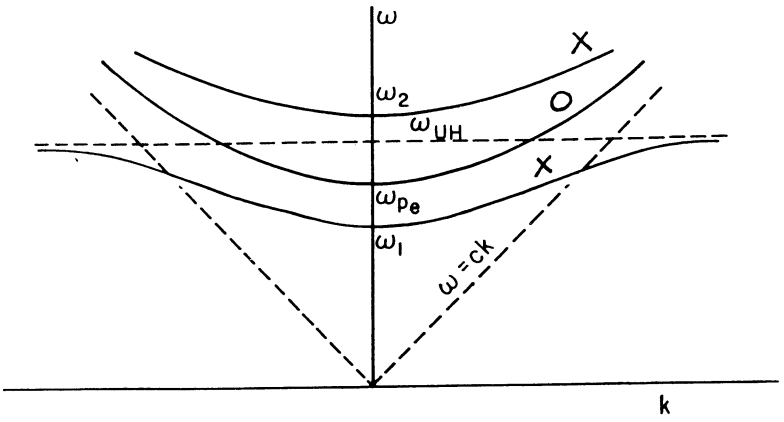


Fig.3 Dispersion relation for waves propagating perpendicular to the equilibrium magnetic field (O denotes the O-mode and X the X-mode)

The hybrid wave has two branches, one of which has a resonance at $\omega = \omega_{UH}$. It is worth noting that compared with the field-free case, the magnetic field has removed the degeneracy between the transverse and longitudinal waves which occurred at $\omega = \omega_{pe}$.

4. LOW-FREQUENCY WAVES IN A MAGNETIZED PLASMA

For frequencies in the vicinity of the ion cyclotron frequency the two-fluid model is still required. However, for $\omega << \Omega_i$ $(\Omega_i \equiv eB_0/m_i)$ and wavelengths long compared with the ion Larmor radius the ions and electrons move together maintaining charge neutrality. Under these circumstances it is a good approximation to describe the plasma as a single fluid. This is the MHD model (ideal or non-ideal depending on whether the plasma resistivity is zero or non-zero). We shall restrict ourselves to the ideal model which is the simplest model to describe the three low-frequency waves of a magnetized plasma. The equations of ideal MHD are

$$\frac{\partial \rho}{\partial t} + \underset{\sim}{\nabla}\cdot(\rho \underset{\sim}{v}) = 0 \tag{4.1}$$

$$\rho\left(\frac{\partial}{\partial t} + \underset{\sim}{v}\cdot\underset{\sim}{\nabla}\right)\underset{\sim}{v} = -\underset{\sim}{\nabla}P + \underset{\sim}{J} \times \underset{\sim}{B} \tag{4.2}$$

$$\underset{\sim}{E} + \underset{\sim}{v} \times \underset{\sim}{B} = 0 \tag{4.3}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{E} = -\frac{\partial \underset{\sim}{B}}{\partial t} \tag{4.4}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{H} = \underset{\sim}{J} \tag{4.5}$$

where ρ, $\underset{\sim}{v}$ and P are the mass density, velocity and pressure of the fluid. The first of these equations expresses the continuity of mass flow and the second is the momentum equation. Notice that there is no force due to the electric field due to the absence of space charge. Apart from the force due to pressure gradients the only other force is due to the flow of current in the presence of a magnetic field. Equation (4.3) results from the assumption of infinite conductivity so that there can be no electric field parallel to the magnetic field. The magnetic field is "frozen in" to the plasma which always moves so that in the plasma rest frame the electric field is zero. The final pair of equations are of course Maxwell's equations but with the displacement current neglected due to the low frequency assumption. The neglect of the displacement current is also related to the absence of space charge.

It is convenient to reduce the above set of five equations to three by substituting for $\underset{\sim}{J}$ from (4.5) into (4.2) and eliminating $\underset{\sim}{E}$ from (4.4) with the aid of (4.3). The linearized version of these equations can then be written

$$\rho_0 \frac{\partial \underset{\sim}{v}_1}{\partial t} = -\underset{\sim}{\nabla}P_1 + (\underset{\sim}{\nabla} \times \underset{\sim}{B}_1) \times \frac{\underset{\sim}{B}_0}{\mu_0} \tag{4.6}$$

$$\frac{\partial \underset{\sim}{B}_1}{\partial t} = \underset{\sim}{\nabla} \times (\underset{\sim}{v}_1 \times \underset{\sim}{B}_0) \tag{4.7}$$

$$\frac{\partial \rho_1}{\partial t} + \underset{\sim}{\nabla}\cdot(\rho_0 \underset{\sim}{v}_1) = 0 \quad . \tag{4.8}$$

In order to close this set we must add an equation of state relating P_1 and ρ_1. We can use the adiabatic or isothermal model and in both cases obtain the relation

$$P_1 = c_s^2 \rho_1 \tag{4.9}$$

where, for the isothermal model $c_s^2 = P_0/\rho_0$. Assuming a constant uniform magnetic field pointing in the z-direction we look for wave solutions varying as

$$\exp i(\underset{\sim}{k}\cdot\underset{\sim}{x} - \omega t)$$

and again take $\underset{\sim}{k} = (0,k_y,k_z)$. We will not go through the details but simply state the main conclusions. First, on writing equations (4.6)-(4.8) in their cartesian components we find that v_{1x} and B_{1x} are independent of all other variables. Solving for these two variables we obtain the dispersion relation

$$\omega^2 = c_A^2 k_z^2 \tag{4.10}$$

where $$c_A^2 = B_0^2/\rho_0\mu_0 \quad .$$

This is the SHEAR ALFVÉN wave which propagates at the Alfvén speed c_A. It will not propagate perpendicularly to the magnetic field and will only transport energy along the field.

Solving for the remaining variables we obtain the dispersion relation for the other two modes

$$(\omega^2 - k_z^2c_s^2)(\omega^2 - k_y^2c_s^2 - k^2c_A^2) - k_y^2k_z^2c_s^4 = 0 \quad . \tag{4.11}$$

For most laboratory plasmas $c_s^2/c_A^2 << 1$. With this assumption the solutions of equation (4.11) are

$$\omega^2 \approx k^2c_A^2\left(1 + \frac{c_s^2}{c_A^2}\right) \tag{4.12}$$

$$\omega^2 \approx k_z^2c_s^2\left(1 + \frac{c_s^2}{c_A^2}\right) \quad . \tag{4.13}$$

The first solution is called the fast magnetosonic wave (sometimes the compressional Alfvén wave) and the second one the slow magnetosonic wave. In general both these waves are hybrid modes but for the parallel propagation the fast wave is a transverse mode and the slow wave a longitudinal mode. The fast wave is the only one of the three low-frequency waves which can propagate perpendicularly to the magnetic field. Apart from its lower velocity the propagation behaviour of the slow wave is similar to the shear wave.

5. RAMAN SCATTERING

Let us now go beyond the linear theory discussed so far and consider the effect of keeping the nonlinear terms contained in equations (2.1)-(2.7). We shall discuss one of the simplest nonlinear effects which is the interaction of two light waves to produce a Langmuir wave. Since all three waves are high frequency only the electrons are involved. Furthermore, the coupling can take place in one dimension.

In order to describe this effect mathematically we return to the electron fluid equations (2.1) and (2.2), Maxwell's equations (2.5) and (2.6) and the equation for the current but with the ion contribution neglected. In addition, we also add a dissipation term $\nu_e \underset{\sim}{v}_e$ to the left hand side of equation (2.2) where ν_e can be thought of either as the electron-ion collision frequency or more loosely as a term which simulates both collisional and collisionless dissipation for the electrons.

Before proceeding with the analysis we make one important observation. The energy density of an electromagnetic wave in a plasma with an electric field amplitude $\underset{\sim}{E}$ is $\varepsilon_0 \left|E\right|^2/2$. The ratio of this quantity to the energy density $n_0 \kappa T_e$ of the plasma is a measure of the size of perturbation the wave produces. It is fortunate from the analytical point of view that for the threshold fields of many interesting nonlinear effects this ratio is very much less than unity and hence such interactions can be described with the aid of perturbation analysis.

Let us now write the dynamical equations with this in mind. Again splitting the fields into an equilibrium part and an oscillating part we write the equations with the linear terms on the left and the nonlinear terms on the right.

$$\frac{\partial \underset{\sim}{v}_{e1}}{\partial t} + \gamma_e \frac{\kappa T_e}{n_0 m_e} \underset{\sim}{\nabla} n_{e1} + \nu_e \underset{\sim}{v}_{e1} + \frac{e}{m} \underset{\sim}{E}_1 = -\frac{e}{m_e} \underset{\sim}{v}_{e1} \times \underset{\sim}{B}_1 - (\underset{\sim}{v}_{e1} \cdot \underset{\sim}{\nabla}) \underset{\sim}{v}_{e1} + \gamma_e \frac{\kappa T_e}{n_0^2 m_e} n_{e1} \underset{\sim}{\nabla} n_{e1} \tag{5.1}$$

$$\frac{\partial n_{e1}}{\partial t} + n_0 \underset{\sim}{\nabla} \cdot \underset{\sim}{v}_{e1} = -\underset{\sim}{\nabla} \cdot (n_{e1} \underset{\sim}{v}_{e1}) \tag{5.2}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{B}_1 - \frac{1}{c^2} \frac{\partial \underset{\sim}{E}_1}{\partial t} + e\mu_0 n_0 \underset{\sim}{v}_{e1} = -e\mu_0 n_{e1} \underset{\sim}{v}_{e1} \tag{5.3}$$

$$\underset{\sim}{\nabla} \times \underset{\sim}{E}_1 + \frac{\partial \underset{\sim}{B}_1}{\partial t} = 0 \quad . \tag{5.4}$$

We know that the linear terms on the left hand side will describe the linear, high frequency waves of a field-free plasma so that the quadratic terms on the right hand sides which are to be treated as perturbations will provide coupling between the modes. We shall begin by assuming that, initially, there is an incident, finite-amplitude electromagnetic

wave present whose electric field is $\underset{\sim}{E}_{T0}$, magnetic field $\underset{\sim}{B}_{T0}$ and wave vector $\underset{\sim}{k}_{T0}$. We choose the following linear polarization

$$\underset{\sim}{E}_{T0} = (0, E_{T0}, 0)$$

$$\underset{\sim}{B}_{T0} = (0, 0, B_{T0})$$

$$\underset{\sim}{k}_{T0} = (k_{T0}, 0, 0) \quad .$$

The frequency ω_{T0} is then given by

$$\omega_{T0}^2 = \omega_{pe}^2 + c^2k_{T0}^2 \quad .$$

The incident wave is a travelling wave varying as

$$\exp i(k_{T0}x - \omega_{T0}t) \quad .$$

The electron fluid velocity associated with this wave is obtained from the linearized version of equation (5.1) and is

$$v_{T0} = eE_{T0}/i\omega_{T0}m_e \tag{5.5}$$

where we have neglected ν_e (the effect of ν_e will come in later). So far this is just the usual linear theory.

We must now obtain the modified equations for a Langmuir wave and another electromagnetic wave of a freqency different from ω_{T0} due to the presence of the finite amplitude wave $(\omega_{T0}, \underset{\sim}{k}_{T0})$. We shall find that these waves are no longer independent but are coupled together due to the presence of the finite amplitude wave.

First let us obtain the equation of a transverse perturbation which we assume to have the same polarization as the incident wave. Taking the y-components of (5.1) and (5.3) and the z-component of (5.4) and assuming the perturbation varies as $\exp i(kx - \omega t)$, we obtain

$$\begin{aligned}(\omega^2 - \omega_{pe}^2 - c^2k^2)E_{1y} = {} & -i\nu_e \frac{\omega_{pe}^2}{\omega} E_{1y} - \omega_{pe}^2 v_{1x}B_{1z} \\ & + \frac{en_0}{\varepsilon_0} v_{1x} \frac{\partial v_{1y}}{\partial x} + \frac{ie\omega}{\varepsilon_0} n_{e1}v_{1y} \quad .\end{aligned} \tag{5.6}$$

The collisional term can now be seen to produce a linear damping on the transverse wave and the three quadratic terms on the right hand side of (5.6) are coupling terms. We shall treat the damping and coupling terms as small perturbations to the linear waves whose dispersion relation, in the absence of the pump wave, is given by the left hand side of

(5.6). Let us now write the electric field of the transverse perturbation as E_{T1} to distinguish it from the incident wave. Since the transverse perturbation is subject to coupling to other modes we cannot assume its amplitude will remain constant. However, since the coupling is assumed to be weak we may assume that E_{T1} can be written as

$$E_{T1}(x,t) = \mathrm{Re}\{A_{T1}(t) \exp i(k_{T1}x - \omega_{T1}t)\} \tag{5.7}$$

where $$\omega^2_{T1} = \omega^2_{pe} + c^2k^2_{T1}$$

and $A_{T1}(t)$ varies slowly in time in comparison with the rapidly varying linear phase. We must also be careful to take only real parts since the wave equation now contains products of complex amplitudes.

It is convenient to express all transverse quantities in terms of the electric field. Since we are performing a perturbation analysis we can relate B_T and v_T to E_T by means of the linear equations. We then obtain

$$B_T = \frac{k_T}{\omega_T} E_T \tag{5.8}$$

$$v_T = -\frac{ieE_T}{\omega_T m_e} \tag{5.9}$$

where B_T and v_T are written instead of B_{1z} and v_{1y} and may represent either the incident or perturbed transverse wave. Since we are describing a process in which an incident transverse wave (T) is transformed into a scattered transverse wave (T′) and a Langmuir wave L

i.e. $$T = T' + L$$

we shall refer to the perturbed electromagnetic wave as the scattered wave. The coupling terms on the right hand side of (5.6) must consist of products of a transverse field and a Langmuir field. It is clear that the Langmuir fields are v_{1x} and n_{e1}. We shall represent the Langmuir wave amplitude by its electric field E_{1x} which we denote by E_L and again expressing v_{1x} and n_{e1} in terms of E_L by means of the linear equations we obtain

$$v_L = -\frac{i\omega_L\varepsilon_0}{en_0} E_L \tag{5.10}$$

$$n_L = -\frac{ik_L\varepsilon_0}{e} E_L \quad . \tag{5.11}$$

We may now obtain a non-linear differential equation for the scattered electromagnetic wave by expanding (5.6) about the linear solution and identifying ω with $i\partial/\partial t$ to obtain

$$\exp\left\{i(k_{T1}x - \omega_{T1}t)\right\}\frac{\partial A_{T1}}{\partial t} = -i\nu_e\frac{\omega_{pe}^2}{\omega_{T1}^2}E_{T1} + \frac{i}{2}\frac{e\omega_{T1}k_L}{\omega_{T0}m_e}E_L^*E_{T0} \quad . \tag{5.12}$$

In order to obtain the final form of this equation we put

$$E_{T0} = \mathrm{Re}\left\{A_{T0}\exp i(k_{T0} - \omega_{T0}t)\right\} \tag{5.13}$$

and write E_L as the product of a slowly varying amplitude and the linear phase

$$E_L(x,t) = \mathrm{Re}\left\{A_L(t)\exp i(k_Lx - \omega_Lt)\right\} \quad . \tag{5.14}$$

The only significant nonlinear coupling terms will be those which satisfy the frequency and wave number matching conditions

$$\omega_{T0} = \omega_{T1} + \omega_L \tag{5.15}$$

$$k_{T0} = k_{T1} + k_L \tag{5.16}$$

whose physical interpretation is the conservation of energy and momentum. Using equations (5.13)-(5.16) in equation (5.12) we obtain the equation for A_{T1}

$$\frac{\partial A_{T1}}{\partial t} + \frac{\nu_e}{2}\frac{\omega_{pe}^2}{\omega_{T1}^2}A_{T1} = \frac{1}{4}\frac{ek_L}{\omega_{T0}m_e}A_L^*A_{T0}\,e^{-i\psi t} \tag{5.17}$$

where we have imposed perfect k-matching but have allowed for a small frequency mis-match $\psi = \omega_{T0} - \omega_{T1} - \omega_L$.

The equation for E_L is obtained in a similar way. Again assuming an $\exp i(kx - \omega t)$ dependence and taking the x-components of (5.1) and (5.3) together with equation (5.2) we obtain

$$(\omega^2 - \omega_{pe}^2 - \gamma_e k^2v_{Te}^2)E_L = -i\nu_e\omega E_L + \omega_{pe}^2v_{1y}B_{1z} \quad . \tag{5.18}$$

In this case the only coupling term comes from the $\underset{\sim}{v} \times \underset{\sim}{B}$ force in the momentum equation since this is the only term which consists of a product of transverse fields. Expanding (5.18) about the linear solution, using (5.10) and (5.11) and imposing the matching relations (5.15) and (5.16) we obtain the equation for $A_L(t)$

$$\frac{\partial A_L}{\partial t} + \frac{\nu_e}{2}A_L = \frac{1}{4}\frac{e\omega_{pe}^2}{m_e\omega_{T0}\omega_{T1}}\frac{k_L}{\omega_L}A_{T0}A_{T1}^*\,e^{-i\psi t} \quad . \tag{5.19}$$

Equations (5.17) and (5.19) describe the coupling between high frequency transverse and longitudinal perturbations in the presence of a finite amplitude transverse wave. Under these conditions we have $A_{T0} >> A_{T1}$ and $A_{T0} >> A_L$. We may therefore linearize

equations (5.17) and (5.19) by assuming that the incident wave amplitude remains constant. Using A_{T1} and $A_L^* e^{-i\psi t}$ as the amplitudes the coupled equations then have constant coefficients so that a solution proportional to $\exp(-i\Omega t)$ may be assumed. The dispersion relation for the coupled waves is

$$(\Omega + i\gamma_T)(\Omega + i\gamma_L - \psi) + C_{0L}C_{01}\left|A_{T0}\right|^2 = 0 \qquad (5.20)$$

where

$$\gamma_L = \frac{\nu_e}{2}, \quad \gamma_T = \frac{\nu_e}{2}\frac{\omega_{pe}^2}{\omega_{T1}^2}, \quad C_{01} = \frac{e\omega_{pe}^2 k_L}{4m_e \omega_{T0}\omega_{T1}\omega_L}$$

and

$$C_{0L} = \frac{ek_L}{4m_e\omega_{T0}} .$$

If we neglect the damping terms and put $\psi = 0$ we obtain a growing solution with growth rate

$$\gamma = (C_{0L}C_{01})^{1/2}\left|A_{T0}\right| . \qquad (5.21)$$

Putting $C_{0L}C_{01}\left|A_{T0}\right|^2 \equiv K$ we obtain the threshold condition for instability

$$K = \gamma_T\gamma_L + \frac{\psi^2\gamma_T\gamma_L}{(\gamma_T + \gamma_L)^2} . \qquad (5.22)$$

Clearly, the minimum threshold occurs, as expected, for perfect frequency matching. For an incident wave whose amplitude exceeds the above threshold a scattered electromagnetic wave and a Langmuir wave would grow out of the background noise. In terms of the plasma parameters the threshold can be written as

$$\frac{v_0^2}{v_{Te}^2} = \frac{4}{k_L^2\lambda_{De}^2}\frac{\nu_e}{\omega_{T1}}\frac{\nu_e}{\omega_L} \qquad (5.23)$$

where $\lambda_{DE} \equiv v_{Te}/\omega_{pe}$ and v_0 is the velocity of the electron fluid in the field of the incident wave. For parameters typical of laser plasmas $v_0^2/v_{Te}^2 \ll 1$ in keeping with our perturbation analysis.

The above interaction is just one example of a whole class. We can use this example to illustrate some important conservation relations which such interactions satisfy. To do this we must first obtain the nonlinear equation for the incident wave A_{T0} whose amplitude is no longer assumed constant as we must relax the constraint that $A_{T0} \gg A_{T1}$ and A_L. The equation for A_{T0} can be obtained from equation (5.6) and is

$$\frac{\partial A_{T0}}{\partial t} + \frac{\nu_e}{2} \frac{\omega_{pe}^2}{\omega_{T0}^2} A_{T0} = \frac{1}{4} \frac{ek_L}{\omega_{T1} m_e} A_L A_{T1} e^{i\psi t} \quad . \tag{5.24}$$

Equations (5.17), (5.19) and (5.24) now form a closed system of coupled nonlinear differential equations for the three waves, two transverse and one Langmuir. Using the result that the wave energy densities of transverse and Langmuir waves are given by

$$\rho_T = \frac{1}{2} \varepsilon_0 \left| E_T \right|^2 \tag{5.25}$$

$$\rho_L = \frac{1}{2} \varepsilon_0 \left| E_L \right|^2 \frac{\omega_L^2}{\omega_{pe}^2} \tag{5.26}$$

we normalize the wave amplitudes to the total energy density carried by each mode as follows

$$a_{T0,1} = \left(\frac{\varepsilon_0}{2}\right)^{1/2} A_{T0,1} \tag{5.27}$$

$$a_L = \left(\frac{\varepsilon_0}{2}\right)^{1/2} \frac{\omega_L}{\omega_{pe}} A_L \quad . \tag{5.28}$$

Neglecting the damping terms and assuming perfect matching the equations for the interacting waves are

$$\frac{\partial a_{T0}}{\partial t} = -\Gamma \omega_{T0} a_{T1} a_L \tag{5.29}$$

$$\frac{\partial a_{T1}}{\partial t} = \Gamma \omega_{T1} a_{T0} a_L^* \tag{5.30}$$

$$\frac{\partial a_L}{\partial t} = \Gamma \omega_L a_{T0} a_{T1}^* \tag{5.31}$$

where

$$\Gamma = \left(\frac{2}{\varepsilon_0}\right)^{1/2} \frac{ek_L \omega_{pe}}{4 m_e \omega_{T1} \omega_{T0} \omega_L} \quad .$$

With the aid of equations (5.29)-(5.31) it is straightforward to show that

$$\frac{\partial}{\partial t} \left(\left| a_{T0} \right|^2 + \left| a_{T1} \right|^2 + \left| a_L \right|^2 \right) = 0 \tag{5.32}$$

which corresponds to the conservation of energy for the interacting waves.

A more revealing result concerns the wave action density $\left|a_n\right|^2/\omega_n$. Calculating the rate of change of this quantity from equations (5.29)-(5.31) we obtain the Manley-Rowe relations (first discussed in the field of electronics)

$$-\frac{1}{\omega_{T0}}\frac{\partial}{\partial t}\left|a_{T0}\right|^2 = \frac{1}{\omega_{T1}}\frac{\partial}{\partial t}\left|a_{T1}\right|^2 = \frac{1}{\omega_L}\frac{\partial}{\partial t}\left|a_L\right|^2 \qquad (5.33)$$

which shows in what proportion the energy from one wave is divided between the other two - the lowest-frequency wave receives the least energy. The coupled equations can also be used, with the aid of perfect wave-number matching, to demonstrate the conservation of momentum in the interaction. These general conservation relations are characteristic of all three wave interactions.

5.1 The beat-wave accelerator

To conclude this discussion we will consider the application of this interaction to the acceleration of electrons. We noted earlier that the longitudinal Langmuir wave could propagate with a wide range of phase velocities both greater and much less than the velocity of light. Since the Langmuir wave is longitudinal it can accelerate electrons provided it has a finite amplitude since an electron travelling close to the phase speed of the wave will be trapped in the potential well of the wave. An electron with the correct phase will then be accelerated from the bottom of the potential well to the top gaining energy from the wave. If two laser signals are launched into a plasma such that they differ in frequency by the Langmuir wave frequency their beating will generate a Langmuir wave. The energy transferred to the Langmuir wave will be subject to the constraint given by the Manley-Rowe relation. The three wave interaction can also be represented on an ω-k diagram which is shown in figures 4a,b.

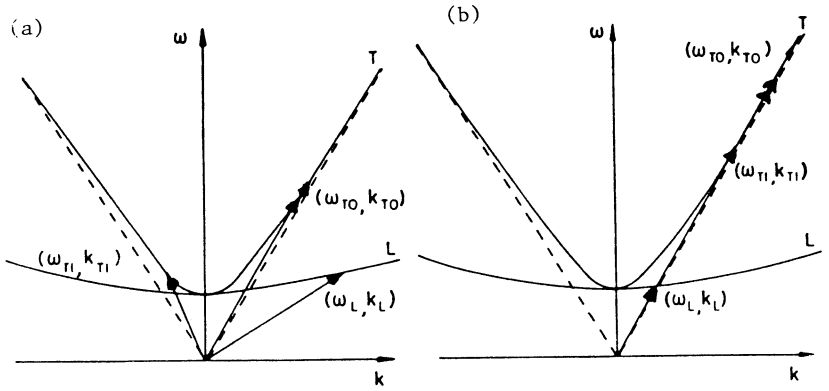


Fig. 4 Nonlinear coupling of two transverse waves with Langmuir wave (a) $\omega_{T1} \sim \omega_{pe}$, (b) $\omega_{T1} >> \omega_{pe}$

Figure 4a is for the case where the lower frequency laser is comparable to the plasma frequency and figure 4b where the lower laser frequency is much larger than the plasma frequency. These two cases correspond to high or low-density plasmas respectively. In the first case more energy would be coupled to the Langmuir wave but it would be excited at phase velocities below the speed of light whereas for the second case the Langmuir wave would travel at or near the speed of light but would receive proportionately less of the laser energy and so would require more power to generate the same strength Langmuir wave.

Despite the constraint imposed by the Manley-Rowe relation the second case is the favoured one for the beat wave accelerator since all three waves travel close to the speed of light. In any case the practical situation is more complicated since both laser fields can act as pumps thus giving rise to further waves. In addition, since $\omega_L << \omega_T$, not only the down-shifted wave but also the up-shifted one, $\omega_T + \omega_L$, $k_T + k_L$, must be included. The division of energy is now more complicated. However, no matter how many additional couplings are included the basic interaction is still given by the previous discussion of Raman scattering.

To give some idea of the possibilities offered by the beat wave accelerator we use present neodymium-glass lasers as an example. These lasers can presently deliver $\sim 5 \times 10^{13}$ watts. For two such lasers with beam widths $\sim$ 1mm the intensity $\sim 10^{16}$ watts / cm^2. The generated Langmuir wave can then have longitudinal electric fields $\sim 10^8$ volts/cm which assumes a perturbed electron density $\delta n/n_o \sim 0.1$ in a plasma where n_o is in the range $10^{16} - 10^{18} cm^{-3}$. If the beat wave could be maintained over a length $\sim$ 5m then an electron injected with 1 MeV energy could, in principle, be accelerated up to 50 GeV in the above distance. However, present experiments are only in the very early stages. In particular, Langmuir wave electric fields $\sim 10^8$ volts/cm have been generated but only over a distance of 1mm. The reader is referred to the original work for a detailed discussion of the above points.

* * *

REFERENCES

The following are a selection of references on waves in plasmas:

T.H.Stix, The theory of plasma waves, McGraw-Hill, New York, (1962).

T.E.Stringer, Plasma Phys. (J.Nucl.Energy, Part C) 5, 89 (1963).

V.D.Shafranov, in Rev. Plasma Physics (Ed. M.A. Leontovich) Consultants Bureau, New York, (1967). vol.3, p.1.

A.I.Akhiezer, I.A.Akhiezer, R.V.Polovin, A.G.Sitenko and K.N. Stepanov, Plasma Electrodynamics Vol.1 Linear Theory Pergamon Press, Oxford, (1975).

C.N.Lashmore-Davies, Plasma Physics and Instabilities, CERN 81-13 Chapters II and V.

The original idea of the beat wave accelerator is due to:

T.Tajima and J.M. Dawson, Phys. Rev. Letters 43, 267, (1979).

Other references on the subject are:

C.Joshi, T.Tajima, J.M. Dawson, H.A.Baldis and N.A.Ebrahim, Phys. Rev. Letters, 47, 1285 (1981).

C.Joshi, W.B.Mori, T. Katsouleas, J.M.Dawson, J.M.Kindel and D.W.Forslund, Nature, 311, 525, (1984).

R.Bingham, Possible Instabilities in the Beat Wave Accelerator, Rutherford Appleton Laboratory Report RL-83-058.

LANDAU DAMPING

H.G. Hereward
United Kingdom

ABSTRACT

Landau damping is a process that arises when one considers a whole collection of particles or other systems which have a spectrum of natural frequencies, and interact in some way. In accelerators we are usually concerned with an interaction of a kind that may make the beam unstable, and we want to find out whether or not the spread of natural frequencies will stabilize it.

1. SPECTRUM OF LINEAR OSCILLATORS

I shall start by looking at a simple model, a collection of harmonic oscillators with a common driving force that comes from their average displacement. This is in fact quite a good model of the transverse coasting beam instability of the kind seen in the ISR.

$$\ddot{x}_k + \omega_k^2 x_k = f(t) \tag{1a}$$

$$f(t) = c\langle x \rangle \ . \tag{1b}$$

The triangular bracket means average over all k and I use c to stand for some complex constant. Later we shall look at the meaning of its real and imaginary parts.

The obvious method of attack is to assume that everything depends on time like exp $j\omega t$ and look for values of ω, real or complex, that enable Eq. (1) to be satisfied. Put

$$x_k = X_k \exp j\omega t \tag{2}$$

Then Eq. (1) becomes

$$X_k = c\langle X \rangle \frac{1}{\omega_k^2 - \omega^2} \ . \tag{3}$$

For simplicity I shall assume that all ω_k and ω are close to some positive ω_0, and make the approximation

$$\omega_k^2 - \omega^2 = 2\omega_0(\omega_k - \omega) \tag{4}$$

so Eq. (3) becomes

$$X_k = \frac{c}{2\omega_0} \langle X \rangle \frac{1}{\omega_k - \omega} \tag{5}$$

and now average this over k, with $\rho(\omega_k)$ the normalized spectrum of particles' resonant frequencies

$$1 = \frac{c}{2\omega_0} \int \frac{\rho(\omega_k)\, d\omega_k}{\omega_k - \omega} \quad . \tag{6}$$

The ω_k are real and you have to integrate over all the ω_k where ρ is not zero. If ω is real and falls within that domain, the denominator goes to zero and the integral is singular. This is the difficulty that makes Landau damping a non-trivial problem.

There is little use in struggling against this singularity by mathematical tricks, already from Eq. (3) you can see that there is no finite solution of the postulated form Eq. (2) when ω is real and falls upon one of the ω_k. There are several ways of proceeding[1]. For our purposes the most simple is to remind ourselves that what we want to know is whether the system is unstable or not, so we are interested in the possibility of exponentially growing solutions:

$$\omega = \text{Re } \omega + j \text{ Im } \omega \,, \quad \text{Im } \omega < 0 \tag{7}$$

That removes the singularity from Eq. (6), but it remains quite difficult to solve it for ω given ρ and c. The procedure initiated at CERN[2] is a mapping technique. We write Eq. (6) as

$$\frac{c}{2\omega_0} = \left(\int \frac{\rho(\omega_k)\, d\omega_k}{\omega_k - \omega} \right)^{-1} \tag{8}$$

and evaluate the right hand side for all relevant values of Re ω with Im $\omega = -|\varepsilon|$. The result is plotted in the complex plane and gives a curve which is the locus of values of $c/2\omega_0$ at which the system is just unstable.

The real part of the integral tends to

$$\text{P.V.} \int \frac{\rho(\omega_k)\, d\omega_k}{\omega_k - \text{Re } \omega} \quad . \tag{9}$$

You then have the resonance dominator and the spectrum, Fig. 1, and for each value of Re ω you combine them and integrate over the particles. P.V. (principal value) means that the $\pm \infty$ at Re ω are cut out. The denominator is an odd function of ω_k - Re ω, so it is the part of ρ that is antisymmetrical about Re ω that contributes to this real part.

The imaginary part of the integral comes from passing close to the pole at ω, and tends to

$$-j\, \pi\, \rho(\text{Re } \omega) \quad . \tag{10}$$

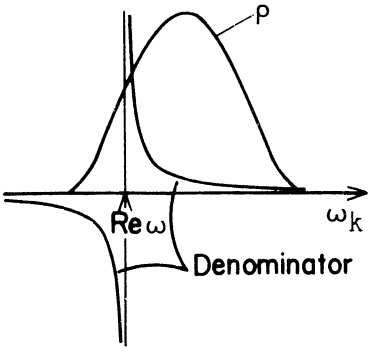


Fig. 1 Evaluation of a point on the stability diagram.

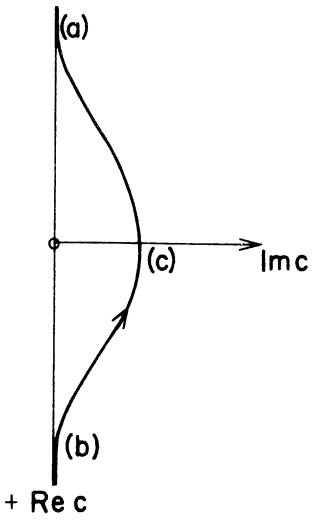


Fig. 2 Stability diagram for a bell-shaped spectrum of oscillators.

When it is drawn with certain normalizations and conventions, this threshold curve Eq. (8) in the complex plane is known as the stability diagram. It is a property of the shape of the spectrum ρ. For reasonable and simple spectra ρ the stability diagram is fairly simple, like Fig. 2[3], and it is easy to see that the side containing the origin is stable, the "far" side is unstable. If you want to know something about growth rates in the unstable region, there is no problem in computing Eq. (8) for all Re ω with some chosen negative Im ω, to get the contour corresponding to that growth rate in the $c/2\omega_0$ plane.

Some particular features of these stability diagrams should be pointed out. Since ρ is positive and $|\varepsilon|$ is positive, the threshold curve lies entirely in Im $c > 0$. This corresponds in Eq. (1) to a force component in phase with the velocity $\langle \dot{x} \rangle$, such as one needs to make antidamping for a single particle or delta-function spectrum. When the spectrum ρ does not have infinite tails, there are two points (a) and (b) on the threshold curve where it comes down almost to Im $c = 0$. They correspond to putting ω at the top or bottom of the spectrum, and for typical bell-shaped spectra one finds them at

$$\frac{c}{2\omega_0} \approx \mp\, 0.8\ \Delta\omega_{HWB} + j|\varepsilon| \quad , \tag{11}$$

where $\Delta\omega_{HWB}$ is the base half width of this spectrum .

This has a simple physical interpretation: If we go back to Eq. (1) and forget about the spread of frequencies, we see that a real c just corresponds to a real ω-shift of $- c/2\omega_0$. So Eq. (11) is saying that an $\langle x \rangle$

tune-shifting term a little less than the base halfwidth of the spectrum is just enough to move the coherent frequency ω outside the spectrum of the particles, and this kills the Landau damping, in the sense that then an infinitesimal antidamping component of force will produce growth.

Then there is the point (c) where this threshold curve crosses the imaginary axis. Consider Re ω on the peak of a simple symmetrial spectrum, the real part of the integral vanishes by symmetry, the imaginary part comes entirely from passing close to the pole at ω, giving

$$\begin{aligned} \frac{c}{2\omega_0} &= \frac{j}{\pi\,\hat{\rho}} \\ &\approx 0.3j\ \Delta\omega_{FWHM} \\ \text{or} \qquad &\approx 0.6j\ \Delta\omega_{HWHM}\ , \end{aligned} \tag{12}$$

where $\Delta\omega_{FWHM}$ and $\Delta\omega_{HWHM}$ are the full and half widths of the spectrum respectively at half maximum.

This means that a finite amount of antidamping component (positive imaginary, in phase with $\langle\dot{x}\rangle$) is needed in c to produce even an infinitesimal growth rate at a frequency within the spectrum, and this is because the response of the spectrum of particles comes only from the small fraction of them that are close to resonance. This the key to a physical understanding of what goes on in Landau damping situations, so it is worth while checking over the orders of magnitude.

A growth rate α corresponds to a bandwith α, so the fraction of the spectrum that responds as though on resonance is of the order of $\alpha/\Delta\omega$. We can approximate Eq. (6) by

$$1 \approx \frac{c}{2\omega_0}\,\frac{\alpha/\Delta\omega}{j\alpha} \tag{13}$$

so

$$\frac{c}{2\omega_0} \approx j\Delta\omega \tag{14}$$

which confirms Eq. (12). One needs that sort of c to produce a growth rate α however small. Smaller growth rate of a resonant system requires smaller positive feedback, but this is provided by the smaller number of particles which count as being on resonance, without appreciable change of c.

Before leaving this model it is worth while to relate the complex coefficient c to the quantities U and V that you will meet in the literature[3] of transverse instabilities. The betatron modes that are potentially unstable correspond to the so-called slow waves which have frequency

$$\omega_n = \Omega_{rev}(n - Q) \tag{15}$$

and for these waves one has the correspondence*)

$$\begin{aligned} \mathrm{Re}\ \frac{c}{2\omega_0} &= -\ U \\ \mathrm{Im}\ \frac{c}{2\omega_0} &= \quad V \end{aligned} \tag{16}$$

where I use always the exp(+ jωt) convention.

The calculated stability diagrams for different shapes of spectrum can be made universal*) by plotting normalized values,

$$\begin{aligned} U' &= \frac{U}{\Delta\omega_{HWHM}} \ , \\ V' &= \frac{V}{\Delta\omega_{HWHM}} \ . \end{aligned} \tag{17}$$

2. LONGITUDINAL INSTABILITY

Landau damping of longitudinal instability in a coasting beam is rather similar to the oscillator problem just discussed, but there are important differences which make it worth while handling the problem by use of the Vlasov equation. This is

$$\frac{\partial\psi}{\partial t} = -\ \dot{\theta}\ \frac{\partial\psi}{\partial\theta} - \dot{W}\ \frac{\partial\psi}{\partial W} \ , \tag{18}$$

where $\psi(\theta, W, t)$ is the distribution function in longitudinal phase-space, θ, W.

By first order perturbation and Fourier expansion in θ one can get the linearized equation of motion for the n'th azimuthal harmonic of the perturbation[5])

$$\frac{\partial\xi_n(W,t)}{\partial t} - n\ j(\omega_0 + k_0 W)\ \xi_n(W,t) = -\ eV_n\ \frac{\partial\psi_0}{\partial W} \tag{19}$$

with

$$V_n = -\ e\ Z_n\omega_0 \int \xi_n(W,t)\ dW \ . \tag{20}$$

*) Sometimes U + V is written where we have U, see Ref. 4.

Although this pair of equations looks much more complicated than Eq. (1), there are important similarities: the left hand side of Eq. (19) is a differential operator with characteristic frequency $n(\omega_0 + k_0 W)$ which varies with W across the spectrum of particles present, and the driving term on the right of Eq. (19) involves V which is an integral (just as in Eq. (1) it was an average) over the spectrum, of the responses ξ. Z_n is the longitudinal coupling impedance.

We use the same method as before, looking for oscillatory solutions with a slow exponential growth

$$\xi_n(W,t) = X_n(W) \exp j\omega t \qquad (21)$$

$$\omega = \mathrm{Re}\ \omega + j\ \mathrm{Im}\ \omega, \quad \mathrm{Im}\ \omega < 0 \qquad (22)$$

and Eq. (19, 20) becomes

$$1 = e^2 Z_n \omega_0 \int \frac{d\phi_0/dW}{j\omega - n\, j(\omega_0 + k_0 W)}\, dW \ . \qquad (23)$$

The mapping technique then consists in evaluating that integral (called dispersion integral) for all relevant values of Re ω, with Im $\omega = -|\varepsilon|$, and so obtaining the locus of values of Z_n at which the system is just unstable. A typical shape for a simple bell-shaped spectrum is shown in Fig. 3, and many different spectra have been calculated [2,3,6].

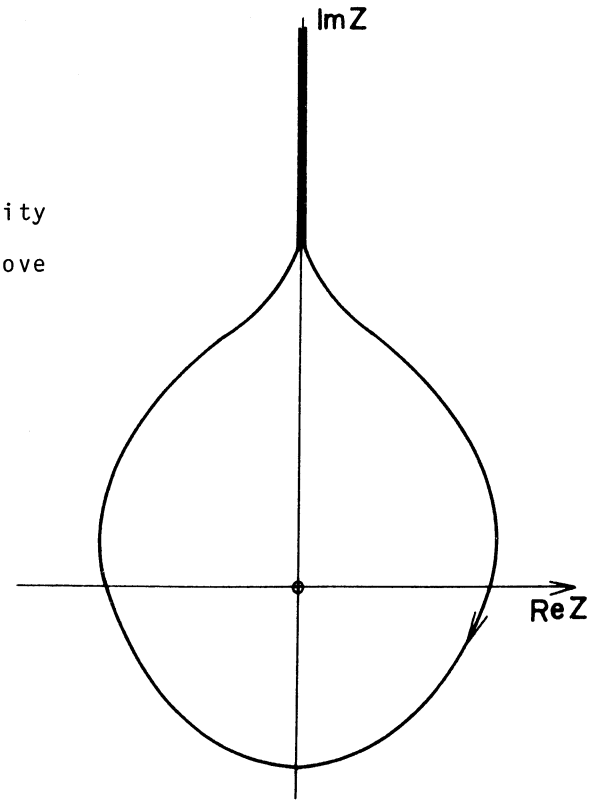


Fig. 3 Longitudinal stability diagram for a bell-shaped spectrum (above transition)

The main difference between this longitudinal, Fig. 3, and the transverse, Fig. 2, is that the longitudinal stability limit passes both to the right and to the left of the origin, showing that instability can be caused by either positive or negative resistivity in Z_n. This is due to the occurence of $d\phi_0/dW$ in Eqs (23) and (19), which in turn is because W is both the quantity that determines the particle's natural frequency in the left-hand side of (19) and in the denominator of (23), and also one of the canonical variables in (18). Equation (18) can be written as a total differential

$$\frac{d\psi}{dt} = 0$$

which implies that phase-space density in the θ, W plane is conserved (Liouville's theorem). When one has conservation of phase-plane density, the particles move in the phase-plane like an incompressible fluid, and any perturbing force can only produce a change in the distribution by acting on particles that are where there is some gradient in the unperturbed phase-space density. Hence the appearance of $d\phi_0/dW$ in Eqs (19) and (23), and the left- and right-hand sides of Fig. 3 are associated with the top and bottom flanks respectively of the spectrum $\phi_0(W)$, where $d\phi_0/dW$ is negative and positive respectively.

The relationship between longitudinal coupling impedance and the quantities U and V of the literature are sufficiently discussed in Refs. 6 and 7.

3. NONLINEAR OSCILLATIONS

The third kind of Landau damping that can arise in accelerator beams is due to nonlinearity of the oscillations. On the face of it, this nonlinearity combines with the spread of unperturbed oscillation amplitudes to produce a spread of the natural frequencies which might be dealt with by the method of section 1, but this is wrong. There is a discussion of the details of why this is so in Ref. 8, but a fairly straight-forward way to get the right answer is to proceed via the Vlasov equation. One can make some simplifications[9] compared with the treatment of Ref. 4.

Consider a collection of particles in a nearly harmonic potential well,

$$\ddot{x} + \omega_0^2 x + \text{nonlinear term} = 0 \quad . \qquad (24)$$

In terms of amplitude and phase, the orbits are

$$\begin{aligned} x &= a \cos \phi + \text{small terms} \\ \dot{x} &= \omega_0 \, a \sin \phi + \text{small terms} \end{aligned} \tag{25}$$

or $\dot{a} = 0$, $\dot{\phi} = -\omega(a)$ if we neglect the small terms.

The Vlasov equation

$$\frac{\partial \psi}{\partial t} + \dot{x} \frac{\partial \psi}{\partial x} + \ddot{x} \frac{\partial \psi}{\partial \dot{x}} = 0 \tag{26}$$

for the density ψ in the phase-plane can be written as

$$\frac{\partial \psi}{\partial t} + \dot{a} \frac{\partial \psi}{\partial a} + \dot{\phi} \frac{\partial \psi}{\partial \phi} = 0$$

or

$$\frac{\partial \psi}{\partial t} - \omega(a) \frac{\partial \psi}{\partial \phi} = 0 \, . \tag{27}$$

A distribution $\psi_0(a)$ of only amplitudes is stationary.

Now add a driving term f to the RHS of Eq. (24), and the corresponding term $f \frac{\partial \psi}{\partial \dot{x}}$ to the LHS of (26) and (27). Small oscillations about the stationary solution

$$\psi = \psi_0(a) + \psi_1(a, \phi) \, e^{j\omega t} \tag{28}$$

are described by the linearized Vlasov equation

$$j\omega\psi_1 - \omega(a) \frac{\partial \psi_1}{\partial \phi} + f \frac{\sin\phi}{\omega_0} \frac{d\psi_0}{da} = 0 \tag{29}$$

where the identity

$$\frac{\partial \psi_0}{\partial \dot{x}} = \frac{\sin \phi}{\omega_0} \frac{\partial \psi_0}{\partial a} \tag{30}$$

has been used.

If we neglect the $\exp(-j\phi)$ part of $\sin \phi$, and assume that the driving term $f_{(x)}$ is approximately constant in the region of the beam, then the ϕ-dependence of ψ_1 is $\psi_1(a, \phi) = \psi_1(a) \, e^{j\phi}$. The neglected part of $\sin \phi$ produces only rapidly oscillating terms which are negligible when ω is close to $\omega(a)$

$$\psi_1(a) = \frac{f}{2\omega_0} \frac{d\psi_0/da}{\omega - \omega(a)} \, . \tag{31}$$

From this ψ_1 the barycentre motion can be obtained by integration and the loop closed by finding the electrodynamic properties of the beam environment in the usual way[4].

This nonlinear oscillator case resembles the longitudinal case in that the beam response comes from the part of the spectrum where there is some phase-space density gradient, and that ϕ_0 gradient appears in the dispersion integral not ϕ_0 itself. On the other hand, it does not imitate the longitudinal stability diagram by going both left and right of the origin for an ordinary simple distribution, because an ordinary simple $\phi_0(a)$ starts at some peak value at $a = 0$ and only has negative d/da, unlike a simple longitudinal $\phi_0(W)$ which has positive and negative d/dW. Some examples of these nonlinear transverse stability diagrams are given in Ref. 3.

* * *

REFERENCES

1) H.G. Hereward, CERN 65-20 (1965).

See also J.D. Lawson, CERN 76-09, section 5.4. and references therein (1976).

2) A.G. Ruggiero and V.G. Vaccaro, CERN ISR-TH/68-33 (1968).

3) K. Hübner, A.G. Ruggiero and V.G. Vaccaro, Proc. VIIth Int. Conf. on High Energy Accelerators, Yerevan, 1969 (Acadamy of Sciences of Armenian SSR, Yerevan, 1970).

Also Hübner at al. CERN ISR-TH-RF/69-23 (1969), CERN ISR-RF-TH/70-2, CERN ISR-TH/70-44 (1970).

4) L.J. Laslett, V.K. Neil and A.M. Sessler, Rev. Sci. Inst. 36, 437, (1965).

5) V.K. Neil and A.M. Sessler, Rev. Sci. Inst. 36, 429, (1965).

6) B. Zotter, CERN ISR-GS/76-11 (1976).

7) E. Keil and W. Schnell, CERN ISR-TH-RF/69-48 (1969).

8) H.G. Hereward, CERN/MPS/DL 69-11 (1969).

9) F. Sacherer, private communication 1976.

10) H.G. Hereward, CERN 77-13 (1977) p. 219.

BUNCHED BEAM COHERENT INSTABILITIES

J.L. Laclare
Laboratoire National Saturne, 91191 Gif-sur-Yvette Cedex, France

ABSTRACT

In this chapter, we will deal with coherent longitudinal and transverse instabilities. It is a collective phenomenon which prevents one from increasing the current circulating in an accelerating device without losing the beam or spoiling its characteristics.

1. INTRODUCTION

The origin of the mechanism is the electromagnetic field created by the beam itself. This self-field is proportional to the beam intensity. Furthermore, like any solution of Maxwell's equations, because of boundary conditions, it depends strongly on the geometry and the electromagnetic properties of the environment. When the intensity gets large enough, it becomes sizeable in the sense that it cannot be neglected anymore when compared to the external guide field.

Dealing with coherent instabilities consists in solving the equation of motion of a population of particles while adding the self-field effect. Obviously, the self field perturbs the single particle motion, but this is not the remarkable effect. The important point is that under certain conditions the beam as a whole is unstable.

In literature, there is a countless list of contributions to the subject. The first to come have been written in the fifties. Nowadays, the subject is still in fashion. Many reports per year are being produced. This shows how difficult and important the subject is.

During these two chapters, I will review the fundamentals of coherent instabilities. In this respect, F. Sacherer's work is certainly the basic source. The main material for this chapter is drawn from the numerous reports he wrote about ten years ago. Numerous developments are derived from B. Zotter and G. Besnier's contributions.

In the following, we will only study bunched beams in circular machines; first longitudinal and then transverse motion.

2. LONGITUDINAL INSTABILITIES

2.1 Single particle longitudinal motion

With respect to the synchronous particle that circulates at the angular revolution frequency

$$\omega_0 = \frac{\beta c}{R} \tag{1}$$

and crosses the Radio Frequency gap when the RF phase is φ_s, we describe the single particle motion with a pair of conjugate coordinates

$$\tau \qquad \text{and} \qquad \dot{\tau} = \frac{d\tau}{dt} \ . \tag{2}$$

For a fixed observer located at azimuthal position θ around the machine, τ expressed in seconds represents the time interval between the reference particle passing and the test particle passing. The second coordinate

$$\dot{\tau} = \eta \frac{dp}{p_{//}} = - \frac{d\omega}{\omega_0} \tag{3}$$

measures the instantaneous momentum deviation of the test particle. The parameter η is negative below the transition energy.

$$\eta = \frac{1}{\gamma_t^2} - \frac{1}{\gamma^2} \ . \tag{4}$$

In smooth machines γ_t is of the order of Q_x (horizontal wave number).

We assume a purely linear synchrotron oscillation around the synchronous particle at frequency ω_{s_0}.

$$\ddot{\tau} + \omega_{s_0}^2 \tau = 0 \tag{5}$$

$$\omega_{s_0}^2 = \frac{-\eta\, \omega_0\, V_{RF}\, h \cos\varphi_s}{2\pi R \left(\frac{p_{//}}{e}\right)} \tag{6}$$

e is the elementary charge
R the machine radius
$p_{//}$ the synchronous particle momentum
V_{RF} the peak RF voltage
h the RF harmonic number

In $(\tau, \dot{\tau})$ phase space the trajectory associated with the unperturbed motion is an ellipse (Fig. 1a).

$$\tau^2 + \frac{\dot{\tau}^2}{\omega_{s_0}^2} = \hat{\tau}^2 = \frac{\eta^2}{\omega_{s_0}^2}\left(\frac{\hat{dp}}{p_{//}}\right)^2 . \tag{7}$$

a) b)

Fig. 1 Synchrotron motion in phase space. a) Phase space $(\tau, \dot{\tau})$ b) Normalised phase space $(\tau, \dot{\tau}/\omega_{s0})$ or $(\hat{\tau}, \psi)$

This ellipse becomes a circle when using a set of normalized coordinates like τ, $\frac{\dot{\tau}}{\omega_{s_0}}$ or phase and amplitude coordinates ψ, $\hat{\tau}$ (Fig. 1b).

As stated from the beginning, we are mainly concerned with the electromagnetic field induced by the beam. This field has electric and magnetic components. It modifies the differential equation of motion (5) of the single particle by adding a term in the right hand side of the equation

$$\ddot{\tau} + \omega_{s_0}^2 \tau = \frac{\eta}{p_{//_0}} \frac{dp_{//}}{dt} = \frac{\eta}{p_{//_0}} e \left[\vec{E} + \vec{\beta c} \wedge \vec{B}\right]_{//}(t, \theta) . \tag{8}$$

2.2 Single-particle longitudinal signal

With the object of writing down Maxwell's equations the solution of which leads to the beam self-field, one needs to express the local charge and current at time t and position θ in the machine. Machine physicists are used to observing the intensity on an oscilloscope by looking at the signal drawn from longitudinal P U electrodes. These electrodes are non-destructive diagnostic equipments which measure the electromagnetic field locally induced by the beam.

Let us assume a perfect P U electrode located at angular azimuth θ, and analyse the intensity signal when a single test particle oscillates in the external guide field (no self field added). This will help us to get more familiar with time domain and frequency domain.

In the time domain, the elementary intensity signal $s_{//}(t, \theta)$ is a series of nearly periodical impulses delivered at each passage through the electrode.

$$s_{//}(t, \theta) = e \sum_{h=-\infty}^{h=+\infty} \delta\left(t - \tau - \frac{\theta}{\omega_0} - 2h\frac{\pi}{\omega_0}\right) \quad \text{Ampere} \tag{9}$$

δ is the Dirac function.

By using the following relations,

$$u = t - \tau - \frac{\theta}{\omega_0} \qquad \sum_{k=-\infty}^{k=+\infty} \delta\left(u - 2\frac{k\pi}{\omega_0}\right) = \frac{\omega_0}{2\pi} \sum_{p=-\infty}^{p=+\infty} e^{jp\omega_0 u} \tag{10}$$

$$\tau = \hat{\tau} \cos(\omega_{s_0} t + \psi_0) \tag{11}$$

$$e^{-jp\omega_0 \hat{\tau} \cos(\omega_{s_0} t + \psi_0)} = \sum_{m=-\infty}^{m=+\infty} j^{-m} J_m(p\omega_0 \hat{\tau})\, e^{jm(\omega_{s_0} t + \psi_0)} \tag{12}$$

where ψ_0 is the synchrotron phase at time $t = 0$
it can be rewritten in the equivalent form

$$s_{//}(t,\theta) = \frac{e\omega_0}{2\pi} \sum_{p,m=-\infty}^{p,m=+\infty} j^{-m} J_m(p\omega_0 \hat{\tau})\, e^{j(\omega_{pm} t - p\theta + m\psi_0)} \tag{13}$$

in which $\omega_{pm} = p\omega_0 + m\omega_{s0}$. (14)

The Fourier transform of the elementary signal is given by

$$s_{//}(\omega,\theta) = \frac{1}{2\pi} \int_{t=-\infty}^{t=+\infty} s_{//}(t,\theta)\, e^{-j\omega t}\, dt \ . \tag{15}$$

With the Fourier transform we can pass on to frequency domain. The actual spectrum of our single particle moving in the external guide field is a line spectrum at frequencies ω_{pm}

$$s_{//}(\omega,\theta) = \frac{e\omega_0}{2\pi} \sum_{p,m=-\infty}^{p,m=+\infty} j^{-m} J_m(p\omega_0 \hat{\tau})\, e^{-j(p\theta - m\psi_0)}\, \delta(\omega - \omega_{pm}) \ . \tag{16}$$

Around every harmonic of the revolution frequency, there is an infinite number of synchrotron satellites. The spectral amplitude of the mth satellite is given by $J_m(p\omega_0\hat{\tau})$ (Bessel function of order m). The spectrum is centered at the origin and spreads in the negative and in the positive frequency domain. Because the argument of the Bessel functions is proportional to $\hat{\tau}$, the width of the spectrum behaves like $(\hat{\tau})^{-1}$. As a consequence, when considering a bunch, particles with small (large) synchrotron amplitude contribute to the high (low) frequency part of the spectrum. At the limit $\hat{\tau} \longrightarrow 0$, the synchrotron satellites disappear. The synchronous particle resumes the observation point periodically. Its spectrum is a line spectrum at harmonics of the revolution frequency.

$$s_{//}(\omega,\theta) = \frac{e\omega_0}{2\pi} \sum_{p=-\infty}^{p=+\infty} J_0(p\omega_0 \hat{\tau})\, e^{-jp\theta}\, \delta(\omega - p\omega_0) \tag{17}$$

2.3 Distribution of particles

The next step consists in gathering particles to form a bunch (a single bunch for the moment). Therefore we have to choose a distribution function $\Psi(\psi_0, \hat{\tau}, t)$ which will

indicate the particle density in phase space.

The signal or the electromagnetic field induced on the P U electrode by the entire beam is obtained by summing up the elementary signal of individual particles over the distribution.

$$S_{//}(t,\theta) = N \int_{\psi_0=0}^{\psi_0=2\pi} \int_{\hat{\tau}=0}^{\hat{\tau}=+\infty} \Psi(\psi_0, \hat{\tau}, t)\, s_{//}(t,\theta)\, \hat{\tau}\, d\hat{\tau}\, d\psi_0 \qquad (18)$$

where N is the number of particles per bunch.
For obvious reasons of normalization, Ψ has to satisfy equation (19)

$$\int \Psi \, dv = 1 \qquad (19)$$

2.3.1 Stationary distribution of particles

First, let us consider a stationary distribution, that is to say a distribution that depends on $\hat{\tau}$ solely.

$$\Psi(\psi_0, \hat{\tau}, t) = q_0(\hat{\tau}) \, . \qquad (20)$$

For such a distribution, as it was already the case for the synchronous particle, the frequency spectrum of the electromagnetic field induced on the P U electrode is a line spectrum at harmonics of the revolution frequency.

$$S_{//_0}(\omega,\theta) = 2\pi I \sum_{p=-\infty}^{p=+\infty} \delta(\omega - p\omega_0)\, e^{-jp\theta} \left(\int_{\hat{\tau}=0}^{\hat{\tau}=+\infty} J_0(p\omega_0\hat{\tau})\, q_0(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \right) \qquad (21)$$

where I is the intensity in one bunch.

$$I = \frac{N e \omega_0}{2\pi} \, . \qquad (22)$$

The amplitude of the spectrum at frequency $p\omega_0$ is given by

$$\sigma_0(p) = \int_{\hat{\tau}=0}^{\hat{\tau}=+\infty} J_0(p\omega_0\hat{\tau})\, q_0(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \, . \qquad (23)$$

There is no synchrotron satellite and therefore no evidence at all of an internal synchrotron motion. Turn after turn, each small volume of synchrotron phase space that rotates at fixed distance from the synchronous particle is replaced by an equivalent volume with the same density.

The stationary distribution gives the average density. It can be adapted to simulate different types of bunches.

A few typical examples are given hereunder. In these examples, τ_L represents the full bunch length (or 4 standard deviations in the gaussian case). To make the writing easier, the dimensionless variables

$$z = \frac{\tau}{\frac{\tau_L}{2}} \qquad \hat{z} = \frac{\hat{\tau}}{\frac{\tau_L}{2}} \qquad \text{and} \qquad B = \frac{\omega_0 \tau_L}{2\pi} \ \text{(bunching factor)} \qquad (24)$$

are used.

- Parabolic amplitude density $(0 < \hat{z} < 1)$

$$q_0(\hat{z}) = \frac{2}{\pi\left(\frac{\tau_L}{2}\right)^2}\left(1-\hat{z}^2\right) \; ; \; \lambda(z) = \frac{8}{3\pi\left(\frac{\tau_L}{2}\right)}\left(1-z^2\right)^{3/2} \; ; \; S_{//_0}(\omega,\theta) = 8I\sum_p \delta(\omega-p\omega_0)\, e^{-jp\theta}\frac{J_2(p\pi B)}{(p\pi B)^2} \qquad (25)$$

- Parabolic line density suitable for proton bunches $(0 < \hat{z} < 1)$

$$q_0(\hat{z}) = \frac{3}{2\pi\left(\frac{\tau_L}{2}\right)^2}\left(1-\hat{z}^2\right)^{1/2} \; ; \; \lambda(z) = \frac{3}{4\left(\frac{\tau_L}{2}\right)}\left(1-z^2\right) \; ; \; S_{//_0}(\omega,\theta) = 3I\sum_p \delta(\omega-p\omega_0)\, e^{-jp\theta}\sqrt{\frac{\pi}{2}}\,\frac{J_{3/2}(p\pi B)}{(p\pi B)^{3/2}} \qquad (26)$$

- Gaussian amplitude density suitable for electron bunches $(0 < \hat{z} < \infty)$

$$q_0(\hat{z}) = \frac{2}{\pi\left(\frac{\tau_L}{2}\right)^2}\, e^{-2\hat{z}^2} \; ; \; \lambda(z) = \sqrt{\frac{2}{\pi}}\,\frac{1}{\left(\frac{\tau_L}{2}\right)}\, e^{-2z^2} \; ; \; S_{//_0}(\omega,\theta) = I\sum_p \delta(\omega-p\omega_0)\, e^{-jp\theta}\, e^{-\frac{(p\pi B)^2}{8}} \qquad (27)$$

- "Water bag" bunch $(0 < \hat{z} < 1)$

$$q_0(\hat{z}) = \frac{1}{\pi\left(\frac{\tau_L}{2}\right)^2} \; ; \; \lambda(z) = \frac{2}{\pi\left(\frac{\tau_L}{2}\right)}\left(1-z^2\right)^{1/2} \; ; \; S_{//_0}(\omega,\theta) = 2I\sum_p \delta(\omega-p\omega_0)\, e^{-jp\theta}\,\frac{J_1(p\pi B)}{p\pi B} \qquad (28)$$

The line density $\lambda(\tau)$ is the projection of the distribution $q_0(\hat{\tau})$ on the τ axis.

$$\lambda(\tau) = \int q_0(\hat{\tau})\,\frac{d\tau}{\omega_{s_0}} \qquad \int \lambda(z)\, dz = 1 \qquad (29)$$

The corresponding power spectra $|S_{//_0}(\omega,\theta)|^2$ (Fig. 2a) and single pass signals (Fig. 2b) are drawn for comparison. The spectrum is peaked at zero frequency and extends $\pm 2\pi/\tau_L$ rad/sec.

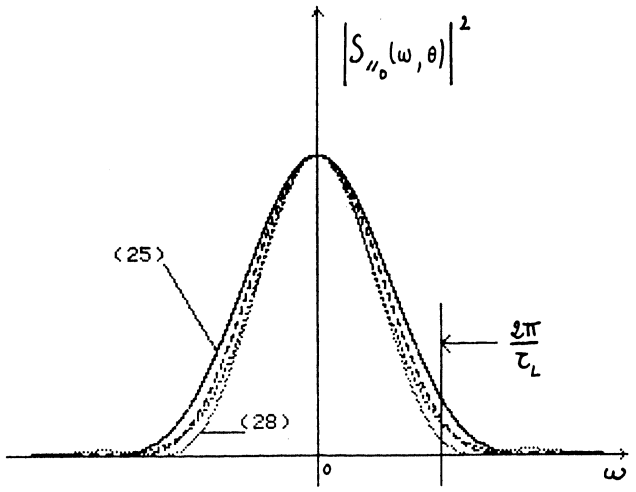


Fig. 2(a) Power spectra for various particle distributions

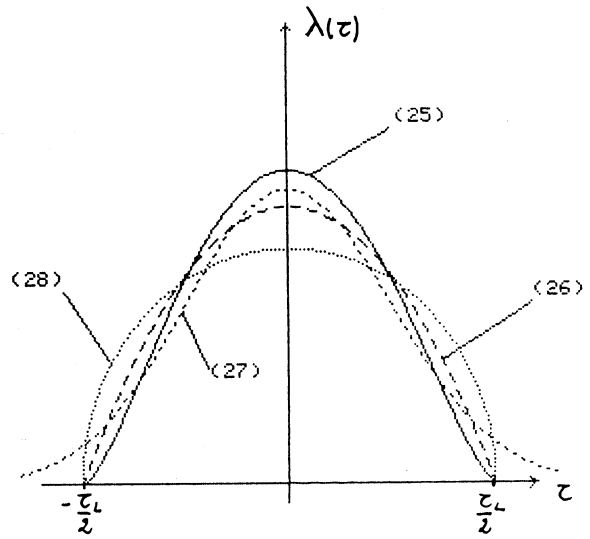


Fig. 2(b) Line densities for the various particle distributions of Fig. 2(a)

As will be pointed out later on, in Vlasov's equation, for the longitudinal case, the quantity of interest is $dg_0/d\hat{\tau}$ or $d\lambda/d\tau$. In this respect, the choice of distributions like the "water bag" distribution must be avoided. As a matter of fact, it shows a sharp discontinuity at its edge that leads to unrealistic results since it attaches too much importance to large amplitude particles.

In the rest of this lecture on longitudinal instabilities, we will mainly use the distribution with parabolic amplitude density. On one hand, it is a good compromise between proton and electron bunches. On the other hand, according to my own experience, the three realistic distributions listed above lead to results that do not deviate by more than 20%.

The electromagnetic field induced by the stationary distribution is at harmonics of the revolution frequency. As a consequence, it acts on the beam as the external RF system does. First it is responsible for synchronous phase shift. Then, it can depress or increase the focusing. It can also introduce non linear terms because of the rich harmonic content of the spectrum.

Nevertheless, the internal synchrotron motion of individual particles is hidden and one cannot expect any excitation of coherent motion without any force at harmonics of the synchrotron frequency. So, we need a new ingredient in order to get a scenario of initial conditions that can lead to instability.

2.3.2 Perturbation

This ingredient consists in a distribution $\Delta\Psi(\psi_0, \hat{\tau}, t)$ which has the property of introducing some electromagnetic field at harmonics of the synchrotron frequency. It is a density perturbation which represents the difference between the actual beam and the stationary distribution. As a matter of fact, in terms of physics, a stationary beam does not exist. On one hand, there is always some density modulation that remains from previous beam manipulations such as injection, bunching, etc... On the other hand, the bunch is composed of individual particles, each of them gives a rich spectrum including synchrotron sidebands. On an average, we get a stationary bunch. Nevertheless, there is always some remaining noise at synchrotron satellite frequencies.

The form that can be given to $\Delta\Psi$ is suggested by the single particle signal (13) and (16). In phase space, on a given orbit $\hat{\tau}$, one can choose the initial ψ_0 dependence of the charge density in order to enhance the signal amplitude at a given harmonic of the synchrotron frequency.

For instance, with

$$\Delta\Psi(\psi_0, \hat{\tau}, t) = g_m(\hat{\tau})\, e^{-jm\psi_0}\, e^{j\Delta_{cm}t} \qquad m \neq 0 \qquad (30)$$

(the assumed time dependence of the perturbation will be justified at the end of this section) one gets a perturbation signal

$$\Delta S_{/\!/_m}(\omega,\theta) = 2\pi I \sum_p \delta(\omega-(p\omega_o+m\omega_s+\Delta\omega_{cm}))\, j^{-m} e^{-jp\theta} \int_{\hat{\tau}=0}^{\hat{\tau}=+\infty} J_m(p\omega_o\hat{\tau})\, q_m(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \,. \qquad (31)$$

Therefore, owing to the $e^{-jm\psi_o}$ charge density, the spectral amplitude is maximum for satellite number m and null for all other satellites.

We have defined a perturbation which is coherent with respect to satellite number m. The amplitude of the perturbation spectrum at frequency $\omega = p\omega_o + m\omega_s + \Delta\omega_{cm}$ is given by

$$\sigma_m(p) = j^{-m} \int_{\hat{\tau}=0}^{\hat{\tau}=+\infty} J_m(p\omega_o\hat{\tau})\; q_m(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \qquad (32)$$

In order to give more physical content to the expression of the perturbation (30), let us imagine an injection scheme in which a bunch is transferred from a booster to a main ring.

After transfer, when ideal conditions are fulfilled, the bunch is perfectly matched and individual particles continue their motion on the "same" phase space orbit.

Now, let us assume a RF phase error or an energy error at transfer. Then, in phase space, the beam center of mass rotates around the main ring synchronous orbit (Fig. 3a). Triggerred by the RF clock, the PU signal shows that the bunch is oscillating forwards and backwards (Fig. 3b). When compared to the ideal transfer (stationary bunch) a density perturbation of $\cos\psi_o$ ($e^{-jm\psi_o}$ with $m=1$) type has been added. It corresponds to an excess of charges at one bunch edge, exactly balanced by a lack of charges at the opposite edge.

Using instability terminology, a coherent dipolar perturbation has been induced.

If one neglects the electromagnetic self-field influence, then, this figure rotates at frequency ω_{so} in phase space. The power spectrum of the signal has two components, the spectrum of the stationary distribution at harmonics of the revolution frequency and the spectrum induced by the perturbation at $p\omega_o + \omega_{so}$. This last component is peaked at higher frequency $\omega \sim 2\pi/\tau_L$ since it represents details with short wavelength occurring during the bunch passing τ_L .

If one takes the electromagnetic self-field into account, then, two major effects are expected.

- Firstly, we will get a shift of the synchrotron frequency ω_{so} . The field induced by the stationary distribution at $p\omega_o$ acts on the beam like an RF system does. It modifies the external focusing. This shift

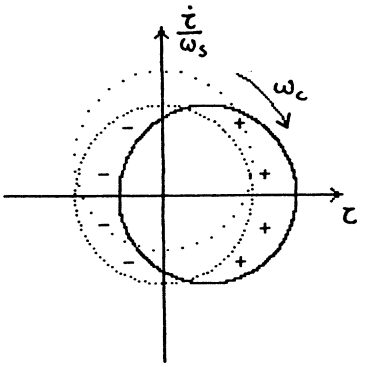
$$\Delta\omega_i = \omega_s - \omega_{so} \qquad (33)$$

will be called incoherent synchrotron frequency shift.

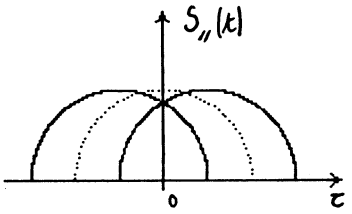
- Secondly, the perturbation will not rotate at the incoherent synchrotron frequency ω_s but at the coherent frequency ω_c , $\Delta\omega_{c1}$ apart.

$$\Delta\omega_{c1} = \omega_c - \omega_s \,. \qquad (34)$$

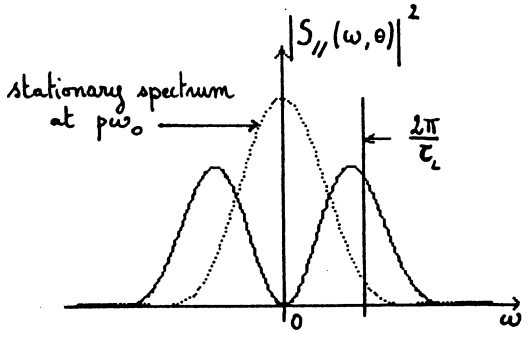
This is the reason why the $e^{j\Delta\omega_{c1}t}$ term has been introduced in equation (30) already. At the same time ω_{so} has been replaced by ω_s .



a)

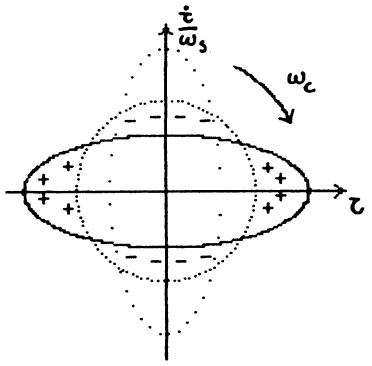


b)

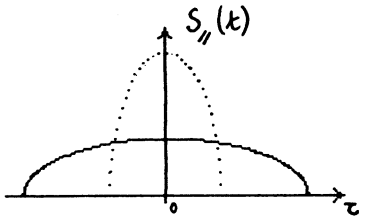


c)

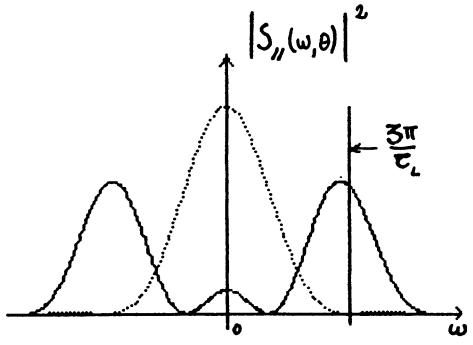
Fig. 3 Dipole mode of longitudinal bunch oscillations



a)



b)



c)

Fig. 4 Quadrupole mode of longitudinal bunch oscillations

a) Phase-plane diagram
b) Line density as seen on a pick-up
c) Frequency spectrum for the dipole/quadrupole mode

$\Delta\omega_{c1}$ is called coherent frequency shift.

The line spectrum of the perturbation which was at $p\omega_0 + \omega_{s0}$ for zero current, is now at $p\omega_0 + \omega_c$.

The goal we aim at consists in finding out the expression of both quantities $\Delta\omega_i$ and $\Delta\omega_{cm}$. We will assume that $\Delta\omega_{cm}$ is a complex number. Its real part will give us the real coherent frequency shift. The sign of its imaginary part will tell us whether the initial perturbation will grow ($I_m(\Delta\omega_{cm}) < 0$ instability) or will be damped ($I_m(\Delta\omega_{cm}) > 0$ stability).

As a second example, let us describe a coherent quadrupolar perturbation. One can assume an error of RF voltage in the main ring. In this case, because of the focusing mismatching, the transferred beam has an elliptical shape in main ring phase (Fig. 4a). This corresponds to a $\cos 2\psi_0$ ($e^{-jm\psi_0}$ with $m = 2$) density perturbation. The ellipse has a twofold symmetry in phase space. It repeats at the coherent frequency ω_c , that is to say at twice the incoherent synchrotron frequency ω_s plus the coherent frequency shift $\Delta\omega_{c2}$.

$$\Delta\omega_{c2} = \omega_c - 2\omega_s \quad . \tag{35}$$

Since we have to describe more and more details during the bunch passing τ_L , the perturbation signal is peaked at higher frequency again $\omega \sim 3\pi/\tau_L$ (Fig. 4c). The electromagnetic field induced by the perturbation is at

$$p\omega_0 + \omega_c \qquad -\infty < p < +\infty \tag{36}$$

$\omega_c = 2\omega_{s0}$ for zero current and $\omega_c = 2\omega_s + \Delta\omega_{c2}$ for $I \neq 0$.

There is an infinite number of possible coherent perturbations

m=1	dipole
m=2	quadrupole
m=3	sextupole, etc...

For zero intensity,

$$\omega_{cm} = m\,\omega_{s0} \tag{37}$$

the frequency separation between two adjacent coherent perturbations is ω_{s0} .

For $I \neq 0$,

$$\omega_{cm} = m\omega_s + \Delta\omega_{cm} \tag{38}$$

one can distinguish two regimes.

In the low intensity regime, the basic frequency separation is not ω_{s0} but ω_s the incoherent synchrotron frequency. It has been slightly changed by the effect of the electromagnetic field at $p\omega_0$ arising from the stationary distribution $\Delta\omega_i = \omega_s - \omega_{s0} \ll \omega_{s0}$. In parallel, the coherent frequency shift remains small $\Delta\omega_{cm} \ll \omega_{s0}$. Therefore, the coherent frequencies of two adjacent perturbations are still well separated and one can study each of the perturbations separately while restricting to a single value of m.

$$\Delta\Psi = g_m(\hat{\tau})\, e^{-jm\psi_0}\, e^{j\Delta\omega_{cm}t} \tag{30}$$

In the high intensity regime, the coherent and incoherent shifts get large and the frequencies associated with two adjacent perturbations m and m+1 can get very close. As a consequence, one has to sum up several values of m in the $\Delta\Psi$ expression.

$$\Delta\Psi = \sum_m q_m(\hat{t})\, e^{-j m \psi_0}\, e^{j \Delta\omega_{cm} t} \,. \tag{39}$$

2.4 Longitudinal coupling impedance

All the properties of the electromagnetic response of a given machine to the beam passings are gathered into a key parameter, the so called longitudinal coupling impedance $Z_{//}(\omega)$. It allows one to predict the self field acting along the beam axis in terms of signal.

$$2\pi R \left[\vec{E} + \vec{\beta c} \wedge \vec{B} \right]_{//} (t,\theta) = - \int_{\omega=-\infty}^{\omega=+\infty} Z_{//}(\omega)\, S_{//}(\omega,\theta)\, e^{j\omega t}\, d\omega \tag{40}$$

$Z_{//}$ is expressed in ohm.

If $Z_{//}(\omega)$ were a constant independent of frequency, then, the self-field would be proportional to the signal. This explains the close connection that has been made between signal and self field up to now. We will avoid undertaking a detailed justification of equation (40) since it has been done in the basic lecture given last year already. We will limit ourselves to a qualitative discussion of the solution of Maxwell's equation applied to the crude model of a round beam of radius a travelling on axis in a circular pipe of radius b (Fig. 5).

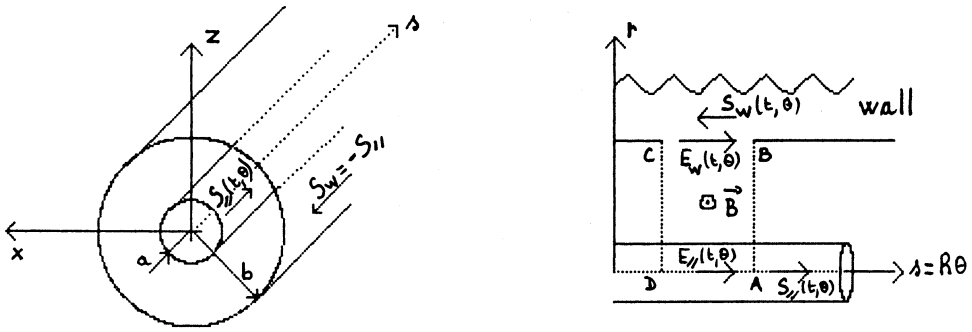


Fig. 5 Boundary conditions for an on-axis beam in a perfectly-conducting circular pipe

At time t and angular position θ in the machine, the local beam current is given by the signal $S_{//}(t,\theta)$.

Let us assume that the beam is completely screened by the pipe wall. Then, there is no electromagnetic field outside the chamber. When applied to a 2π path enclosing the chamber cross section, Ampere's theorem leads to

$$\oint \vec{H} . \vec{dl} = I_{total}(t,\theta) \tag{41}$$

where $I_{total}(t,\theta)$ is the total current through the path plane. Since we already have the beam current $S_{//}(t,\theta)$ flowing downstream, it has to be exactly balanced by a return current or image current flowing upstream in the wall thickness.

$$S_w(t,\theta) = - S_{//}(t,\theta) \tag{42}$$

For perfectly conducting walls, the electric field in the wall E_w that has to be associated with the image current S_w is null. Well below the pipe cut-off frequency, that stands in the GHz region

$$\omega_{cut\text{-}off} = \frac{c}{b} , \tag{43}$$

one can apply Faraday's law in the longitudinal cross section of the pipe, along a contour like ABCD (Fig. 5).

$$\oint \vec{E} \cdot \vec{dl} = -\frac{\partial}{\partial t} \int \vec{B} \cdot \vec{d\sigma} \quad . \tag{44}$$

The result is the standard longitudinal space charge field set up by the direct current $S_{//}(t,\theta)$.

In terms of wall current, this space charge field can be written

$$2\pi R\, E_{//s.c.}(t,\theta) = \frac{-Z_0 g}{2\beta\gamma^2\omega_0} \frac{\partial}{\partial t} S_w(t,\theta) \tag{45}$$

$Z_0 = 377$ ohm is the space free impedance
$g = 1 + 2\, Ln\left(\frac{b}{a}\right)$

Therefore, the space charge impedance is a negative inductance,

$$L_{s.c.} = -\frac{Z_0 g}{2\beta\gamma^2\omega_0} \quad . \tag{46}$$

Large for low-β particles, it vanishes at ultrarelativistic energies. In the form of equation (40), in terms of signal, it can be written

$$2\pi R\, E_{//s.c.}(t,\theta) = -\int_{\omega=-\infty}^{\omega=+\infty} \frac{Z_0 g \omega}{j 2\beta\gamma^2\omega_0} S_{//}(\omega,\theta)\, e^{j\omega t}\, d\omega \quad . \tag{47}$$

For resistive walls, the electric field in the wall E_w is different from zero. Due to the skin effect impedance an additional electric field appears on beam axis.

$$2\pi R\, E_{//R.W.}(t,\theta) = -\int_{\omega=-\infty}^{\omega=+\infty} Z_{//R.W.}(\omega)\, S_{//}(\omega,\theta)\, e^{j\omega t}\, d\omega \quad . \tag{48}$$

The resistive wall impedance is given by

$$Z_{//R.W.}(\omega) = (1+j)\frac{Z_0\beta}{2b}\delta_0^* \left(\frac{\omega}{\omega_0}\right)^{1/2} \qquad \delta_0^* = \sqrt{\frac{2\rho}{\mu_0\omega_0}} \tag{49}$$

where δ^*_o is the skin depth at revolution frequency.
ρ = resistivity $\sim 1.\,10^{-6}\,\Omega m$ for stainless steel, $\mu_o = 4\pi.10^{-7}$.

In these two examples of fields the environment enters via its geometry (boundary conditions imposed at distance b) and via its electromagnetic properties (resistivity ρ in δ^*_o for instance). Obviously different machines have different wall geometries and different wall electromagnetic properties. Accordingly, the impedance varies from one machine to another.

Since the impedance is a key parameter that rules instability threshold current, many attempts have been made to measure, to understand and to minimize machine impedances.

The quantity of interest is not the longitudinal impedance itself but the impedance divided by the frequency $Z_{//}(\omega)/\omega$ or even better, $Z_{//}(p)/p$ the impedance divided by the harmonic number p of the revolution frequency $\omega = p\omega_o$. In a diagram with $Z_{//}(p)/p$ along the vertical axis and ω running from $-\infty$ to $+\infty$ along the horizontal axis, a pure inductance $jL\omega$ is associated with a constant $I_m(Z_{//}(p)/p) = L\omega_o$, a pure resistance R is represented by an hyperbola $R_e(Z_{//}(p)/p) = R\omega_o/\omega$.

The main components of the impedance that can be found in a standard circular machine are listed hereunder and sketched in Fig. 6.

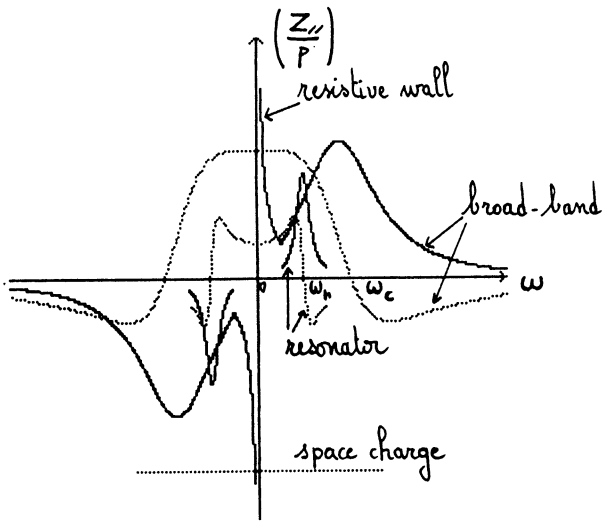


Fig. 6 Qualitative description of longitudinal impedance for various components (full line = real part, dotted line = imaginary part)

a) Resistive wall component

It is peaked at low frequency.

$$\frac{Z_{//R.W.}}{p} = (1+j)\frac{Z_o\beta}{2b}\delta^*_o\frac{1}{\sqrt{p}} \qquad \text{(thick wall assumption)} \qquad (50)$$

We will see that longitudinal modes have no spectral amplitude at low frequencies. As a consequence, it acts very weakly on longitudinal motion and can be disregarded.

b) Narrow band resonators

One of them, at $\omega_r = h\omega_o$, is the necessary RF system that keeps the beam bunched. One can also find parasitic high Q resonators due to higher order modes in RF cavities for instance.

c) Broad band component (BB)

The averaged effect of numerous changes in vacuum chamber cross section (step changes, bellows, tanks, electrodes, etc) can be approximated by a low $Q \sim 1$ resonator with the resonance ω_r at pipe cut-off frequency (43).

The origin of the mechanism is the electromagnetic field created by the beam itself

$$Z_{/\!/\,B.B}(\omega) = \frac{R_s}{1 + j Q\left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega}\right)} \tag{51}$$

R_s is the shunt impedance in ohm
$Q \sim 1$ the quality factor.

It acts like a pure inductance at low frequencies,

$$\frac{Z_{/\!/\,B.B.}^{(p)}}{p} = j R_s \frac{\omega_0}{\omega_r} \tag{52}$$

like a pure resistance at resonance,

$$\frac{Z_{/\!/\,B.B.}^{(p)}}{p} = R_s \frac{\omega_0}{\omega_r} \quad . \tag{53}$$

It acts like a capacitance at high frequencies,

$$\frac{Z_{/\!/\,B.B.}^{(p)}}{p} = -j R_s \frac{\omega_r \omega_0}{\omega^2} \quad . \tag{54}$$

This broad band model is in quite good agreement with experimental results. For those machines carefully designed to lower the BB impedance, the peak value of $\left| Z_{/\!/}(p)/p \right|$ can be as low as a fraction of an ohm. It can reach 50 Ω when no care at all is taken.

d) Space charge component

$$\frac{Z_{/\!/\,s.c.}^{(p)}}{p} = -j \frac{Z_0 g}{2 \beta \gamma^2} \quad . \tag{55}$$

It can be very large for low - β particles.

~ 1.5 kΩ for $\beta = .3$ (50 MeV protons)
~ 7 Ω for (7 GeV protons).

The actual impedance seen by the beam is the sum of the components listed above.

2.5 Effect of the stationary distribution

In the first part of this section, we are still assuming a single bunch in the machine.

The effect of the electromagnetic field induced by the stationary distribution η_0 is included in the single particle differential equation of motion

$$\ddot{\tau} + \omega_{s0}^2 \tau = F_0 = \frac{\eta e}{P_{/\!/}} \left[\vec{E} + \vec{\beta} c \wedge \vec{B} \right]_{/\!/} \left(t, \theta = \omega_0 (t - \tau) \right) \tag{56}$$

$\theta = \omega_0 (t - \tau)$ when following the particle.
The right hand side of equation (56) can be developped by using equations (40), (21), (23), (6) successively.

$$F_0 = \frac{2\pi I \omega_{s0}^2}{\omega_0 V_{R.F.} h \cos \varphi_s} \sum_p Z_{/\!/}(p)\, \sigma_0(p)\, e^{j p \omega_0 \tau} \tag{57}$$

In the following, we will be concerned with small amplitudes. Therefore we can expand the exponential in series:

$$\ddot{\tau} + \omega_{s_o}^2 \tau = \frac{2\pi I\, \omega_{s_o}^2}{\omega_o V_{RF} h \cos\varphi_s} \sum_p Z_{/\!/}(p)\, \sigma_o(p) \left\{ 1 + j p \omega_o \tau - \frac{p^2 \omega_o^2 \tau^2}{2} + \cdots \right\} . \qquad (58)$$

Let us analyze equation (58). In the absence of self-field, a linear motion has been assumed. The external RF voltage varies linearly with time during the bunch passing and all particles oscillate at the same frequency ω_{s_o} whatever their amplitude $\hat{\tau}$ is.

The stationary self field introduces :

- a zero order term that changes the synchronous phase,
- a first order term responsible of the incoherent frequency shift,
- non-linear terms which make the frequency shift amplitude dependent. Some synchrotron frequency spread appears.

2.5.1 Synchronous phase shift "higher-mode loss"

The constant term in equation (58) moves the stable fixed point

$$\Delta\varphi_s = h \omega_o \Delta\tau = \frac{2\pi I}{V_{RF} \cos\varphi_s} \sum_p R_e\left(Z_{/\!/}(p)\right) \sigma_o(p) . \qquad (59)$$

Because of the power dissipated in the wall by the return current the beam loses energy. The synchronous phase is displaced so that this loss can be restored by the RF system.

Equation (59) can be used to probe the resistive part of the impedance.

We will see later on that this effect tends to lower the synchrotron frequency and contributes to bunch lengthening.

2.5.2 Incoherent frequency shift - Potential well distortion

When gathering first order terms of equation (58), we get

$$\omega_s^2 = \omega_{s_o}^2 \left(1 - \frac{2\pi I}{V_{RF} h \cos\varphi_s} \sum_p j Z_{/\!/}(p)\, p\, \sigma_o(p) \right) . \qquad (60)$$

$p\,\sigma_o(p)$ is an odd function with respect to p, therefore resistance does not modify the synchrotron frequency.

When the bunch is long enough, $2\pi/\tau_L \ll \omega_{cut\text{-}off}$, the major part of the stationary spectrum stands in the low frequency region where space-charge impedance and broad-band impedance are constant. Under these assumptions, equation (60) can be rewritten

$$\Delta = \frac{\omega_s^2 - \omega_{s_o}^2}{\omega_{s_o}^2} = - \frac{2\pi I}{V_{RF} h \cos\varphi_s} \left[j \frac{Z_{/\!/}(p)}{p} \right]_{\substack{s.c. \\ B.B.}} \sum_p p^2 \sigma_o(p) . \qquad (61)$$

For the bunch with parabolic amplitude (25) with the help of

$$\sum_{p=-\infty}^{p=+\infty} J_2(p x) = \frac{2}{x}$$

we get

$$\Delta = \frac{\omega_s^2 - \omega_{s_0}^2}{\omega_{s_0}^2} = - \frac{16\, I}{\pi^3 B^3 V_{RF} h \cos\varphi_s} \int \frac{Z_{//}(p)}{p} = \frac{V_T}{V_{RF}} - 1 \quad . \tag{62}$$

Below the transition energy $\cos\varphi_s > 0$ the synchrotron frequency is reduced by space charge $\int Z_{//}(p)/p > 0$ and increased by the B B inductance $\int Z_{//}(p)/p < 0$. The reverse applies above transition. The same qualitative remarks are valid for any stationary distribution.

The change in RF slope corresponds to an effective voltage V_T given by

$$\frac{V_T}{V_{RF}} = \left(\frac{\omega_s}{\omega_{s_0}}\right)^2 \quad . \tag{63}$$

If we consider a parabolic line density bunch (26) interacting with a constant $I_m(Z_{//}(p)/p)$, then, without any approximation, equation (56) can be reduced to

$$\ddot{\tau} + \omega_{s_0}^2 \tau = \frac{3\, I\, \omega_{s_0}^2}{\pi^2 V_{RF} h \cos\varphi_s B^3} \int Z_{//}(p)\, \tau \quad . \tag{64}$$

In this particular case, the focusing force is purely linear. The corresponding incoherent frequency shift is given by

$$\Delta = \frac{\omega_s^2 - \omega_{s_0}^2}{\omega_{s_0}^2} = \frac{3\, I}{\pi^2 V_{RF} h \cos\varphi_s B^3} \int \frac{Z_{//}(p)}{p} \tag{65}$$

which is about two times less than (62).

In Fig. 7 the potential-well distortion corresponding to a long bunch interacting with a B.B. above transition is sketched.

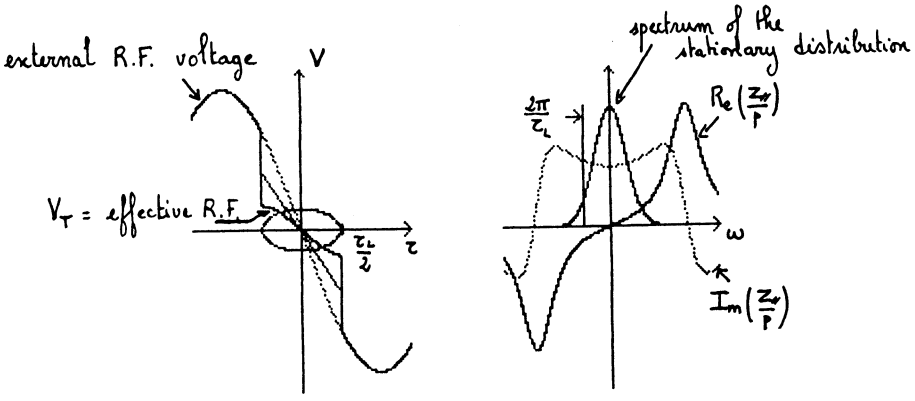


Fig. 7 Distortion of the RF wave form (left) due to the interaction with an inductive broad-band resonance (right)

2.5.3 Bunch lengthening

A direct consequence of the shifts of the synchronous phase and of the incoherent frequency is that the bunch length and momentum spread depend on intensity.

For electrons, the equilibrium momentum spread is imposed by radiation. Then,

$$\frac{B}{B_0} = \frac{\omega_{s0}}{\omega_s}\left(\frac{\cos\varphi_s}{\cos(\varphi_s+\Delta\varphi_s)}\right)^{1/2} \tag{66}$$

in which, B_0 is the bunching factor for zero intensity.

It is a transcendental equation since B appears in ω_s and in $\Delta\varphi_s$.

The synchronous phase shift lowers the incoherent synchrotron frequency. Above transition, with long bunches $2\pi/\tau_L \ll \omega_{cut\text{-}off}$ the inductance of the broad-band impedance has the same effect (with very short bunches, the stationary spectrum reaches the B B capacitance and this last contribution can be reversed).

In most cases the synchronous phase shift can be neglected. This allows us to rewrite equation (66) in the following form

$$\frac{B}{B_0} = \left(\frac{B}{B_0}\right)^3 + \Delta_0 \tag{67}$$

in which $\Delta_0 = \left((\omega_s^2-\omega_{s0}^2)/\omega_{s0}^2\right)_0$ is the normalized deviation of frequencies squared one would obtain with $B = B_0$ (nominal current associated with zero current bunch length).

For protons, we assume that the emittance is invariant. Therefore,

$$\left(\frac{B}{B_0}\right)^2 = \frac{\omega_{s0}}{\omega_s}\left(\frac{\cos\varphi_s}{\cos(\varphi_s+\Delta\varphi_s)}\right)^{1/2} \tag{68}$$

holds.

When comparing the electron and the proton case, $(B/B_0)^2$ replaces B/B_0. Therefore the bunch is less affected by the self field of the stationary distribution.

If one neglects the synchronous phase shift again, equation (68) can be written

$$\left(\frac{B}{B_0}\right)^{-1} = \left(\frac{B}{B_0}\right)^3 + \Delta_0 \quad . \tag{69}$$

For Δ_0 the same definition as in equation (67) has to be used.

Since the emittance is constant for protons, the momentum spread has to be readjusted.

In Fig. 8, the solutions of equations (67) and (69) are drawn.

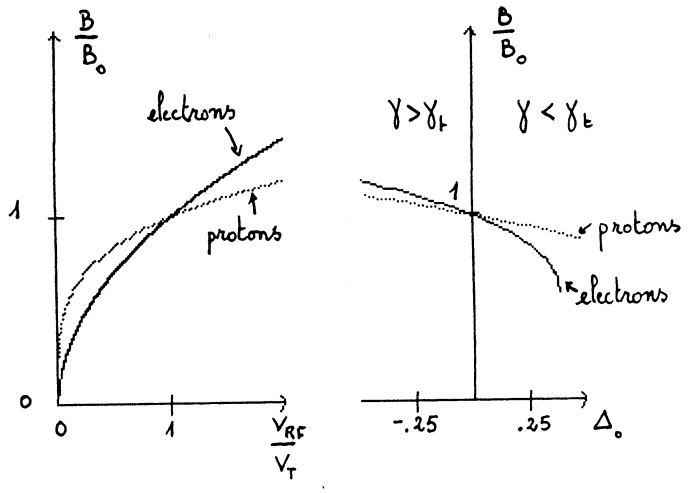


Fig. 8 Bunch lengthening as a function of current

2.5.4 Multibunch case

Now, we assume that the machine is operated in a multibunch mode. There are M identical bunches equally spaced corresponding to a total current MI. The spectrum is a line spectrum again, with the same form factor as the single bunch one. The spacing between lines is the bunch repetition frequency $M\omega_0$, that is to say M times larger than it was. On the other hand the spectrum amplitude is multiplied by M.

Equations (59) and (60) become

$$\Delta\varphi_s = \frac{2\pi I}{V_{RF}\, h \cos\varphi_s}\, M \sum_p R_e\left(Z_{/\!/}(n_p)\right)\, \sigma_0(n_p) \qquad (59)^*$$

$$\omega_s^2 = \omega_{s_0}^2 \left(1 - \frac{2\pi I}{V_{RF}\, h \cos\varphi_s}\, M \sum_p j\, Z_{/\!/}(n_p)\, M_p\, \sigma_0(n_p)\right). \qquad (60)^*$$

If the attenuation of the wake field between two successive bunches is weak, in other words, if the environment can memorize the passing of a bunch for a time larger than the bunch repetition period (abrupt change of the impedance within a frequency interval of the order of the repetition frequency), then, the results for the multibunch case (59)* and (60)* can be very different from the single bunch case (59) and (60).

On the other hand, if the interaction is a local one (delta function wake) such as space charge or inductive wall (constant $Z_{/\!/}(p)/p$), bunches ignore each other, then, multibunch and single bunch cases lead to the same result. This is exactly the case for equations (61), (62) and (65).

This ends up our study of the electromagnetic field induced by the stationary distribution. For a given current, we have defined a new set of matched conditions that takes into account the external RF field and the self field of the stationary distribution. As long as parameters remain realistic, the stationary bunch is stable. In the rest of this lecture we will assume that all the above effects (synchronous phase shift, incoherent frequency shift, modified bunch length and momentum spread) have been taken into account.

Around the new fixed point, the linear equation of motion of the single particle is

$$\ddot{\tau} + \omega_s^2 \tau = 0 \tag{70}$$

and

$$\omega_s^2 = \frac{-\eta e \omega_o V_T h \cos(\varphi_s + \Delta\varphi_s)}{2\pi R p_{//}} . \tag{71}$$

2.6 Effect of the perturbation

Now, as suggested in section 2.3.2 we want to add a coherent perturbation that rotates at frequency ω_c in phase space and therefore induces some field at $p\omega_o + \omega_c$.

2.6.1 Vlasov's equation

First we have to introduce the necessary equations to allow the evolution of the distribution to be followed.

The basic equation that rules the time evolution of the local density distribution is the collision-free Boltzmann's equation.

$$\frac{\partial \Psi}{\partial t} + \operatorname{div}\left(\Psi \vec{v}\right) = 0 \tag{72}$$

where

$$\vec{v} \begin{pmatrix} \dot{\tau} \\ \ddot{\tau} \end{pmatrix} .$$

It can be developed as follows

$$\frac{d\Psi}{dt} = -\Psi \operatorname{div} \vec{v} . \tag{73}$$

If one uses a set of canonical conjugate variables like $(\tau, \dot{\tau})$, then $\operatorname{div} \vec{v} = 0$ and an equivalent form of equation (70) is

$$\frac{d\Psi}{dt} = 0 = \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi}{\partial \tau}\dot{\tau} + \frac{\partial \Psi}{\partial \dot{\tau}}\ddot{\tau} . \tag{74}$$

In this form, it is called Vlasov's equation. It expresses that phase-space density does not vary with time when following the motion in canonical variables.

2.6.2 Equation of coherent motion

We consider a distribution that sums up a stationary distribution and a coherent perturbation $\Delta\Psi$ as described in equation (39).

$$\Psi\left(\psi_o, \hat{\tau}, t\right) = q_0(\hat{\tau}) + \sum_m q_m(\hat{\tau})\, e^{-jm\psi_o}\, e^{j(\omega_c - m\omega_s)t} . \tag{75}$$

We aim at finding out ω_c to study the stability via its imaginary part. We rewrite Vlasov's equation (74) with ψ_o and $\hat{\tau}$ as new variables.

$$0 = \frac{\partial \Delta\Psi}{\partial t} + \left(\frac{\partial q_o}{\partial \hat{\tau}} + \frac{\partial \Delta\Psi}{\partial \hat{\tau}}\right)\frac{d\hat{\tau}}{dt} + \frac{\partial \Delta\Psi}{\partial \psi_o}\frac{d\psi_o}{dt} \tag{76}$$

$\partial\Delta\Psi/\partial\psi_0 \cdot d\psi_0/dt$ and $\partial\Delta\Psi/\partial\hat{\tau} \cdot d\hat{\tau}/dt$ are dropped since they represent second order terms with respect to the perturbation.

$$j e^{j\omega_c t} \sum_m (\omega_c - m\omega_s) q_m(\hat{\tau}) e^{-jm\psi} = -\frac{\partial q_0}{\partial \hat{\tau}} \frac{d\hat{\tau}}{dt} \qquad (77)$$

$\psi = \omega_s t + \psi_0$ will be used to make the writing easier.

In the right-hand side of the linearized Vlasov's equation (77), we have to express the product

$$\frac{\partial q_0}{\partial \hat{\tau}} \frac{d\hat{\tau}}{dt} \quad .$$

The stationary distribution comes in via its derivative with respect to $\hat{\tau}$. As pointed out before in section 2.3.1, distributions like the water bag distribution (28) with infinite $\partial q_0/\partial\hat{\tau}$ must be avoided.

The expression of $d\hat{\tau}/dt$ can be drawn from a single particle equation of motion.

$$\ddot{\tau} + \omega_s^2 \tau = F_c = \frac{\eta e}{p_{/\!/}} \left[\vec{E} + \vec{\beta c} \wedge \vec{B} \right]_{/\!/} \left(t, \Theta = \omega_0 (t - \tau) \right) \qquad (78)$$

$$\frac{d\hat{\tau}}{dt} = \frac{d}{dt} \left(\tau^2 + \frac{\dot{\tau}^2}{\omega_s^2} \right)^{1/2} = -\frac{F_c}{\omega_s} \sin\psi \qquad (79)$$

F_c is the coherent electromagnetic "force". Let us recall that the stationary distribution effect has been already taken into account (70) and (71).

By means of equations (40), (31), (32) and (71), F_c can be written :

$$F_c = \frac{2\pi I \omega_s^2 e^{j\omega_c t}}{\omega_0 V_T h \cos(\varphi_s + \Delta\varphi_s)} \sum_p Z_{/\!/}(p) e^{jp\omega_0 \tau} \sum_k \sigma_k(p) \quad . \qquad (80)$$

It is interesting to notice that F_c does not depend on $\dot{\tau}$, therefore τ and $\dot{\tau}$ are canonical congugate variables.

Now, we expand the product $\sin\psi \, e^{jp\omega_0 \tau}$ in series

$$\sin\psi \, e^{jp\omega_0 \tau} = \sum_r j^r e^{-jr\psi} \frac{r}{p\omega_0 \hat{\tau}} J_r(p\omega_0 \hat{\tau}) \quad . \qquad (81)$$

Equation (12) and

$$J_{r+1}(x) + J_{r-1}(x) = \frac{2r}{x} J_r(x) \qquad (82)$$

have been used.

Finally we get

$$\hat{\tau}\,\frac{\partial g_0}{\partial \hat{\tau}}\,\frac{d\hat{\tau}}{dt} = -\frac{2\pi I\,\omega_s\,e^{j\omega_c t}}{\omega_0^2 V_T\,h\cos(\varphi_s+\Delta\varphi_s)}\,\frac{\partial g_0}{\partial \hat{\tau}}\sum_{p,r}\frac{Z_{//}(p)}{p}\,j^r e^{-jr\psi}\,J_r(p\omega_0\hat{\tau})\sum_k \sigma_k(p)\,. \quad (83)$$

We gather both sides of equation (77) and notice that it splits into an infinite number of equations, one for each m. As a matter of fact, on an average, the component g_m of the perturbation is driven by the term $r=m$ essentially.

The final form of the equation of coherent motion of a single bunch is

$$j(\omega_c - m\omega_s)\,j^{-m}\hat{\tau}\,g_m = \frac{2\pi I\,m\,\omega_s}{\omega_0^2 V_T\,h\cos(\varphi_s+\Delta\varphi_s)}\,\frac{\partial g_0}{\partial \hat{\tau}}\sum_p \frac{Z_{//}(p)}{p}\,J_m(p\omega_0\hat{\tau})\,\sigma(p) \quad (84)$$

$\sigma(p)$ is the resultant spectrum amplitude at frequency $p\omega_0+\omega_c$.

$$\sigma(p) = \sum_m \sigma_m(p)\,. \quad (85)$$

2.6.3 Coherent modes of oscillation

Because there are two degrees of freedom $(\psi_0, \hat{\tau})$, the general solution of equation (84) is a twofold infinity of coherent modes of oscillation. We will use the subscripts m and q $(-\infty < m,q < \infty)$ to label these modes.

Each of them is characterized by

- a coherent frequency ω_{cmq}. Its imaginary part will tell us whether this mode is stable or unstable

- a particular perturbation $\Delta\Psi$. That is to say a twofold infinity of functions $g_{mq}(\hat{\tau})$ the sum of which gives the detailed density pattern that rotates in phase space at frequency ω_c.

- a spectrum $\sigma(p)$ which is obviously peaked in the frequency region where this mode is driven by the impedance. At low intensity the self force responsible for the frequency shift is small when compared to the external force responsible for the synchrotron frequency. The coherent frequencies of the infinity of modes with the same m value cluster near the unperturbed frequency $m\omega_s$. Therefore it is logical to study each family m separately and to ignore the coupling between different m.

When equation (84) is restricted to a single value of m,

$$j(\omega_{cmq} - m\omega_s)\,j^{-m}\hat{\tau}\,g_{mq} = \frac{2\pi I\,m\,\omega_s}{\omega_0^2 V_T\,h\cos(\varphi_s+\Delta\varphi_s)}\,\frac{\partial g_0}{\partial \hat{\tau}}\sum_p \frac{Z_{//}(p)}{p}\,J_m(p\omega_0\hat{\tau})\,\sigma_{mq}(p) \quad (86)$$

there are several methods to get the solution.

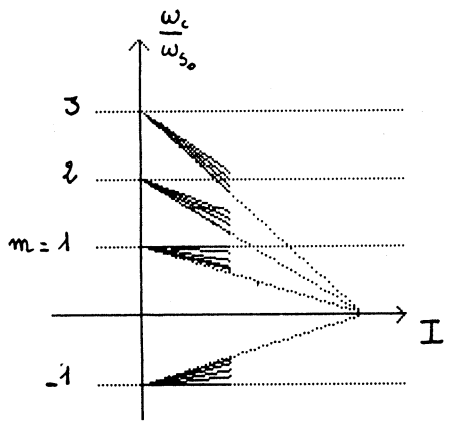


Fig. 9 Schematic representation of coherent-mode frequencies versus current

2.6.4 Sacherer's approach

The functions $g_{mq}(\hat{\tau})$ are solutions of Sacherer's integral equation.

$$\left(\omega_{c_{mq}} - m\omega_s\right) g_{mq}(\hat{\tau}) + \frac{1}{\hat{\tau}} \frac{\partial g_0}{\partial \hat{\tau}} \int_0^\infty G_m(\hat{\tau}, \hat{\tau}')\, g_{mq}(\hat{\tau}')\, \hat{\tau}'\, d\hat{\tau}' = 0 \quad . \tag{87}$$

The definition of σ_{mq} (32) has been used.
$G_m(\hat{\tau}, \hat{\tau}')$ is the synthetic kernel.

$$G_m(\hat{\tau}, \hat{\tau}') = \frac{2\pi I\, m\omega_s}{\omega_o^2 V_T h \cos(\varphi_s + \Delta\varphi_s)} \sum_p \frac{Z_{//}(p)}{p} J_m(p\omega_o \hat{\tau}) J_m(p\omega_o \hat{\tau}') \quad . \tag{88}$$

In search of solutions g_{mq}, let us study the properties of these functions.
We define an adjoint function

$$g^+_{mq} = \frac{g_{mq}}{\frac{1}{\hat{\tau}} \frac{\partial g_0}{\partial \hat{\tau}}} \tag{89}$$

that satisfies

$$\left(\omega_{c_{m\ell}} - m\omega_s\right) g^+_{m\ell} + \int_0^\infty G_m(\hat{\tau}, \hat{\tau}')\, g_{m\ell}(\hat{\tau}')\, \hat{\tau}'\, d\hat{\tau}' = 0 \quad . \tag{90}$$

Then we define the scalar product of two solutions $g_{m\ell}$ and g_{mq} as follows

$$\langle g_{m\ell}, g_{mq} \rangle = \langle g_{mq}, g_{m\ell} \rangle = \int_0^\infty g_{mq}(\hat{\tau})\, g^+_{m\ell}(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \tag{91}$$

and we use equations (10) and (12) to express it:

$$(\omega_{c\,mq} - m\omega_s)\langle g_{ml}\,,\,g_{mq}\rangle = (\omega_{c\,ml} - m\omega_s)\langle g_{mq}\,,\,g_{ml}\rangle$$

$$= \int_{\hat{\tau}}\int_{\hat{\tau}'} G_m(\hat{\tau},\hat{\tau}')\, g_{ml}(\hat{\tau})\, g_{mq}(\hat{\tau}')\, \hat{\tau}\hat{\tau}'\, d\hat{\tau}d\hat{\tau}' = \frac{2\pi I\, m\omega_s}{\omega_o^2 V_T h \cos(\varphi_s+\Delta\varphi_s)} \sum_p \left(j\frac{Z_{//}(p)}{p}\right)\sigma_{mq}(p)\,\sigma_{ml}(p) \quad (92)$$

We can conclude that :

- If $q \neq l$, the scalar product has to be null. Therefore, the solutions g_{mq} are orthogonal according to equation (91). This is the only information we have in order to get the solutions. A priori, these solutions which depend on q_o via $\frac{1}{\hat{\tau}}\partial q_o/\partial\hat{\tau}$ are not necessarily expressible in terms of functions that exist in the standard library of functions.

- If the g_{mq} were known exactly, then, with $q=l$ in equation (92) one could get the coherent frequencies $\omega_{c\,mq}$.

$$\omega_{cmq} = m\omega_s\left(1 + \frac{2\pi I}{\omega_o^2 V_T h\cos(\varphi_s+\Delta\varphi_s)} \frac{\sum_p j\frac{Z_{//}(p)}{p}\sigma_{mq}^2(p)}{\int \frac{g_{mq}^2}{\frac{1}{\hat{\tau}}\frac{\partial g_o}{\partial\hat{\tau}}}\hat{\tau}\,d\hat{\tau}}\right) . \quad (93)$$

- Furthermore, for a local interaction like space charge or inductive walls (constant $Z_{//}(p)/p$) $Z_{//}(p)/p$ can be removed out of the summation over p values. The kernel is a Dirac function and the spectra σ_{mq} are self orthogonal.

This last property allowed F. Sacherer to find out a simple exact solution with Jacobi and Legendre polynomials $P_m^{(\alpha,\beta)}$ and P_m . It is valid for a constant $Z_{//}(p)/p$ and corresponds to the bunch with parabolic line density (26).

$$\omega_{cmq} = m\omega_s\left(1 + \frac{2\pi I}{\omega_o^2 V_T h\cos(\varphi_s+\Delta\varphi_s)}\left(j\frac{Z_{//}(p)}{p}\right) C_{mq}\right) \quad (94)$$

$$g_{mq}(\hat{\tau}) = \int_0^{2\pi} e^{jq\theta} P_m\left(\frac{2\hat{\tau}}{\tau_L}\cos\theta\right) d\theta \quad (95)$$

$$\Delta\lambda_{mq}(\tau) = P_m\left(\frac{2\tau}{\tau_L}\right) . \quad (96)$$

2.6.5 G. Besnier expansion in orthogonal polynomials

We assume there exists a complete set of normalized functions $f_\mu(\hat{\tau})$ which fulfil the orthogonality relation (90).

$$\langle f_\mu(\hat{\tau})\,,\,f_\nu(\hat{\tau})\rangle = \begin{cases} 0 & \mu\neq\nu \\ 1 & \mu=\nu \end{cases} . \quad (97)$$

We expand the solutions $g_{mq}(\hat{\tau})$ and the kernel $G_m(\hat{\tau},\hat{\tau}')$ in terms of the f_μ 's.

$$g_{mq}(\hat{\tau}) = \sum_{\mu=0}^{+\infty} a_\mu f_\mu(\hat{\tau}) \tag{98}$$

$$G_m(\hat{\tau},\hat{\tau}') = \sum_{\mu,\nu} M^m_{\mu\nu} f^+_\mu(\hat{\tau}) f^+_\nu(\hat{\tau}') \quad . \tag{99}$$

Applying the orthogonality condition (97) several times we get an infinite system of linear equations.

$$\left(\omega_{cmq} - m\omega_s\right) a_\mu + \sum_\nu M^m_{\mu\nu} a_\nu = 0 \tag{100}$$

$M^m_{\mu\nu}$ is the interaction matrix.

$$M^m_{\mu\nu} = \frac{2\pi I m\omega_s}{\omega_o^2 V_T h \cos(\varphi_s+\Delta\varphi_s)} (-1)^m \sum_p \sigma_{mf_\mu}(p)\, \sigma_{mf_\nu}(p) \left(j \frac{Z_\ell(p)}{p}\right) \tag{101}$$

with $$\sigma_{mf_\mu}(p) = j^{-m} \int_0^\infty J_m(p\omega_o\hat{\tau})\, f_\mu(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \quad .$$

Numerical solutions of the truncated equation (100) yield eigenvalues $\omega_{cmq} - m\omega_s$ and the eigenvectors a_μ which allow to express g_{mq} (equation 98).

For the bunch with parabolic line density orthogonal functions with the required weight-function and a range which extends from 0 to $\tau_l/2$ can be obtained from the Jacobi polynomials. Laguerre polynomials can be used for a gaussian bunch.

2.6.6 J.L. Laclare's eigenvalue problem

We simply multiply both sides of equation (86) by $J_m(l\omega_o\hat{\tau})$ and integrate over $\hat{\tau}$ values to get

$$\left(\omega_{cmq} - m\omega_s\right) \sigma_{mq}(l) = \sum_p K^m_{lp}\, \sigma_{mq}(p) \tag{102}$$

$$K^m_{lp} = -\frac{2\pi I\, m\omega_s \left(j \frac{Z_\ell(p)}{p}\right)}{\omega_o^2 V_T h \cos(\varphi_s+\Delta\varphi_s)} \int_0^\infty \frac{\partial g_o}{\partial \hat{\tau}}\, J_m(p\omega_o\hat{\tau})\, J_m(l\omega_o\hat{\tau})\, d\hat{\tau} \quad . \tag{103}$$

For q running from $-\infty$ to $+\infty$,

$\omega_{cmq} - m\omega_s$ is the q^{th} eigenvalue

$\sigma_{mq}(p)$ the q^{th} eigenvector

of the infinite matrix K^m_{lp} the elements of which are defined by equation (103) (one column and one row per frequency line). For numerical reasons, the matrix needs be truncated. It explores a finite frequency domain (smooth impedances over a large frequency range,

small details over a reduced frequency range). For a given problem its dimension have to be larger when compared to those of Besnier's matrix. On the other hand each coefficient is easier and faster to compute. As a matter of fact Besnier's coefficients are given by slowly converging series.

2.6.7 J. Wang and C. Pellegrini approach

This last method is based essentially on the same analysis as other's but the authors do not define the mode number m in their perturbation.

It has been developed to explain fast blow up of the beam in a time interval that could be less than the synchrotron period.

2.7 Low intensity coherent modes of oscillation

We are assuming a low intensity bunched beam which interacts with the different components of a standard circular impedance 2.4. The modes of oscillation are solutions of equation (102).

2.7.1 Local interaction space charge or inductive walls

Because $j\frac{Z_{/\!/}(p)}{p}$ is constant and real,

$$j\frac{Z_{/\!/}(p)}{p} = \frac{Z_0 g}{2\beta\gamma^2} - L\omega_0 \qquad \omega \leq \omega_c \tag{104}$$

the matrix is real and symmetrical. The eigenvectors σ are orthogonal, the eigenvalues $\omega_{cmq} - m\omega_s$ are real. There is no instability. Once started, a pure eigenmode rotates indefinitely at coherent frequency in phase space.

Figure 10 illustrates the results for the bunch with parabolic amplitude (25).

The coherent frequency shifts can be written

$$\Delta\omega_{cmq} = \omega_{cmq} - m\omega_s = \frac{4I\,m\,\omega_s}{\pi^2 B^3 V_T h \cos(\varphi_s+\Delta\varphi_s)}\, C_{mq}\left(j\frac{Z_{/\!/}(p)}{p}\right) . \tag{105}$$

The largest C_{mq} values are listed in Fig. 10 in a decreasing order for the lowest values of m. Given m, the largest C_{mq} takes the subscript $q=m$, the next one $q=m+2$, $\ldots$, $q=m+2k$, $\ldots$ $0\leq k<\infty$.

The spectrum of mode mq is peaked at $\omega \sim (q+1)\pi/\tau_L$ and extends $\pm 2\pi/\tau_L$ rad/sec .

Often, only the most coherent modes, with $q=m$ are referred to $m=1$ and $q=1$ dipole $m=2$ and $q=2$ quadrupole $m=3$ and $q=3$ sextupole etc... It should be noted that two modes with same q are peaked at the same frequency (same line density). They also have roughly the same C_{mq} and accordingly the same sensitivity in the same frequency range. Nevertheless they correspond to entirely different density patterns in phase space (compare g_{13} , S_{13} and g_{33} , S_{33} for instance).

Provided the bunch is long enough and the q value is not too large so that the main part of the mode spectrum stands well below the cutoff frequency, equation (105) gives a good approximation of space charge and broad band inductance effects.

In order to understand the physical meaning of the coherent frequency shift let us take the dipole mode $m=q=1$ and let us compare the coherent frequency ω_{c11} with the zero intensity incoherent frequency ω_{s_0} .

The quantity $\omega_{c11} - \omega_{s_0}$ is obtained

$$\omega_{c11} - \omega_{s_0} = \omega_{c11} - \omega_s + \omega_s - \omega_{s_0} = \Delta\omega_{c11} + \Delta\omega_i$$

by adding the coherent (105) and the incoherent (62) shifts.

$$\omega_{c11} - \omega_{s_0} = \frac{4I\,\omega_s}{\pi^2 B^3 V_T h \cos(\varphi_s+\Delta\varphi_s)}\left(j\frac{Z_{/\!/}(p)}{p}\right)\left(C_{11} - \frac{2}{\pi}\right) . \tag{106}$$

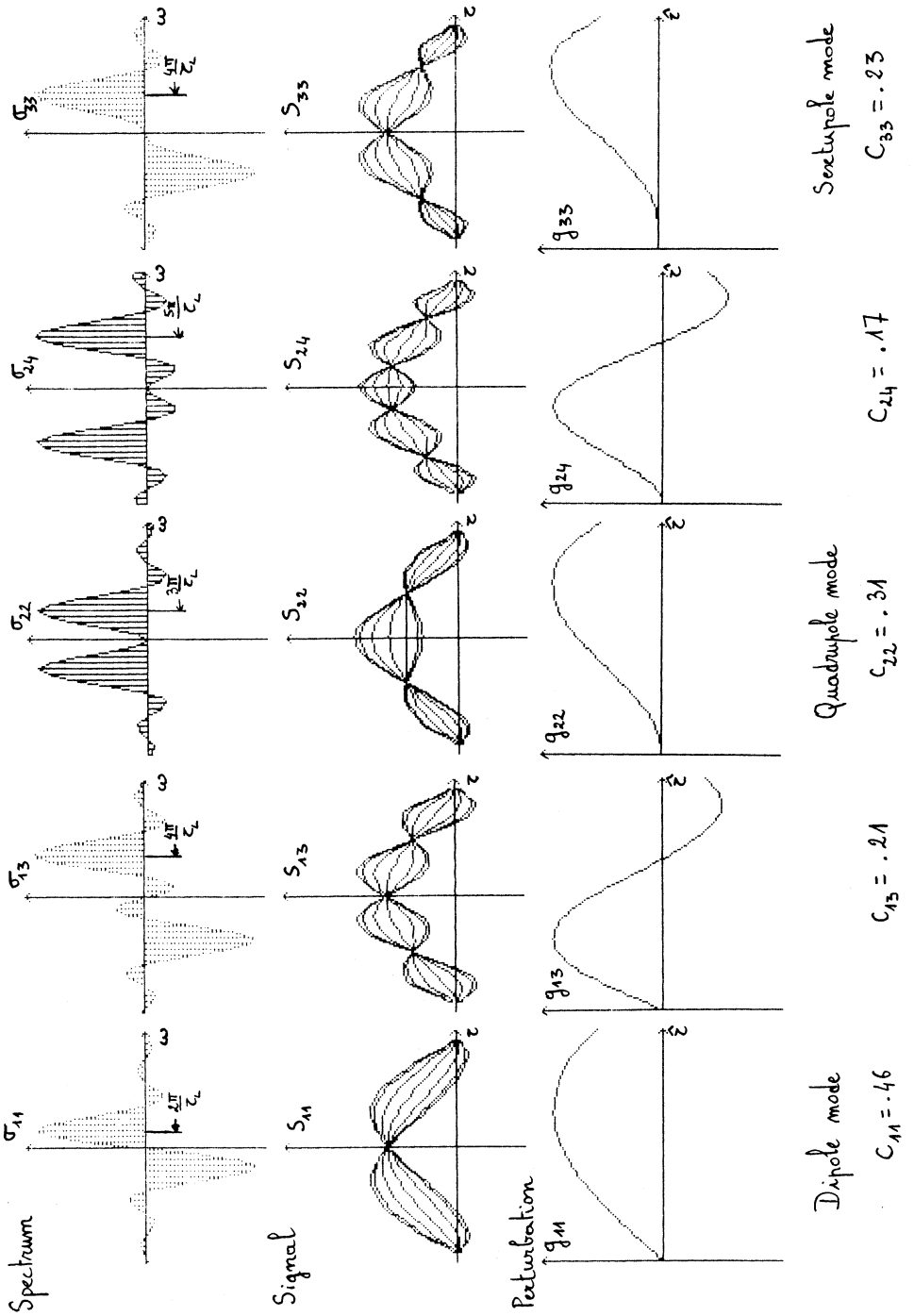


Fig. 10 Coherent modes of oscillation for a parabolic bunch distribution in an inductive impedance

Therefore, as far as frequencies are concerned, coherent and incoherent effects subtract. The frequency of the dipole mode is pushed back towards ω_{s_o} . We would obtain $\omega_{cm} = \omega_{s_o}$ exactly if a bunch with parabolic line density were assumed.

Depending on the stationary distribution, the dipole mode corresponds more or less to an off-centered rigid motion of the bunch in phase space. For a pure rigid motion (Fig. 11), the potential well deformation (change of focusing) induced by the stationary distribution is a pure quadrupole attached to the bunch center of mass position. Therefore, the bunch center of mass does not see any electric field change. It keeps oscillating in the external RF guide field and oscillates at frequency ω_{s_o} .

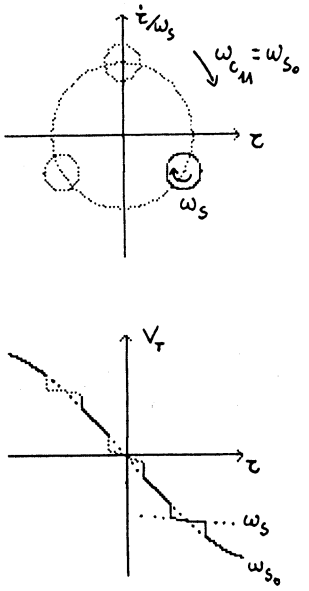


Fig. 11 Rigid dipole motion in phase space and induced potential inside the bunch

The coherent frequency shifts $\Delta\omega_{cmq}$ get smaller and smaller with increasing values of m and q . Nevertheless the overall tendancy is to push back the coherent frequency towards $m\omega_{s_o}$ (see Fig. 9).

2.7.2 Parastitic high Q resonators - Coupled bunch modes

Let us imagine a very narrow band object, so narrow that the coherent motion is driven by a single line at $p\omega_o + \omega_c$. Going back to the matrix equation (102), (103) valid at low intensity for a single value of m , we can write

$$\left(\omega_{cm} - m\omega_s\right) \sigma_m(l) = K_{lp}^{m} \; \sigma_m(p) \tag{107}$$

The sum over p has disappeared. A single coherent mode of m type is solution of equation (107).

- coherent frequency

$$\omega_{cm} = m\omega_s \left(1 - \frac{2\pi I}{\omega_o^2 V_T h \cos(\varphi_s + \Delta\varphi_s)} \, j \, \frac{Z_n(p)}{p} \int_0^{\infty} \frac{\partial q_o}{\partial \hat{\tau}} \, J_m^2(p\omega_o\hat{\tau}) \, d\hat{\tau} \right) \tag{108}$$

- spectrum

$$\sigma_m(p) = 1 \tag{109}$$

$$\sigma_m(l) = \frac{\int_0^\infty \frac{\partial g_0}{\partial \hat{\tau}} J_m(l\omega_0\hat{\tau})\, J_m(p\omega_0\hat{\tau})\, d\hat{\tau}}{\int_0^\infty \frac{\partial g_0}{\partial \hat{\tau}} J_m^2(p\omega_0\hat{\tau})\, d\hat{\tau}} \tag{110}$$

- perturbation (see equation (39))

$$\hat{\tau}\, g_m(\hat{\tau}) \propto \frac{\partial g_0}{\partial \hat{\tau}}\, J_m(p\omega_0\hat{\tau}) \quad . \tag{111}$$

Let us examine the stability of such a system

$$I_m(\omega_c) \propto \frac{-m}{V_T \cos(\varphi_s+\Delta\varphi_s)}\, R_e\left(\frac{Z_{//}(p)}{p}\right) \int_0^\infty \frac{\partial g_0}{\partial \hat{\tau}}\, J_m^2(p\omega_0\hat{\tau})\, d\hat{\tau} \quad . \tag{112}$$

A resistance is needed in order to get an imaginary frequency shift.

A priori, g_0 is a monotonic decreasing function null at bunch edge $\hat{\tau} = \tau_L/2$. Therefore the integral in (112) is negative.

On the other hand, $V_T \cos(\varphi_s + \Delta\varphi_s)$ is negative above transition and $R_e(Z_{//}(p)/p)$ is positive at positive frequencies.

The conclusion is the following. Above transition, upper (lower) synchrotron sidebands as seen by a spectrum analyser have a destabilizing (stabilizing) effect, with the opposite below transition.

This result can be qualitatively explained by means of Fig. 12.

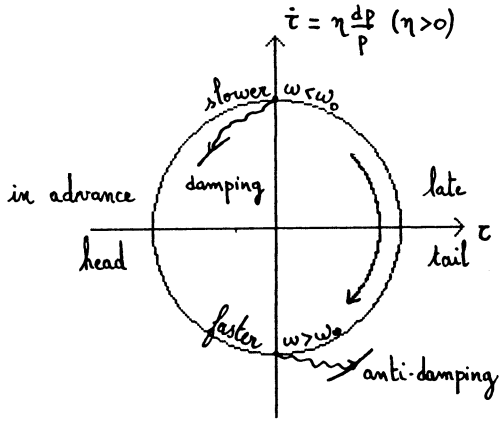


Fig. 12 Qualitative description of the synchrotron motion in phase-space

Above transition ($\eta > 0$), a positive dp/p is associated with a positive $\dot{\tau}$ and a lower revolution frequency. Therefore, the upper part of the phase-space diagram coincides with contribution to the lower sidebands of the spectrum.

If we assume a rigid dipole mode oscillation, when the bunch center of mass is at noon in phase space, it loses energy by interacting with the resistance at lower sideband frequencies. It spirals towards the synchronous particle, the coherent amplitude is damped.

On the other hand, the resistance associated with upper sidebands will have a destabilizing effect in the lower part of phase space diagram.

Up to now, a single bunch has been assumed and by means of a coherent perturbation $\Delta\Psi$ (see equation (39)), particles within this bunch have been arranged in order to produce a coherent signal at $p\omega_0 + \omega_s$.

If M equidistant bunches are present in the ring, one can go further and coherently arrange particles within successive bunches to obtain a coupled bunch coherent motion.

With the index n running from 0 to $M-1$, coupled bunch mode number n will correspond to a phase shift of $n\,2\pi/M$ between the coherent perturbations of two successive bunches.

The Fourier component of the signal is M times larger but only every M^{th} line occurs.

With the above convention, the spectrum of coupled bunch mode is at frequencies

$$\omega_n = (n + pM)\,\omega_0 + m\omega_s \quad . \tag{113}$$

The equations of coherent coupled bunch motion are essentially the same as the equations of coherent single bunch motion (84), (86), (102), (103) except that, I, current in one bunch becomes MI current in M bunches.

Indeed, the summation over the spectrum is restricted to the coupled bunch mode spectrum.

B remains the bunching factor of the single bunch (24).

An example is given in Fig. 13. We are considering $M = 4$ bunches performing a coherent coupled bunch dipole mode. The four possibilities of coupling the bunches in phase space are shown in the upper part of the figure. In the lower part, the line spectrum is drawn as seen by a spectrum analyser.

The spacing between upper (destabilizing above transition) and lower (stabilizing) sidebands is minimum $(\Delta\omega = 2\omega_c)$ for coupled bunch mode $n = 0$ and $n = M/2$. It is maximum $(\Delta\omega = \frac{M}{2}\omega_0 + 2\omega_c)$ for $n = \frac{M}{4}$ and $n = \frac{3M}{4}$.

As a consequence, modes $n = M/4$ and $n = 3M/4$ are very sensitive to narrow band resonators. When mode n is damped (antidamped) the complementary mode $M-n$ is antidamped (damped). The maximum frequency shift is obtained when the resonator frequency coincides with one of the frequency lines in the coupled bunch mode signal $\omega_r = \omega_n$.

Provided the spacing between two adjacent lines of the spectrum is larger than the resonator bandwidth $\delta\omega$,

$$\delta\omega = \omega_r / 2Q \tag{114}$$

a single line drives the mode. Equation (108) applies. It can be adapted to the bunch with parabolic amplitude (25).

$$\frac{\omega_c - m\omega_s}{m\omega_s} = \frac{4\,MI}{(m+1)\pi^2 B^2 V_T h \cos(\varphi_s + \Delta\varphi_s)}\, 2 \frac{R_s \omega_0}{\omega_r} F_m\left(\omega_1 \frac{\tau}{2}\right) D \tag{115}$$

$$F_m(X) = \frac{2(|m|+1)}{X^2} \int_0^X J_m^2(u)\, u\, du \tag{116}$$

R_s is the shunt impedance (see equation (51) for instance).

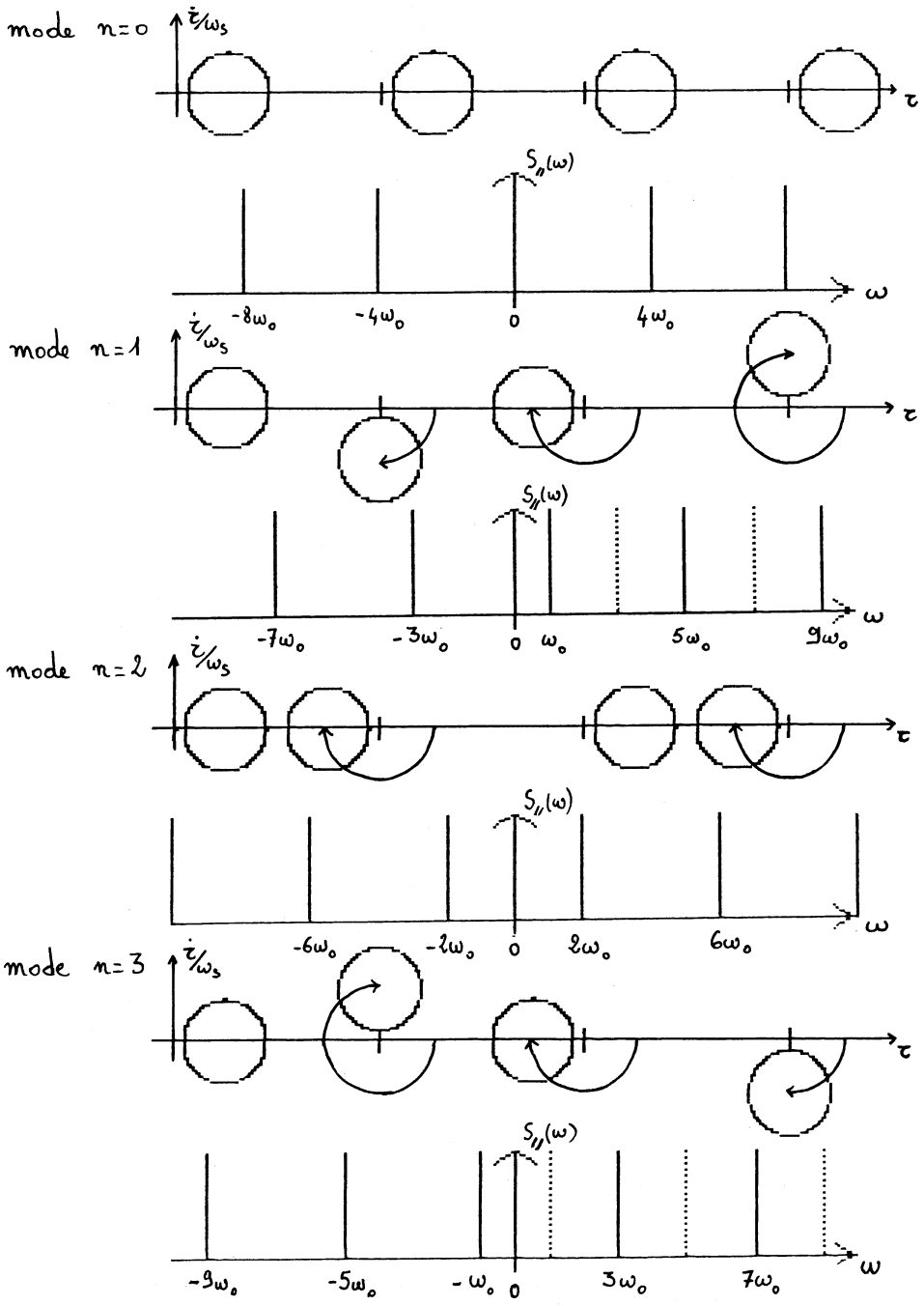


Fig. 13 Coherent coupled-bunch modes of four bunches

The form factor F_m is plotted in Fig. 14.

For large bandwidths (broad band impedance for instance), more than one line must be included in the frequency shift calculation. Cancellation occurs between upper and lower sidebands. The frequency shift is reduced by the factor D shown in Fig. 15.

$$D = \frac{\alpha}{\sinh \alpha} \tag{117}$$

The quantity

$$\alpha = \frac{\omega_r}{2Q} \frac{2\pi}{\omega_0} \frac{1}{M} \tag{118}$$

is the attenuation of the wake between successive bunches. There is no instability for wakefields that decay appreciably before the next bunch arrives.

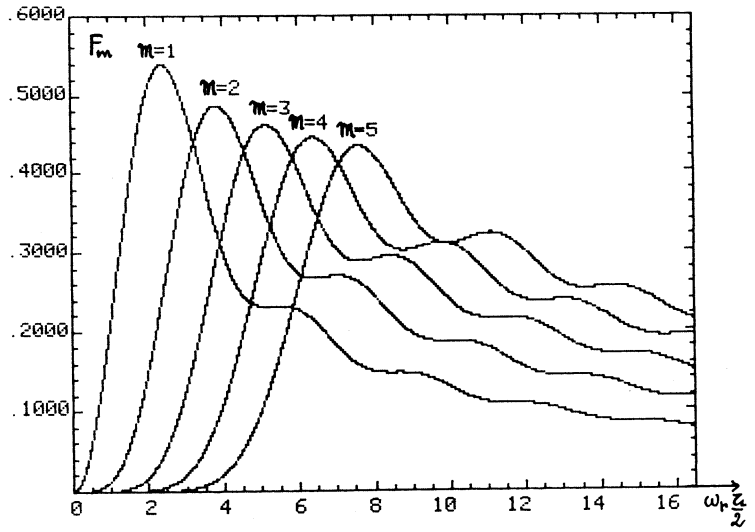


Fig. 14 Form factors F_m for different modes m of longitudinal bunch oscillation

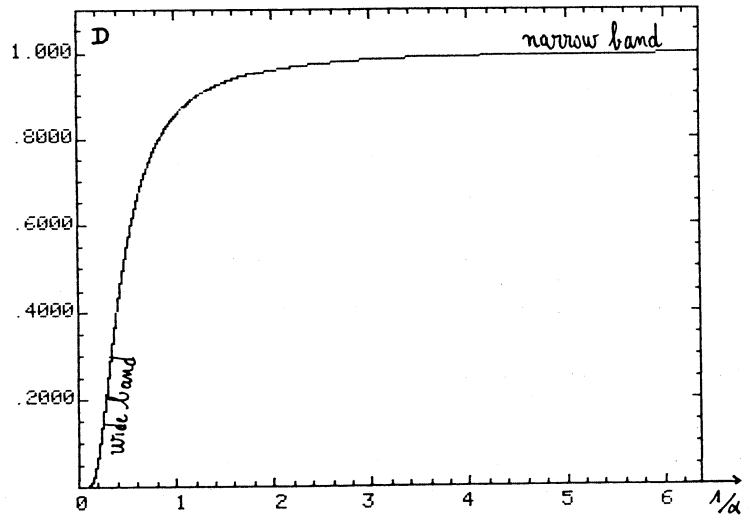


Fig. 15 Attenuation factor for the frequency shift in the case of a resonator impedance

Coupled bunch coherent instability is very harmful and difficult to fight.

Non-linearities can Landau damp the mechanism. When assuming some spread in incoherent synshrotron frequency, the threshold current is defined by

$$\Delta\omega_c \leq \frac{\sqrt{m}}{4}\,\Delta\omega_{spread} \quad . \tag{119}$$

In the above expression, $\Delta\omega_{spread}$ is the full spread in ω between center and edge of the bunch due to non-linear synchrotron forces.

A spread in incoherent frequencies of individual bunches can also help in decoupling. This spread can be provided by a low frequency RF system. It can come from a spread in bunch population as well, in which case the required r.m.s spread is given by

$$\left\langle \frac{\delta I}{I} \right\rangle_{r.m.s} > \frac{I_m(\Delta\omega_c)}{R_e(\Delta\omega_c)} \quad . \tag{120}$$

Indeed, one can imagine a feed back system to damp this type of instability. Nevertheless, the number of coupled bunch modes and consequently the number of bunches has to be limited in order to keep the feed back system bandwidth down.

2.7.3 RF cavity - Robinson's effect

For a single bunch or two bunches, upper and lower sidebands belong to the same couple bunch mode and therefore tend to cancel unless the impedance is very narrow band as in an RF cavity.

The RF is tuned to ω_{RF} very close to $h\omega_0$. Given m, the imaginary part of the frequency shift can be obtained by adding the effect of the upper and lower coherent sidebands (at $h\omega_0 + m\omega_s$ and $h\omega_0 - m\omega_s$ when the real part of the frequency shift is neglected $R_e(\omega_c - m\omega_s) \ll \omega_s$).

Let $R(\omega)$ be the resistance at frequency ω, we apply equation (108) and find

$$I_m(\omega_c) = -\frac{2\pi I\, m\omega_s}{\omega_0^2 V_T h^2 \cos(\varphi_s + \Delta\varphi_s)} \left[R(h\omega_0 + m\omega_s) - R(h\omega_0 - m\omega_s)\right] \int_0^\infty \frac{\partial g_0}{\partial \hat{\tau}}\, J_m^2(h\omega_0 \hat{\tau})\, d\hat{\tau} \quad . \tag{121}$$

It is important to notice that the result is different from zero essentially because $R_e(Z_{/\!/}(p)/p)$ is evaluated at $h\omega_0 + m\omega_s$ while $R_e(Z_{/\!/}(-p)/-p)$ is evaluated at $h\omega_0 - m\omega_s$.

Trouble is avoided in this case by tuning the cavity to overlap the stable sideband according to Robinson's criterion.

2.7.4 Resistive wall impedance

The resistive wall impedance is peaked at low frequencies were the spectral amplitude of longitudinal modes is very weak. In most cases this effect can be disregarded.

2.7.5 Conclusions about coherent instabilities at low intensity

The theory that has been developed here above is in perfect agreement with observations.

The frequency shift formula looks reliable. It has been tested on several low order modes with space charge dominated beams as well as ultrarelativistic beams in inductive walls. Nowadays, it is often used to measure the reactive part of the impedance at low frequencies.

The predicted bunch shape oscillations (up to sextupole mode) are observable in many machines. In Fig. 16 top, photographs of a dipole mode (left) and of a quadrupole mode (right) are presented.

Coupled bunch modes are frequently observed in machines operated in multibunch mode. In Fig. 16 bottom, one can observe coupled bunch mode $n = 1$ with three bunches in SATURNE (Saclay).

Lastly, the theory gives a good picture of Robinson's instability.

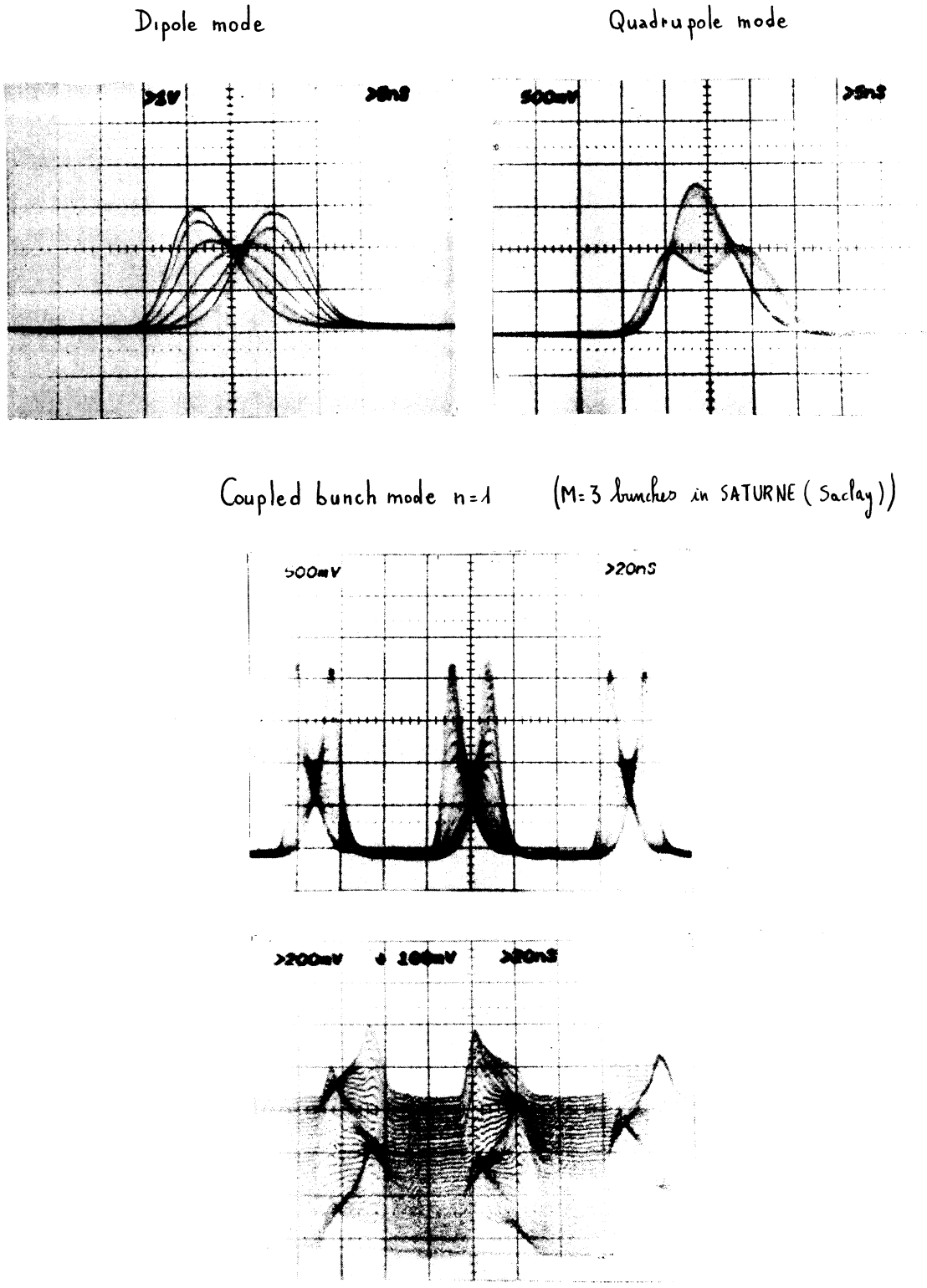


Fig. 16 Experimental observation of bunch oscillations (top) and coupled-bunch mode (bottom) in SATURNE (Saclay)

2.8 Single-bunch instability at high intensity

As stated before, the original theory of F. Sacherer can quite well explain low intensity phenomena that have been observed with bunched beams in a circular machine.

However, in this form, this theory cannot explain other types of single bunch instabilities that have been found.

For instance the "microwave instability", the signature of which is a high frequency signal, shows up in most machines (long proton bunches and short electron bunches), above some threshold current. Boussard's criterion derived from the Keil-Schnell coasting beam stability requirement by using local values of current and momentum spread can predict this threshold quite accurately.

In addition, with increasing current the rate of bunch lengthening due to potential well effect changes suddenly above a threshold current. This "anomalous" bunch lengthening often called "turbulent" bunch lengthening is accompanied by a bunch widening (increase of momentum spread).

An attempt to explain some of the above effects with theory consists of introducing "mode coupling".

2.8.1 Analysis of the instability mechanism

In order to explain the way mode coupling acts, let us go back to the low intensity regime and consider a single perturbation again, the dipole mode for example.

As represented in Fig. 10, once the stationary signal has been subtracted, the single pass signal of the perturbation is an odd function of time. As a consequence, the perturbation spectrum is an odd function of frequency.

When considering a local interaction (2.7.1) the force is exactly the derivative of the signal (signal as seen through a pure inductance). Therefore, the force is an even function of time. It can efficiently drive the perturbation and it leads to a frequency shift proportional to the impedance, according to equation (105).

The same result would apply with a capacitance as well. The force would be the integral of the signal, that is to say an even function of time again and it would drive efficiently the perturbation.

On the other hand, with a pure resistance the force is exactly the signal. This time it has the wrong parity to drive the perturbation (weak differential effect).

Conclusion : an odd perturbation is essentially driven by an even force. Indeed, an even perturbation is essentially driven by an odd force.

Now, let us add two adjacent perturbations, an odd one (dipole $m=1$) and an even one (quadrupole $m=2$) interacting with a general impedance (inductance or capacitance plus resistance).

Each perturbation will interact with the forces that have the right parity to drive them, namely, the reactive part of its self field plus the resistive part of the adjacent perturbation field.

Conclusion : two adjacent perturbations cannot couple via a pure reactance. A resistance is necessary.

Let us notice that resistance contributes an imaginary frequency shift and may cause instability.

2.8.2 Matrix equation of single-bunch modes at high intensity

The starting point is equation (84). Both sides can be multiplied by $J_m(l\omega_0\hat{\tau})$ and integrated over $\hat{\tau}$ values.

$$\left(\omega_c - m\omega_s\right)\sigma_m(l) = \frac{-2\pi I m\omega_s}{\omega_0^2 V_T h\cos(\varphi_s+\Delta\varphi_s)} \sum_p \frac{Z_{//}(p)}{p} \int_0^\infty \frac{\partial g_0}{\partial\hat{\tau}} J_m(p\omega_0\hat{\tau}) J_m(l\omega_0\hat{\tau})\, d\hat{\tau}\ \sigma(p) . \quad (122)$$

Then, using the definition (85), one can sum up over m values and express the spectral amplitude at $l\omega_0 + \omega_c$.

$$\sigma(\ell) = \frac{-2\pi I}{\omega_o^2 V_T h \cos(\varphi_s+\Delta\varphi_s)} \sum_p \frac{Z_{//}(p)}{p} \left\{ \sum_m \frac{m\omega_s}{\omega_c - m\omega_s} \int_0^\infty \frac{\partial g_o}{\partial \hat{\epsilon}} J_m(p\omega_o\hat{\epsilon})\, J_m(\ell\omega_o\hat{\epsilon})\, d\hat{\epsilon} \right\} \sigma(p) \quad (123)$$

Let us introduce a matrix element $M_{\ell p}$.

$$\sigma(\ell) = \frac{4I}{\pi^2 B^3 V_T h \cos(\varphi_s+\Delta\varphi_s)} \sum_p \frac{Z_{//}(p)}{p} M_{\ell p}\, \sigma(p) \quad . \quad (124)$$

When the bunch has a stationary distribution with parabolic amplitude (25), $M_{\ell p}$ takes the following form

$$M_{\ell p} = 2B \sum_m \frac{m}{\frac{\omega_c}{\omega_s} - m} \int_0^1 J_m(p\pi B x)\, J_m(\ell\pi B x)\, x\, dx \quad . \quad (125)$$

Let us examine the method that can be used to solve the above matrix equation (124).

Assume a real coherent frequency ω_c measured in incoherent frequency unit, ω_c/ω_s.

Look for the eigenvalues of the matrix

$$\left[\frac{Z_{//}(p)}{p} \right] \left[M_{\ell p} \right] \quad (126)$$

where $\left[Z_{//}(p)/p \right]$ is the diagonal matrix associated with the impedance.

Scale the intensity parameter

$$\xi = \frac{4I}{\pi^2 B^3 V_T h \cos(\varphi_s+\Delta\varphi_s)} \quad (127)$$

in order to adjust the eigenvalue to unity.

Assume that the bunch is long enough and use the low frequency imaginary part of the impedance to calculate V_T / V_{RF} and $\Delta = (\omega_s^2 - \omega_{s_o}^2)/\omega_{s_o}^2$ (equation 62).

Finally, by means of equation (67) or (69) find out the bunch length.

In the following, we apply this method to three special types of impedances, inductive wall or space-charge, narrow-band resistance, broad-band resonator.

2.8.3 Space-charge and inductive-wall modes at high intensity

As it has been already pointed out in section 2.8.1, odd and even perturbations cannot couple since there is no resistive component in the assumed impedance. There is a very weak coupling between perturbations with the same parity. This coupling modifies slightly the C_{mq}'s.

The results are presented in Fig. 17.

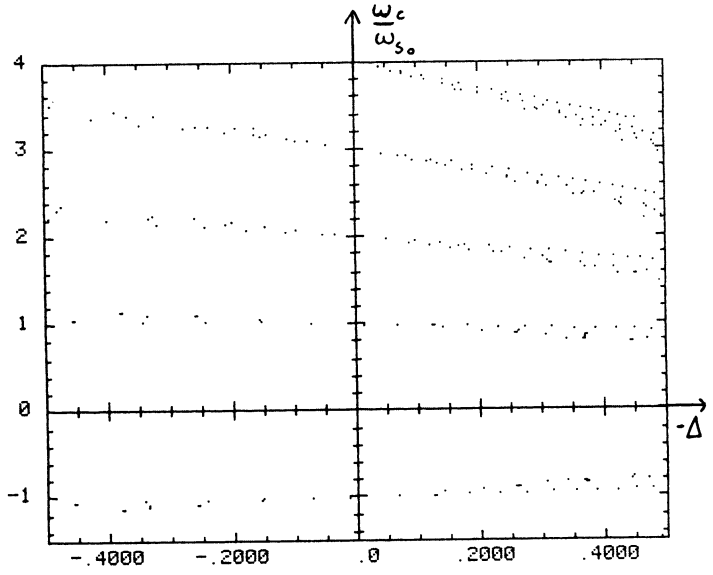


Fig. 17 Coherent frequency shift as a function of the incoherent shift for an inductive impedance

2.8.4 Narrow band resistance

In this section the goal we aim at is to illustrate the resistive coupling. To this end we assume a narrow-band resistance at coherent frequency $p\omega_o + \omega_c$.

$$j\frac{Z_{//}(p)}{p} = 0 \quad \text{except at} \quad p = \pm p_c \quad \text{and} \quad j\frac{Z_{//}(p_c)}{p_c} = -j\frac{Z_{//}(-p_c)}{-p_c} = j\frac{R_s}{p_c} \ . \tag{128}$$

Furthermore, we assume a pair of adjacent perturbations m and $m+1$ for instance.

By using some elementary algebra and the following definitions

$$k_m = \int_0^{+\infty} J_m^2(p_c\omega_o\hat{\tau})\,\hat{\tau}\,d\hat{\tau} > 0 \quad , \quad \sigma_m(-p_c) = (-1)^m \sigma_m(p_c)$$

the matrix equation (124) can be reduced to

$$\begin{pmatrix} \sigma_m(p_c) \\ \sigma_{m+1}(p_c) \end{pmatrix} = 4B\varepsilon j\frac{R_s}{p_c} \begin{pmatrix} 0 & \dfrac{m k_m}{\frac{\omega_c}{\omega_s} - m} \\ \dfrac{(m+1)k_{m+1}}{\frac{\omega_c}{\omega_s} - (m+1)} & 0 \end{pmatrix} \begin{pmatrix} \sigma_m(p_c) \\ \sigma_{m+1}(p_c) \end{pmatrix} . \tag{129}$$

The eigenvalues are solutions of a quadratic equation

$$\left(\frac{\omega_c}{\omega_s} - m\right)\left(\frac{\omega_c}{\omega_s} - (m+1)\right) = -C^2\varepsilon^2 \tag{130}$$

with

$$C^2 = 16\, m(m+1) R_s^2 B^2 \frac{k_m k_{m+1}}{p_c^2} \ .$$

The instability threshold is obtained for $\varepsilon_{th} = 1/2C$. This threshold depends strongly on the value of p_c . The instability is efficiently driven when p_c is such that the product $k_m \cdot k_{m+1}$ is maximum, that is to say $p_c \omega_0$ about $(2m+3)\pi/\tau_L$.

Below threshold, the two coherent frequencies are real. The motion is stable. Above threshold, the real part of the coherent frequency ω_c is constant and the imaginary part is positive for one mode (stability) and negative for the other one (instability). Since one mode is unstable, the bunch is unstable.

Results are summarized in Fig. 18a for $m = 1$.

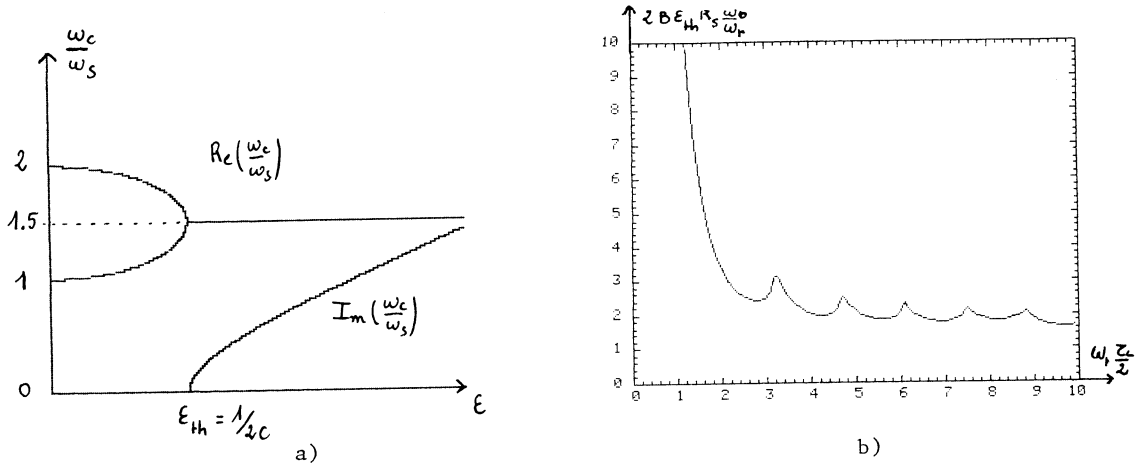


Fig. 18 a) Coherent frequency shift versus intensity for a narrow-band impedance
b) Threshold intensity as a function of the bunch length

At low intensity, we find two modes. The first one at coherent frequency $\omega_c = \omega_s$ is essentially composed of a dipole perturbation. The second one at $\omega_c = 2\omega_s$ consists of a rather pure quadrupole perturbation.

For increasing current, the spacing between the two coherent frequencies gets smaller. Each of the two modes is a mixing of dipole and quadrupole perturbations.

At threshold, the two modes are identical. They are associated with the same coherent frequency $\omega_c = 3/2\,\omega_s$. Maximum stable coupling is reached.

For a narrow-band resonator, when including an infinite number of perturbations, the intensity parameter at threshold ε_{th} is given by

$$2B\varepsilon_{th}R_s\frac{\omega_0}{\omega_r} = \left\{ -\left(\sum_{n=1}^{+\infty} \frac{(2n)^2 k_{2n}}{\left(\frac{\omega_c}{\omega_s}\right)^2 - (2n)^2}\right)\left(\sum_{n=0}^{+\infty} \frac{(2n+1)^2 k_{2n+1}}{\left(\frac{\omega_c}{\omega_s}\right)^2 - (2n+1)^2}\right)\right\}^{-1/2} .$$

In Fig. 18b the above quantity is represented as a function of $\omega_r \tau_L$.

2.8.5 Broad band resonator

Now, we give the results one would obtain with an impedance that consists of a pure resonator of Broad-Band type (Fig. 20). The resonant frequency ω_r (Fig. 19) lies between the maxima of the spectra of modes 22 and 33 ($\omega_{22} \sim 3\pi/\tau_L$ and $\omega_{33} \sim 4\pi/\tau_L$) as obtain with a constant $Z_{//}(p)/p$ (see Fig. 10).

The solution of the matrix equation of coherent motion (124) is a twofold infinity of modes as shown in Fig. 20. The upper graph gives the coherent frequency ω_c of the modes in ω_{s0} unit (vertical axis) as a function of $-\Delta$ (equation (62)) that measures the normalized shift of incoherent frequency squared (horizontal axis) for a long bunch $\omega_r \tau_L \gg 1$. The lower graph gives the same results with $-\varepsilon R_s \omega_0/\omega_r$ along the horizontal axis (equation (127)).

$$\Delta = \frac{\omega_s^2 - \omega_{s0}^2}{\omega_{s0}^2} = \frac{16\, I}{\pi^3 B^3 V_{RF} h \cos\varphi_s} \frac{R_s \omega_0}{Q \omega_r} = \frac{4}{\pi} \frac{V_T}{V_{RF}} \varepsilon \frac{R_s \omega_0}{Q \omega_r} \quad . \tag{131}$$

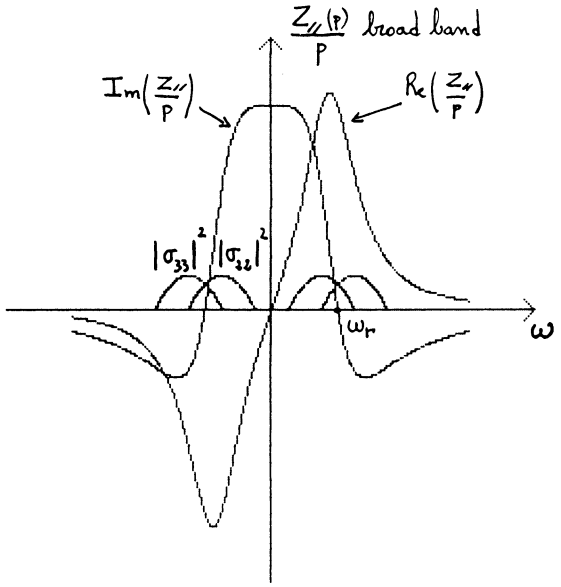


Fig. 19 Broad-band resonator impedance and mode spectra

Below transition "mode coupling" cannot lead to instability. This is the reason why we restrict ourselves to working pcints with a positive abscissa $(-\Delta > 0)$. We are above the transition and because of the low frequency inductance the incoherent frequency decreases for increasing intensity. At very low intensity $(-\Delta \to 0)$, there is no instability, the coherent frequencies are real, they cluster near the integer values of ω_c/ω_{s0}. For increasing intensity, some of the coherent frequencies are pulled up, other are pulled down. At sufficiently high intensity for two coherent frequencies to merge, instability appears. In the present example, the lowest threshold occurs at $-\Delta_{th} = .54$. In the interval, $2 < \omega_c/\omega_{s0} < 3$ when following the solid curve, we can see that the comments already made in the previous section apply again. At low intensity, we get two modes on the curve. The first one, at coherent frequency $\omega_c \sim 2\,\omega_{s0}$ is essentially composed of quadrupole perturbation g_{22} . Its spectrum is shown in Fig. 20a. The second one at $\omega_c \sim 3\,\omega_{s0}$, consists in a rather pure sextupole perturbation g_{33} (Fig. 20c). For increasing current, the spacing between the two coherent frequencies gets smaller and smaller. On one hand, the coherent frequency of the first mode is pulled up because its spectrum overlaps the inductive part of the impedance. On the other hand, the coherent frequency of the second mode is pulled down because its spectrum overlaps the capacitive part of the impedance. Each of the two modes is a mixing of quadrupole and sextupole perturbation. At threshold $-\Delta = -\Delta_{th}$ (Fig. 20b) the two modes are at the same coherent frequency $\omega_c \sim 1.7\,\omega_{s0}$. Maximum stable coupling is reached.

Obviously, the instability threshold depends on the resonant frequency of the resonator and on its bandwidth.

The example of Fig. 20 is a broad band with a bandwidth $\Delta f \cdot \tau_L = 1$ about as wide as the mode spectra. In other words the wakefield decay in about a bunch length τ_L . This is the worst situation leading to the lowest thresholds.

For very small bandwidths, only a few lines contribute and the threshold is high.

In Fig. 21, with $-\varepsilon R_s \omega_0/\omega_r$ along the horizontal axis, the result corresponding to a very narrow band resonator at frequency $\omega_r \tau_L = 7.85$ is shown.

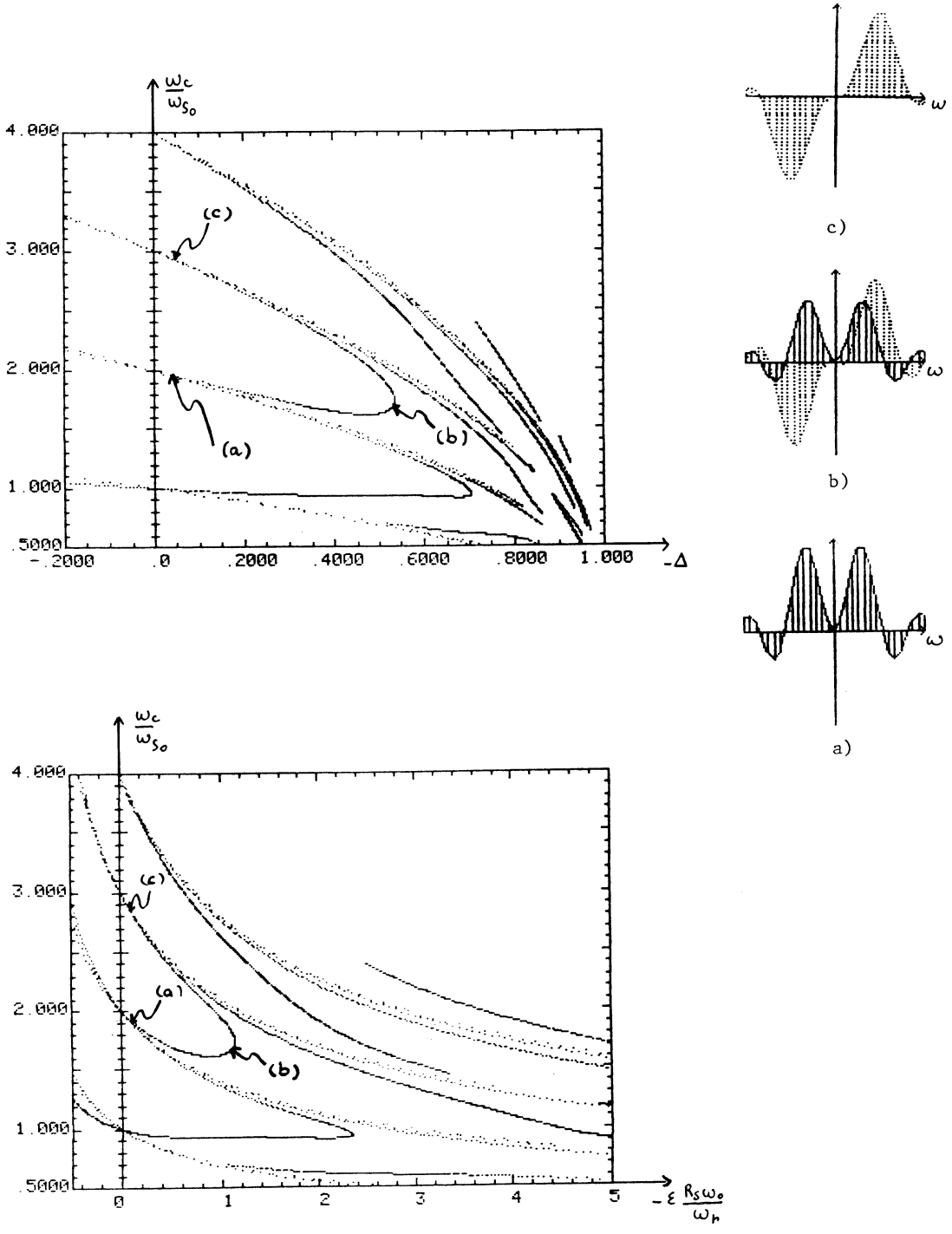


Fig. 20 Coherent-mode frequencies (m = 1 to 4) versus incoherent frequency shift (upper) and intensity parameter (lower)
a) Spectrum of the lowest radial quadrupole mode g_{22}
b) Coupling between quadrupole and sextupole modes at threshold
c) Spectrum of the lowest radial sextupole mode g_{33}

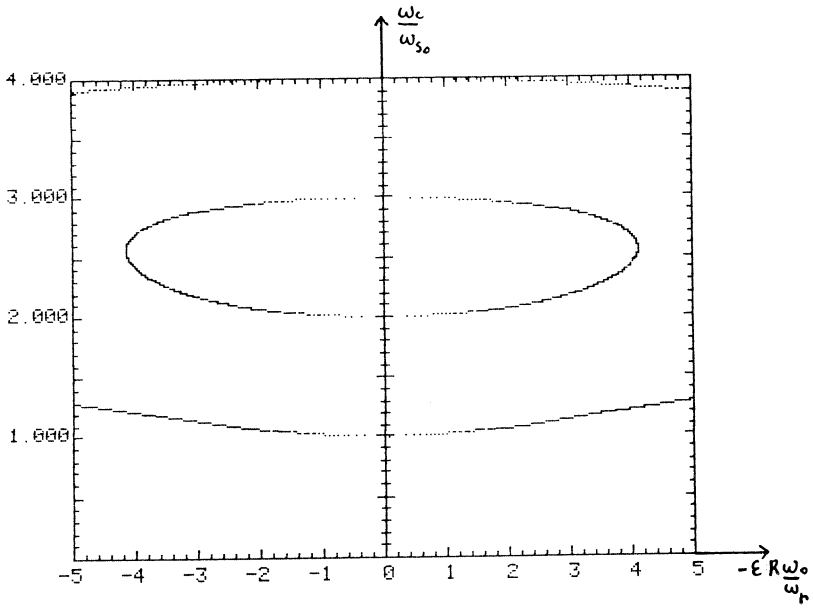


Fig. 21 Coherent-mode frequencies versus intensity for a narrow-band resonator impedance

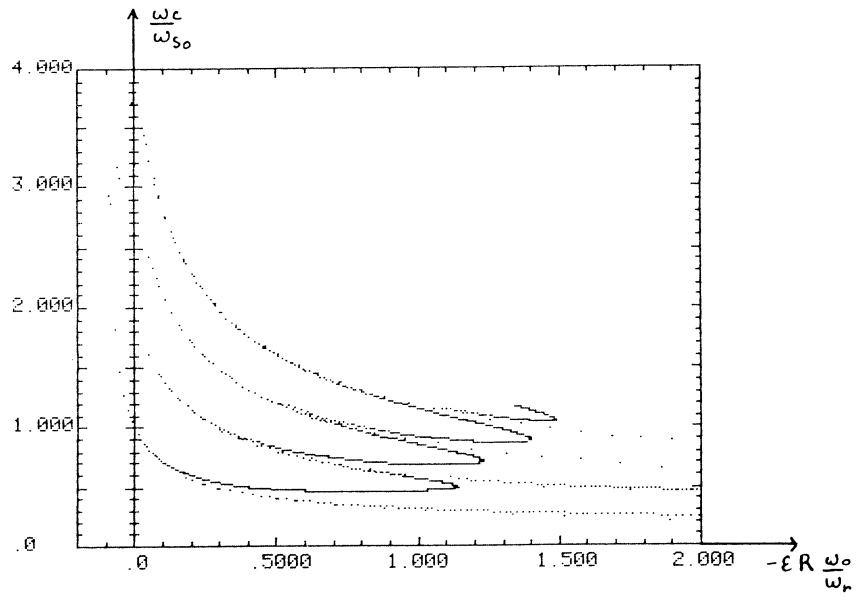


Fig. 22 Coherent-mode frequencies versus intensity for a wide-band resonator impedance

For very large bandwidths, more and more modes lie under the resonance curve and the threshold is higher again (see Fig. 22).

Above transition energy, results corresponding to different resonant frequencies and different bandwidths of the resonator are summarized in Fig. 23.

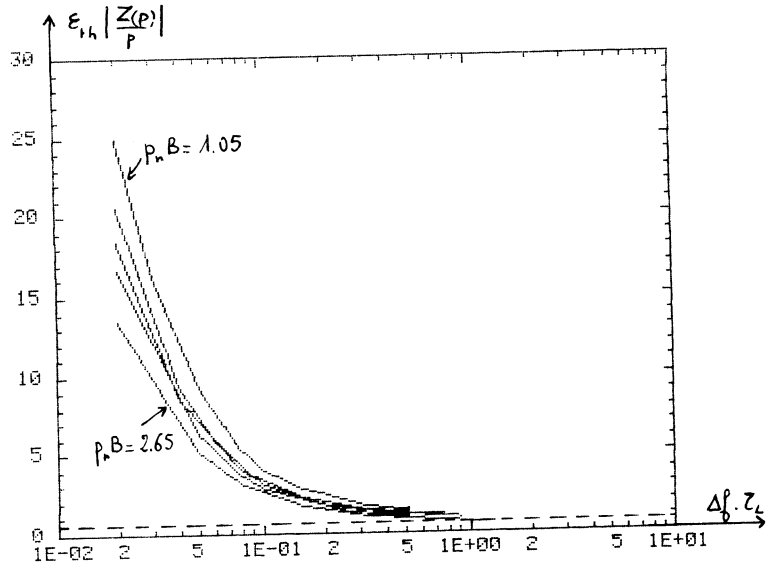


Fig. 23 Normalised threshold intensity parameter versus bandwidth for different resonant harmonic numbers $p_r = \omega_r/\omega_c$

If one defines the peak current $\hat{I}$ and the full-width half-height momentum spread as follows,

$$\hat{I} = \frac{3I}{2B} \tag{132}$$

$$\left(\frac{\Delta p}{p}\right)^2_{F.W.H.H.} = \frac{\omega_s^2 \tau_L^2}{2\eta^2} \tag{133}$$

then stability requires

$$\varepsilon_{th}\left|\frac{Z(p)}{p}\right| < 0.7 \quad \text{or} \quad \frac{8}{3\pi(0.7)} \frac{\hat{I}}{\omega_0 \eta R\left(\frac{p}{e}\right)} \left|\frac{Z(p)}{p}\right| \leq \left(\frac{\Delta p}{p}\right)^2_{F.W.H.H.} \tag{134}$$

It is interesting to notice that the lowest threshold is equivalent ($8/3\pi \sim 0.7$) to the coasting beam threshold for peak current (D. Boussard criterion). Presumably, the bunch lengthens and widens to remain just below threshold.

2.8.6 Bunch lengthening and widening

For given RF conditions (V_{RF}, h, $\cos\varphi_s$) and given impedance, we plan to define the dimensions of the bunch required for stability. Let us recall the following notations

$$\Delta = \frac{\omega_s^2 - \omega_{s0}^2}{\omega_{s0}^2} \quad \text{and} \quad \Delta_0 = \left(\frac{\omega_s^2 - \omega_{s0}^2}{\omega_{s0}^2}\right)_0 = \left(\frac{B}{B_0}\right)^3 \Delta \ . \tag{135}$$

- Proton case

Let B_0 be the zero intensity bunch ..g factor. Matching fixes the peak momentum spread and therefore the emittance.

We start increasing current. Because of the potential well effect, the incoherent synchrotron frequency changes (decreases for example). According to equation (69), new matched conditions can be found, same emittance, smaller momentum spread, longer bunch.

$$\left(\frac{B}{B_0}\right)^{-1} = \left(\frac{B}{B_0}\right)^3 + \Delta_0 \quad . \tag{69}$$

This remains valid till we stay below the instability threshold $\Delta < \Delta_{th}$.

Once the threshold is reached, the emittance of the bunch cannot be kept constant. The bunch has to lengthen in order to maintain Δ at threshold. The momentum dimension has to follow to satisfy matched conditions. The bunch widens out. Equations (136), (137) and (138) apply.

$$\left(\frac{B}{B_0}\right)^3 = \frac{\Delta_0}{\Delta_{th}} \tag{136}$$

$$\left(\frac{\varepsilon}{\varepsilon_0}\right)^6 = (1+\Delta_{th})^3 \left(\frac{\Delta_0}{\Delta_{th}}\right)^4 \tag{137}$$

$$\left(\frac{\Delta p/p}{(\Delta p/p)_0}\right)^6 = (1+\Delta_{th})^3 \left(\frac{\Delta_0}{\Delta_{th}}\right)^2 \quad . \tag{138}$$

- Electron case

B_0 Radiation damping determines the momentum spread. At zero intensity, matching fixes and consequently the emittance.

We start increasing current. Potential well decreases the focusing, the momentum spread remains the same, the bunch lengthens according to equation (67), the emittance grows.

$$\frac{B}{B_0} = \left(\frac{B}{B_0}\right)^3 + \Delta_0 \tag{67}$$

$$\frac{\varepsilon}{\varepsilon_0} = \frac{B}{B_0} \quad . \tag{139}$$

This remains valid till we stay below the instability threshold $\Delta < \Delta_{th}$.

When the threshold is reached, the bunch starts lenghening differently to remain stable. The same equations as for protons (136), (137) and (138) are valid.

Figure 24 gives a qualitative description of the evolution of beam parameters below and at threshold.

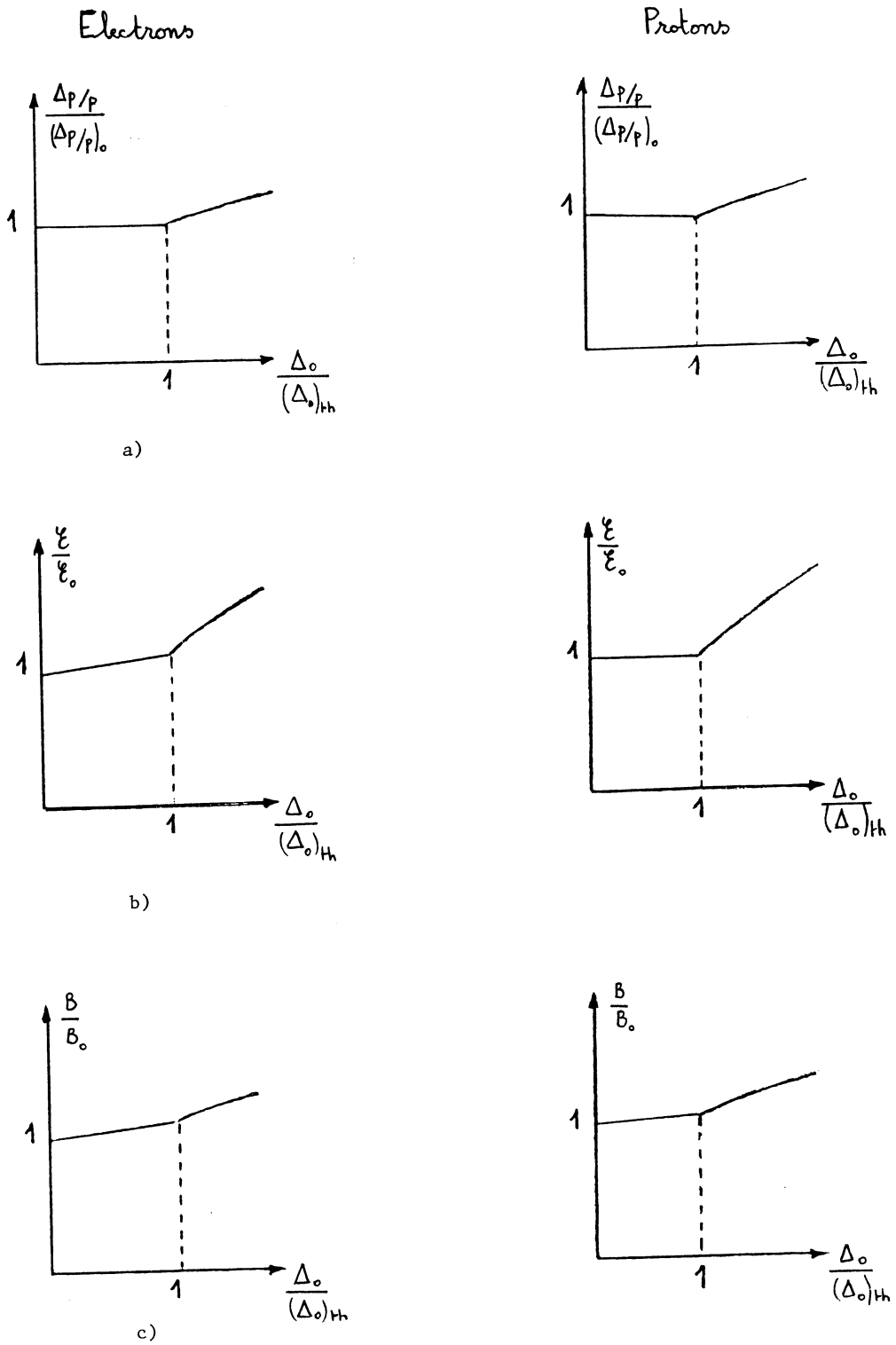


Fig. 24 a) Relative momentum, b) emittance and c) bunch length, versus normalized frequency shift for electrons and protons

3. TRANSVERSE INSTABILITIES

3.1 Preliminary remarks

To deal with transverse instabilities we will follow exactly the same procedure as before. Again we will work in the frequency domain with transverse signal, stationary distribution and perturbation. We will assume a transverse impedance and by means of Vlasov's equation we will describe the coherent motion which consists in eigenmodes of oscillation and coherent frequencies.

Indeed, most of the comments already made in the previous section apply. Therefore, in order to avoid repetitions we will mainly insist on differences between longitudinal and transverse cases.

In the longitudinal case, the instability process is initiated by a perturbation of particle density which creates an electromagnetic field across the beam. Remembering the example of a circular pipe, this electromagnetic field was associated with a return or image current S_w flowing upstream in the wall and uniformly distributed around the beam axis.

In the transverse case, the perturbation consists of a slight transverse displacement of the beam which oscillates from side to side in the external focusing guide field. This time, the wall current S_w is not uniformly distributed around the pipe axis. It is a differential current which flows in opposite directions on either side of the vacuum chamber. This requires a longitudinal electric field $E_{//}$ which varies in strength across the aperture and a transverse dipole magnetic field B_z as shown in Fig. 25. This magnetic field B_z deflects the beam. It can increase the initial displacement (instability) or decrease it (stability).

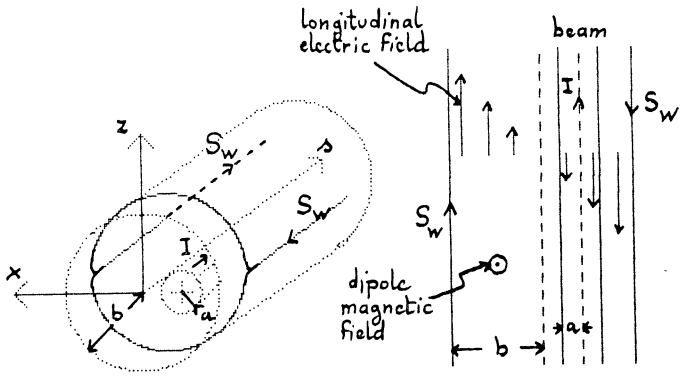


Fig. 25 Beam transversely oscillating in a conducting pipe

Indeed, we could have considered higher-order defects induced by an on-axis beam the transverse dimensions of which change periodically with time (throbbing-beam modes).

In the present lecture, we will restrict ourselves to the study of the effect of the dipolar magnetic field. This is the simplest transverse interaction between the beam and the surroundings.

3.2 Single particle motion

In order to describe the transverse motion of particles in a bunch, one needs four coordinates. Two of them are related to the projection on the longitudinal phase space. We will use again

(τ , $\dot{\tau}$) or (ψ_0 , $\hat{\tau}$) (see section B.1)

$$\tau = \hat{\tau} \cos(\omega_s t + \psi_0) \quad . \tag{140}$$

The second pair of coordinates
$(x , \dot{x})$ or $(\varphi , \hat{x})$

corresponds to the projection on one of the two transverse phase spaces (horizontal for instance).

The solution of the equation of unperturbed motion in the transverse plane, will be written

$$x(t) = \hat{x} \cos \varphi(t) \quad . \tag{141}$$

A very important point is that the betatron phase advance per unit of time or betatron frequency

$$\dot{\varphi} = Q \omega \tag{142}$$

depends on the instantaneous momentum deviation of the particle in first order. It can be expanded as follows

$$\dot{\varphi} = Q_0 \left(1 - \xi \frac{dp}{p}\right) \omega_0 \left(1 - \eta \frac{dp}{p}\right) = Q_0 \omega_0 \left(1 - \dot{\tau}\left(1 - \frac{\xi}{\eta}\right)\right) \quad . \tag{143}$$

Q_0 and ω_0 are the transverse wavenumber and the revolution frequency evaluated at synchronous energy.

$$\xi = \frac{dQ}{Q} \Big/ \frac{dp}{p} \quad \text{is the machine chromaticity} \tag{144}$$

$\eta = -d\omega/\omega$ has been defined in section 2.1 already.

According to the above definitions, the transverse coordinate satisfies

$$\ddot{x} - \frac{\ddot{\varphi}}{\dot{\varphi}} \dot{x} + \dot{\varphi}^2 x = 0 \quad . \tag{145}$$

The second order, Q and ω depend on transverse amplitudes and momentum spread. In the following, we will neglect this dependence. Furthermore, we will assume that the motion in longitudinal plan is stable (no coherent effect).

Here again, we will be mainly interested in the electromagnetic field induced by the beam. Taking this field into account will modify the differential equation of motion (145) of the single particle.

$$\ddot{x} - \frac{\ddot{\varphi}}{\dot{\varphi}} \dot{x} + \dot{\varphi}^2 x = \frac{e}{m_0 \gamma} \left[\vec{E} + \vec{\beta c} \wedge \vec{B} \right]_{\perp} \left(t, \Theta = \omega_0 (t - \tau) \right) \quad . \tag{146}$$

3.3 Single particle signal

The transverse displacement of a beam can be detected by means of transverse P U electrodes. These are diagnostic equipments allowing us to measure the local beam center of mass position. The signal drawn from an electrode is the product of the longitudinal signal times the transverse position.

Let us assume a perfect P U electrode at angular position Θ in the ring and let us observe the transverse signal $s_{\perp}(t,\theta)$ induced by an off-centered single particle.

By definition we get

$$s_{\perp}(t,\theta) = s_{//}(t,\Theta)\, x(t) = s_{//}(t,\theta)\, \hat{x} \cos \varphi(t) \quad . \tag{147}$$

In the time domain, $s_\perp(t,\theta)$ consists of a series of impulses, the amplitude of which $X(t)$ changes at each passage through the electrode. Integrating equation (143) to get

$$\varphi(t) = Q_0 \omega_0 (t-\tau) + \omega_\xi \tau + \varphi_0 \tag{148}$$

with $\omega_\xi = Q_0 \omega_0 \frac{\xi}{\eta}$ and $\varphi_0 = \varphi(t)$ at time $t = 0$ (149)

and using equation (9) to express $s_{//}(t,\theta)$, one can successively obtain

$$s_\perp(t,\theta) = e\hat{X} \cos\varphi(t) \sum_k \delta\left(t - \tau - \frac{\theta}{\omega_0} - \frac{2k\pi}{\omega_0}\right) \tag{150}$$

$$s_\perp(t,\theta) = \frac{e\omega_0}{2\pi} \hat{x} \frac{(e^{j\varphi(t)} + e^{-j\varphi(t)})}{2} \sum_p e^{jp(\omega_0(t-\tau)-\theta)} \tag{151}$$

when using equations (10) and (12)
and finally

$$s_\perp(t,\theta) = \frac{e\omega_0}{4\pi} \hat{X} e^{j(Q_0\omega_0 t + \varphi_0)} \sum_{m,p} j^{-m} J_m\left(((p+Q_0)\omega_0 - \omega_\xi)\hat{\tau}\right) e^{j(\omega_{pm} t + m\psi_0 - p\theta)} + C.C. \tag{152}$$

with $\omega_{pm} = p\omega_0 + m\omega_s$. (14)

With the Fourier transform we can pass on to frequency domain.

$$s_\perp(\omega,\theta) = \frac{e\omega_0}{4\pi} \hat{X} e^{j\varphi_0} \sum_{m,p} j^{-m} J_m\left(((p+Q_0)\omega_0 - \omega_\xi)\hat{\tau}\right) \delta\left(\omega - ((p+Q_0)\omega_0 + m\omega_s)\right) e^{j(m\psi_0 - p\theta)} + C.C. \tag{153}$$

The single particle spectrum is a line spectrum at frequencies $(p+Q_0)\omega_0 + m\omega_s$. Around every betatron line $(p+Q_0)\omega_0$ there is an infinite number of synchrotron satellites the amplitude of which is given by the Bessel function $J_m(((p+Q_0)\omega_0 - \omega_\xi)\hat{\tau})$. It is important to notice that the spectrum is centered around ω_ξ . This is one of the fundamental differences between transverse and longitudinal cases. For standard machines, the uncorrected chromaticity ξ is negative. Therefore, according to definition (149), ω_ξ is a negative frequency above transition energy and a positive frequency below.

3.4 Distribution of particles

The next step consists of gathering particles to form a bunch. To this end, we introduce a distribution function $\Psi(\psi, \hat{\tau}, \varphi, \hat{x}, t)$ that is split into two different parts, a stationary distribution and a perturbation.

3.4.1 Stationary distribution

In the absence of perturbation $\hat{x}$ and $\hat{\tau}$ are constant during the motion. Therefore the stationary part is a realistic function of the two peak amplitudes $\Psi_0(\hat{x},\hat{\tau})$. We will assume no correlation between transverse and longitudinal planes and we will write the stationary part as the product

$$\Psi_0 = g_0(\hat{\tau})\, f_0(\hat{x}) \tag{154}$$

of two stationary distributions, one for each phase space.

$g_o(\hat{\tau})$ and $f_o(\hat{x})$ can be one of the functions already suggested in the previous section ((25) through (28)) for instance

$$\int g_o(\hat{\tau})\,\hat{\tau}\,d\hat{\tau} = \frac{1}{2\pi} \qquad \int f_o(\hat{x})\,\hat{x}\,d\hat{x} = \frac{1}{2\pi} \quad . \tag{155}$$

In the following, our numerical examples will assume a water bag distribution (28) in the longitudinal plane.

Since on an average, the beam center of mass is on axis, the transverse signal as well as the transverse dipole magnetic field induced by the stationary distribution are null.

$$S_\perp(t,\theta) = N \int s_\perp(t,\theta)\,\Psi_o(\hat{\tau},\hat{x})\,\hat{\tau}\,\hat{x}\,d\hat{\tau}\,d\hat{x}\,d\varphi_o\,d\psi_o = 0 \tag{156}$$

N is the number of particles per bunch.

Indeed, because of the current, there is some space charge quadrupolar magnetic field that adds to the external guide field and that changes transverse focusing. The incoherent Q value is different from Q_o (from now on, we will assume that the effective transverse tune is Q, and we will change Q_o into Q in the equations). Nevertheless this is not the type of field we are concerned with in this section.

3.4.2 Perturbation

In order to get some dipole field, we introduce a density perturbation $\Delta\Psi$ that simulates a beam center-of-mass displacement along the bunch. The mathematical form of the perturbation is suggested by the single particle signal. As a matter of fact, because of the integral over φ_o and ψ_o, the signal induced would be null unless one introduces the complex conjugate of $\exp j(\varphi_o + m\psi_o)$ in the expression of $\Delta\Psi$.

So, in considering a single value of m, let us write the perturbation as follows :

$$\Delta\Psi = h_m(\hat{\tau},\hat{x})\; e^{-j(\varphi_o+m\psi_o)}\; e^{j\Delta\omega_{cm}t} \tag{157}$$

with $\Delta\omega_{cm} = \omega_c - m\omega_s$.

Then, in the time domain, the signal takes the following form:

$$S_\perp(t,\theta) = \frac{e\omega_o}{4\pi} N \int \hat{x} \sum_p j^{-m} J_m\left[((p+Q)\omega_o - \omega_\xi)\hat{\tau}\right] h_m(\hat{\tau},\hat{x})\, e^{-jp\theta}\, e^{j((p+Q)\omega_o+\omega_\xi)t}\, \hat{\tau}\,\hat{x}\, d\hat{\tau}d\hat{x}\, d\varphi_o d\psi_o \quad . \tag{158}$$

In the frequency domain,

$$S_\perp(\omega,\theta) = \frac{4\pi^2 I}{\ell} \sum_p e^{-jp\theta}\,\sigma_m(p)\,\delta\left(\omega - (\omega_c + (p+Q)\omega_o)\right) \tag{159}$$

with

$$\sigma_m(p) = j^{-m} \int h_m(\hat{\tau},\hat{x})\, J_m\left[((p+Q)\omega_o - \omega_\xi)\hat{\tau}\right] \hat{x}^2\hat{\tau}\, d\hat{x}\, d\hat{\tau} \tag{160}$$

we obtain a line spectrum at frequencies $\omega = (p+Q)\omega_o + m\omega_s$. In comparison with the rich spectrum of the single particle, a single synchrotron satellite remains. The perturbation is coherent with respect to satellite number m .

By means of the perturbation, we have arranged the transverse initial conditions of the particles in the bunch. Particles on a given synchrotron orbit $\hat{\tau}$ have the same peak betatron amplitude $\hat{x}$. This is included in $h_m(\hat{\tau},\hat{x})$. Furthermore, their betatron phase and their synchrotron phase are chosen in order to satisfy

$$\varphi_o + m\psi_o = \text{constant} \quad .$$

The result of this perturbation is that the position of the center of mass changes along the bunch.

We have also assumed that, in phase space, the distribution does not rotate at incoherent frequency $Q\omega_0 + m\omega_s$ exactly but at frequency $Q\omega_0 + \omega_c$. The goal we aim at consists in finding out the values of the coherent betatron frequency shift $\Delta\omega_{cm}$. Its imaginary part will tell us whether the perturbation will increase or will be damped.

We will come back to the physical meaning of $h_m(\hat{z},\hat{x})$ when we describe the coherent modes of oscillation later on.

As was the case for longitudinal instabilities, we will distinguish two regimes.

At low intensity, the coherent betatron frequency shift will remain small when compared to the incoherent synchrotron frequency ω_s .

$$\Delta\omega_{cm} << \omega_s \quad . \tag{161}$$

Therefore, two adjacent perturbations associated with two successive values of m oscillate independently. One can study them separately (equation (157)).

On the other hand, for increasing current, the frequency shifts get larger and it is necessary to sum up several elementary perturbations

$$\Delta\Psi = \sum_m h_m(\hat{z},\hat{x})\, e^{-j(\varphi_0 + m\psi_0)}\, e^{j(\omega_c - m\omega_s)t} \quad . \tag{162}$$

3.5 Transverse coupling impedance

Because of the initial slight displacement of the beam which oscillates from side to side in the external focusing field, we get a differential current which creates a transverse dipole field. This field perturbs the particle motion.

3.5.1 Definition of the transverse coupling impedance

As previously in the case of the longitudinal plane, we introduce a key parameter of the machine, the transverse coupling impedance $Z_\perp(\omega)$ which gathers all the characteristics of the electromagnetic response of the machine to a passing particle. The coupling impedance allows us to express the transverse field in terms of transverse signal.

$$\left[\vec{E} + \vec{\beta c} \wedge \vec{B}\right]_\perp (t,\theta) = \frac{-j\beta}{2\pi R} \int Z_\perp(\omega)\, S_\perp(\omega,\theta)\, e^{j\omega t}\, d\omega \tag{163}$$

$Z_\perp(\omega)$ is expressed in ohm/m.

3.5.2 Transverse coupling impedance of a circular machine

The main components of the transverse impedance of a standard circular machine are listed hereunder and sketched in Fig. 26.

a) resistive wall component
Peaked at low frequencies, it is the principal source of transverse instabilities. When assuming a thick wall vacuum chamber, the impedance can be written

$$Z_{\perp\,R.W.}(\omega) = (1+j)\,\frac{R}{b^3}\,Z_0\,\delta_0^*\left(\frac{\omega_0}{\omega}\right)^{1/2} \tag{164}$$

with the same notations as in equation (49).

b) parasitic resonators
High Q resonances can be found in RF cavities, septum tanks and kicker tanks.

c) broad band impedance

The broad band component of the impedance takes into account the numerous cross section changes of the vacuum chamber. It can be roughly derived from the longitudinal broad-band impedance (equations (51) through (54)) as follows.

$$Z_{\perp\,B.B.}(\omega) = \frac{2c}{b^2}\,\frac{Z_{/\!/}(\omega)}{\omega} \quad . \tag{165}$$

In fact, the above convenient relationship between $Z_\perp(\omega)$ and $Z_{/\!/}(\omega)/\omega$ is strictly valid for the resistive wake in a round pipe, at frequencies well below the pipe cut-off frequency. It is a quiet good approximation as far as the broad-band impedance is concerned. It does not apply for other components of a machine impedance.

Because of equation (165), $Z_\perp(\omega)$ has the same form as $Z_{/\!/}(\omega)/\omega$ (see Fig. 26). It is interesting to notice that the transverse BB impedance varies like b^{-2}. In other words, it gets larger when the pipe radius is reduced.

d) space charge component

When a round beam of radius a circulates in a round pipe of radius b, the space charge component is given by

$$Z_{\perp\,S.C.}(\omega) = -\frac{jRZ_0}{\beta_0^2\gamma_0^2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right) \quad . \tag{166}$$

It is a negative inductance that can be large for low β particles.

The actual impedance seen by the beam is the sum of the components listed above.

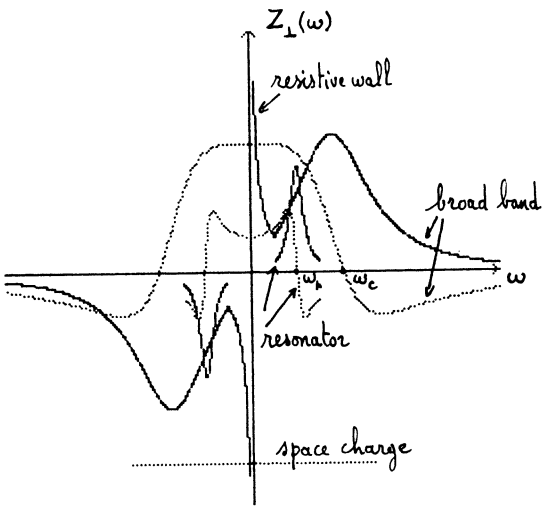


Fig. 26 Qualitative description of transverse impedances for various components (full line = real part, dotted line = imaginary part)

3.6 Equation of coherent motion

We consider a distribution function that sums up a stationary part Ψ_0 (equation (154)) and a perturbation $\Delta\Psi$ (equation (162)).

$$\Psi = g_0(\hat{z})\,f_0(\hat{x}) + \sum_m h_m(\hat{z},\hat{x})\,e^{-j(\varphi_0+m\psi_z)}\,e^{j(\omega_c - m\omega_s)t} \quad . \tag{167}$$

Then, we use Vlasov's equation

$$\frac{\partial \Psi}{\partial t} + \frac{\partial \Psi}{\partial \hat{\tau}}\dot{\hat{\tau}} + \frac{\partial \Psi}{\partial \psi_0}\dot{\psi}_0 + \frac{\partial \Psi}{\partial \hat{x}}\dot{\hat{x}} + \frac{\partial \Psi}{\partial \varphi_0}\dot{\varphi}_0 = 0 \qquad (168)$$

and we drop second order terms with respect to the perturbation in order to end up with

$$j \sum_m (\omega_c - m\omega_s)\, h_m(\hat{\tau},\hat{x})\, e^{-j(\varphi_0 + m\psi_0)}\, e^{j(\omega_c - m\omega_s)t} = -g_0(\hat{\tau})\frac{\partial f_0}{\partial \hat{x}}\dot{\hat{x}} \quad . \qquad (169)$$

In the right hand side of the linearized Vlasov's equation, it is $g_0(\hat{\tau})$ but not $\partial g_0/\partial \hat{\tau}$ that appears. This is the reason why a water-bag bunch can be assumed in the present case.

The expression of $\dot{\hat{x}}$ can be drawn from single particle equation of motion (146).

$$\dot{\hat{x}} = -\frac{\sin\varphi}{\dot{\varphi}} F_c = -\frac{\sin\varphi}{\dot{\varphi}} \frac{e}{m_0 \gamma} \left[\vec{E} + \beta c \wedge \vec{B}\right]_{\perp}(t, \theta = \omega_0(t-\tau)) \quad . \qquad (170)$$

In the right hand side, the coherent transverse electromagnetic "force" F_c can be written

$$F_c = -\frac{j\beta e}{2\pi R m_0 \gamma} \sum_{m,p} Z_{\perp}(p)\, \sigma_m(p)\, e^{-jp\omega_0(t-\tau)}\, e^{j((p+Q)\omega_0 + \omega_c)t} \qquad (171)$$

when equations (163), (159) and (160) are used.

After F_c has been combined with $\sin\varphi$ and $\exp j((p+Q)\omega_0 - \omega_\xi)\tau$ has been developped according to equation (12), one gets

$$\dot{\hat{x}} = -\frac{e\pi I}{2 m_0 c \gamma Q} \sum_{m,p} Z_{\perp}(p)\, \sigma_m(p)\, j^m J_m\left[((p+Q)\omega_0 - \omega_\xi)\hat{\tau}\right] e^{-j(\varphi_0 + m\psi_0)}\, e^{j(\omega_c - m\omega_s)t} \quad . \qquad (172)$$

This allows us to rewrite equation (169)

$$j\, h_m(\hat{\tau},\hat{x})(\omega_c - m\omega_s) = \frac{e\pi I}{2 m_0 c \gamma Q} \sum_p Z_{\perp}(p)\, \sigma(p)\, j^m J_m\left[((p+Q)\omega_0 - \omega_\xi)\hat{\tau}\right] g_0(\hat{\tau}) \frac{\partial f_0}{\partial \hat{x}} \qquad (173)$$

where $$\sigma(p) = \sum_m \sigma_m(p) \qquad (174)$$

is the amplitude of the spectrum of the signal induced by the perturbation at frequency $(p+Q)\omega_0 + \omega_c$.

Both sides of equation (173) can be multiplied by $\hat{x}$ and integrated over $\hat{x}$ values. In the right hand side the integral is easy to perform

$$\int_0^\infty \frac{\partial f_0}{\partial \hat{x}} \hat{x}^2 d\hat{x} = -2\int_0^\infty f_0(\hat{x})\, \hat{x}\, d\hat{x} = \frac{-1}{\pi} \quad . \qquad (175)$$

It leads to

$$j(\omega_c - m\omega_s)\int_0^\infty h_m(\hat{\tau},\hat{x})\,\hat{x}^2\,d\hat{x} = \frac{-eI}{2m_0c\gamma Q}\sum_p Z_\perp(p)\,\sigma(p)\,j^m J_m\left[\left((p+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau}) \quad . \quad (176)$$

It is important to notice that the transverse stationary function $f_0(\hat{x})$ has disappeared. Only the longitudinal function $g_0(\hat{\tau})$ remains in the equation. The problem we are dealing with concerns the beam center of mass essentially. On an average, the detailed betatron motion of individual particles around the center of mass can be disregarded.

The same comments must apply to the perturbation $h_m(\hat{\tau},\hat{x})$. As a matter of fact, the quantity of interest is the averaged peak betatron amplitude $\hat{X}_m(\hat{\tau})$ associated with a given synchrotron orbit $\hat{\tau}$.

$\hat{X}_m(\hat{\tau})$ can be defined as follows.

$$g_0(\hat{\tau})\,\hat{X}_m(\hat{\tau}) = \int_0^\infty h_m(\hat{\tau},\hat{x})\,\hat{x}^2\,d\hat{x} \quad . \quad (177)$$

The final form of the equation of coherent motion of a single bunch is

$$j(\omega_c - m\omega_s)\,\hat{X}_m(\hat{\tau}) = \frac{-eI}{2m_0c\gamma Q}\sum_p Z_\perp(p)\,\sigma(p)\,j^m J_m\left[\left((p+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] \quad . \quad (178)$$

From equation (160), the new form of $\sigma_m(p)$ (see equation (174)) is

$$\sigma_m(p) = j^{-m}\int_0^\infty J_m\left[\left((p+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau})\,\hat{X}_m(\hat{\tau})\,\hat{\tau}\,d\hat{\tau} \quad (179)$$

For transverse motion, equation (178) replaces equation (84).

3.7 Coherent modes of oscillation

The general properties of the solutions as well as the different ways to solve equation (178) have been already described in section 2.6. We will develop the eigenvalue method for instance.

Then, both sides of equation (178) are multiplied by $j^{-m} J_m\left[\left((l+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau})\,\hat{\tau}\,d\hat{\tau}$ and integrated over $\hat{\tau}$ values.

$$j(\omega_c - m\omega_s)\,\sigma_m(l) = \frac{-eI}{2m_0c\gamma Q}\sum_p Z_\perp(p)\,\sigma(p)\int_0^\infty J_m\left[\left((p+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] J_m\left[\left((l+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau})\,\hat{\tau}\,d\hat{\tau} \quad . \quad (180)$$

3.8 Low intensity coherent modes of oscillation

We are assuming a low intensity bunch interacting with the different components of the impedance of a ring. A single value of m is retained in the perturbation (157).

We define the matrix element

$$K_{lp}^m = \frac{eI}{2m_0c\gamma Q}\left(j\,Z_\perp(p)\right)\int_0^\infty J_m\left[\left((l+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] J_m\left[\left((p+Q)\omega_0-\omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau})\,\hat{\tau}\,d\hat{\tau} \quad (181)$$

and rewrite equation (180):

$$\left(\omega_{c_m} - m\omega_s\right) \sigma_m(l) = \sum_p K^m_{lp}\, \sigma_m(p) \quad . \tag{182}$$

When comparing the K^m_{lp}'s associated with transverse (181) and longitudinal (103) coherent motions, one can draw two conclusions.

First, the value $m = o$ is allowed now.

Second, in the integral $\partial g_o / \partial \hat{z}$ is replaced by $g_o(\hat{z})$, which means that in the transverse case the water-bag bunch (28) leads to results very similar to the ones obtained for the bunch with parabolic amplitude (25) in the longitudinal case. This is the main reason for choosing these two particular distributions.

3.8.1 Local interaction space charge or inductive walls

In this example $\int Z_\perp(p)$ is real and constant.

Most of the comments already made in section 2.7.1 apply again. Results for the water bag bunch are shown in Fig. 27. At the top, the spectra of the low order coherent modes mq are centered around ω_ξ; below are the corresponding PU signals (several passings superimposed) for a finite chromaticity ($\omega_\xi \tau_L = 3\pi/4$) and for zero chromaticity. At the bottom the coherent peak betatron amplitude (equation (178)) as a function of $\hat{z}$.

The coherent frequency shifts can be written

$$\left(\omega_{c_{mq}} - m\omega_s\right) = \frac{eI}{4\pi m_o c \gamma Q B}\, C_{mq} \left(\int Z_\perp(\omega_\xi)\right) \quad . \tag{183}$$

The values of the C_{mq} are listed in Fig. 27. The largest one corresponds to mode oo ($C_{oo} = 1.09$, dipole mode) that consists of a quite rigid oscillation with the same amplitude $\hat{x}$ for all the in the bunch. The quadrupole mode $m=q=1$ ($C_{11} = .46$) is about two times less sensitive. This is due to the fact that particles at bunch center have no coherent response $\hat{x}_{11}(o) = o$ and do not contribute to the motion.

Because of the special form of the impedance, the shifts are real and motion is stable. A resistance is needed to drive an instability.

Equation (183) is not strictly limited to space charge or inductive walls. It applies to any type of interaction (reactance plus resistance) provided the impedance is sufficiently smooth over the mode spectrum.

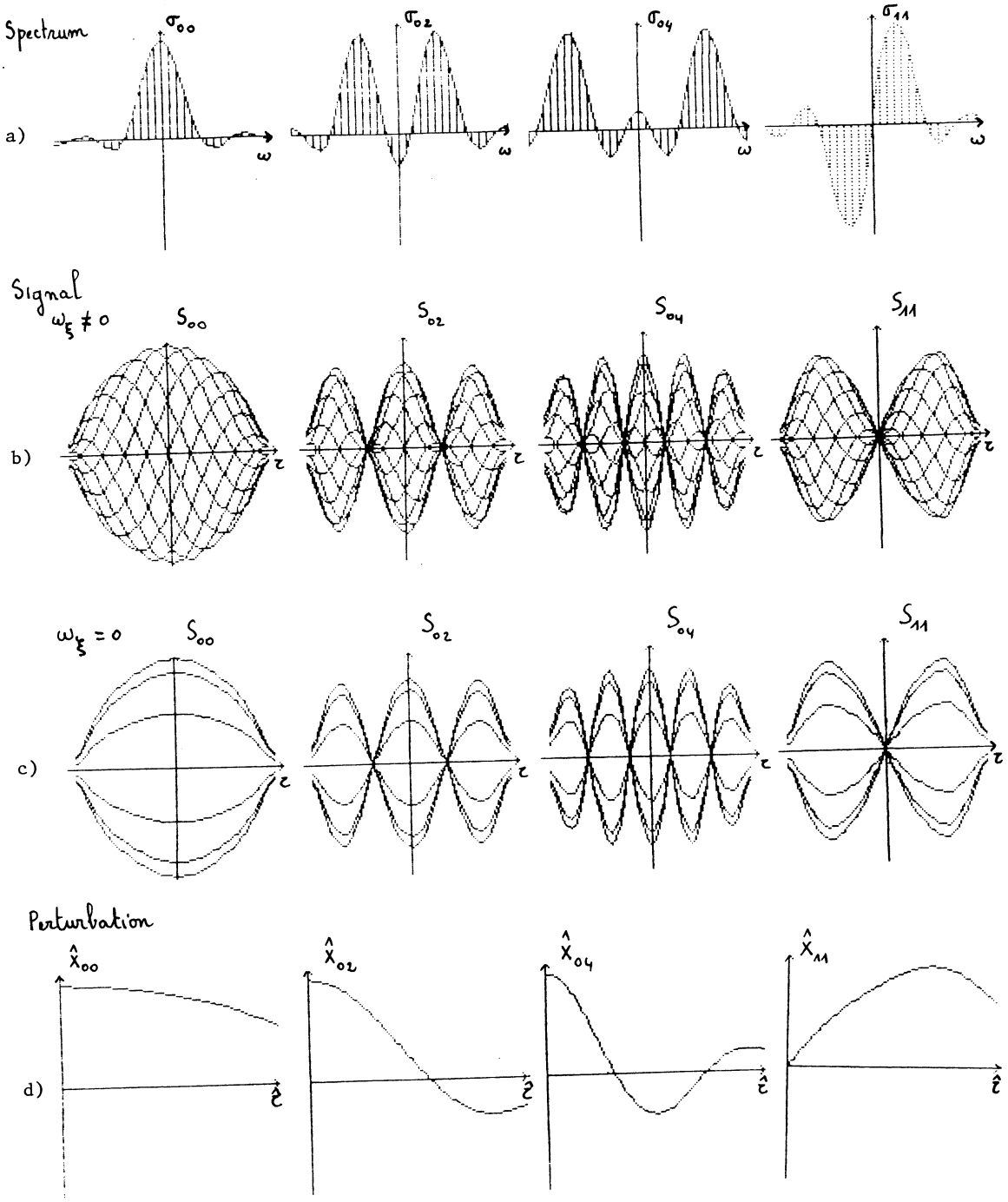


Fig. 27 Transverse oscillations with mode numbers m,n for a parabolic bunch in an inductive impedance; a) spectrum, b) and c) signals with and without chromaticity respectively, and d) perturbation amplitude

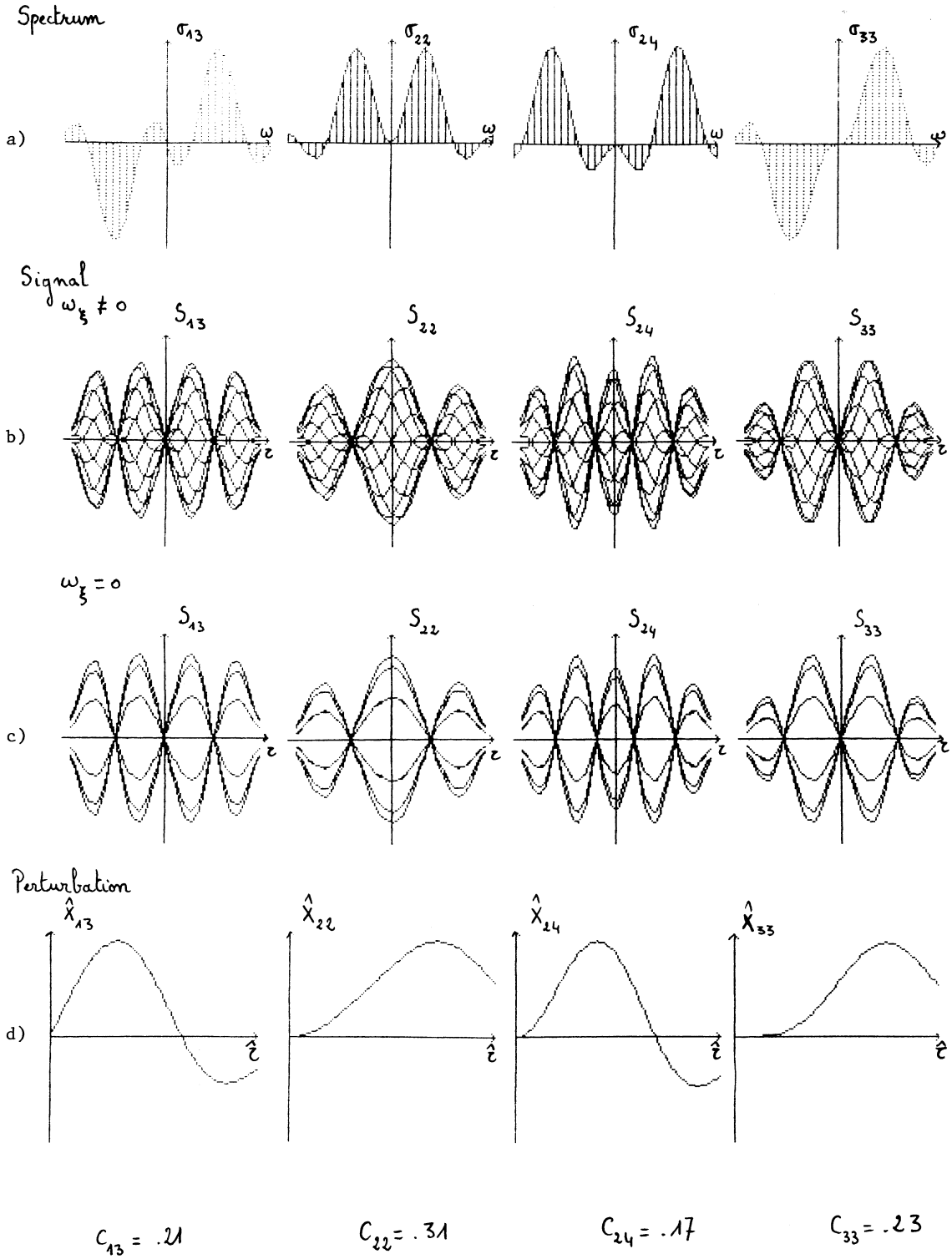


Fig. 27 (continued)

In Fig. 28, we have sketched an example of stable interaction for which equation (183) is valid. The spectrum of mode $o\,o$ is drawn with a rather large and positive value of ω_ξ ($\omega_\xi \gg 2\pi/\tau_L$) so that it does not overlap the low frequency region where the resistive wall impedance is peaked. We are above transition with $\xi > 0$ or below transition with $\xi < 0$.

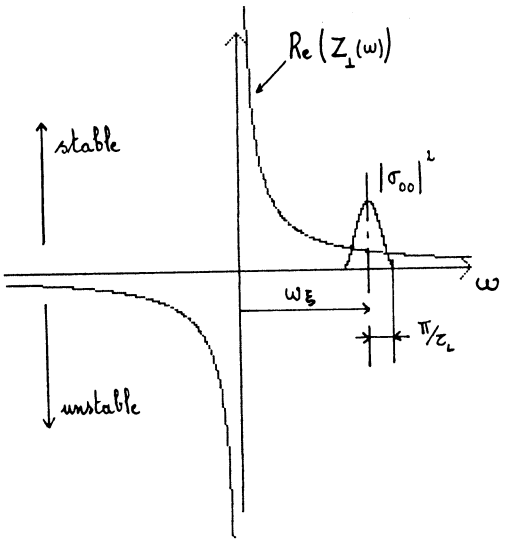


Fig. 28 Stabilisation of resistive-wall instability by chromaticity (for m = 0 mode)

A negative ω_ξ would lead to instability.

3.8.2 Parasitic high Q resonators - coupled bunch modes

Let us imagine a very narrow band object, so that the motion is driven by a single line at $(p+Q)\omega_o + \omega_c$. Then, the matrix equation reduces to

$$\left(\omega_c - m\omega_s\right) \sigma_m(l) = K^m_{lp}\ \sigma_m(p) \quad . \tag{184}$$

The sum over p values has disappeared. For every m , a single mode is a solution of equation (184):

coherent frequency

$$\omega_{c_m} = m\omega_s + \frac{eI}{2m_o c\gamma Q}\left(j Z_\perp(p)\right) \int_0^\infty J_m^2\left[\left((p+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] g_o(\hat{\tau})\ \hat{\tau}\ d\hat{\tau} \quad , \tag{185}$$

spectrum

$$\sigma_m(p) = 1 \tag{186}$$

$$\sigma_m(l) = \frac{\int_0^\infty J_m\left[\left((l+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] J_m\left[\left((p+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] g_o(\hat{\tau})\ \hat{\tau}\ d\hat{\tau}}{\int_0^\infty J_m^2\left[\left((p+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] g_o(\hat{\tau})\ \hat{\tau}\ d\hat{\tau}} \quad , \tag{187}$$

perturbation (see equation (178))

$$\hat{x}_m(\hat{\tau}) \propto J_m\left[\left((p+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] \quad . \tag{188}$$

A resistance is needed in order to get an imaginary frequency shift

$$I_m(\omega_c) \propto I \; R_e\left(Z_\perp(p)\right) \int_0^\infty J_m^2\left[\left((p+Q)\omega_o - \omega_\xi\right)\hat{\tau}\right] g_o(\hat{\tau})\, \hat{\tau}\, d\hat{\tau} \quad . \tag{189}$$

The integral is always positive. As a consequence a negative resistance (negative frequency region) leads to unstable motion.

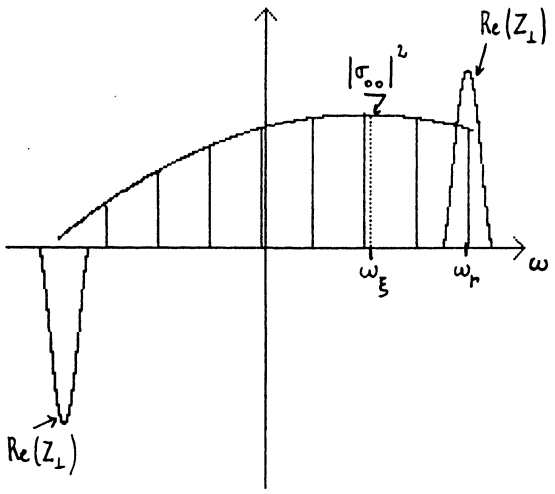


Fig. 29 Stabilisation of transverse instability, due to narrow-band impedance, by chromaticity (for m = 0 mode)

In Fig. 29, the picture of a stable scenario is shown.

Let ω_r be the resonant frequency, then, for the water-bag bunch, an equivalent form of equation (185) is

$$\left(\omega_{c_m} - m\omega_s\right) = \frac{1}{(|m|+1)} \frac{eI}{4\pi m_o c \gamma Q} \left(j Z_\perp(p)\right) F_m\left(\left(\omega_\xi - (p+Q)\omega_o - m\omega_s\right)\frac{\tau_\ell}{2}\right) \quad . \tag{190}$$

The form factor F_m is plotted in Fig. 30

$$F_m(x) = \frac{2(|m|+1)}{x^2} \int_0^x J_m^2(u)\, u\, du \tag{191}$$

Up to now, a single bunch has been assumed. As shown previously in the longitudinal case, if M equidistant bunches are present in the ring, one can consider M possibilities of coupled bunch coherent motion. With the index n running from 0 to $M-1$, coupled bunch mode n corresponds to a phase shift of $n 2\pi / M$ between the coherent perturbations of two successive bunches.

Every M^{th} line occurs at

$$\omega = \left(n + pM + Q\right)\omega_o + \omega_c \tag{192}$$

but with an amplitude M times larger. As a consequence, the frequency shift (190) must be multiplied by M .

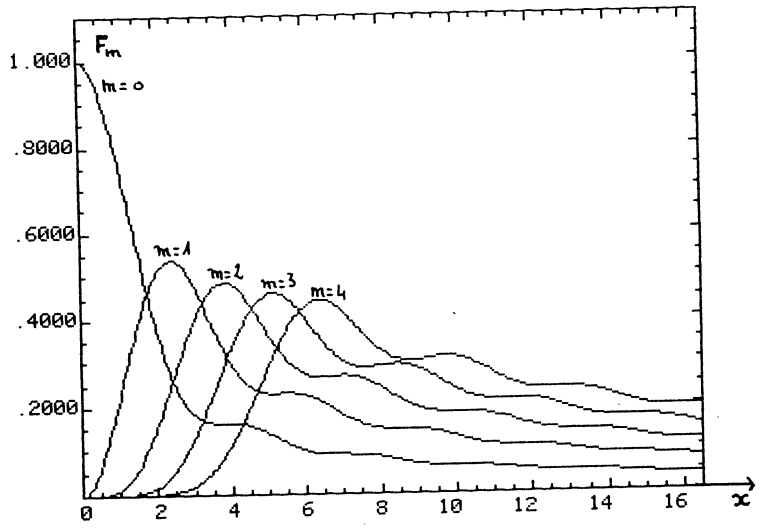


Fig. 30 Form factors F_m for different modes m of transverse-bunch oscillation

3.8.3 Resistive wall component

As we have seen previously, the resistive part of $Z_\perp(p)$ causes instability for negative resistance (negative frequency region) and damping for positive resistance (positive frequency region).

The resistive wall interaction can be split into two parts.

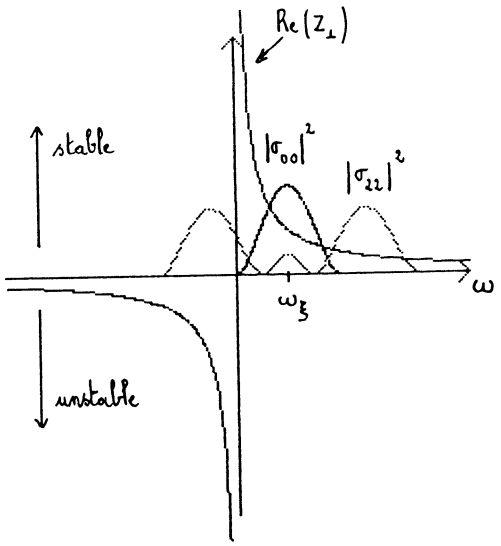


Fig. 31 Stabilisation of resistive-wall instability by chromaticity for modes m = 0 and m = 2

a) First, one can consider the effect of the short range wake (broad band type effect). This is equivalent to replacing the discrete line spectrum by a continuous spectrum. A slightly positive value of ω_ξ is sufficient in order to stabilize mode **oo**, but it can be seen that mode **22** for instance is unstable and the reason is evident in Fig. 31. For ω_ξ sufficiently large, both modes are damped. On the other hand, a more realistic picture of the impedance would take into account the additional damping provided by the broad band resistance. This last contribution is very efficient for stabilizing high-order modes with short bunches. As a conclusion, a slightly positive value of ω_ξ is required to get rid of the short-range wake effect.

b) Then, one can consider the effect of the long-range wake (narrow band type effect). In Fig. 32, mode **oo** is shown with a small ω_ξ value and Q just below an integer. Here a single line, namely the line $\omega_{R.W.} = (p+Q)\omega_0$ nearest the origin in the narrow band region of $Z_\perp(\omega)$, contributes to antidamping. From equation (190), one can calculate the effect.

$$\omega_c = \frac{eI}{4\pi m_0 c \gamma Q} \, j Z_\perp(\omega_{R.W.}) \, F_0\left((\omega_\xi - \omega_{R.W.})\frac{\tau_L}{2}\right) \qquad (193)$$

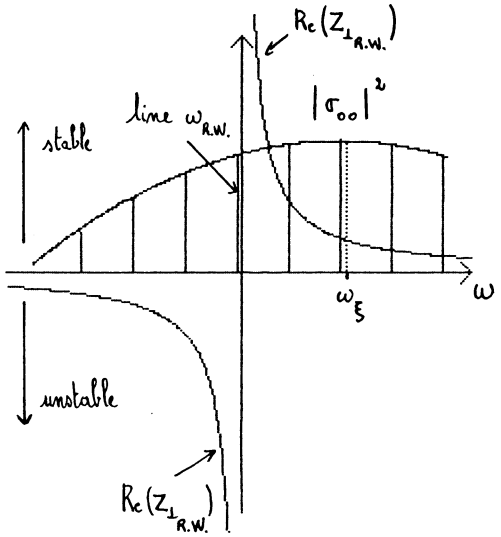


Fig. 32 Long-range effect of resistive-wall impedance for a short bunch

On the other hand, when the transverse tunes are chosen just above an integer, the long range wake has a stabilizing effect. Indeed, for multibunch operation, coupled bunch modes must be considered. The frequency shift (193) is M times larger.

Above transition, the necessity of correcting the chromaticity of the machine (strong sextupoles for sophisticated optics), gives rise to the difficult problem of dynamic aperture.

3.8.4 Stability criterion

A spread in betatron frequencies can Landau damp the instability. Let $\delta S_\perp$ be the within bunch full width half high spread, then, stability requires that

$$\delta S_\perp \geq |\Delta\omega_{c_m}| . \qquad (194)$$

3.9 Single-bunch instability at high intensity

The theory that has been developed in the previous sections is in quite good agreement with experiments. At low intensity, the frequency shift formulae have been extensively checked and the predicted bunch shape oscillations are commonly observed.

However, at this stage, this theory cannot predict the fast single-bunch instability that has been found in some storage rings (in PETRA, Hamburg, first).

In order to explain "Petra instability", and in analogy to the longitudinal "turbulence", R.D. Kohaupt has implemented the theory by considering transverse "mode coupling".

The basic mechanism is of the same type as the one described for the longitudinal case. We consider the most general perturbation (equation (162)) with several values of m. When the intensity gets sufficiently large for the coherent frequencies of two adjacent modes to merge, instability appears.

3.9.1 Matrix equation of single-bunch modes at high intensity

The starting point is equation (180). Both sides are divided by $j(\omega_c - m\omega_s)$ and summed over m values:

$$\sigma(l) = \frac{eI}{2m_0c\gamma Q} \sum_p \left(j Z_\perp(p)\right) \sigma(p) \sum_m \frac{1}{\omega_c - m\omega_s} \int_0^\infty J_m\left[\left((l+Q)\omega_0 - \omega_\xi\right)\hat{\tau}\right] J_m\left[\left((p+Q)\omega_0 - \omega_\xi\right)\hat{\tau}\right] g_0(\hat{\tau})\,\hat{\tau}\,d\hat{\tau} . \quad (195)$$

Let us use matrix notation.

$$\sigma(l) = \varepsilon \sum_p \left(j Z_\perp(p)\right) M_{lp}\, \sigma(p) . \quad (196)$$

When assuming a water-bag bunch, the matrix element M_{lp} has the following form,

$$M_{lp} = 2B \sum_m \frac{1}{\frac{\omega_c}{\omega_s} - m} \int_0^1 J_m\left[\left((l+Q)\omega_0 - \omega_\xi\right)\frac{\tau}{2}u\right] J_m\left[\left((p+Q)\omega_0 - \omega_\xi\right)\frac{\tau}{2}u\right] u\,du \quad (197)$$

while the intensity parameter ε is given by

$$\varepsilon = \frac{eI}{4\pi m_0 c \gamma Q B \omega_s} . \quad (198)$$

In order to solve equation (196) one can proceed as follows :

Assume a real coherent betatron frequency shift measured in incoherent synchrotron frequency unit, ω_c/ω_s.

Look for the eigenvalues of the matrix

$$\left[j Z_\perp(p)\right] \cdot \left[M_{lp}\right] \quad (199)$$

where $\left[j Z_\perp(p)\right]$ is the diagonal matrix of the impedance.

Scale the intensity parameter ε in order to adjust the eigenvalue to unity.

In the next two sections, we apply this method to two types of impedances, inductive wall or space charge, and broad band resonator.

3.9.2 Space-charge and inductive-wall modes at high intensity

Since there is no resistance in the assumed impedance, on one hand the frequency shifts are real (stability), on the other hand, odd and even perturbations cannot couple as already explained in section 2.8.1.

Results are gathered in Fig. 33 with ω_c / ω_s along the vertical axis and $-\varepsilon \left| Z_\perp(p) \right|$ along the horizontal axis. With such a convention, inductive wall (space charge) corresponds to positive (negative) abscissas.

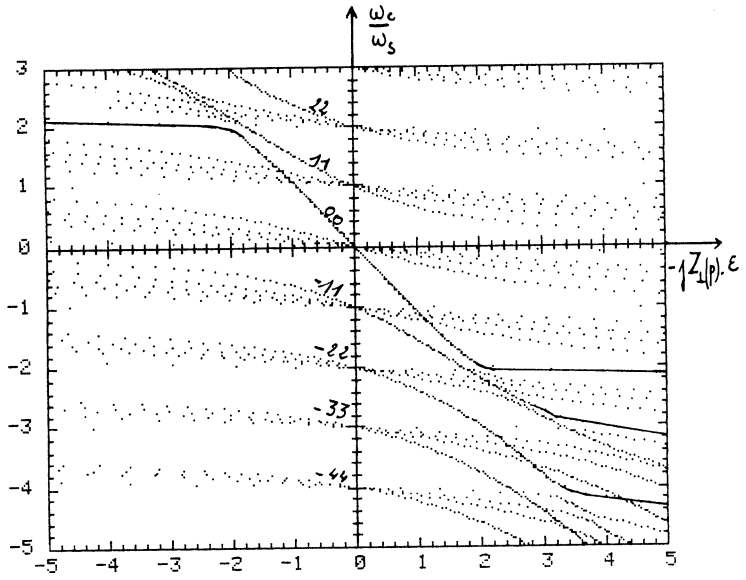


Fig. 33 Transverse coherent-frequency shift versus intensity parameter for a long bunch in an inductive impedance

At low intensity, the infinity of modes with the same m value cluster near the unperturbed frequency $m \omega_s$. The slope of the different curves is directly related to the C_{mq} 's (equation (183)).

For increasing current, let us follow continuously the curve corresponding to mode $o\,o$. Till ω_{co}/ω_s reaches ± 2 , the frequency shift is proportional to the current. Then, at higher current, because of coupling with modes of the same parity, $m = 2$ or $m = -2$, the coherent frequency remains constant.

The same qualitative remarks are valid for mode 11 when considering a space charge impedance. On the contrary, when considering inductive wall, a very non linear dependence with intensity is observed. The effective high intensity frequency shift is much less than expected when extrapolating low intensity formula.

3.9.3 Broad band impedance

We assume a broad band transverse impedance

$$Z_\perp(\omega) = \frac{R_\perp \frac{\omega_r}{\omega}}{1 + j Q^* \left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega} \right)} \qquad \left(\text{ohm/m} \right) \tag{200}$$

with $Q^* = 1$, $\omega_r = \frac{c}{b}$ $\left(f_r \sim 1\,\text{GHz} \right)$

and $$R_\perp \sim \frac{2c R_s}{b^2 \omega_r} \tag{201}$$

when (165) is assumed.

Furthermore, we suppose that the chromaticity has been corrected $\omega_\xi = 0$ as was already the case in the previous section.

Then, provided the bunch length is not extremely short ($p_r B = \omega_r \tau_L / 2\pi \gtrsim 1$), the coherent frequency of mode oo is pushed down by the inductive part of the broad-band impedance. A good estimate of $\omega_{c\,oo}$ can be found by using equation (183) and by assuming that the inductance is constant over the spectrum of mode oo.

$$\frac{\omega_{coo}}{\omega_s} = \frac{eI}{4\pi m_o c \gamma Q B \omega_s} C_{oo}\left[j Z_\perp(o)\right] = \varepsilon\, C_{oo}\left[j Z_\perp(o)\right] \qquad (202)$$

with $j Z_\perp(o) = -\frac{R_\perp}{Q^*} = -R_\perp$ and $C_{oo} = 1.1$.

On the other hand, the coherent frequencies of the modes mq with m = -1 and q = 1, 3,5, etc... are pushed up or pushed down depending whether their spectrum, peaked at $\omega \sim (q+1)\pi/\tau_\perp$, overlaps a frequency region where the broad band impedance is mostly capacitive or mostly inductive.

The lowest instability threshold will occur when the coherent frequency of mode oo and the coherent frequency of one of the modes mq (with m = -1 and q = 1,3, 5...) merge. Since the frequency shifts of the modes belonging to the m = -1 family are weak, their coherent frequencies remain clustered near $-\omega_s$ and the lowest threshold can be approximated by

$$\frac{\omega_{coo}}{\omega_s} = -1 .$$

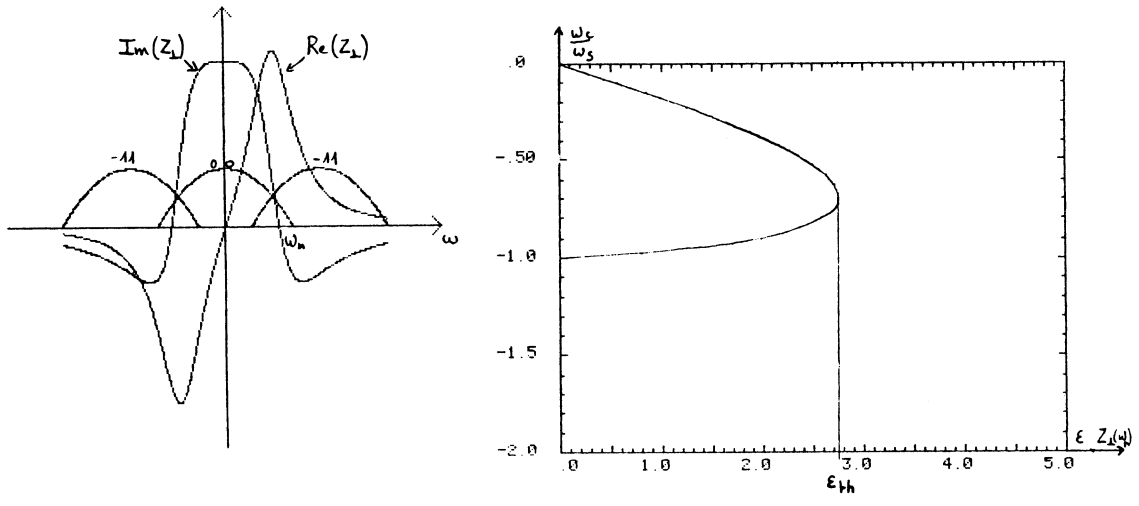
That is to say

$$\varepsilon_{th}\, C_{oo}\, \frac{R_\perp}{Q^*} = 1 \quad \text{or} \quad \varepsilon_{th} Z_\perp(\omega_n) = \frac{Q^*}{C_{oo}} \sim .9$$

or

$$I_{th} = \frac{4\pi B Q \omega_s E/e}{c\, R_\perp \frac{C_{oo}}{Q^*}} = \frac{4\pi B Q \omega_s E/e}{c\, R_\perp} . \qquad (203)$$

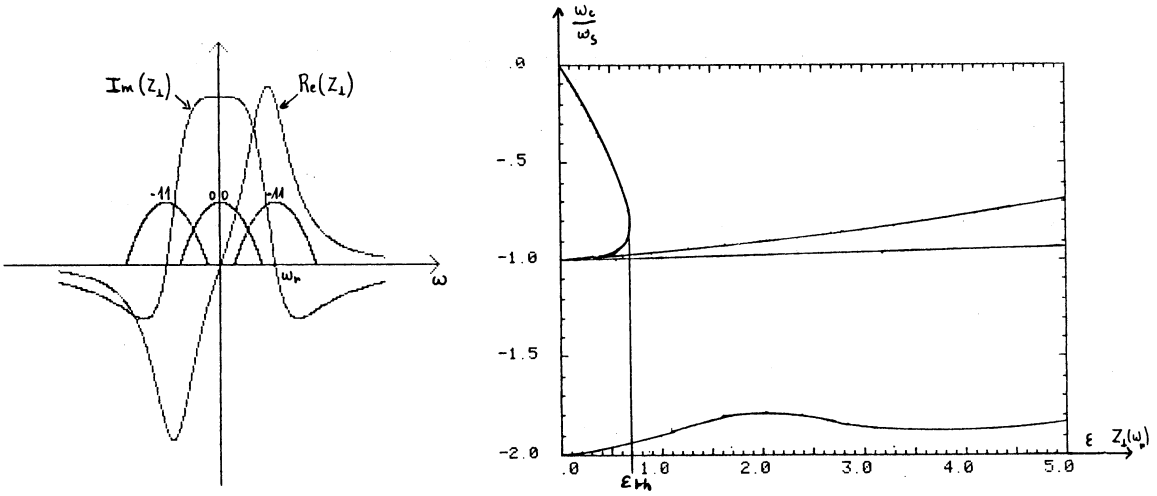
Three examples with $Q^* = 1$ are given hereunder.



a) b)

Fig. 34 Mode coupling for a very short bunch in a broad-band resonator impedance (normalized bunch length $p_r B = 0.2$)

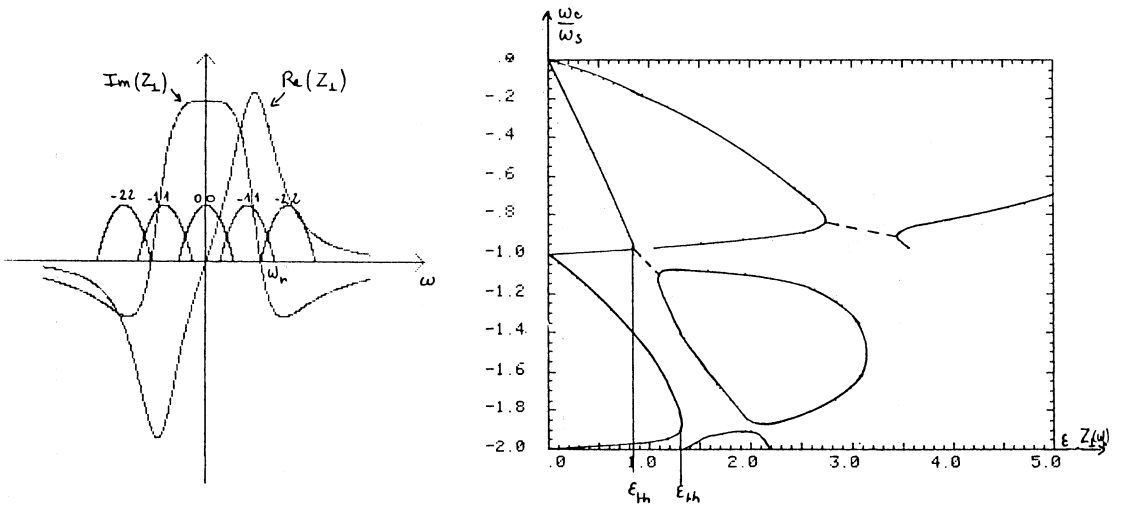
In the first case, we are considering a very short bunch $p_r B = .2$ (Fig. 34a). With respect to the assumption made above, the effective C_{oo} ($C_{oo}^* = .2$) is much smaller than 1.1 because the tails of the frequency spectrum of mode oo extend in the capacitive region. Accordingly, the threshold is higher $\varepsilon_{th} \sim 2.75$ (Fig. 34b). It corresponds to a coupling between mode oo and mode -1 1.



a) b)

Fig. 35 Mode coupling for a bunch in a broad-band resonator impedance (normalized bunch length $p_r B = 0.8$)

In the second case, $p_r B = .8$ the effective C_{oo} is about 1.1 (Figs. 35a and 35b). On the other hand, the effective C_{-11} is about zero. Coupling between modes oo and -1 1 occurs again at $\varepsilon_{th} \sim .7$ this time. This example corresponds to the worst situation. It leads to a threshold somewhat lower than foreseen before simply because the frequency shift due to the resistance has been neglected in our previous estimation.



a) b)

Fig. 36 Mode coupling for a long bunch in a broad-band resonator impedance (normalized bunch length $p_r B = 1.9$)

In the third case (Figs. 36a and 36b), the bunch is long $p_r B = 1.9$, and several thresholds are found. When compared to the previous cases, this time the coherent frequency of mode -1 1 is pushed down towards $2\omega_s$ and mode -11 does not couple with mode oo but with mode -22 at $\varepsilon_{th} \sim 1.3$. This coupling between the two most coherent modes of two successive families m = -1 and m = -2 is strong as could be seen when considering the imaginary part of ω_c . A lower threshold is found at $\varepsilon_{th} \sim .8$. It corresponds to the coupling between mode oo and mode -13. The instability is rather weak since there is little overlap between these two modes in the region where the resistance is large.

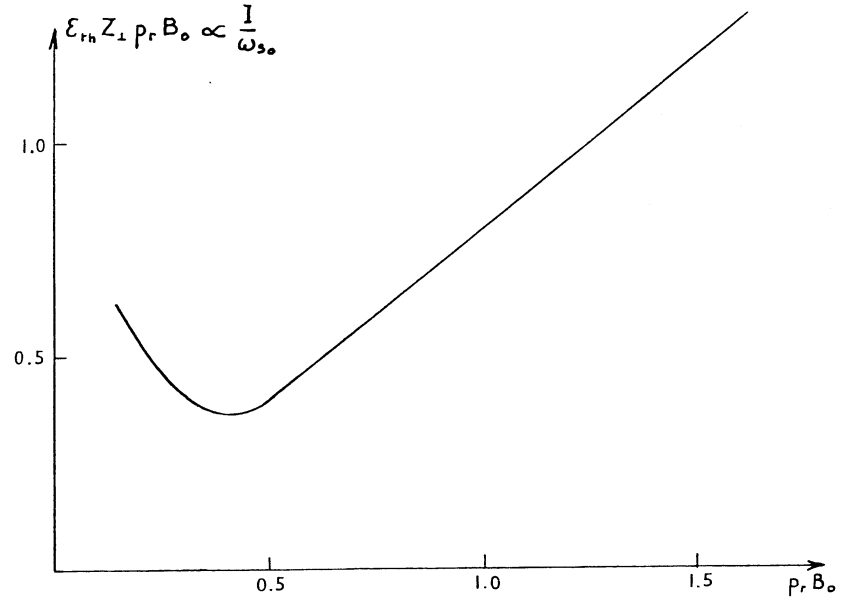


Fig. 37 Normalized threshold intensity as a function of the normalized bunch length

Results are summarized in Fig. 37. A broad band ($Q^* = 1$) impedance at resonant frequency $\omega_r = p_r \omega_o$ is assumed.

For extremely short bunches ($p_r B \approx .2$) the threshold is high. The minimum ($\varepsilon_{th} \sim .7$) is found for $p_r B \sim .8$, in rather good agreement with equation (203).

For long bunches ($p_r B > 1$) the lowest threshold is still at about the same level ($\varepsilon_{th} \sim .8$). Nevertheless, it corresponds to a rather weak coupling. Strong mode coupling occurs at higher intensity.

4. CONCLUSION

This lecture covers most of our present understanding of longitudinal and transverse coherent instabilities of bunched beams.

It seems that experimental results can be qualitatively explained by the extended F. Sacherer theory.

Quantitatively, we have insufficient knowledge of the effective coupling impedance of a machine.

ACKNOWLEDGEMENTS

This manuscript has been written in close collaboration with T. Aniel and J.P. Garnier from L.N. Saturne. A large part of the material presented here is directly drawn from the thesis prepared by J.P. Garnier.

I am very grateful to both of them for their efficient help.

I wish to thank G. Besnier for reading the manuscript and for raising several problems.

As far as references are concerned, I have proposed some of the basic ones, that look quite fundamental in order to illustrate F. Sacherer's theory. In advance, I apologize for having omitted many other valuable contributions to the subject.

REFERENCES

Longitudinal instability

1) F. Sacherer, Ph.D Thesis UCRL-18454 (1968).
2) F. Sacherer, CERN/SI/BR 72-5 (1972).
3) F. Sacherer, Trans. IEEE NS-20 (1973).
4) F. Sacherer, Trans. IEEE NS-24 (1977).
5) F. Sacherer, CERN/PS/BR/ 77-5 (1977).
6) B. Zotter, CERN/ISR/TH/ 78-9, 16(1978).
7) B. Zotter, CERN/ISR/TH/ 80-03 (1980).
8) B. Zotter, CERN/SPS/ 81-14, 18, 19, 20 (1981).
9) G. Besnier, B. Zotter, CERN/ISR/TH/ 82-17 (1982).
10) G. Besnier, Nucl. Inst. Meth. 164 (1979) 235, also Thesis, Faculté de Rennes (1978).
11) K. Robinson, report CEAL 1010 (1964).
12) E. Keil, W. Schnell, CERN/ISR/TH/RF/ 69-48 (1969).
13) D. Boussard, CERN/LAB II/RF/ 75-2 (1975).
14) T. Suzuki, Y. Chin and K. Satoh, Part. Accelerators, 13, 179 (1983).
15) A. Hofmann, Proc. of the Int. School of Part. Acc., Erice (1976) p. 139, CERN 77-13.
16) A. Hofmann. Proc. of the SLAC Summer School on Phys. of High Energy Particle Accelerators, Stanford, 1982, (AIP, New York, 1983).
17) A.W. Chao, Proc. of the SLAC Summer School on Phys. of High Energy Particle Accelerators (1984) p. 353.
18) S. Chattopadhyay, Proc. of the BNL/SUNY Summer School on Phys. of High Energy Particle Accelerators (1983) p. 467.
19) P.J. Channel. Proc. of the 12th Int. Conference on High Energy Accelerators, Batavia, 1983 (FNAL, Batavia, IH, 1983), p. 276.
20) J.L. Laclare, Proc. of the 11th Int. Conference on High Energy Accelerators, Geneva, 1980 (Birkhäuser, Basle, 1980), p. 525.

Transverse instability

1) C. Pellegrini, Nuovo Cimento 64A, 477 (1969).
2) M. Sands, SLAC-TN-69/8 and 69/10 (1969).
3) F. Sacherer, CERN/SI/BR/ 72-5 (1972).
4) F. Sacherer, CERN/PS/BR/ 76-21 (1976).
5) F. Sacherer, Proc. of the Int. School of Part. Acc., Erice (1976).
6) B. Zotter, CERN/ISR/TH/ 82-10 (1982).
7) B. Zotter, CERN/LEP/TH/ 85-06 (1985).
8) D. Brandt, B. Zotter, CERN/LEP/TH/ 83-42 (1983).
9) G. Besnier, D. Brandt, B. Zotter, CERN/LEP/TH/ 84-11 (1984).
10) Y.H. Chin, CERN/SPS/ 85-02 (1985).
11) R.D. Kohaupt, DESY report 80/22 (1980) and Proc. of the 11th Int. Conference on High Energy Accelerators, Geneva, 1980 (Birkhäuser, Basle, 1980), p. 562.

SPACE CHARGE DOMINATED BEAM TRANSPORT

I. Hofmann
GSI Darmstadt, P.O. Box 110 541, 6100 Darmstadt, Germany

ABSTRACT

In beam transport systems where space charge forces are comparable in size with the external focusing force, space charge plays an important role for emittance growth. In this lecture we introduce an approach which relates emittance directly to nonlinear electrostatic field energy. Analytic results and estimates for emittance growth are compared with computer simulation.

1. INTRODUCTION

In high-current linear accelerators and in transport systems for protons or heavier ions the repulsive force due to the space charge carried by the beam itself plays an important role for the design of the focusing system and for conservation of beam emittance. In proton or heavy ion linear accelerators the actual space charge bottle-neck is at injection, where the ion is slow and correspondingly the space charge density and resulting force large. More recently transport experiments[1,2,3] have been performed to study emittance growth under stationary conditions, i.e. no acceleration. Interest in such transport experiments has largely been stimulated by the idea of using heavy ions for inertial confinement fusion[4], which requires transport of large currents over long distances. We emphasize, however, that the role of space charge for emittance growth is equally important for present-day high-current proton or heavy ion linacs. Here we present an approach to the problem, which is based on the more recently developed idea of relating emittance to field energy[5,6] and therefore kept general enough to be of broader interest. While the analytical theory has recently been generalized to 3-D bunched beams[6] we confine this lecture to the 2-D problem of beam transport ignoring the longitudinal degree of freedom.

2. BASIC PROPERTIES

2.1 Incoherent Effect of Space Charge

For comparison we recall that in circular accelerators space charge is always a relatively weak perturbation described by the betatron tune shift $\Delta\nu$

$$\frac{d^2 x}{d\theta^2} = - (\nu_o - \Delta\nu)^2 x = - \nu^2 x \qquad (1)$$

where ν_o is the betatron tune in the absence of space charge and $\Delta\nu < \frac{1}{4} << \nu_o$ in order to use a gap free of machine resonances. The corresponding slight increase in betatron wavelength (Fig. 1a) is contrasted by the opposite case in a linear accelerator (Fig. 1b), where one tries to compensate the external focusing force (given by ν_o^2) as much as possible by space charge, hence $\nu^2 < \nu_o^2$.

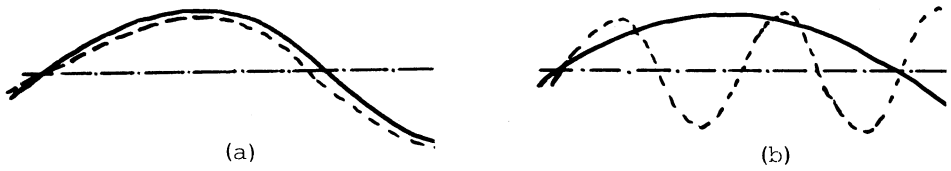


Fig. 1 Space charge effect upon betatron oscillation
(a) circular accelerator (b) linear accelerator.

2.2 Collective Effects

In practice the ideal space charge limit $\nu = 0$ with an exact balance between the applied focusing and the defocusing force due to space charge can be approached only to a limited extent. With decreasing ν^2/ν_o^2 the beam is dominated by collective behaviour, which can lead to emittance growth. Coherent modes of oscillation can be calculated analytically only for the highly specialized Kapchinskij-Vladimirskij-distribution[7]. For more realistic beam models one depends a great deal on computer simulation, which has shown various possibilities of emittance growth due to such coherent oscillations[8,9]. A practically important conclusion has been that in a periodic focusing system harmful "structure resonances" of coherent modes of oscillation with the focusing period can be avoided for certain values of the phase advance σ_o (per focusing period), for instance $\sigma_o \leq 60^0$.

2.3 Space Charge Limited Current

The relationship between beam current and σ/σ_o for a periodic quadrupole channel can be expressed in the following way[10]

$$I = 3.66 \cdot 10^6 \left(\frac{A}{Z}\right)^{1/3} B_o^{2/3} (\beta\gamma)^{7/3} \varepsilon^{2/3} F(\sigma/\sigma_o) \qquad (2)$$

with B_o the quadrupole pole-tip field, ε the transverse emittance and $F(\sigma/\sigma_o)$ a factor, which depends only weakly on the type of focusing (see Fig. 2 for a symmetric quadrupole

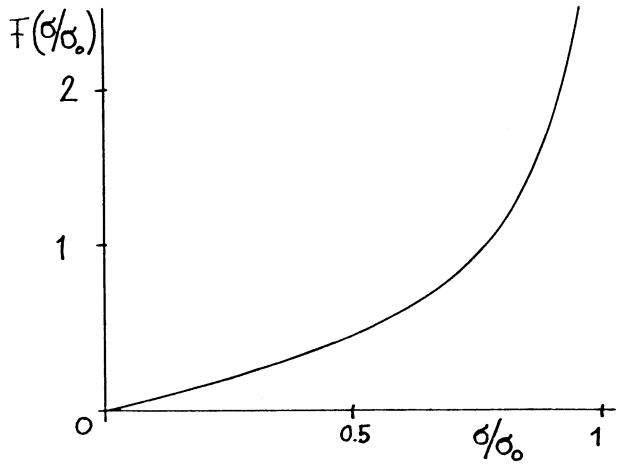


Fig. 2 Current transmission as function of σ/σ_o.

channel). Equation (2) is nothing but the condition for a matched stationary beam. The dependence of beam current on σ for a given channel can be seen more directly from the approximately valid formula[11]

$$I \sim a^2 \sim \frac{\varepsilon}{\sigma} \tag{3}$$

where a is the channel radius. Knowledge of the minimum σ for stable transport is thus essential for optimum current transmission.

3. EMITTANCE AND FIELD ENERGY

3.1 General Principle

We are interested to understand under what conditions and to what extent the emittance can grow. Obviously, it would be desirable to predict emittance growth from general principles without actually computing the dynamical process responsible for it. Here we present an approach, which allows a qualitative discussion of the various sources of emittance growth, as well as quantitative estimates for the expected growth in a number of cases. The general idea of this approach is to consider emittance growth as increase of thermal energy, which must be taken from the overall available beam energy. If we succeed in estimating the available source of energy we can also estimate the maximum possible emittance growth.

We assume the total energy of the beam can be written in the following way

$$E_{total} = E_o + E_{th,x} + E_{th,z} + E_{ex} + E_{field} \tag{4}$$

where

x,z	= transverse coordinates
E_o	= forward kinetic energy
$E_{th;x,z}$	= thermal energy in x,z (equivalent to emittance)
E_{ex}	= potential energy due to externally applied focusing force
E_{field}	= energy of space charge induced electric field

A growth of thermal energy (in x or z) and thus emittance is possible at the expense of either E_o or E_{field}. Due to the fact that the external potential is (at least in time average) focusing, E_{ex} is not a source of energy available for emittance growth, i.e. we expect E_{ex} also to increase during emittance growth. In addition, it is possible that thermal energy is transferred between the transverse plane (x,z) and the longitudinal direction (y), if there is a significant initial imbalance. For this exchange obviously E_{field} must be taken into account as well. For this more general case the reader is referred to Ref. 6.

It should be noted here that in a constant focusing system $E_{th} + E_{ex} + E_{field}$ is constant, hence we have no coupling with E_o. In periodic focusing, on the other hand, such a coupling can be significant as will be shown below, depending on the values of σ_o and σ.

3.2 Generalized Emittance Equation

It can be shown that the emittances ε_x, ε_z for the two transverse planes x,z are related to the electrostatic self-field energy by the following equation, valid for constant or periodic focusing (the derivation is given in Appendix A)

$$\frac{1}{\overline{x^2}} \frac{d}{ds} \varepsilon_x^2 + \frac{1}{\overline{z^2}} \frac{d}{ds} \varepsilon_z^2 = - 4K \frac{d}{ds} \frac{W-W_u}{w_o} \tag{5}$$

with

$$K = \frac{Nq^2}{2\pi\varepsilon_o m\gamma^3 v_o^2} \quad \text{(generalized perveance)}$$

$\overline{x^2}$, $\overline{z^2}$ mean squares of x, z

W actual field energy

W_u field energy of equivalent uniform density beam

w_o normalization constant (= field energy inside the actual beam boundary of a uniform beam)

The significance of this equation is that the change of emittance is related to a change of the "nonlinear field energy" $W-W_u$, i.e. the excess field energy due to the non-uniformity of the density.

3.3 Minimum Field Energy of Uniform Density Beam

A practically important property of the field energy is that it adopts its minimum for a uniform density beam. This can be shown by defining a variational expression

$$S \equiv W + \alpha_1 x^2 + \alpha_2 z^2 \tag{6}$$

with α_1, α_2 Lagrangian multipliers to keep the r.m.s. size constant. For a minimum we require the variation of S to vanish, hence

$$\delta S = \iint [\varepsilon_o E \delta E + N^{-1}(\alpha_1 x^2 + \alpha_2 z^2)\delta n]\, dx\, dz\,. \tag{7}$$

The density perturbation is defined by an arbitrary displacement

$$\delta n(x,z) = \nabla n \cdot \delta x, \tag{8}$$

thus we obtain

$$\delta S = \iint [\phi + N^{-1}(\alpha_1 x^2 + \alpha_2 z^2)](\nabla n \delta x) dx dz = 0 \tag{9}$$

requiring either $\phi = -N^{-1}(\alpha_1 x^2 + \alpha_2 z^2)$ (interior of beam, hence uniformly charged ellipsoid) or n = const. = 0 (exterior).

3.4 Field Energy of Different Beam Models

Calculation of the field energy for a parabolic density profile

$$n = \frac{2N}{\pi ab}\left(1 - \frac{x^2}{a^2} - \frac{z^2}{b^2}\right) \tag{10}$$

leads to the following result

$$W = \frac{N^2 q^2}{16\pi\varepsilon_o}\left(\frac{11}{6} - 4\ln\sqrt{6} + 4\ln\frac{2R}{\tilde{x}+\tilde{z}}\right) \tag{11}$$

with the r.m.s. envelope $\tilde{x} = a/\sqrt{6}$ and $\tilde{z} = c/\sqrt{6}$. Due to the minimum field energy property of a uniform beam of same r.m.s. size it is convenient to calculate the difference energy in normalized units:

$$\frac{W-W_u}{W_o} = \frac{5}{6} - 4\ln\frac{\sqrt{6}}{2} \simeq 0.0224\,. \tag{12}$$

This normalized "nonlinear field energy" hence is always positive according to the previous section. For a Gaussian beam (truncated at 4 standard deviations) it is 0.154.

3.5 Stationary Distributions and Debye Shielding

In computer calculations it has been observed that distribution functions with non-uniform real density in the low current limit lead to nearly uniform density in the limit of high current (Fig. 3).

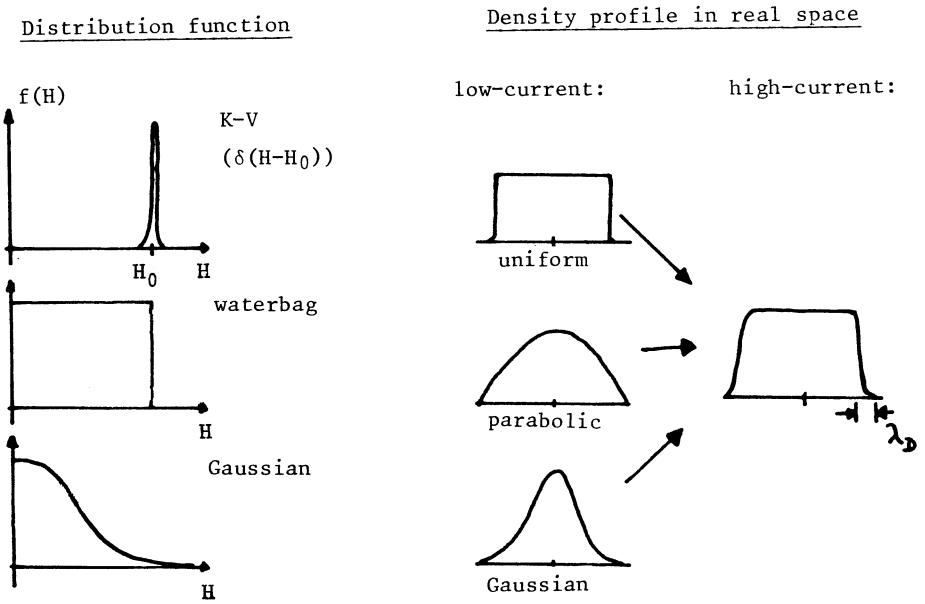


Fig. 3 Density profiles for different distribution functions. A boundary sheath of thickness ~ λ_D is obtained in the high-current limit.

This is readily shown for a Gaussian distribution

$$f \sim \exp\left[-\left(\frac{x'^2+y'^2}{2} + \frac{k}{2}r^2 + \frac{q}{m\gamma^3 v_o^2}\phi\right)/\mu\right] \tag{13}$$

which, after insertion into Poisson's equation and expansion of the exponential yields

$$n \simeq n_o[1-(\frac{a}{r})^{1/2}\exp((r-a)/\lambda_D)] \tag{14}$$

provided that the total potential in equation (13) is small as compared with the average transverse energy μ. This condition can be expressed in terms of the Debye-length λ_D, i.e.

$$\frac{\lambda_D}{a} << 1 \tag{15}$$

where λ_D is given by

$$\lambda_D^2 = \frac{\mu}{\omega_p} \simeq 2\frac{x'^2}{\omega_p^2} \simeq \frac{v_{th}^2}{\omega_p^2} \quad . \tag{16}$$

We can also express λ_D/a in terms of the tune depression

$$\frac{\lambda_D}{a} \simeq \frac{1}{\sqrt{8}}\frac{\nu}{\nu_o} \quad . \tag{17}$$

The physical meaning of the Debye-length is the one familiar in plasma physics: a high-current beam shields the external focusing force from its interior by buildung up a uniform density with a space charge force that cancels the focusing force. The shielding is ineffective in the boundary layer of a thickness given by λ_D (Fig.3). Hence, for small ν/ν_o there is practically no restoring force in the interior of the beam and particles are 'reflected' by the uncompensated force in the boundary layer.

4. APPLICATION TO EMITTANCE GROWTH

Using equation (5) we can now discuss the possible sources for emittance growth. The relationship between emittance and thermal rsp. kinetic energy suggests the analogy with a ball in a potential trough. The following situations can be envisaged:

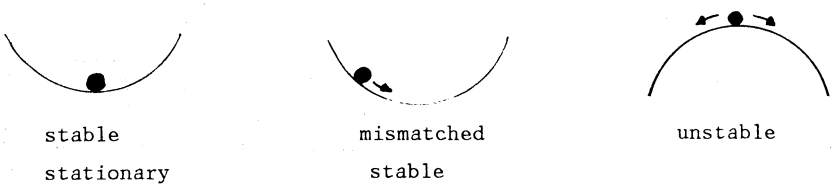


Fig. 4 Analogy with ball in a potential trough

Stable Stationary Beam

The emittances ε_x and ε_z are conserved along the channel and the envelopes x and z are allowed to vary periodically. According to equation (5) this requires that $W-W_u$ is a con-

stant in this case. A sufficient criterion therefore is that the beam density profile remains self-similar (this has been confirmed in numerical simulation; a rigorous analytical proof for beams with non-uniform density and periodic focusing has not yet been found).

Mismatched Beam

An initial excess field energy is transferred into emittance according to equation (5). For a round beam and constant focusing we can integrate equation (5) and find

$$\frac{\Delta\varepsilon}{\varepsilon}^2 = -\frac{1}{2}\left(\frac{\nu_o}{\nu}^2 - 1\right) \Delta U \qquad (18)$$

with $U \equiv (W-W_u)/w_o$ the normalized nonlinear field energy.

As an example we inject a beam with Gaussian profile into a channel with $\nu_o/\nu = 6$. Since we know from equation (14) that close to the space charge limit a stationary solution has nearly uniform density, nearly all of the excess nonlinear field energy (0.15 in normalized units) is found as emittance growth during less than a betatron oscillation. Hence we can estimate the predicted emittance growth by ignoring the final nonlinear field energy and thus obtain

$$\frac{\Delta\varepsilon}{\varepsilon}^2 \leq 2.6 \qquad (19)$$

which is equivalent to an emittance growth of 90 % and agrees very well with the result of simulation. This is shown in Fig. 5a for a periodic solenoid channel. We observe that the mismatch emittance growth formula applies equally well to periodic focusing also - growth factors found in simulation are practically identical with those in constant focusing. It is only the ratio ν_o/ν, which is important.

A different result is obtained if a strictly uniform beam is injected, hence we start at the minimum field energy. Such a beam has the tendency of smoothing its sharp boundary, which requires a small amount of field energy leading to a small decrease in emittance. This is again confirmed by simulation (Fig. 5b). We observe that such an r.m.s. emittance decrease is not violating Liouville, which holds for the exact phase space volume.

We thus conclude that injection of as uniform a density as possible minimizes emittance growth of high-current beams.

Coherent Instabilities

In linear beam transport coherent instabilites occur due to direct interaction of neighbouring particles via their space charge force. This is in contrast with beams in circular accelerators, where coupling to the surrounding structure or to other bunches plays a dominant role.

In the present framework coherent instabilites can be discussed qualitatively by writing equation (5) for a round beam in periodic solenoidal focusing ($k_x = k_z$).

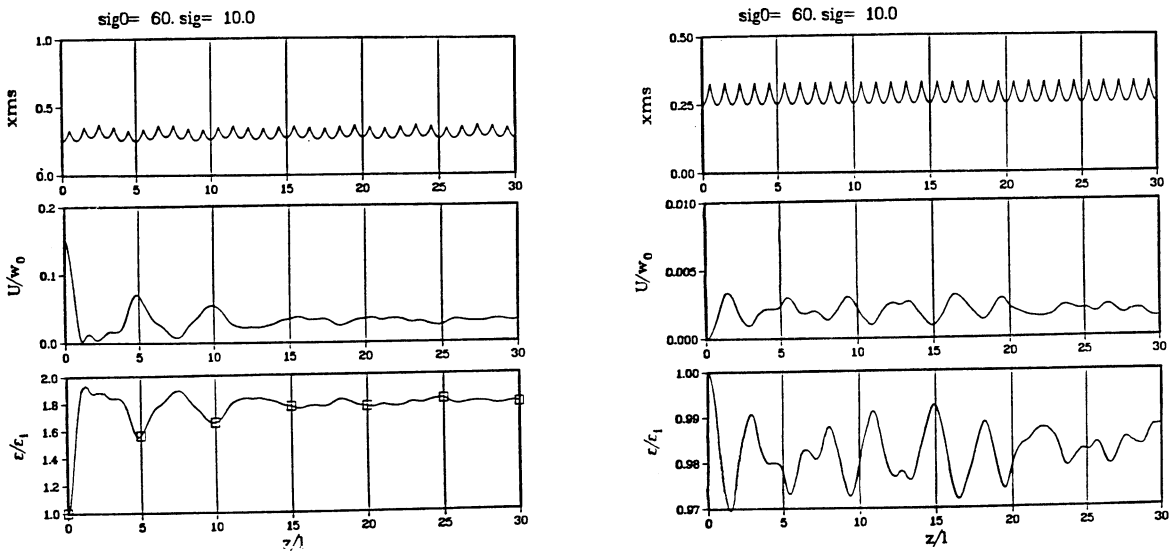


Fig. 5 Nonlinear field energy U (normalized) and emittance growth factors ($\varepsilon/\varepsilon_{initial}$) versus number of focusing periods in a periodic solenoid channel with $\sigma_o = 60^0$, $\sigma = 10^0$ (equivalent to $\nu_o/\nu = 6$)

(a) Gaussian distribution

(b) Semi-Gaussian distribution (i.e. uniform real density, Gaussian in velocity space).

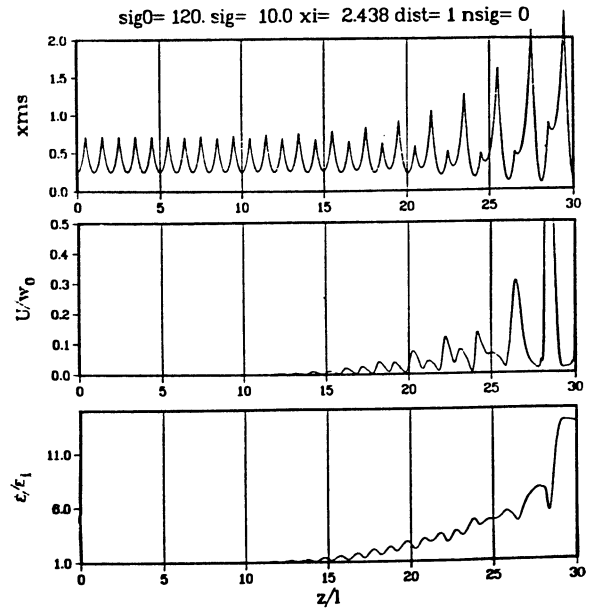


Fig. 6 Structure resonance in periodically interrupted solenoid focusing for initially uniform beam (KV-distribution)

$$\frac{d}{ds}\varepsilon^2 = -2K x^2(s) \frac{d}{ds} \frac{W-W_u}{w_o} . \tag{20}$$

We assume a matched envelope, hence x^2 oscillates with the focusing periodicity. Furthermore, we assume a small coherent oscillation of the beam, such that $W-W_u$ oscillates with the period of oscillation of the particular mode. We now anticipate a steadily growing ε^2, if $x^2(s)$ and $W-W_u$ oscillate with the same frequency (phase shifted by $\sim 90^0$). In this case the r.h.s. of equation (20) has a non-oscillatory part, which is integrated to a finite emittance growth until the resonance condition is detuned. The resulting "structure resonance" is thus a coupling between the zero-order oscillatory envelope and a non-uniform coherent mode of oscillation as illustrated by the example in Fig. 6 of a round beam in periodic focusing with $\sigma_o = 120^0$ and $\sigma = 10^0$.

A direct calculation of the emittance growth is not possible from equation (20); this would need determination of the actual time dependence of $W-W_u$, which is not possible analytically. For illustration of the principle of a structure resonance we can, however, derive a scaling expression for the emittance growth by making a simple harmonic approximation

$$x^2(s) \simeq x^2 + \delta x^2 \sin(\omega s) \tag{21}$$

$$U \simeq U + \delta U \cos(\omega s) \tag{22}$$

where $\omega = 2\pi/L$ and L the focusing period. We then find from equation (20)

$$\frac{d}{ds}\varepsilon^2 \simeq K \delta x^2 \omega \delta U . \tag{23}$$

In smooth approximation K follows from the envelope equation (with ε_i the matched initial emittance)

$$K = \frac{1}{4} \frac{\varepsilon_i^2}{\hat{x}^2} \left(\frac{\sigma_o^2}{\sigma^2} - 1\right) . \tag{24}$$

We then obtain, ignoring the time-dependence of $\overline{\delta x^2}$ and δU,

$$\frac{\varepsilon}{\varepsilon_i} \leq \left[1 + \frac{\pi}{2} \left(\frac{\sigma_o^2}{\sigma^2} - 1 \right) \frac{\delta x^2}{x^2} \delta U \frac{s}{L} \right]^{1/2} . \tag{25}$$

From this expression we conclude that in an advanced stage the emittance rise is weaker than linear in time. It is enhanced by small ν and a large flutter of the unperturbed envelope. In the example of Fig. 6 the excited mode has an octupole type field perturbation, which grows from noise on the initially uniform beam (KV-distribution) and saturates between focusing period 15 and 20. In this region the average normalized nonlinear field energy δU is estimated as 0.02 and equation (25) yields $\varepsilon/\varepsilon_o \simeq 3$, which agrees reasonably well with the numerical result. After focusing period 20 the situation becomes very nonlinear: a different mode with half the oscillation frequency starts growing and leads to further strong emittance growth.

This particular mode can be avoided if $\sigma_o \leq 90^0$, in which case it is out of resonance. Another type of mode can appear for $\sigma_o = 90^0$ and $\sigma \simeq 45^0$ (sextupole type field perturbation), whereas $\sigma_o = 60^0$ has been found free of emittance growth for values of σ as small as a few percent. This result has been confirmed also by experiment[1].

ACKNOWLEDGEMENT

The author has benefitted from the results of computer simulation by K. Crandall (LANL).

* * *

REFERENCES

1) A. Faltens et. al., Proc. of the 1984 Linear Accelerator Conference, Seeheim, Germany, p. 312 (1984)
2) J. Klabunde et. al., ibid., p. 315 (1984)
3) M. Reiser et. al., ibid., p. 309 (1984)
see for instance the proceedings of the 1984 workshop on heavy ion fusion: Proc. of the INS International Symposium on Heavy Ion Accelerators and Their Applications to Inertial Fusion, INS, Tokyo (1984)
4) T.P. Wangler et. al., IEEE Trans. Nucl. Sci NS-32, p. 2196 (1985)
5) I. Hofmann, Proc. 1986 Linear Accelerator Conference, Stanford, June 2-6, 1986, see also I. Hofmann and J. Struckmeier, GSI-Report GSI 86-11 (1986)
6) I. Hofmann, L.J. Laslett, L. Smith and I. Haber, Particle Accelerators 13, 145 (1983)
7) I. Hofmann in "Applied Charged Particle Optics" (ed. A. Septier), part C, Academic Press N.Y., p. 99 (1983)
8) I. Hofmann, Proc. of the 1984 Linear Accelerator Conference, Seeheim, Germany, p. 304 (1984)
9) A.W. Maschke, Proc. ERDA Summer Study of Heavy Ions for Inertial Fusion, Berkeley, 1976;
10) G.R. Lambertson et. al., IEEE Trans. Nucl. Sci. NS-24, 933 (1977)
11) M. Reiser, Part. Accel. 8, 167 (1978).

A P P E N D I X

A. Derivation of Generalized Emittance Equation

In order to derive analytically a relationship between emittances and field energy we consider x and z as transverse coordinates and start from the single-particle equation of motion in x (similar for z)

$$x'' + k_x(s)x - \frac{q}{m\gamma^3 v_o^2} E_x(x,z,s) = 0 \tag{A.1}$$

where k_x describes the periodically varying focusing force and $\underline{E}$ follows from Poisson's equation

$$\underline{\nabla}\cdot\underline{E} = \frac{q}{\varepsilon_o} n(x, z, s) \tag{A.2}$$

and n is the density obtained by projecting a 4-D phase space distribution into real space:

$$n = \iint f(x,z,x',z',s)\, dx'\, dz' \tag{A.3}$$

The distribution function f is subject to Vlasov's equation ($v \equiv dx/ds$).

$$\frac{\partial f}{\partial s} + (\underline{x}'\cdot\underline{\nabla}) f - (\underline{k} - \frac{q}{m\gamma^3 v_o^2}\underline{E})\cdot\underline{\nabla}_{x'} f = 0 \tag{A.4}$$

Solution of equations (A.2)-(A.4) is fraught with the usual difficulties in solving a partial integro-differential equation. Keeping in mind that the r.m.s. emittance is defined through second-order moments of the distribution function, we make an attempt to convert equation (A.4) into ordinary differential equations involving second-order moments only. We thus define the moments, where we assume that N is the total number of particles per unit length of beam, according to

$$\overline{x^2} = N^{-1} \int\ldots\int x^2 f\, dx\ldots dz', \tag{A.5}$$

similar with x'^2, xx' and likewise in z. The r.m.s. emittance is then defined as

$$\varepsilon_x \equiv 4\,(\overline{x^2}\,\overline{x'^2} - \overline{xx'}^2)^{1/2} \tag{A.6}$$

and similar in z. We then obtain from equation (A.4) by multiplying it with x^2 and integrating over all phase space

$$\frac{d}{ds}\overline{x^2} - 2\overline{x\,x'} = 0 \tag{A.7}$$

$$\frac{d}{ds}\overline{xx'} - \overline{x'^2} + k_x\,\overline{x^2} - \frac{q}{m\gamma^3 v_o^2}\,\overline{xE_x} = 0 \tag{A.8}$$

$$\frac{d}{ds}\overline{x'^2} + 2k_x\overline{xx'} - \frac{2q}{m\gamma^3 v_o^2}\overline{x'E_x} = 0 \tag{A.9}$$

By following Sacherer's procedure we readily obtain the r.m.s. envelope equation

$$\frac{d^2}{ds^2}\tilde{x} + k_x(s)\tilde{x} - \frac{\varepsilon_x^2(s)}{16\tilde{x}^3} - \frac{q}{m\gamma^3 v_o^2}\frac{\overline{xE_x}}{\tilde{x}} = 0 \tag{A.10}$$

(similar in z) with $\tilde{x} \equiv \overline{x^2}^{\frac{1}{2}}$ the r.m.s. envelope. This requires, however, knowledge of ε_x to be of real use. Clearly, for constant ε_x it allows a straightforward integration of the r.m.s. envelope, if $\overline{xE}_x$ can be calculated.

We readily derive from equation (10) the differential equation

$$\frac{d}{ds}\varepsilon_x^2 = \frac{32q}{m\gamma^3 v_o^2}\left(\overline{x^2}\ \overline{x'E_x} - \overline{xx'}\ \overline{xE_x}\right) \tag{A.11}$$

and similar for z. The practical difficulty now lies in the fact that we had to introduce the moments $\overline{xE}_x$, $\overline{x'E}_x$; for E_x other than a linear function of x (or a constant) these moments are of higher order, hence the above differential equations are not a closed set. By going to higher order moments we would obtain an infinite set of coupled equations, in general.

In the following we show that it is possible to transform the terms involving the electric field in such a way that only the energy of the electric field appears explicitly. The latter still includes higher-order moments, but we benefit from the fact that it is a quantity of direct physical meaning and amenable to estimates.

For this purpose we re-write

$$\overline{x'E_x} = N^{-1}\int\ldots\int x'E_x f\,dx\ldots dz' = N^{-1}\iint E_x n v_x\,dx\,dz \tag{A.12}$$

where we have introduced v_x as local averaged velocity of beam particles (in a frame, where the beam is at rest).

With the local current

$$\underline{j} = qnv_o\underline{v} \tag{A.13}$$

we obtain

$$\overline{x'E_x} = (Nqv_o)^{-1}\iint E_x j_x\,dx\,dz \tag{A.14}$$

and similar for z. We thus find ($\underline{E} = -\underline{\nabla}\phi$)

$$\iint \underline{E}\cdot\underline{j}\,dx\,dz = \iint \phi\underline{\nabla}\cdot\underline{j}\,dx\,dz = -qv_o\iint \phi\frac{\partial n}{\partial s}\,dx\,dz \tag{A.15}$$

where we have used the continuity equation

$$\frac{\partial n}{\partial s} + (qv_o)^{-1}\underline{\nabla}\cdot\underline{j} = 0 \tag{A.16}$$

derived from equation (A.4) by integration. The integration in equation (A.15) is performed over the cross section F, which contains the beam in its interior, hence we may neglect a boundary integral. Using Poisson's equation we obtain by partial integration:

$$\iint \underline{E} \cdot \underline{j} \, dx \, dz = - v_o \frac{d}{ds} W - \varepsilon_o v_o \int \phi \frac{\partial}{\partial s} E_n \, d\sigma \tag{A.17}$$

with E_n the normal component of $\underline{E}$ on the boundary of F and

$$W = \frac{\varepsilon_o}{2} \iint \underline{E}^2 \, dx \, dz \tag{A.18}$$

the electric field energy within F.

The next step is to add the three equations (A.11) and express $\overline{x' E_x}$ by the electric field energy according to equations (A.14, A. 17). We thus find the relationship

$$\frac{1}{\overline{x^2}} \frac{d}{ds} \varepsilon_x^2 + \frac{1}{\overline{z^2}} \frac{d}{ds} \varepsilon_z^2 = \frac{32q}{m\gamma^3 v_o^2} \left[- \frac{1}{Nq} \frac{dW}{ds} - \frac{\varepsilon_o}{Nq} \int \phi \frac{\partial}{\partial s} E_n \, d\sigma - I \right] \tag{A.19}$$

with

$$I = \frac{1}{2} \left(\frac{1}{\overline{x^2}} \frac{d\overline{x^2}}{ds} \overline{xE}_x + \frac{1}{\overline{z^2}} \frac{d\overline{z^2}}{ds} \overline{zE}_z \right). \tag{A.20}$$

Equ. (A.19) holds exactly, but we need some approximation to evaluate the term I on the r.h.s. and obtain a practically useful equation:

Firstly, it can be shown that for a uniform beam with elliptic cross section the field energy per unit length calculated within a large circle of radius R is given by

$$W_u = \frac{N^2 q^2}{16\pi\varepsilon_o} \left(1 + 4 \ln \frac{2R}{a+b} \right) \tag{A.21}$$

where a and b are the semi-axi in x and z. We then also find for the uniform beam that I is related to W_u according to

$$I = - \frac{1}{Nq} \frac{d}{ds} W_u \tag{A.22}$$

Next we use the result found by Sacherer that $\overline{xE}_x$ and $\overline{zE}_z$ are independent of the density profile as long as elliptical symmetry is given:

$$n(x,z,s) = n\left(\frac{x^2}{a^2} + \frac{z^2}{b^2}, s \right) \tag{A.23}$$

Hence, equation (A.22) is true for all beams satisfying equation (A.23) and we re-write equation (A.19) as

$$\frac{1}{\overline{x^2}} \frac{d}{ds} \varepsilon_x^2 + \frac{1}{\overline{z^2}} \frac{d}{ds} \varepsilon_z^2 = - 4K \frac{d}{ds} \frac{W - W_u}{W_o} . \tag{A.24}$$

Here we have neglected the boundary integral (which is justified for large R) and introduced the generalized perveance

(A.25) $$K \equiv \frac{Nq^2}{2\pi\varepsilon_o m\gamma^3 v_o^2} \tag{A.25}$$

and the field energy normalization constant

$$w_o \equiv \frac{N^2 q^2}{16\pi\varepsilon_o} \tag{A.26}$$

which gives the field energy of a uniform beam within the actual beam boundary.

We note that the generalized emittance equation (A.24) derived here for 2-D beams can be derived for 3-D bunched beams as well[6]. We have found by computer simulation that equation (A.24) can be applied to distributions more general than elliptically symmetric, with an error that is negligible in practice.

WAKE FIELDS, IMPEDANCES AND GREEN'S FUNCTION

L. Palumbo
Dip. Energetica, Univ. "La Sapienza", Roma
INFN, Laboratori Nazionali di Frascati

V.G. Vaccaro
Dip. Fisica, Università degli Studi, Napoli
INFN, Sezione di Napoli

ABSTRACT

The concept of 'coupling impedance' was introduced in the middle sixties in the early studies of the instabilities arising in the ISR at CERN. Since then this impedance has been successfully used to schematize the interaction between the beam and the machine allowing simple stability criteria in the frequency domain to be established. The concept of 'wake potential' has been applied to beam dynamics more recently. It was orginally used in the study of the energy spectrum due to beam loading in the SLAC linear accelerator. Of particular interest nowadays is its application in the developments of new high-gradient acceleration mechanisms such as the wake field transformers. In this chapter we will introduce these quantities showing their main properties and their relationship with the Green function.

1. LONGITUDINAL WAKE POTENTIAL AND IMPEDANCE

1.1 Introduction

The behaviour of a system consisting of a large number of charged particles under the action of electro-magnetic (e.m.) forces is described by the Vlasov equation[1] when radiation is neglected. In this chapter we will focus our attention on the fields created by the charged particles, which act back on the beam perturbing the particle motion. If a small perturbation appears in the beam density, it will interact with the surrounding equipment generating e.m. fields. These will always have in their spectrum a wave component which rides in synchronism with the perturbation and which may enhance it leading, in the absence of damping, to an avalanche phenomenon. The mechanism is schematized in Fig. 1 where we see that perturbation and fields have at the same time the role of cause and effect and we can understand why circular accelerators, where the loop of Fig. 1 may work over a long time, are potentially unstable feed-back devices.

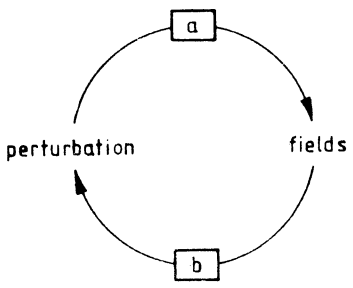


Fig. 1 Circular accelerator as a feedback device

In the following sections we will deal only with the transfer function branch "a" of the loop and will introduce the basic concepts of impedances, δ-function wake potential and their relationship with the Green's function. We shall see that the wake potential is the time response of the system to an impulsive excitation and that concerning the linearity of the system (Fig. 2) its transfer function in the frequency domain is strictly related to the impedance of the system $Z(\omega)$.

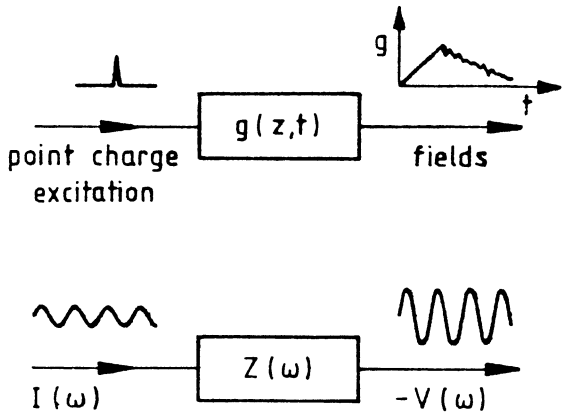


Fig. 2 Time and frequency response of a linear system

We will use the following Fourier transform definitions:

$$f(z,t) = \int_{-\infty}^{\infty} \tilde{F}(z,\omega)\, e^{-i\omega t}\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\kappa \int_{-\infty}^{\infty} d\omega\, \tilde{\tilde{F}}(\kappa,\omega)\, e^{i(\kappa z-\omega t)}$$

$$\tilde{F}(z,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(z,t)\, e^{i\omega t}\, dt = \mathcal{F}_t[f(z,t)]$$

$$\tilde{\tilde{F}}(\kappa,\omega) = \int_{-\infty}^{\infty} \tilde{F}(z,\omega)\, e^{-i\kappa z}\, dz = \frac{1}{2\pi} \iint_{-\infty}^{\infty} dz\, dt\, f(z,t)\, e^{-i(\kappa z-\omega t)} = \mathcal{F}_z \mathcal{F}_t[f(z,t)]$$

(Note: the Fourier transform is usually normalised by $1/\sqrt{2\pi}$, but it is conventional in the accelerator world to normalise $\mathcal{F}_t$ by $1/2\pi$ and to leave $\mathcal{F}_z$ unmodified).

1.2 Basic concepts

Consider two point charges travelling on the axis z at constant velocity $v = \beta c$, the trailing charge being at a distance $\beta c\tau$ from the leading one (Fig. 3), and let $t = 0$ be the time when the first charge crosses the origin $z = 0$. Consider the longitudinal electrical field generated by the interaction of the leading charge with the surroundings. We define the δ-function wake potential, the potential per unit charge seen by the trailing charge over a distance $\Delta\ell$ as[2-5]:

$$W_e(\tau) = -\frac{1}{Q} \int_0^{\Delta Z} E_z\, [z = \beta c(t-\tau), t]\, dz; \quad [\Omega s^{-1}] \tag{1.1}$$

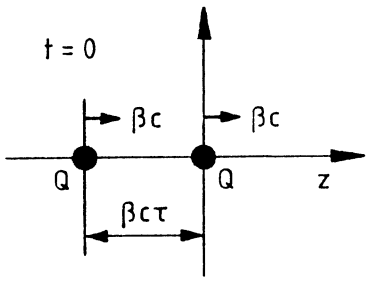


Fig. 3 The source and test charges in the relevant reference frame

Applying the Fourier transform over the variable τ we get:

$$\mathcal{F}_\tau[W(\tau)] = -\frac{1}{2\pi Q}\int_0^{\Delta z} dz \int_{-\infty}^{\infty} E_z[z = \beta c(t-\tau), t]\, e^{i\omega\tau}\, d\tau \equiv Z(\omega); \quad [\Omega] \tag{1.2}$$

which defines the longitudinal coupling impedance. From the frequency spectrum of the electrical field due to the charge at z = 0 we obtain

$$Z(\omega) = -\frac{1}{Q}\int_0^{\Delta z} \tilde{E}_z(\omega)\, e^{+i\kappa_0 z/\beta}\, dz = \frac{V(\omega)}{I(\omega)} \tag{1.3}$$

where $\kappa_0 = \omega/c$, and $\tilde{E}_z(\omega)$ is the spectrum of the field at the point z = 0. Noting that the associated current $I_z(t) = +Q\beta c\delta(-\beta ct)$ has the spectrum $I(\omega) = Q$, Eq. (1.3) can be interpreted as the ratio between the harmonic of the potential $V(\omega)$ over a distance ΔZ due to a current $I(\omega)$, and the current itself[6]. In the case of longitudinal e.m. homogenity of the wall properties the field configuration travels with the source and the trailing charge sees a constant field; without loss of generality we set t = 0 and introduce the wake per unit length as:

$$\frac{dW(\tau)}{dz} = -\frac{E_z(-\beta c\tau)}{Q} \tag{1.4}$$

and the corresponding impedance per unit length:

$$Z(\omega) = -\frac{\tilde{E}_z(\omega)}{\beta cQ} \quad . \tag{1.5}$$

The wake potential is defined as a real function of the time. We may write it as the inverse Fourier transform of its spectrum, i.e. the impedance:

$$W(\tau) = \int_{-\infty}^{\infty} Z(\omega)\, e^{-i\omega\tau}\, d\omega$$

$$W(\tau) = \int_{-\infty}^{\infty} [Z_R(\omega)\cos\omega t + Z_I(\omega)\sin\omega\tau]\, d\omega - i\int_{-\infty}^{\infty} [Z_R(\omega)\sin\omega\tau - Z_I(\omega)\cos\omega\tau]\, d\omega \tag{1.6}$$

Now the imaginary part of Eq. (1.6) must vanish; this implies that the real part of the impedance $Z_R(\omega)$ is an even function of the frequency while the imaginary part is an odd function of it. Accordingly we may write

$$W(\tau) = 2 \int_0^\infty [Z_R(\omega) \cos \omega\tau + Z_I(\omega) \sin \omega\tau] \, d\omega \, . \tag{1.7}$$

For the case of ultrarelativistic charges $\gamma \to \infty$, the velocity of the leading charge is practically the light velocity. Therefore, we expect the wake to vanish in the region ahead of the charge, i.e. $\tau < 0$

$$W(-\tau) = 2 \int_0^\infty [Z_R(\omega) \cos \omega\tau - Z_I(\omega) \sin \omega\tau] \, d\omega = 0 \tag{1.8}$$

$$\int_0^\infty Z_R(\omega) \cos \omega\tau \, d\omega = \int_0^\infty Z_I(\omega) \sin \omega\tau \, d\omega \tag{1.9}$$

and

$$W(\tau) = 4 \int_0^\infty Z_R(\omega) \cos \omega\tau \, d\omega \, . \tag{1.10}$$

For $t \to 0$ Eq. (1.10) gives a result which is twice the result given by Eq. (1.6) at $\tau = 0$. The self wake seen by the particle is then half of the limit value of $w(\tau)$ for $t \to 0$. This property is often referred to as the "fundamental theorem on the beam loading" and is a consequence of the causality principle.

2. LONGITUDINAL IMPEDANCE AND WAKE POTENTIAL FOR SIMPLE STRUCTURES

In this section we will explain how to derive the analytical expression of the e.m. fields excited by a point charge in simple structures of interest in accelerator physics. The wake potential and the impedance will be found in the longitudinal case. A general relationship between longitudinal and transverse impedance will be given in a later section.

2.1 Perfectly conducting cylindrical pipe

Consider a unit point charge travelling at constant velocity βc along the z direction at a distance r_0 from the axis of a perfectly conducting (p.c.) cylindrical pipe of radius b (Fig. 4). The scalar potential V is a solution of the wave equation

$$\nabla^2 V - \frac{1}{c^2} \frac{\partial^2 V}{\partial t^2} = - \frac{\rho}{\varepsilon_0} \tag{2.1}$$

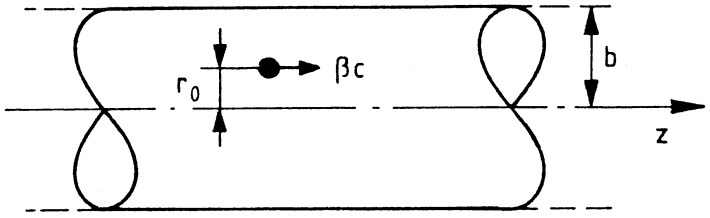


Fig. 4 An off-axis point charge moving within a perfectly conducting pipe

and satisfies the boundary condition V = 0 at r = b. Because of the longitudinal homogeneity of the boundary, potentials and fields propagate at the same velocity βc as the source. Therefore, Eq. (2.1) may be rewritten for an impulsive source as:

$$\nabla_t^2 g + \frac{1}{\gamma^2}\frac{\partial^2 g}{\partial z^2} = -\frac{1}{\varepsilon_0}\frac{\delta(r-r_0)}{r_0}\delta(\theta-\theta_0)\delta(z-\beta ct) \tag{2.2}$$

where g(z-βct) is Green's function of the problem, and ∇_t is the "nabla" operator* on the transverse cylindrical coordinates r and θ. The electrical field is obtained from:

$$\underline{E} = -\nabla_t g - \frac{1}{\gamma^2}\frac{\partial g}{\partial z}\hat{z}\ . \tag{2.3}$$

Since the field configuration travelling with the charge does not change with the time, without loss of generality we may set t = 0. It is understood that a time-space representation is obtained by replacing z by (z-βct).

Making use of the following relations:

$$\delta(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\kappa z}\, d\kappa \tag{2.4}$$

$$\delta(\theta-\theta_0) = \frac{1}{2\pi}\sum_{0}^{\infty}{}_m\, \alpha_m \cos[m(\theta-\theta_0)] \tag{2.5}$$

where $\alpha_m = \begin{cases} 1 & , m = 0 \\ 2 & , m \neq 0 \end{cases}$

and assuming for the potential g(r,z,θ) the Fourier representations:

$$g(z,r,\theta) = \sum_{0}^{\infty}{}_m\, g_m(z,r)\cos[m(\theta-\theta_0)] \tag{2.6}$$

---

* Introduced by Sir William Hamilton and named after an ancient Assyrian harp whose shape it resembles.

with

$$g_m(z,r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{G}_m(\kappa,r) e^{i\kappa z} d\kappa \tag{2.7}$$

we obtain from Eq. (2.6) for each component $\tilde{G}_m$ in cylindrical coordinates:

$$\frac{\partial^2 \tilde{G}_m}{\partial r^2} + \frac{1}{r} \frac{\partial \tilde{G}_m}{\partial r} - \left(\frac{m^2}{r^2} + q^2\right) \tilde{G}_m = - \frac{\delta(r-r_0)}{2\pi\varepsilon_0 r_0} \alpha_m \tag{2.8}$$

where $q = \kappa/\gamma$.

The above differential equation has a twofold solution[7)]

$$\tilde{G}_m(\kappa,r) = \frac{\alpha_m}{2\pi\varepsilon_0} \begin{bmatrix} K_m(qr) I_m(qr_0) \\ K_m(qr_0) I_m(qr) \end{bmatrix} ; \quad \begin{matrix} r > r_0 \\ r < r_0 \end{matrix} \tag{2.9a, 2.9b}$$

where I_m and K_m are the modified Bessel functions of first and second kind respectively. Equations (2.9) are the solutions in the free space. In our problem we require the potential to vanish at $r = b$. This implies that the general expression of the potential is given by a superposition of Eqs. (2.9) plus a term accounting for the image currents induced on the pipe wall; since the images are at $r = b$ and we are looking for the potential within the region $r < b$, the additive term must behave like (2.9b):

$$\tilde{G}_m(\kappa,r) = \frac{\alpha_m}{2\pi\varepsilon_0} \left\{ \begin{bmatrix} K_m(qr) I_m(qr_0) \\ K_m(qr_0) I_m(qr) \end{bmatrix} + B_m I_m(qr) \right\} \begin{matrix} r > r_0 \\ r < r_0 \end{matrix} \tag{2.10}$$

The boundary conditions are satisfied when $\tilde{G}_m = 0$ for any m, namely:

$$B_m = - \frac{I_m(qr_0)}{I_m(qb)} K_m(qb) \ . \tag{2.11}$$

Consider now the longitudinal wake and impedance relative to a charge travelling on axis ($r_0 = 0$); in this case Eq. (2.10) simplifies to

$$\tilde{G}_0(\kappa,r) = \frac{1}{2\pi\varepsilon_0} \left[K_0(qr) - \frac{K_0(qb)}{I_0(qb)} I_0(qr)\right] \ . \tag{2.12}$$

Since we have assumed a point charge as the source, the above expression diverges logarithmically when $r \to 0$ and cannot be used in Eq. (1.3) to compute the impedance.

To this end we shall assume that the charge has a uniform radial distribution within a δ-disk of radius "a" so that the potential can be obtained by convolution of Eq. (2.12) over the distribution; at r = 0; in the limit q ≪ 1 we get:

$$\tilde{V}(k) \simeq \frac{1}{4\pi\varepsilon_0}\left[1 + 2\ \ln\frac{b}{a}\right] \quad . \tag{2.13a}$$

Now from (2.3) we have:

$$\tilde{E}(\kappa) = -\frac{ik}{\gamma^2}\ \tilde{V}(\kappa) \ . \tag{2.13b}$$

The longitudinal impedance is obtained applying Eq. (1.5) remembering that $\kappa = \omega/\beta c$, $c\varepsilon_0 = Z_0^{-1}$ and $\tilde{E}(\kappa) = \beta c\tilde{E}(\omega)$:

$$Z(\omega) = \frac{i\omega Z_0}{4\pi c(\beta\gamma)^2}\left[1 + 2\ \ln\frac{b}{a}\right] \ ; \quad [\Omega/m] \tag{2.14}$$

(Z_0 = 377 Ω is the free-space impedance), which is the impedance per unit length commonly used[8]. It is worth noting that the impedance is purely reactive, as we would have expected, since there is no energy loss during the charge motion. An extension of Eq. (2.14) to an elliptic vacuum chamber is found in Refs. 9 and 10.

We consider now Eq. (2.2). In the limit $\gamma \to \infty$ we get the approximate expression:

$$\tilde{G}(\kappa,r) \sim \frac{-1}{2\pi\varepsilon_0}\ \ln\,(r/b) \quad . \tag{2.15}$$

By inverse Fourier transformation we obtain in the time domain:

$$g(r,z,t) = \frac{-1}{2\pi\varepsilon_0}\ \ln\,(r/b)\ \delta(z - \beta ct) \tag{2.16a}$$

$$E_z = O(\frac{1}{\gamma^2}) \tag{2.16b}$$

$$E_r = \frac{1}{2\pi\varepsilon_0 r}\ \delta(z - \beta ct) \tag{2.16c}$$

$$H_\phi = \frac{1}{2\pi r Z_0}\ \delta(z - \beta ct) \ . \tag{2.16d}$$

In the limit $\beta \to 1$, the fields are the same as in the free space, all the field components go to zero both ahead and behind the charge (Fig. 5), there are no forces acting on the test charge and according to Eqs. (1.4) and (2.3) the wake potential vanishes as $1/\gamma^2$ with the energy:

$$W(\tau) = \frac{g(\tau)}{\gamma^2} \ ; \quad \tau = (z\text{-}ct). \tag{2.17}$$

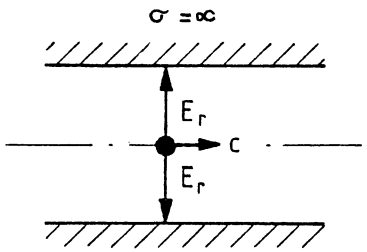


Fig. 5 The electric field in the limiting case $\beta = 1$

2.2 Cylindrical pipe with resistive walls

We consider now the more realistic case of walls with finite conductiviy σ [2,4,11-15]. If the skin depth is much smaller than the thickness of the pipe wall, the latter can be considered to be infinite (Fig. 6). We will see that due to the diffusion of the fields within the conductor, during the motion of the charge a non-zero wake field will appear behind the charge.

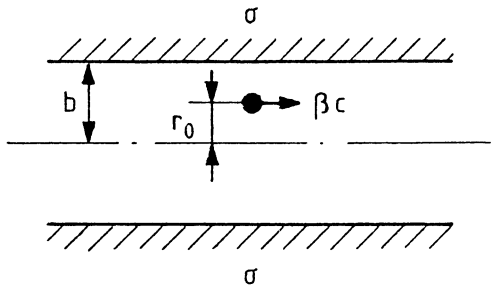


Fig. 6 Circular pipe of finite conductivity σ

In the region $r < b$, the equations satisfied by the potential and its spectrum are still (2.8) and (2.9); in the region $r > b$, which we label I, the scalar potential satisfies the diffusion equation:

$$\nabla^2 g^I - \frac{1}{v^2}\frac{\partial^2 g^I}{\partial t^2} + \mu\sigma\frac{\partial g^I}{\partial t} = 0 \tag{2.18}$$

Assuming also for g^I a multipole expansion (2.6), the Fourier transform $\tilde{G}^I_m(k)$ is a solution of the equation:

$$\frac{\partial^2 \tilde{G}^I_m}{\partial r^2} + \frac{1}{r}\frac{\partial \tilde{G}^I_m}{\partial r} - \left(\frac{m}{r^2} + \bar{q}^2\right)\tilde{G}^I_m = 0 \tag{2.19}$$

where $\bar{q}^2 = q^2 - i\kappa\beta Z_0\sigma$.

The boundary conditions require in this case the continuity of the field component $\tilde{E}_z$ and $\tilde{H}_\phi$ tangent to the pipe surface; within the conductor they are given by:

$$\tilde{E}_z(\kappa) = \frac{\bar{q}^2}{i\kappa}\tilde{G}^I \tag{2.20a}$$

$$\tilde{H}_\phi(\kappa) = \frac{\sigma - i\kappa v\varepsilon}{i\kappa v\varepsilon} \frac{\partial \tilde{G}^I}{\partial r} \,. \tag{2.20b}$$

The general solution of Eq. (2.19) is:

$$\tilde{G}^I_m(\kappa) = \frac{\alpha_m}{2\pi\varepsilon_0} C_m K_m(\bar{q}r) \tag{2.21}$$

Matching the fields at $r = b$, the constants B_m and C_m are uniquely determined. It is convenient to express the solution in the region $r < b$ as:

$$\tilde{G}_m(\kappa) = \tilde{G}^p_m(\kappa) + \tilde{G}^\sigma_m(\kappa) \tag{2.22}$$

where $\tilde{G}^p_m(\kappa)$ is the perfectly conducting solution given by Eq. (2.10) with Eq. (2.11). We obtain:

$$\tilde{G}^\sigma_m(\kappa) = -\frac{\alpha_m}{2\pi\varepsilon_0} \frac{\xi\,\bar{q}}{q^2 b I_m(qb)} \left[\frac{K_m(\bar{q}b) I_m(qr_0)}{\xi \frac{\bar{q}}{q} K_m(\bar{q}b) I'_m(qb) - I_m(qb) K'_m(qb)} \right] I_m(qr) \tag{2.23}$$

where $\xi = \left[1 + \frac{i\sigma}{\kappa v \varepsilon_0}\right]^{-1}$

Our attention is now focused on the term $m = 0$ needed to compute longitudinal wake and impedance. In the approximation $\kappa \ll \beta\gamma^2\sigma Z_0$ and $\kappa \ll \sigma Z_0/\beta$, i.e high conductivity and low frequency (or high energy), we have

$$\tilde{G}^\sigma_0(\kappa) \sim -\frac{\xi\bar{q}}{2\pi\varepsilon_0 q^2 b} \frac{I_0(qr)}{I^2_0(qb)} \cdot I_0(qr) \tag{2.24}$$

Using Eqs. (2.20a, 2.24, 1.3) in the limit $q \ll 1$, we get the further contribution to the impedance per unit length due to the conductivity σ:

$$Z^\sigma(\omega) = \frac{(1-i)}{2\pi b} \sqrt{\frac{\kappa\beta Z_0}{2\sigma}} \; ; \; [\Omega/m] \,. \tag{2.25}$$

The potential for $r_0 = 0$ in the time domain is obtained by transforming Eq. (2.24):

$$g^\sigma_0(z-vt) = \frac{1}{4\pi^2\varepsilon_0 b} \gamma^2\beta \sqrt{\frac{i\beta}{Z_0\sigma}} \int_{-\infty}^{\infty} \frac{e^{i\kappa(z-vt)} d\kappa}{\sqrt{\kappa}\, I^2_0(qb)} \,. \tag{2.26}$$

In the limit $\gamma \to \infty$ the integral may be solved analytically obtaining:

$$g^\sigma_0(z-ct) \sim \frac{\gamma^2}{2\pi\varepsilon_0 b} \sqrt{\frac{1}{\pi\sigma Z_0}} (z-ct)^{-1/2} \,. \tag{2.27}$$

The longitudinal wake potential per unit length is then (from Eq. (2.17)):

$$\frac{dW^\sigma_0}{dz} = \frac{1}{\gamma^2} \frac{\partial g^\sigma_0(z-ct)}{\partial z} = \frac{1}{4\pi\varepsilon b} \sqrt{\frac{1}{\pi\sigma Z_0}} |z-ct|^{-3/2} \tag{2.28}$$

where $z < ct$. Behind the charge the wake potential is independent of the energy; the trailing charge experiences an accelerating field and gains energy from the leading one. This is true only beyond a critical distance d_0 from the leading charge[2,4]

$$d_0 = \left(\frac{b^2}{Z_0\sigma}\right)^{1/3} . \tag{2.29}$$

At shorter distances the field changes sign and its behaviour is shown in Fig. 7.

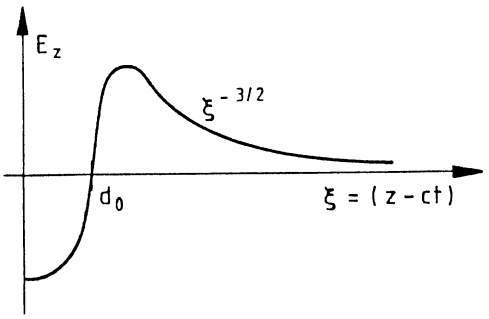


Fig. 7a The longitudinal field behind the source charge for a lossy pipe ($\beta = 1$)

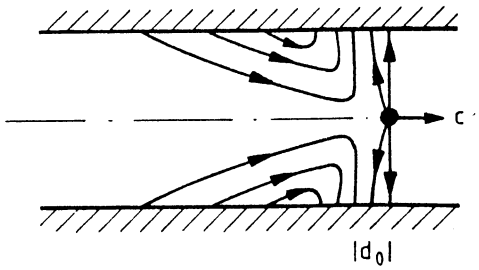


Fig. 7b Circular pipe of finite conductivity σ

Beyond the distance (2.29) the wake may be easily computed by asymptotic evaluation of the integral in Eq. (2.26). For a charge which is not necessarily ultrarelativistic we get the simple result:

$$g_0^{\sigma}(z-vt) \sim \frac{1}{2\pi^2\varepsilon_0 b}\,\beta\gamma^2\sqrt{\frac{\beta\gamma}{Z_0\sigma\, b}}\sum_0^{\infty} n\,\frac{G^n(o)}{n!}\,\mathrm{In}(\Omega) \tag{2.30}$$

where

$$\Omega = \frac{\gamma|z-vt|}{b} \tag{2.31}$$

$$G^n(o) = \left[\frac{d^n}{dx^n}\,\frac{1}{I_0^2(-x^2)}\right]_{x=0} \tag{2.32}$$

$$\mathrm{In}(\Omega) = \begin{cases} \dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\Omega\left(\frac{n-1}{2}\right)} & n = \text{even} \\ 0 & n = \text{odd} \end{cases} \tag{2.33}$$

where $\Gamma(x)$ is Euler's function. The first term of Eq. (2.30) for $n = 0$ gives again Eq. (2.27); the next non-zero term corresponding to $n = 2$ contributes to the wake behaving as $1/\gamma^2$ with the energy on $|z-\beta ct|^{-7/2}$ with the distance. This makes Eq. (2.28) a very good approximation also for non-ultrarelativistic charges.

2.3 RF cavity: impedance and wake potential

One of the main contributions to the impedance of an accelerator is due to the accelerating cavities[2-4]). These are usually schematized by the resonant circuit of Fig. 8 whose impedance is given by

$$Z(\omega) = \frac{R}{1 + iQ_r \left(\frac{\omega_r}{\omega} - \frac{\omega}{\omega_r}\right)} \tag{2.34}$$

where $\omega_r = (LC)^{-1/2}$ is the rsonant frequency, $Q_r = R(C/L)^{1/2}$ is the quality factor and R is the shunt impedance whose meaning will be explained hereafter.

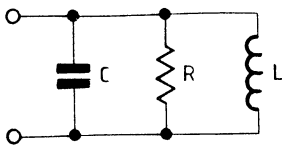


Fig. 8 Parallel resonant circuit

When Q_r is very high, the impedance is of narrow band type with a bandwidth given approximately by $\Delta\omega \sim \omega_r/2Q_r$. The impedance modulus and phase of a typical resonator are shown in Fig. 9. Below (above) the resonance, the impedance has an inductive (capacitive) imaginary part while at the resonance it is purely resistive. A real cavity exhibits many such sharp resonances corresponding to all the resonant modes of the structure. The fundamental mode is the "accelerating" one which furnishes energy to the beam; all the others being "parasitic modes". When the beam crosses the cavity it gains energy from the accelerating voltage but it will also loose energy by exciting the resonances of the cavity. The energy left by a charge Q is often expressed as $U = kQ^2$ where the parameter k is called the "loss parameter" and is different for each resonant mode. The cavity then stores e.m. energy which can influence the dynamics of a forthcoming bunch.

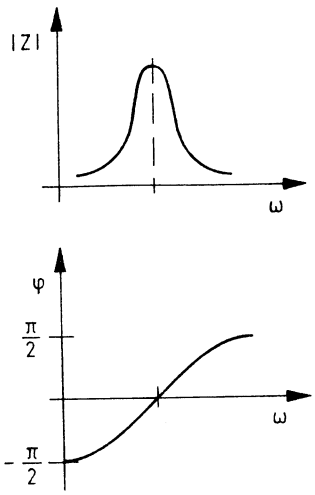


Fig. 9 Impedance amplitude and phase of a resonator

At frequencies beyond the cut-off value of the pipe, the fields excited will flow into the vacuum chamber and the impedance is then better represented by a continuous wider spectrum[2] (see Fig. 10).

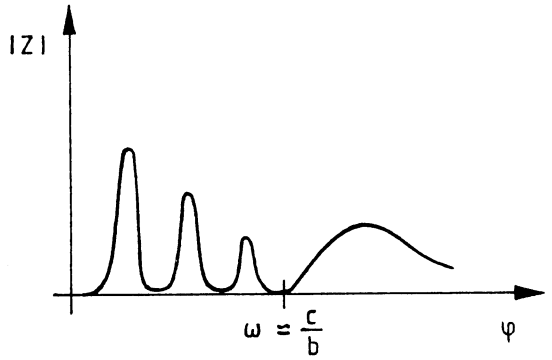


Fig. 10 Typical behaviour of the impedance of an RF cavity

For the accelerating mode the shunt impedance is usually defined as

$$R = \frac{V_a^2}{P_d} \tag{2.35}$$

where V_a is the applied voltage and P_d is the power dissipated into the cavity walls. The above definition, expressed in terms of the voltage induced by the beam, gives the shunt impedance for the parasitic modes. The power lost can be calculated as the power dissipated on the surface resistance

$$Rw = \sqrt{\frac{\omega\mu}{2\sigma}} \tag{2.36}$$

by the current flowing on the cavity walls. Expressing the currents in terms of the magnetic field H_t tangential to the surface we get:

$$P_d = \frac{1}{2} Rw \int_{surface} H_t^2 \, ds \, . \tag{2.37}$$

The shunt impedance is sometimes calculated by considering the voltage across the gap to be constant, ignoring the variation of the field acting on the crossing particle during its flight:

$$V_g = \int_{gap} E_z \, dz \, . \tag{2.38}$$

In this case we need to introduce a corrective factor called the "transit-time factor" defined by

$$T = \frac{V_a}{V_g} = \frac{\left| \int_{gap} E_z \, e^{i\kappa z} \, dz \right|}{\int_{gap} E_z \, dz} \tag{2.39}$$

so that the shunt impedance is expressed as:

$$R = \frac{V_g^2 T}{P_d} . \tag{2.40}$$

If the field across the gap of length "ℓ" is uniform we get:

$$T = \frac{1}{\ell} \int_{-\ell/2}^{\ell/2} e^{i\kappa z}\, dz = \frac{\sin(\kappa\ell/2)}{(\kappa\ell/2)} . \tag{2.41}$$

Note that the transit time approaches unity for those frequencies whose wavelengths are much bigger than the gap.

Finally it is interesting to find the wake potential by calculating the Fourier anti-transform of the impedance (Fig. 11)

$$W(\tau) = W_0\, e^{-\omega_r \tau/2Q_r}\left[\cos \tilde{\omega}\tau - \frac{1}{2Q_r} \sin \tilde{\omega}\tau\right] H(\tau) \tag{2.42}$$

where $H(\tau)$ is the Heaveside step function, $W_0 = \omega_0 R/2Q_r$ and

$$\tilde{\omega} = \omega_r \sqrt{1 - (2Q_r)^{-2}} . \tag{2.43}$$

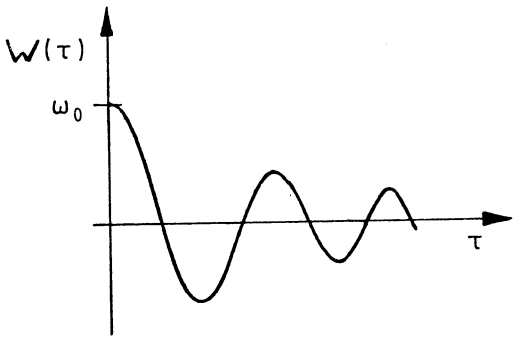


Fig. 11 Wake potential for a resonator

After the passage of a bunch one has to wait a number Q_r/π of periods $T_r = 2\pi/\omega_r$ before the wake potential reduces to W_0/e. Since Q_r is usually very high ($\sim 10^4$) it is clear that the resonator is the main cause of coupling among the bunches circulating around the ring.

Assuming a very high Q_r mode so that $\tilde{\omega} \sim \omega_r$ we get:

$$W(\tau) \sim \frac{\omega_r R}{2Q_r} \cos \omega_r \tau . \tag{2.44}$$

To this wake corresponds a δ-function impedance at $\omega = \omega_r$:

$$Z(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\tau) e^{i\omega\tau} \, d\tau = \frac{\omega_r R}{4Q_r} \delta(\omega - \omega_r) \ . \tag{2.45}$$

The total energy lost into this impedance is given by [see Eq. (3.17)]

$$U = \int_{-\infty}^{\infty} |I_0(\omega)|^2 Z(\omega) \, d\omega \ . \tag{2.46}$$

For a point charge $|I_0(\omega)| = Q$ therefore, we get:

$$U = Q^2 \cdot \frac{\omega_r R}{4Q_r} \tag{2.47}$$

and a loss factor which is:

$$\kappa_r = \frac{U}{Q^2} = \frac{\omega_r R}{4Q_r} \ . \tag{2.48}$$

Finally from Eq. (2.44) we have:

$$W(\tau) = 2\kappa_r \cos \omega_r \tau \ . \tag{2.49}$$

Extending over all the modes of the resonant structure we have the following expression of the wake potential derived from the loss parameter and the resonant frequencies only[16-20]

$$W(\tau) = 2 \sum_r \kappa_r \cos \omega_r \tau \ . \tag{2.50}$$

To compute the wake potential we need the values of κ_n and ω_n for as many modes as possible. This can be achieved by means of numerical programs[22,25,26]. The programs KN7C and SUPERFISH are commonly used to this end but because of computer time consumed they are limited in the total number of modes that can be calculated. For the higher modes the optical resonator model[21-24] is generally adopted and predicts a loss parameter per unit frequency

$$\frac{d\kappa}{d\omega} = \frac{A_0}{\omega^{3/2}} \ . \tag{2.51}$$

The wake potential is then calculated by:

$$W(\tau) = 2 \sum_{1}^{m} \kappa_r \cos \omega_r \tau + A_0 \int_{\omega_m}^{\infty} \frac{\cos \omega\tau}{\omega^{3/2}} \, d\tau \ . \tag{2.52}$$

or

$$W(\tau) = 2 \sum_{1}^{m} \kappa_r \cos \omega_r \tau + \frac{4A_0}{\omega_m^{3/2}} \left\{ \cos x - \frac{\pi x}{2} \left[1 - 2S\left(\frac{2x}{\pi} \right) \right] \right\}_{x = \omega_m \tau} \tag{2.53}$$

where S(x) is the integral sinus function.

For Gaussian bunches, one can define a wake potential per unit charge which is an integrated effect at a time t over the charge distribution:

$$\bar{W}(t) = \int_{-\infty}^{t} W(t-\tau)\, I(\tau)\, d\tau \ . \tag{2.54}$$

In this case a direct calculation of the wake $\bar{W}(t)$ in the time domain can be performed by using the code BCI [27,28].

2.4 Multipole longitudinal wake potential

So far we have defined the longitudinal wake as the potential seen by a test particle trailing a leading one, both moving on the same trajectory with a velocity βc, and at a distance $\beta c\tau$ apart. We have then considered the case of boundaries with cylindrical symmetry assuming the charges to be moving on the axis, deriving the expression of the wake potential that is usually referred to as "zero-mode" wake. We remember, in fact, that in the derivation of the wake potential for perfectly conducting and lossy cylindrical pipes, we have made use only of the term m = 0 of the expansions (2.10) and (2.23), the only surviving one when we consider the particle on the axis ($r_0 = 0$).

Nonetheless, we may also talk of longitudinal wake in the general case of two particles travelling on parallel trajectories which are not necessarily coincident with the axis. The wake potential per unit length is then defined in terms of the longitudinal electric field E_z at a point P distinct from P_0 where the source charge is located (Fig. 12).

A charge moving off-axis can be represented by a superposition of multipole moments. One can see by inspection of Eq. (2.5) that the m-th multipole charge density is:

$$\rho_m = \frac{I_m}{\pi r_0^{m+1}}\, \delta(z-\beta ct)\delta(r-r_0)\cos[m(\theta-\theta_0)] \tag{2.55}$$

where I_m is the m-th moment of the charge distribution. In the simple case of point charge (our case) $I_m = \alpha_m r_0^m$. According to the above expression of ρ_m the charge is distributed as an infinitesimally thin ring of radius r_0 with a $\cos[m(\theta-\theta_0)]$ angular dependence.

Because of the linearity of the wave equation (2.2) the scalar potential is expressed by an expansion of multipole terms [see (2.16)] and all the quantities derived from the potential by means of a linear (differential) operator will have a multipole expansion as well.

We have already exploited these properties introducing Eq. (2.6) where the scalar potential $g(P/P_0)$ has been expanded as

$$g(P/P_0) = \sum_{0}^{\infty}{}_m\, g_m(r,r_0,\tau)\cos[m(\theta-\theta_0)] \ . \tag{2.56}$$

Now a similar expansion can be given also for the wake potential:

$$W(P/P_0) = \sum_0^\infty{}_m W_m(P/P_0) \tag{2.57}$$

where the multipole wake term is related to the multipole scalar potential by the simple relation:

$$W_m = (P/P_0) = \frac{1}{\gamma^2} g_m(r,r_0,\tau) \cos[m(\theta-\theta_0)] \ . \tag{2.58}$$

Thus whenever there is a charge distribution term like (2.55) in the beam, there will be a corresponding term in the scalar potential and in the wake potential expansions as well.

The modes m = 0, 1, 2 are called respectively monopole, dipole and quadrupole modes. It is apparent that when the charge is on the axis we will have a pure monopole term m = 0; analogously, a dipole longitudinal wake corresponds to a pure dipole charge distribution m = 1.

Example: Find the longitudinal dipole wake for a lossy pipe.

Consider the term m = 1 of the expansion (2.23). Under the same approximations of high conductivity and high γ values, we get in the frequency domain:

$$\tilde{G}_1^\sigma(\kappa) = \frac{-\xi\bar{q}}{\pi\varepsilon_0 q^2 b} \frac{I_1(qr_0)I_1(qr)}{I_1^2(qb)} \ . \tag{2.59}$$

Assuming that qr and qr_0 are sufficiently small to allow the approximate expressions of the modified Besse functions $I_1(x) \simeq x/2$ we get the following simple expression:

$$\tilde{G}_1^\sigma(\kappa) \simeq \frac{2rr_0}{b^2} \tilde{G}_0^\sigma(\kappa) \tag{2.60}$$

where $\tilde{G}_0^\sigma(\kappa)$ is the term m = 0 of (2.23) already considered and whose Fourier transform into the the real space is given by Eq. (2.27). It is also quite straightforward to transform (2.60) obtaining:

$$g_1^\sigma(r,r_0,\tau) = \frac{2rr_0}{b^2} g_0^\sigma(r,r_0,\tau) \ . \tag{2.61}$$

Applying (2.58) one finds the corresponding longitudinal wake potential:

$$W_1^\sigma(P/P_0) = \frac{2rr_0}{b^2} \cos(\theta-\theta_0) W_0^\sigma(P/P_0) \ . \tag{2.62}$$

It is interesting to note that the dipole wake vanishes at the point r = 0 because of the cylindrical symmetry, and everywhere for a degenerate dipole ($r_0 = 0$).

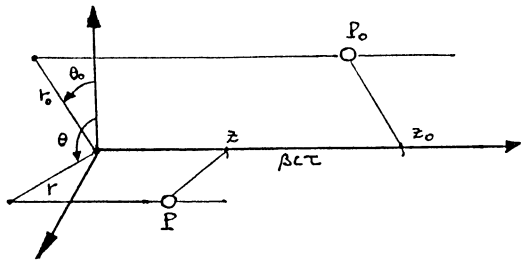


Fig. 12 Cylindrical coordinates of the source point P_0 and test point P

3. GENERAL ANALYSIS

3.1 Piecewise homogeneous pipe with lumped discontinuities

Let us consider a piecewise homogeneous structure with lumped discontinuities in the electromagnetic properties which are located at the coordinates z_i (Fig. 13). This is a more realistic model of an accelerator ring which indeed consists of a pipe of uniform cross section interrupted by discontinuities such as cavities, bellows, pick-ups, etc. We shall show a general method for studying the e.m. interaction of a particle bunch with such a structure.

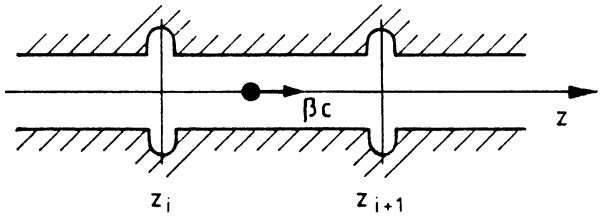


Fig. 13 Discontinuities along the pipe

The current $I(z,t)$ and the corresponding linear charge density $\lambda(z,t)$ are the source of the e.m. interaction. Assuming that all the particles within the bunch have the same velocity βc, current and charge densities may be written in the following way:

$$I(z,t) = I_0(z-\beta ct) \tag{3.1}$$

$$\lambda(z,t) = I_0(z-\beta ct)/\beta c \ . \tag{3.2}$$

The current spectrum in the frequency domain is

$$\tilde{I}(z,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} I_0(z-\beta ct)\, e^{i\omega t}\, dt = \tilde{I}_0(\omega)\, e^{i\kappa_0 z/\beta} \tag{3.3}$$

where

$$\tilde{I}_0(\omega) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} I_0(-\beta ct)\, e^{i\omega t}\, dt \tag{3.4}$$

is the Fourier transform of the current signal measured at $z = 0$, and $k_0 = \omega/c$ is the free-space wave number. It is clear from Eq. (3.3) that to each harmonic of angular frequency ω corresponds one, and only one, space harmonic of wave number κ_0/β and all the spectrum components have the phase velocity βc equal to the group velocity of the particles.

Allowing for only one discontinuity located at $z = z_0$, let us focus attention on the longitudinal electric field $\tilde{E}_z(z,\omega)$ generated from the e.m. interaction of one harmonic Eq. (3.3) with the surrounding structure. The space harmonic content of the e.m. fields will exhibit a continuous spectrum (Fig. 14). In this case, in fact, the fields do not travel solidly with the source but propagate within the pipe.

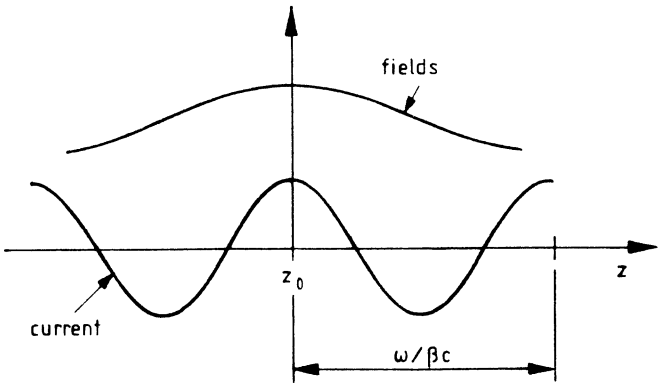


Fig. 14 A monochromatic current perturbation excites fields with a wide space harmonic spectrum

The frequency spectrum structure is peculiar to the characteristics of the discontinuity and of the medium. Because of the assumed linearity the general expression of E_z is:

$$\tilde{E}_z(\omega,z) = \tilde{I}_0(\omega)\, e^{i\kappa_0 z/\beta}\, \tilde{G}_e(z-z_0,\omega); \quad \left(\frac{V}{m}\right) \tag{3.5}$$

where $\tilde{G}_e(z-z_0,\omega)$ is the Green's function for the E_z component in the ω-domain and is measured in Ω/m.

We may now express the Green's function in terms of its Fourier transform in the wave number domain:

$$\tilde{\tilde{G}}(\kappa,\omega) = \mathfrak{F}_z[\tilde{G}(z,\omega)] \; . \tag{3.6}$$

Equation (3.5) then becomes:

$$\tilde{E}_z(\omega,z) = \tilde{I}_0(\omega)\, e^{i\kappa_0 z/\beta} \int_{-\infty}^{\infty} \tilde{\tilde{G}}(\kappa,\omega)\, e^{+i\kappa(z-z_0)}\, d\kappa \; . \tag{3.7}$$

Finally we get in the time-space domain:

$$E_z(z,t) = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\kappa \, \tilde{I}_0(\omega) \, e^{i(\kappa_0/\beta - \kappa)z_0} \, \tilde{\tilde{G}}(\kappa,\omega) \, e^{-i(\omega t - \kappa z)} \tag{3.8}$$

while the analogous expression for the current is:

$$I(z,t) \equiv I_0(z-\beta ct) = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\kappa \, \tilde{I}_0(\omega) \, e^{i(\kappa_0/\beta - \kappa)z} \, \delta(\kappa - \kappa_0/\beta) \, e^{i(\kappa z - \omega t)} \, . \tag{3.9}$$

From Eqs. (3.8) and (3.9) we see that in the $(\omega$-$\kappa)$ domain the function $\tilde{\tilde{G}}(\kappa,\omega)$ can be thought of as the transfer function, i.e. the response to the impuslive function $\delta(\kappa - \omega/\beta c)$. We have to stress that the δ-function has been inserted in Eq. (3.9) in order to get a suitable expression of the current which could be instructively compared with the field spectrum. For the current there is no distinction between the ω and and κ domains as previously mentioned; at each frequency ω corresponds only one wave number $\kappa = \omega/\beta c \; \Delta \; \kappa_0/\beta$ for the current but a continuous spectrum for the fields (Fig. 15).

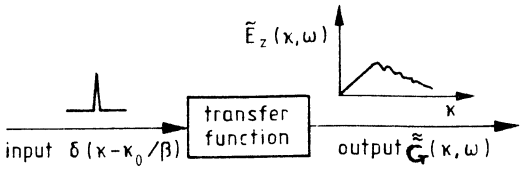


Fig. 15 Frequency response $\tilde{\tilde{G}}(\kappa,w)$ to a monochromatic excitation

In other words we may say that while the current has a spectrum of waves all having the same phase velocity βc, the fields have in their spectrum waves of any phase velocity. This is shown diagrammatically in Fig. 16.

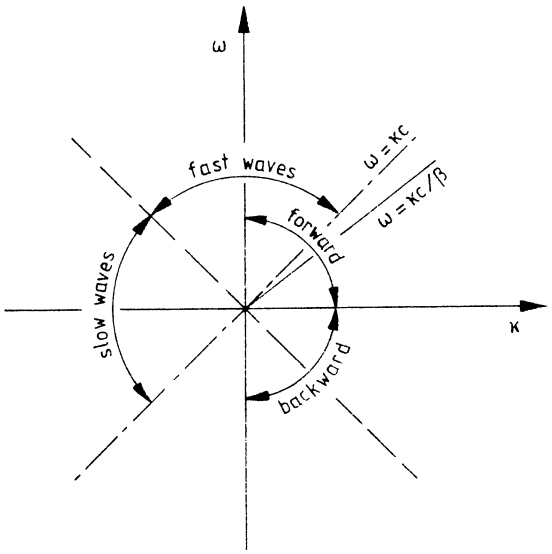


Fig. 16 Wave characteristics in the (ω,κ) plane

Now the following question arises: what is the interaction between the current and the longitudinal electric field, i.e. the amount of the energy losses? Intuitively, among the frequencies of the field spectrum only the wave with the same phase velocity $v_\phi = \beta c$ will contribute to the energy losses. All the other waves will exchange zero mean energy with the current. This will be exactly proved below.

The energy losses may be evaluated as the work done by the charged particles against the e.m. forces. In terms of current and fields we have:

$$U = - \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dz \; I_0(z-\beta ct) \; E_z(z,t) \; . \tag{3.10}$$

On the other hand, the field is generated by the current I_0 and may be obtained as convolution of the Green function over the current distribution. [Note that Eq. (3.11) is the time-space equivalent of Eq. (3.5) and can be achieved from this last.)

$$E_z(z,t) = \int_{-\infty}^{\infty} I(z-\beta ct) \; g(z-z,t) \; dz \; . \tag{3.11}$$

By applying the Fourier transform to the convolution (3.11) and using Eq. (3.10) we obtain the very interesting result:

$$U = - \int_{-\infty}^{\infty} \left|\tilde{I}_0(\omega)\right|^2 \; \tilde{\tilde{G}}(\omega, \kappa = \omega/\beta c) \; d\omega \tag{3.12}$$

which confirms the fact that only the waves with the same phase velocity as the current may interact with it and cause the losses of energy. It is quite straightforward from Eq. (3.12) to define the longtudinal impedance as[34]:

$$Z_\ell(\omega) = -\tilde{\tilde{G}}(\omega, \kappa = \omega/\beta c) \; . \tag{3.13}$$

It is worth noting th t by computing Eq. (1.3) over an infinite interval and noting that $\tilde{E}_z(\omega,z)/Q = \tilde{G}(\omega,z)$ we get the same expressions of Eq. (3.13).

3.2 Parseval theorem

The previous result can be obtained by applying the Parseval theorem to Eq. (3.10); we get:

$$U = - \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\kappa \; \tilde{\tilde{E}}(\kappa,\omega) \; \tilde{\tilde{I}}^*(\kappa,\omega) \; . \tag{3.14}$$

From Eqs. (3.7) and (3.8) we have:

$$\tilde{\tilde{I}}(\kappa,\omega) = \tilde{I}_0(\omega)\ \delta(\kappa-\omega/\beta c) \tag{3.15}$$

$$\tilde{\tilde{E}}(\kappa,\omega) = \tilde{I}_0(\omega)\ \tilde{\tilde{G}}(\kappa,w) \tag{3.16}$$

which substituted into Eq. (3.14) leads to Eq. (3.12)

$$U = \int_{-\infty}^{\infty} d\omega \left|\tilde{I}_0(\omega)\right|^2 Z\ (\omega)\ . \tag{3.17}$$

4. CIRCULAR ACCELERATORS

For circular accelerators it is useful to find definitions which take into account the intrinsic periodicity of the structure and of the sources. The e.m. problem may be schematized as an infinite pipe which indefinitely repeats the same piece of structure.

If R is the nominal radius of the machine, the current is obtained by summing up the contributions of currents of periodicity $2\pi R$ (Fig. 16):

$$I(z,t) = \sum_{-\infty}^{\infty}{}_p\ I_0(z-\beta ct + 2p\pi R)\ . \tag{4.1}$$

The spectrum of the above current distribution is:

$$I(z,\omega) = \tilde{I}_0(\omega)\ e^{i\omega z/\beta c} \sum_{-\infty}^{\infty}{}_p\ e^{i\omega 2p\pi R/\beta c} \tag{4.2}$$

where $I_0(\omega)$ has the same definition as in Eq. (3.4).

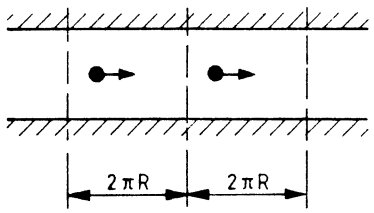


Fig. 17 Unfolding a circular accelerator into a periodic structure

Defining $\omega_0 = \beta c/R$ and noting that

$$\sum_{-\infty}^{\infty}{}_p\ e^{ip2\pi\omega/\omega_0} = \omega_0 \sum_{-\infty}^{\infty}{}_p\ \delta(\omega - p\omega_0) \tag{4.3}$$

the spectrum of the current will be represented by an infinite set of discrete frequencies

$$I(z,\omega) = \tilde{I}_0(\omega)\, e^{i\omega z/\beta c}\, \omega_0 \sum_{-\infty}^{\infty}{}_p\, \delta(\omega - p\omega_0) \ . \tag{4.4}$$

We can go back to the time domain by using the Fourier integral transform:

$$I(z,t) = \omega_0 \int_{-\infty}^{\infty} \tilde{I}_0(\omega)\, e^{-i\omega(t-z/\beta c)} \sum_{-\infty}^{\infty}{}_p\, \delta(\omega - p\omega_0)\, d\omega \ . \tag{4.5}$$

How can we find the fields excited by the above current distribution within a pipe with periodic boundary properties? We shall proceed in two steps:

i) obtain the field E_z^0 for an infinite structure with only one discontinuity but loaded by the periodic source Eq. (4.5),

ii) repeat the field obtained in i) for the real structure.

The first step can be easily achieved by making use of the Green's function $g(z,t)$ introduced in Section 3, and of the convolution theorem applied to the periodic current distribution. In the frequency domain we get:

$$E_z^0(z,\omega) = \tilde{I}_0\, e^{i\omega z_0/\beta c}\, \omega_0 \sum_{-\infty}^{\infty}{}_p\, \delta(\omega - p\omega_0)\, \tilde{g}(z-z_0,\omega) \ . \tag{4.6}$$

Assuming that there is no reciprocal interference between the discontinuities (this assumption is certainly true at frequencies below the cut-off of the pipe) the field excited in the independently repeating structure can be obtained by overlapping Eq. (3.6) for the whole set of discontinuities displaced by a distance $2\pi qR$:

$$E_z(z,\omega) = \tilde{I}_0(\omega)\, \omega_0 \sum_q \sum_p e^{i\omega(z_0+2\pi qR)/\beta c}\, \delta(\omega - p\omega_0)\, \tilde{g}(z-z_0-2\pi qR,\omega) \ . \tag{4.7}$$

Expressing the function g in terms of its integral representation

$$E_z(z,\omega) = \tilde{I}_0(\omega)\, \omega_0\, e^{i\omega z_0/\beta c} \sum_{-\infty}^{\infty}{}_p \left[\frac{\delta(\omega-p\omega_0)}{2\pi} \int_{-\infty}^{\infty} \tilde{g}(\kappa,\omega)\, e^{i\kappa(z-z_0)} \sum_q e^{i(\omega/\beta c-\kappa)2\pi Rq}\, d\kappa \right. \tag{4.8}$$

where we recognize that

$$\sum_q e^{i(\omega/\beta c-\kappa)2\pi Rq} = \frac{1}{R} \sum_{-\infty}^{\infty}{}_q\, \delta\left(\frac{\omega-q\omega_0}{\beta c} - \kappa\right) \ . \tag{4.9}$$

Therefore Eq. (3.8) becomes:

$$\tilde{E}_z(z,\omega) = \frac{\omega_0}{2\pi R}\,\tilde{I}_0(\omega)\; e^{i\omega z_0/\beta c} \sum_{p,q} \delta(\omega - p\omega_0)\; \left[\tilde{\tilde{G}}\left(\frac{\omega - q\omega_0}{\beta c}, \omega\right)\; e^{i(\omega - q\omega_0)(z-z_0)/\beta c}\right] \tag{4.10}$$

and back into the time domain:

$$E_z(z,t) = \frac{\omega_0}{2\pi R} \sum_{-\infty}^{\infty}{}_p\; \tilde{I}_0\,(p\omega_0)\; e^{-ip\omega_0(t - z_0/\beta c)} \sum_{-\infty}^{\infty}{}_q\; \tilde{\tilde{G}}\left[\frac{(p-q)}{\beta c}\,\omega_0, p\omega_0\right]\; e^{i(p-q)\omega_0(z-z_0)/\beta c}\;. \tag{4.11}$$

The field is then represented by a superposition of waves with phase velocity

$$v_\phi = \beta c p/(p-q)\;.$$

However, we have learned that only the component with the same phase velocity of the current may exchange energy with it; therefore we will require that

$$\frac{p}{p-q} = 1 \tag{4.12}$$

which is satisfied for $q = 0$ and any p, namely

$$\langle E_z(z,t)\rangle = \frac{\omega_0}{R} \sum_p \tilde{I}_0(p\omega_0)\; e^{-ip\omega_0(t - z/\beta c)} \cdot \tilde{\tilde{G}}_e\left(\frac{p\omega_0}{\beta c}\,,\; p\omega_0\right) \tag{4.13}$$

where again we can define:

$$Z\,(p\omega_0) = \tilde{\tilde{G}}\left(\frac{p\omega_0}{\beta c},\; p\omega_0\right)\;. \tag{4.14}$$

5. TRANSVERSE-WAKE POTENTIAL AND IMPEDANCE

5.1 Transverse-wake potential

In Section 2.4 we have introduced the multipole longitudinal wake allowing the source and test charges to be on two parallel trajectories, not coincident with the axis (we assume a circular cylindrical symmetry). In this case the trailing charge may also experience a transverse deflecting force during motion. The transverse kicker per unit charge and unit length of the trajectory assumed to be unperturbed defines the transverse wake potential per unit length:

$$\frac{d\bar{W}_\perp(P/P_0)}{dz} = \frac{\bar{E}_\perp(P/P_0) + \bar{v}\times\bar{B}(P/P_0)}{Q}\;. \tag{5.1}$$

The transverse force has to be evaluated in a reference frame which is moving with the particles. In such a frame the Panofsky-Wenzel theorem[36] applies and relates the longitudinal gradient of the transverse force to the transverse gradient of the longitudinal force. In terms of fields it states:

$$\frac{\partial}{\partial z}\left[\bar{E}_{\perp} + \bar{v}\times\bar{B}\right] = \nabla_t E_z \ . \tag{5.2}$$

Thus the transverse wake (5.1) can be given in terms of the longitudinal electric field:

$$\frac{d}{dz}\left[\bar{W}_{\perp}(P/P_0)\right] = \frac{\nabla_t}{Q}\int E_z(P/P_0)dz \tag{5.3}$$

and resorting to the definition of longitudinal wake we get:

$$\frac{d}{dz}\left[\bar{W}_{\perp}(P/P_0)\right] = \nabla_t W(P/P_0) \tag{5.4}$$

which expresses a differential linear relationship between transverse and longitudinal wakes.

Similarly to the longitudinal case, we may write the transverse wake as a sum of multipole terms

$$\bar{W}_{\perp}(P/P_0) = \sum_{0}^{\infty}{}_m \ \bar{W}_{\perp\, m}(P/P_0) \tag{5.5}$$

each of them being related to the corresponding longitudinal term by

$$\frac{d}{dz}\left[\bar{W}_{\perp\, m}(P/P_0)\right] = \nabla_t W_m(P/P_0) \ . \tag{5.6}$$

Example : Find the transverse dipole wake for a lossy pipe

The longitudinal dipole wake for a lossy pipe has been found to be given by Eq. (2.62):

$$W_1^{\sigma}(P/P_0) = \frac{2rr_0}{b^2}\cos(\theta-\theta_0)W_0^{\sigma}(P/P_0) \tag{5.7}$$

where W_0^{σ} is obtained by integrating Eq. (2.28). According to (5.6) we find the transverse dipole wake applying to the transverse gradient to Eq. (5.7) yielding:

$$\frac{d}{dz}\left[\bar{W}_{\perp}{}^{\sigma}{}_1(P/P_0) = \frac{2r_0}{b^2}\left[\hat{r}\cos(\theta-\theta_0) + \hat{\theta}\sin(\theta-\theta_0)\right]W_0^{\sigma}(P/P_0)\right. \tag{5.8}$$

$$\frac{d}{dz}\left[\bar{W}_{\perp}{}^{\sigma}{}_1(P/P_0) = \frac{2\bar{r}_0}{b^2}\ W_0^{\sigma}(P/P_0)\right. \tag{5.9}$$

where $\bar{r}_0$ is the dipole vector.

Returning to Eq. (2.28) one gets:

$$\frac{d}{dz}\left[\bar{W}_{\perp 1}^{\sigma}(P/P_0)\right] = \frac{\bar{r}_0}{\pi\varepsilon_0 b^3}\sqrt{\frac{1}{\pi\sigma Z_0}}\ \left|z-ct\right|^{-1/2} . \tag{5.10}$$

It is interesting to see that the transverse dipole wake decays only as $|z-ct|^{-1/2}$ (weaker than the longitudinal dipole wake). Furthermore, we call the reader's attention to the fact that the transverse wake is everywhere parallel to the dipole vector $\bar{r}_0$. This is a consequence of the particular symmetry of the boundaries. In general, the transverse wake is represented by a vector which is not collinear with the dipole and is connected to the latter by means of a matrix operator[9]. It would be more convenient and elegant to call the transverse dipole wake this matrix which does not depend on the particular dipole direction but is peculiar to the boundary shape. Accordingly, one can formally write:

$$\frac{d}{dz}\left[\bar{W}_{\perp}(P/P_0)\right] = \frac{d}{dz}\left[\bar{\bar{W}}_{\perp}(P/P_0)\right]\cdot\bar{r}_0 . \tag{5.11}$$

For the example presented here the matrix operator is given by the product of the scalar (5.10) by the unity matrix $\bar{\bar{I}}$:

$$\frac{d}{dz}\left[\bar{\bar{W}}_{\perp}(P/P_0)\right] = \frac{1}{\pi\varepsilon_0 b^3}\sqrt{\frac{1}{\pi\sigma Z_0}}\ \left|z-ct\right|^{-1/2}\ \bar{\bar{I}} . \tag{5.12}$$

In rectangular coordinates $\bar{\bar{I}} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$.

The property described in the last part of the example has a general validity as will now be shown. The longitudinal electric field created by a charge slightly displaced from the axis of a quantity $\bar{r}_0$ may be approximated to:

$$E_z(P/P_0) \simeq E_z(P,0) + \left\{\left[\nabla_{to}E_z(P/P_0)\right]\right\}_{\bar{r}_0\to 0}\cdot\ \bar{r}_0 + O(r_0^2) \tag{5.13}$$

where ∇_{to} is the gradient operator acting on the transverse coordinates of P_0 in the chosen reference system.

Substituting Eq. (5.13) into Eq. (5.3) we get:

$$\frac{d}{dz}\left[\bar{W}_{\perp}(P/P_0)\right] \simeq \frac{\nabla_t}{Q}\left\{\int E_z(P/0)dz + \left\{\left[\nabla_{to}\int E_z(P/P_0)dz\right]\right\}_{\bar{r}_0\to 0}\cdot\ \bar{r}_0 + O(r_0^2)\right\} \tag{5.14}$$

or

$$\frac{d}{dz}\left[\bar{W}_{\perp}(P/P_0)\right] \simeq \nabla_t\left\{W(P/0) + \left\{\left[\nabla_{to}W(P/P_0)\right]\right\}_{\bar{r}_0\to 0}\cdot\ \bar{r}_0 + O(r_0^2)\right\} . \tag{5.15}$$

The first term of this expression is the monopole contribution to the transverse force which disappears for particular geometries (circular pipe), the latter is the dipole term, which in the linear approximation may be written as the product of a matrix operator and the dipole vector:

$$\frac{d}{dz}\ [\bar{W}_{\perp 1}(P/P_0)] \simeq \frac{d}{dz}[\bar{\bar{W}}_{\perp 1}(P/P_0)] \cdot \bar{r}_0 \tag{5.16}$$

where the matrix wake is

$$\frac{d}{dz}\ [\bar{\bar{W}}_{\perp 1}(P/P_0)] = \nabla_t\ [\nabla_{to}\ W(P/P_0)]\ \bar{r}_0 \to 0\ . \tag{5.17}$$

This differential relationship between the dipole transverse wake and the longitudinal wake is quite general.

5.2 Transverse impedance

As for the longitudinal case, the transverse impedance is the frequency image of the transverse wake potential. In the literature this quantity is defined as the ratio between the frequency component of the transverse force amplitude due to a transverse dipole and the dipole moment. It is apparent that this definition introduces the impedance as a scalar quantity whereas, analogously to the transverse wake, it would be more convenient to stress its intrinsic matrix behaviour. In fact with reference to Eq. (5.1) the transverse force per unit charge (in the frequency domain) may be written as:

$$\tilde{\bar{E}}_{\perp}(\bar{r},\bar{r}_0) + \bar{v}x\tilde{\bar{B}}(\bar{r},\bar{r}_0) = \tilde{I}_0 \mathcal{F}_\tau\ [\frac{d}{dz}\ \bar{\bar{W}}_{\perp}(P/P_0)] \cdot \bar{r}_0 \tag{5.18}$$

and defining transverse impedance as the quantity

$$\bar{\bar{Z}}_{\perp}(\bar{r},\bar{r}_0) = i\,\mathcal{F}_\tau\ [\ \frac{d}{dz}\ \bar{\bar{W}}_{\perp}\ (P/P_0)]\beta^{-1}\ ;\ [\Omega/m^2] \tag{5.19}$$

we get

$$\bar{\bar{Z}}_{\perp}(\bar{r},\bar{r}_0) \cdot \bar{r}_0 = -i\ [\ \tilde{\bar{E}}_{\perp}(\bar{r},\bar{r}_0) + \bar{v}x\ \tilde{\bar{B}}(\bar{r},\bar{r}_0)]\ \beta^{-1}\ I_0^{-1} \tag{5.20}$$

which states that the impedance is a matrix operator relating the transverse force to the transverse dipole. Obviously for particular geometries such as a circular cross section pipe, the matrix nature of this operator is not apparent since the transverse force is independent of the dipole direction. Using Eqs. (5.17) and (5.19) we may also write:

$$\bar{\bar{Z}}_{\perp}(\bar{r},\bar{r}_0) = \frac{c}{\omega}\ \nabla_t \nabla_{to}\ Z(\bar{r},\bar{r}_0) \tag{5.21}$$

where with $Z(\bar{r},\bar{r}_0)$ we mean a generalized longitudinal impedance which is derived from the longitudinal wake allowing the two points P and P_0 to be on different trajectories. The transverse and longitudinal impedances so defined merge into the commonly used ones by making the limit $\bar{r} \to \bar{r}_0$. The meaning of the relations introduced here will be better clarified by the following applications.

Example: Find the transverse dipole impedance for a lossy circular pipe.

In this case we find the impedance directly by transforming the dipole transverse wake, i.e. by applying (5.19) on (5.9) to get:

$$\overline{\overline{Z}}^{\sigma}_{\perp 1}(\bar{r},\bar{r}_0) = \frac{2i}{b^2} F_\tau \left[W_0^{\sigma}(P/P_0) \right] \cdot \overline{\overline{I}} . \tag{5.22}$$

In the limit $\bar{r} \to \bar{r}_0$ we obtain the well known relation:

$$\overline{\overline{Z}}^{\sigma}_{\perp 1} = \frac{2c}{\omega^2 b} Z_0^{\sigma} \cdot \overline{\overline{I}} . \tag{5.23}$$

Explicitly with Z_0^{σ} given by Eq. (2.25):

$$\overline{\overline{Z}}^{\sigma}_{\perp 1} = \frac{(1-i)\, c}{\pi\, b^3\, \omega} \sqrt{\frac{\kappa\beta\, Z_0}{2\sigma}}\ \overline{\overline{I}} ; \quad [\Omega/m^2] . \tag{5.24}$$

Example: Find the transverse impedance of a perfectly conducting circular pipe

In Section 2 we have found the longitudinal impedances of a perfectly conducting circular pipe. In the approximation $(\kappa/\gamma) \ll 1$ we have obtained the m = 0 potential $\tilde{V}$ [see Eq. (2.16)] which is basically the static solution of a bi-dimensional boundary problem. With these approximations, the longitudinal impedance may be calculated from Eqs. (2.13) and (1.5):

$$Z(\bar{r},\bar{r}_0) = \frac{i\omega}{\beta c \gamma^2} \tilde{V}(\bar{r},\bar{r}_0) \tag{5.25}$$

where $\tilde{V}(\bar{r},\bar{r}_0)$ is the scalar potential at a point P due to a charge distribution in P_0, in the static approximation.

Let us consider a uniform charge distribution of radius a whose "centre of mass" is at the point r_0. The potential $V(\bar{r},\bar{r}_0)$ can be found by applying the image method replacing the walls with an image charge at $r = b^2/r_0$ (Fig. 18). The system of the two charges makes equipotential the surface r = b, and at a point within the charge distribution the scalar potential is[37]:

$$\tilde{V}(\bar{r},\bar{r}_0) = \frac{1}{2\pi\varepsilon_0}\left\{\frac{1}{2}\left[1 - \frac{|r-r_0|^2}{a^2}\right] + \ln\frac{b}{a} + \frac{1}{2}\ln\left[1 - \frac{|r-r_0|^2}{b^2}\right]\right\} \tag{5.26}$$

One can easily verify that the longitudinal impedance found in Eq. (2.14) is obtained by making the limit:

$$Z(\omega) = \lim_{\bar{r}\to\bar{r}_0} Z(\bar{r},\bar{r}_0) \ .$$

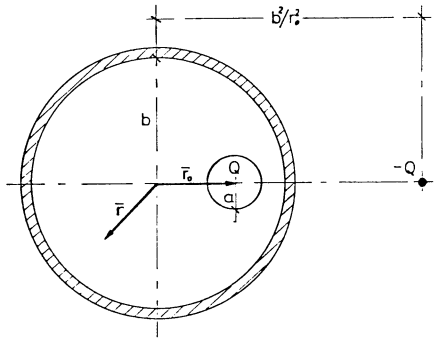


Fig. 18 The image charge of an off-axis beam

Expressing $\tilde{V}(\bar{r},\bar{r}_0)$ in Cartesian coordinates and using the operator

$$\nabla_t\nabla_{to} = \begin{vmatrix} \frac{\partial^2}{\partial x\partial x_0} & \frac{\partial^2}{\partial y\partial y_0} \\ \frac{\partial^2}{\partial y\partial x_0} & \frac{\partial^2}{\partial y\partial y_0} \end{vmatrix} \tag{5.27}$$

we get, according to Eq. (5.21), in the limit $\bar{r} \to \bar{r}_0$:

$$\overline{\overline{Z}}_{\perp} = \lim_{\bar{r}\to\bar{r}_0} \overline{\overline{Z}}_{\perp}(\bar{r},\bar{r}_0) = \frac{iZ_0}{2\pi(\beta\gamma)^2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right) \cdot \overline{\overline{I}} \ ; \ [\Omega/m^2] \tag{5.28}$$

6. REMARKS

It is useful to give here some indications on how to match this chapter with that of J.L. Laclare on instabilities where the concept of impedance is extensively used:

i) Because of a different Fourier transform definiton, one has to replace the imaginary unit "i" with "-j". (This is equivalent to the conjugation).

ii) The total impedance of a machine is obtained by multiplying the impedance for unit length by the average machine length $2\pi R$.

iii) The longitudinal impedance introduced in Laclare's chapter (Z/p) is obtained by multiplying the impedance ii) (ω_0/ω) where ω_0 is 2π times the revolution frequency and $\omega = p\omega_0$.

Example: $\left(\frac{Z_{RW}}{p}\right) = 2\pi R\left[Z_0^{\sigma}\right]^{*}\left(\frac{\omega_0}{\omega}\right)$.

7. CONCLUSIONS

A great deal of work has been done by many authors on the subject presented in this chapter and it has been impossible to cover all the topics related to the impedance and wake potential concepts and their applications. Perhaps what is still missing nowadays is a well organised arrangement of the subject into a well defined framework, but this is not easy because of the continuous developments in this field. However, we hope that the attempt made here to unify definitions and simplify concepts will help to introduce the reader to the field and to provide the required theoretical foundations.

ACKNOWLEDGEMENTS

We wish to thank K. Johnsen and P.J. Bryant for their encouragement and S. Turner for having revised the manuscript.

* * *

REFERENCES

1) J.L. Laclare, These proceedings.

2) A.W. Chao, SLAC-PUB-2946 (1982).

3) P.B. Wilson, SLAC-PUB-2884 (1982).

4) K.L. Bane, P.B. Wilson and T. Weiland, SLAC-PUB-3528 (1984).

5) R.F. Koontz, G.A. Loew, R.H. Miller and P.B. Wilson, IEEE Trans. N. Sci. NS-24, (1977).

6) V.G. Vaccaro, CERN ISR-RF/66-35 (1966).

7) J.D. Jackson, "Classical Electrodynamics", II edition, Wiley, NY (1975).

8) M.H. Blewett, Ed., Proc. First Course of the Int. School of Part. Accelerators of the 'Ettore Majorana' Centre for Scientific Culture, Erice, 1976, CERN 77-13 (1977).

9) L. Palumbo, V.G. Vaccaro and G. Wüstfeld, IEEE Trans. N. Sci. NS-31 (1984).

10) B. Zotter, Nucl. Instrum. and Methods, 129 (1975).

11) R.V. Churchill, Fourier Series and Boundary Value Problems, McGraw Hill, NY (1941).

12) M.J. Lee, F.E. Mills and P.L. Morton, SLAC-Report 76 (1976).

13) P.L. Morton, V.K. Neil and A.M. Sessler, J. App. Phys., 37 (1966).

14) A. Piwinskii, Desy Report 72/72 (1972).

15) A.W. Chao, Proc. AIP Conf. on Physics of High Energy Particle Accelerators, BNL/SUNY Summer School, 1983 (AIP Conf. Proc. No. 127 1983).

16) A. Papiernik, M. Chatard-Moulin and B. Jecko, Proc. 11th Int. Conf. on High-Energy Accelerators, Geneva, 1980 (Birkhäuser, Basle, 1980).

17) A.W. Chao and P.L. Morton, SLAC-Report PEP-105 (1975).

18) B. Zotter, SLAC Report, PEP-310 (1979).

19) E.U. Condon, J. Appl. Physics, 12 (1941).

DIFFUSION DUE TO RF NOISE

G. Dôme
CERN, Geneva, Switzerland

ABSTRACT

The effect of RF noise on bunches in a collider is described by a Fokker-Planck equation. It is shown that this equation reduces to a diffusion equation in phase space; the diffusion coefficient is computed in terms of the spectral density of noise.

1. STATISTICAL PROPERTIES OF RANDOM VARIABLES

We shall consider only random variables which are stationary in time, i.e. random variables for which some statistical properties (the mean and the autocorrelation) are independent of time. Obviously, a random variable f(t) which is stationary in time does not vanish at $t = \pm\infty$. Therefore its Fourier transform can only be computed on a finite time interval T:

$$\tilde{f}(\omega,T) = \int_0^T e^{-i\omega t} f(t)dt \ , \qquad f(t) = \int_{-\infty}^{+\infty} e^{i\omega t} \tilde{f}(\omega,T) \cdot \frac{d\omega}{2\pi} \qquad \text{for } 0 < t < T \tag{1.1}$$

Let brackets represent an ensemble average over all noises of the same kind; it is assumed that

$$\langle f(t) \rangle = 0 \quad \text{whence} \quad \langle \tilde{f}(\omega,T) \rangle = 0$$

The time average over T of the power of a particular noise sample reads

$$\frac{1}{T}\int_0^T \left[f(t)\right]^2 dt = \int_{-\infty}^{+\infty} \frac{1}{T} \left|\tilde{f}(\omega,T)\right|^2 \cdot \frac{d\omega}{2\pi}$$

and the ensemble average of the noise power is

$$\frac{1}{T}\int_0^T \langle \left[f(t)\right]^2 \rangle \, dt = \int_{-\infty}^{+\infty} \frac{1}{T} \langle |\tilde{f}(\omega,T)|^2 \rangle \frac{d\omega}{2\pi} \tag{1.2}$$

If f(t) is stationary in time, $\langle \left[f(t)\right]^2 \rangle$ is independent of t (by definition) and the integrand in the right hand side of (1.2) should approach a definite limit when $T \to \infty$; therefore it is also assumed that the following limit exists:

$$\boxed{\lim_{T\to\infty} \frac{1}{T} \langle |\tilde{f}(\omega,T)|^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \left\langle \left| \int_0^T e^{-i\omega t} f(t)dt \right|^2 \right\rangle = S_f(\omega) \geq 0} \qquad S_f(-\omega) = S_f(\omega) \tag{1.3}$$

$S_f(\omega)$ is called the *spectral power density* of the random variable f at frequency ω.

If $S_f(\omega)$ contains a Dirac peak at ω_n, the variable f(t) has a discrete spectral line at frequency ω_n. Pure noise is characterized by a continuous spectrum with a power density $S_f(\omega)$ free of Dirac peaks. The existence of $S_f(\omega)$ as an ordinary function implies that

$$|\tilde{f}(\omega,T)| = \left| \int_0^T e^{-i\omega t} f(t)dt \right| = 0(f \cdot \sqrt{T}) \qquad \text{when } T \to \infty \tag{1.4}$$

However, for a particular noise of the ensemble, the quantity

$$\frac{1}{T} |\tilde{f}(\omega,T)|^2 \tag{1.5}$$

does not tend to a definite limit when $T \to \infty$ (Ref. 1, p. 222; Ref. 2, p. 305). The quantities (1.5) take different values for each particular (large) T and each particular noise. They are statistically distributed around the ensemble average (1.3), with a standard deviation which is at least equal to the ensemble average itself; this means that fluctuations around the ensemble average (1.3) are quite large. These fluctuations can be reduced by smoothing the spectrum (i.e. replacing each point of the spectrum by some average taken over the neighbouring points) at the cost of reduced resolution (Ref. 2, p. 380; Ref. 3, p. 72).

Another way of estimating the power spectrum $S_f(\omega)$ is to use the *autocorrelation (or autocovariance) function* $R(\tau)$, which is defined as

$$R(\tau) \doteq \langle f(t) \cdot f(t+\tau) \rangle \tag{1.6}$$

For a stationary process, $R(\tau)$ is independent of t and $R(-\tau) = R(\tau)$

It is easy to show (Wiener-Khintchine theorem) that the spectral power density of the random variable f(t) is the Fourier transform of its autocorrelation function:

$$S_f(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega\tau} R(\tau)\, d\tau = 2\int_0^{\infty} R(\tau) \cos(\omega\tau)\, d\tau \tag{1.7}$$

whence

$$R(\tau) = \int_{-\infty}^{+\infty} e^{i\omega\tau} S_f(\omega) \frac{d\omega}{2\pi} = 2\int_0^{\infty} S_f(\omega) \cos(\omega\tau) \cdot \frac{d\omega}{2\pi} \tag{1.8}$$

If $R(\tau)$ is appreciable only for $\tau \leq \tau_c$, $S_f(\omega)$ extends at least up to frequencies ω_c where

$$2\frac{\omega_c}{2\pi} \cdot 2\tau_c \geq 1$$

(uncertainty principle).

By combining either (1.2) and (1.3) or (1.6) and (1.8) we get

$$\langle \left[f(t) \right]^2 \rangle = R(0) = 2 \int_0^\infty S_f(\omega) \cdot \frac{d\omega}{2\pi} \tag{1.9}$$

Very often a stationary process having a continuous power spectrum is assumed to be ergodic, which means that ensemble averages are equal to time averages taken over any single sample of the process. Then the autocorrelation function may also be computed as

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) \cdot f(t+\tau)\, dt \,. \tag{1.10}$$

2. FOKKER-PLANCK EQUATION

Here we follow the derivation given in Ref. 4, p. 98. Let x be a generalized coordinate describing the motion of particles; in the most general case x should be replaced by a complete set of coordinates in phase space. Let $\rho(x_0,x,t)dx$ be the probability of a particle starting at $(x_0, t=0)$ to arrive somewhere between (x,t) and $(x+dx,t)$; obviously $\int \rho\, dx = 1$.

When t increases by Δt, x increases by a Δx which is different for each particle. Let $\psi(\Delta x,x,t)\cdot d(\Delta x)$ be the probability of finding Δx between Δx and $\Delta x + d(\Delta x)$. The basic assumption is that ψ is independent of x_0, which means (by definition) that the stochastic process is a Markoff process. Writing $x_1 = x + \Delta x$ we get

$$\rho(x_0,x_1,t+\Delta t) = \int \rho(x_0,x_1-\Delta x,t) \cdot \psi(\Delta x, x_1-\Delta x,t) d(\Delta x) \tag{2.1}$$

The intregrand is expanded as a Taylor series in Δx about x_1:

$$\rho(x_0,x_1,t)\ \psi(\Delta x,x_1,t) - \Delta x \frac{\partial}{\partial x_1}\left[\rho(x_0,x_1,t)\ \psi(\Delta x,x_1,t)\right] + \frac{(\Delta x)^2}{2!}\frac{\partial^2}{\partial x_1^2}\left[\rho(x_0,x_1,t)\ \psi(\Delta x,x_1,t)\right] - \ldots$$

With

$$\int \psi(\Delta x,x_1,t)\ d(\Delta x) = 1 \quad \text{and} \quad \int (\Delta x)^m\ \psi(\Delta x,x_1,t)\ d(\Delta x) = \langle (\Delta x)^m \rangle \,,$$

Eq. (2.1) becomes

$$\rho(x_0,x_1,t+\Delta t) = \rho(x_0,x_1,t) - \frac{\partial}{\partial x_1}\left[\rho(x_0,x_1,t)\cdot\langle \Delta x \rangle\right] + \frac{1}{2}\frac{\partial^2}{\partial x_1^2}\left[\rho(x_0,x,t)\cdot\langle(\Delta x)^2\rangle\right] - \ldots$$

Let

$$A_1(x_1,t) = \lim_{\Delta t\to 0} \frac{\langle \Delta x \rangle}{\Delta t} \qquad A_2(x_1,t) = \lim_{\Delta t\to 0} \frac{\langle (\Delta x)^2 \rangle}{\Delta t} \tag{2.2}$$

and assume

$$A_\ell(x_1,t) = \lim_{\Delta t \to 0} \frac{\langle (\Delta x)^\ell \rangle}{\Delta t} = 0 \qquad \text{for } \ell > 2 \tag{2.3}$$

then the last equation reduces to

$$\boxed{\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial x} (\rho\, A_1) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\rho\, A_2)} \tag{2.4}$$

which is the Fokker-Planck equation in one-dimensional space.

In the limits (2.2), (2.3), Δt must include many periods of the unperturbed motion (in case of RF noise, many synchrotron oscillations). The limit $\Delta t \to 0$ means that the time scale which is considered in the Fokker-Planck equation is much longer than Δt, and that on the intermediate time scale Δt, $< \Delta x >$ and $< (\Delta x)^2 >$ are small quantities of the order of Δt. In fact, in equations (4.4) and (4.8), Δt will be considered as infinite (with respect to the period of synchrotron oscillations).

With the definitions (2.2), the Fokker-Planck equation (2.4) remains invariant when x is replaced by kx, where k is any constant. Moreover, since A_1 and A_2 are ensemble averages over all noises with the same power spectrum, the Fokker-Planck equation does not represent the evolution of $\rho(x,t)$ under the influence of a particular noise sample; it rather represents an average evolution of $\rho(x,t)$ under the influence of all noises of the ensemble. However, any non-linearity in the RF voltage produces a variation of synchrotron frequency with amplitude, which entails filamentation and mixing of particles in phase space; this mixing is ultimately sufficient to make the description by Eq. (2.4) adequate, even for a single noise sample. The only exception is the case of a linear RF voltage with a noise whose spectrum does not extend so far as the RF frequency; in that particular case the bunch is shaken (by phase noise) or deformed (by amplitude noise) as a whole but it does not diffuse in phase space (Ref. 10).

The one-dimensional Fokker-Planck equation (2.4) may be written as

$$\frac{\partial \rho}{\partial t} = - \operatorname{div} (F\rho) + \operatorname{div} (D \operatorname{grad} \rho) \tag{2.5}$$

where $F = A_1 - \frac{1}{2} \frac{\partial A_2}{\partial x}$ is a *friction* coefficient.

$D = \frac{A_2}{2}$ is a *diffusion* coefficient.

In higher dimensional space, F and D become tensors of rank 1 and 2 respectively.

The Fokker-Planck equation has been used to describe momentum stochastic cooling in the ISR, in the form (Ref. 5, p.93)

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial E} (F\rho) + \frac{\partial}{\partial E} \left(D \frac{\partial \rho}{\partial E} \right) \tag{2.6}$$

where

E is the particle energy

F is a cooling coefficient due to feedback, proportional to -(cooling time)$^{-1}$;

D is a diffusion coefficient due to amplifier noise injected on the kicker, and to particle Schottky noise.

In 1977, an experiment with F = 0 was performed at the ISR (Ref. 6) in order to decide between $\frac{\partial^2}{\partial E^2}(D\rho)$ or $\frac{\partial}{\partial E}(D\frac{\partial \rho}{\partial E})$ in the right-hand side of Eq. (2.6) ; it confirmed the validity of the diffusion equation as written in (2. 6). In fact, it will be shown later that for trapped motion (i.e. libration), the Fokker-Planck equation (2.4) reduces to a diffusion equation (4.10) when the generalized coordinate x in (2.4) is taken as being proportional to the action J. For a coasting beam (i.e. rotation), to first order $\delta J = \delta E/\omega_r$ is proportional to δE; this is the reason why (2.4) reduces to the diffusion equation (2 . 6) when E is taken as the generalized coordinate.

3. DIFFERENTIAL EQUATIONS FOR A STATIONARY BUCKET WITH NOISE (AMPLITUDE AND PHASE)

With noise, the finite difference equations of synchrotron motion [see Eq. (11.1), (11.2) in Ref. 7] become

$$\begin{cases} \delta p_{n+1} - \delta p_n = \dfrac{eV(1 + a_n)}{v_{s,n}} \left[g(\phi_n) - g(\phi_s) \right] \\ \phi_{n+1} - \phi_n = -\dfrac{2\pi h \eta}{N p_{s,n+1}} \cdot \delta p_{n+1} + \varphi_{n+1} - \varphi_n \end{cases} \qquad (3.1)$$

where a_n is the relative amplitude noise

φ_n is the phase noise

ϕ_n is the actual phase with noise

For a stationary bucket, $g(\phi_s) = 0$.

Reduced variables

Let us define

$$g_s(\delta\phi) = \frac{g(\phi_s + \delta\phi)}{g'(\phi_s)} \quad \text{so that} \quad g_s'(0) = 1 \qquad (3.2)$$

and, by analogy with Eq. (4.8) in Ref. 7,

$$G_s(\delta\phi) = -\int g_s(\delta\phi)\, d(\delta\phi) = \frac{G(\phi_s + \delta\phi)}{g'(\phi_s)} . \qquad (3.3)$$

From now on, we shift the origin of phases to ϕ_s, which means that ϕ_n will represent $\delta\phi_n = \phi_n - \phi_s$; similarly φ_n will represent $(\varphi_n - \phi_s)$.

For a linear RF voltage,

$$g_s(\phi) = \phi \qquad \text{and} \qquad G_s(\phi) = -\frac{\phi^2}{2} \qquad (3.4)$$

For a sinusoidal RF voltage,

$$g_s(\phi) = \sin \phi \quad \text{and} \quad G_s(\phi) = \cos \phi \tag{3.5}$$

With (11.4), (11.5) in Ref. 7 and (3.2), the system (3.1) reads, since γ_s is constant:

$$\begin{cases} P_{n+1} - P_n = - K_0(1 + a_n) g_s(\phi_n) \\ \phi_{n+1} - \phi_n = P_{n+1} + \varphi_{n+1} - \varphi_n \end{cases} \tag{3.6}$$

The effect of noise will be considered as a perturbation of the motion without noise. With finite difference equations, the unperturbed motion is known only for a linear RF voltage (see Section 6). For any other RF voltage, we must again use differential equations:

$$\begin{cases} \dot{P} = - \frac{N^2}{T_0^2} K_0 \left[1 + a(t)\right] g_s(\phi) \\ \dot{\phi} = P + \dot{\varphi}(t) \end{cases} \tag{3.7}$$

where

$$P = \frac{N}{T_0} P_n \tag{3.8}$$

We also put

$$\frac{N^2}{T_0^2} K_0 = \Omega^2 \tag{3.9}$$

where, with (11.8) of Ref. 7,

$$\Omega = \frac{2N}{T_0} \sin\left(\frac{\omega_{so} T_0}{2N}\right)$$

Let us note that $\dot{\varphi}(t) = f(t)$ is the frequency noise, and that ω_{so} is the frequency of small synchrotron oscillations for finite difference equations; for differential equations this frequency is Ω.

The system (3.7) can be derived from the time-dependent Hamiltonian

$$H = \frac{1}{2} P^2 + \Omega^2 \left[1 + a(t)\right] \left[G_s(0) - G_s(\phi)\right] + P f(t) \tag{3.10}$$

Action-angle variables

For the unperturbed motion,

$$H_0 = \frac{1}{2} P^2 + U(\phi) \qquad \text{where} \qquad U(\phi) = \Omega^2 \left[G_s(0) - G_s(\phi)\right] \tag{3.11}$$

The action J is defined as

$$2\pi J = \oint P \, d\phi = 2 \int_{\phi_1}^{\phi_2} d\phi \, \sqrt{2H_0 - 2\,U(\phi)} \tag{3.12}$$

where at ϕ_1 and ϕ_2,

$$2H_0 - 2\,U(\phi) = 0 \ .$$

Using (3.12) we get

$$\frac{dJ}{dH_0} = \frac{2}{2\pi}\left[\int_{\phi_1}^{\phi_2} \frac{d\phi}{\sqrt{2H_0 - 2\,U(\phi)}} + \underbrace{\sqrt{2H_0 - 2\,U(\phi_2)}}_{0} \cdot \frac{1}{U'(\phi_2)} - \underbrace{\sqrt{2H_0 - 2\,U(\phi_1)}}_{0} \cdot \frac{1}{U'(\phi_1)}\right]$$

hence

$$\frac{dJ}{dH_0} = \frac{1}{2\pi}\oint \frac{d\phi}{P} = \frac{T_s}{2\pi} = \frac{1}{\omega_s} \quad \text{and} \quad H_0 = \int \omega_s \, dJ \tag{3.13}$$

We now seek a canonical transformation from the variables (P,ϕ) to other canonical variables (J,θ). This is achieved (Ref. 9 , p.239-241) by using the generating function

$$F_2(\phi, J) = \int d\phi \sqrt{2 \int \omega_s \, dJ - 2\,U(\phi)}$$

which is defined in such a way that $P = \frac{\partial F_2}{\partial \phi}$; the variable θ is then given by

$$\theta = \frac{\partial F_2}{\partial J} = \omega_s \int \frac{d\phi}{\sqrt{2 \int \omega_s \, dJ - 2\,U(\phi)}} = \omega_s \int dt \text{ in unperturbed motion} \tag{3.14}$$

while the new Hamiltonian reads $H + \frac{\partial F_2}{\partial t} = H$.

From (3.14) it is seen that θ has period 2π in the unperturbed motion; hence its name "angle variable".

For the motion with noise, from (3.10), (3.11) and (3.13):

$$H = \int \omega_s \, dJ + P\, f(t) + U\, a(t) \tag{3.15}$$

In (3.15) P and U must be expressed in terms of J and θ. Since the unperturbed motion has period 2π in θ, we may write

$$\begin{cases} P = \dfrac{\partial H}{\partial f} = \displaystyle\sum_{m=-\infty}^{+\infty} b_m(J)\, e^{im\theta} & \text{where } b_{-m} = b_m^* \\ U = \dfrac{\partial H}{\partial a} = \displaystyle\sum_{m=-\infty}^{+\infty} c_m(J)\, e^{im\theta} & \text{where } c_{-m} = c_m^* \end{cases} \tag{3.16}$$

4. COMPUTATION OF THE COEFFICIENTS IN THE FOKKER-PLANCK EQUATION

The equations of motion are canonical:

$$\frac{dJ}{dt} = -\frac{\partial H}{\partial \theta} \qquad \frac{d\theta}{dt} = \frac{\partial H}{\partial J}$$

For computing A_1, A_2 as defined in (2.2), we have first to compute (with $\Delta t = T$)

$$\Delta J = J(T) - J(0) = -\int_0^T \frac{\partial H}{\partial \theta}\, dt$$

In order to simplify the writing, from now on we consider only the phase noise; the derivation would be identical for the amplitude noise. From (3.15) and (3.16):

$$J(T) - J(0) = -\int_0^T f(t)\, dt \sum_{m=-\infty}^{+\infty} im\, b_m(t) \cdot e^{im\theta} \tag{4.1}$$

where

$$b_m(t) = b_m(0) + \frac{db_m}{dJ}(0) \cdot [J(t)-J(0)] + \ldots ,$$

whereas

$$\frac{d\theta}{dt} = \omega_s + f(t) \sum_{n=-\infty}^{+\infty} \frac{db_n}{dJ}(t) \cdot e^{in\theta} \tag{4.2}$$

where

$$\omega_s(t) = \omega_s(0) + \frac{d\omega_s}{dJ}(0) \cdot [J(t)-J(0)] + \ldots$$

The relation (1.4) shows that in order to compute quantities of order T, we must keep quantities up to 2nd order in f(t). Moreover, for computing ΔJ up to 2nd order in f(t), we must retain quantities up to first order when using $b_m(t)$ and θ.

Integrate (4.2) up to first order in f(t):

$$\theta(t) = \theta_0 + \omega_s(0)\cdot t + \frac{d\omega_s}{dJ}(0)\cdot \underbrace{\int_0^t dt'[J(t')-J(0)]}_{} + \int_0^t f(t')dt' \sum_{n=-\infty}^{+\infty} \frac{db_n}{dJ}(0)\cdot e^{in\theta_0+in\omega_s(0)t'}$$

$$- \int_0^t dt' \int_0^{t'} f(t'')dt'' \sum_{n=-\infty}^{+\infty} in\, b_n(0)\cdot e^{in\theta_0+in\omega_s(0)t''}$$

and insert into (4.1):

$$J(T) - J(0) = - \sum_{m=-\infty}^{+\infty} \int_0^T f(t)dt\; im\, e^{im\theta_0+im\omega_s(0)t} \cdot$$

$$\cdot \left[b_m(0) - \frac{db_m}{dJ}(0) \int_0^t f(t')dt' \sum_{n=-\infty}^{+\infty} in\, b_n(0) e^{in\theta_0+in\omega_s(0)t'} \right] \cdot$$

$$\cdot \left\{ 1 + im \sum_{n=-\infty}^{+\infty} e^{in\theta_0} \left[- \frac{d\omega_s}{dJ}(0) \cdot \int_0^t dt' \int_0^{t'} f(t'')dt''\, in\, b_n(0) e^{in\omega_s(0)t''} \right.\right.$$

$$\left.\left. + \int_0^t f(t')dt' \frac{db_n}{dJ}(0)\cdot e^{in\omega_s(0)t'} \right] \right\}$$

$$J(T) - J(0) = \sum_{m=-\infty}^{+\infty} e^{im\theta_0} \int_0^T f(t)dt \cdot e^{im\omega_s(0)t} \cdot \left\{ - im\, b_m(0) \right.$$

$$+ im \frac{db_m}{dJ}(0) \sum_{n=-\infty}^{+\infty} e^{in\theta_0}\, in\, b_n(0) \cdot \int_0^t f(t')dt' \cdot e^{in\omega_s(0)t'}$$

$$+ |im|^2 b_m(0) \sum_{n=-\infty}^{+\infty} e^{in\theta_0} \left[- in\, b_n(0) \frac{d\omega_s}{dJ}(0) \int_0^t dt' \int_0^{t'} f(t'')dt'' \cdot e^{in\omega_s(0)t''} \right.$$

$$\left.\left. + \frac{db_n}{dJ}(0) \cdot \int_0^t f(t')dt' \cdot e^{in\omega_s(0)t'} \right] \right\} \quad (4.3)$$

Up to 2nd order in f(t):

$$[J(T)-J(0)]^2 = \left[-\sum_{m=-\infty}^{+\infty} e^{im\theta_0} \cdot im\, b_m(0) \int_0^T f(t)dt\, e^{im\omega_s(0)t} \right]^2$$

$$= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} e^{i(m+n)\theta_0} \cdot im\, b_m(0) \cdot in\, b_n(0) \cdot \int_0^T f(t)dt\, e^{im\omega_s(0)t} \cdot$$

$$\cdot \int_0^T f(t)dt\, e^{in\omega_s(0)t}$$

When taking the ensemble average of this expression we must also average over all initial conditions θ_0; since particles are assumed to be uniformly distributed in θ_0, the only terms which will be left are those with (m+n) = 0:

$$\langle [J(T)-J(0)]^2 \rangle = \sum_{m=-\infty}^{+\infty} \left| im\, b_m(0) \right|^2 \cdot \left\langle \left| \int_0^T f(t)dt\, e^{im\omega_s(0)t} \right|^2 \right\rangle$$

With (1.3):

$$A_2 = \lim_{T\to\infty} \frac{\langle [J(T)-J(0)]^2 \rangle}{T} = \sum_{m=-\infty}^{+\infty} \left| im\, b_m \right|^2 \cdot S_f(m\omega_s) \tag{4.4}$$

In the ensemble average of (4.3), the only terms which are left are the m = 0 term in $\sum_{m=-\infty}^{+\infty}$, and the (m+n) = 0 terms in $\sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty}$; therefore

$$\langle [J(T)-J(0)] \rangle = \sum_{m=-\infty}^{+\infty} \left\{ im \frac{db_m}{dJ}(0) \left[im\, b_m(0) \right]^* \cdot \left\langle \int_0^T f(t)dt\, e^{im\omega_s(0)t} \int_0^t f(t')dt'\, e^{-im\omega_s(0)t'} \right\rangle \right.$$

$$+ |im|^2\, b_m(0) \cdot \left\langle \int_0^T f(t)dt\, e^{im\omega_s(0)t} \left[im\, b_m^*(0) \frac{d\omega_s}{dJ}(0) \cdot \right. \right.$$

$$\left. \left. \cdot \int_0^t dt' \int_0^{t'} f(t'')dt''\, e^{-im\omega_s(0)t''} + \frac{db_m^*}{dJ}(0) \cdot \int_0^t f(t')dt'\, e^{-im\omega_s(0)t'} \right] \right\rangle \Bigg\}$$

$$= \frac{1}{2} \sum_{m=-\infty}^{+\infty} \left\{ |im|^2 \frac{d}{dJ} \left[b_m(0)\, b_m^*(0) \right] \cdot \right.$$

(continued on next page)

$$\cdot \left\langle \int_0^T f(t)dt\, e^{im\omega_s(0)t} \int_0^t f(t')dt'\, e^{-im\omega_s(0)t'} + \text{complex conjugate} \right\rangle$$

$$+ \left| im\, b_m(0) \right|^2 \cdot im \frac{d\omega_s}{dJ}(0) \cdot \left\langle \int_0^T f(t)dt\, e^{im\omega_s(0)t} \cdot \right.$$

$$\left. \cdot \int_0^t dt' \int_0^{t'} f(t'')dt''\, e^{-im\omega_s(0)t''} - \text{complex conjugate} \right\rangle \Bigg\}$$

(4.5)

The complex conjugate terms are obtained by changing m into -m.

Now, with (1.1):

$$\int_0^T f(t)dt\, e^{im\omega_s(0)t} \int_0^t f(t')dt'\, e^{-im\omega_s(0)t'} + \text{complex conjugate}$$

$$= \int_0^T d\left[\tilde{f}(m\omega_s(0),t)\right]^* \tilde{f}(m\omega_s(0),t) + \text{complex conjugate} = \left|\tilde{f}(m\omega_s(0),T)\right|^2 \tag{4.6}$$

whereas

$$\int_0^T f(t)dt\, e^{im\omega_s(0)t} \int_0^t dt' \int_0^{t'} f(t'')dt''\, e^{-im\omega_s(0)t''} - \text{complex conjugate}$$

$$= \int_0^T d\left[\tilde{f}(m\omega_s,t)\right]^* \int_0^t dt'\, \tilde{f}(m\omega_s,t') - \text{complex conjugate}$$

$$= \int_0^T d\left[\tilde{f}(m\omega_s,t)\right]^* \left[t\, \tilde{f}(m\omega_s,t) - \int_0^t t'\, d\tilde{f}(m\omega_s,t') \right] - \text{complex conjugate}$$

$$= \left\{ \int_0^T d\left[\tilde{f}(m\omega_s,t)\right]^* t\, \tilde{f}(m\omega_s,t) - \left[\tilde{f}(m\omega_s,T)\right]^* \int_0^T t'\, d\tilde{f}(m\omega_s,t') \right.$$

$$\left. + \int_0^T \left[\tilde{f}(m\omega_s,t)\right]^* t\, d\tilde{f}(m\omega_s,t) \right\} - \text{complex conjugate}$$

$$= \left\{ \int_0^T t\, d|\tilde{f}(m\omega_s,t)|^2 - [\tilde{f}(m\omega_s,T)]^* \int_0^T t\, e^{-im\omega_s t} \cdot f(t)dt \right\} - \text{complex conjugate}$$

$$= \frac{1}{i} [\tilde{f}(m\omega_s,T)]^* \left[\frac{\partial}{\partial\omega} \tilde{f}(\omega,T) \right]_{\omega=m\omega_s(0)} - \text{complex conjugate}$$

$$= \frac{1}{i} \left[\frac{\partial}{\partial\omega} |\tilde{f}(\omega,T)|^2 \right]_{\omega=m\omega_s(0)} \qquad (4.7)$$

In this derivation, $m\omega_s$ has been written shortly for $m\omega_s(0)$.

With (4.6) and (4.7), (4.5) reads

$$\langle [J(T)-J(0)] \rangle = \frac{1}{2} \sum_{m=-\infty}^{+\infty} \left\{ \frac{d}{dJ} |im\, b_m(0)|^2 \cdot \langle |\tilde{f}(m\omega_s(0),T)|^2 \rangle \right.$$

$$\left. + |im\, b_m(0)|^2 \cdot m \frac{d\omega_s}{dJ}(0) \cdot \left[\frac{\partial}{\partial\omega} \langle |\tilde{f}(\omega,T)|^2 \rangle \right]_{\omega=m\omega_s(0)} \right\}$$

Using (1.3) again:

$$A_1 = \lim_{T\to\infty} \frac{\langle [J(T)-J(0)] \rangle}{T} = \frac{1}{2} \sum_{m=-\infty}^{+\infty} \left\{ \left[\frac{d}{dJ} |im\, b_m|^2 \right] \cdot S_f(m\omega_s) + |im\, b_m|^2 \cdot \frac{d(m\omega_s)}{dJ} \left[\frac{dS_f(\omega)}{d\omega} \right]_{\omega=m\omega_s} \right\} \qquad (4.8)$$

From (4.3),

$$[J(T)-J(0)]^\ell = O([f(t)]^\ell)$$

and

$$A_\ell = \lim_{T\to\infty} \frac{\langle [J(T)-J(0)]^\ell \rangle}{T} = O([f(t)]^\ell) \qquad \text{for } \ell > 2 .$$

Therefore, the conditions (2.3) are satisfied up to 2nd order in f(t).

Comparing (4.8) with (4.4) we get the remarkable relation

$$\boxed{A_1 = \frac{1}{2} \frac{\partial A_2}{\partial J}} \qquad (4.9)$$

which entails that the friction coefficient vanishes in (2.5).

Because of this relation, if the variable x is taken as the action J in (2.4), the Fokker-Planck equation contains a single coefficient A_2 and reduces to

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial J} \left(\frac{A_2}{2} \frac{\partial\rho}{\partial J} \right) \qquad (4.10)$$

which is the general one-dimensional form of a diffusion equation

$$\frac{\partial \rho}{\partial t} = \mathrm{div}(D \ \mathrm{grad} \ \rho) \tag{4.11}$$

where $A_2/2 = D$ is the diffusion coefficient.

If amplitude noise is also present, (4.4) becomes

$$A_2 = \sum_{m=-\infty}^{+\infty} |\mathrm{im}\ b_m|^2 \cdot S_f(m\omega_s) + \sum_{m=-\infty}^{+\infty} |\mathrm{im}\ c_m|^2 \cdot S_a(m\omega_s) \tag{4.12}$$

In this expression, we have assumed that frequency and amplitude noises are uncorrelated; in that case the ensemble average of the crossed terms over all possible noises is zero. From (3.16), it appears that the coefficient of $S_f(m\omega_s)$ is the square of the modulus of the Fourier component of $\partial^2 H/\partial\theta\partial f$ at $m\omega_s$; similarly the coefficient of $S_a(m\omega_s)$ is the square of the modulus of the Fourier component of $\partial^2 H/\partial\theta\partial a$. It should be noticed that noise plays a role only at those frequencies which are present in the unperturbed motion of the particles; moreover, as expected, the zero-frequency component of the noise never plays any role.

Since $S_f(\omega) = \omega^2 S_\varphi(\omega)$ where $S_\varphi(\omega)$ is the spectral density of phase noise, we have

$$A_2 = \sum_{m=-\infty}^{+\infty} (m\omega_s)^2 |\mathrm{im}\ b_m|^2 \ S_\varphi(m\omega_s) + \sum_{m=-\infty}^{+\infty} |\mathrm{im}\ c_m|^2 \ S_a(m\omega_s) \tag{4.13}$$

Remark: The relation (4.9) was first derived in 1980 for the case of finite difference equations and a linear RF voltage (Ref. 8); it was then extended to the case of differential equations and a sinusoidal RF voltage (Ref. 10). In 1982 it was proven to hold for any RF voltage (Ref. 11).

5. CASE OF A SINUSOIDAL RF VOLTAGE

We only need to determine the coefficients $b_n(J)$, $c_n(J)$ which appear in (3.16) for the unperturbed motion. With (3.5) , the unperturbed Hamiltonian (3.11) reads

$$H_0 = \frac{1}{2} P^2 + \Omega^2 \cdot 2 \sin^2 \frac{\phi}{2} = 2 \Omega^2 \left[\left(\frac{P}{2\Omega} \right)^2 + \sin^2 \frac{\phi}{2} \right]$$

We define polar coordinates k, ψ through the relations

$$\frac{P}{2\Omega} = k \cos \psi \ , \quad \sin \frac{\phi}{2} = k \sin \psi \ , \quad k = \sin \frac{\hat{\phi}}{2} \leq 1 \tag{5.1}$$

The unperturbed motion corresponds to $H_0 (= 2 \Omega^2 \cdot k^2)$ being a constant. The action is given by

$$J = \frac{1}{2\pi} \oint P \, d\phi = \frac{4\Omega k^2}{2\pi} \int_0^{2\pi} \cos \psi \ \frac{\cos \psi \ d\psi}{\sqrt{1 - k^2 \sin^2 \psi}} = 2 \Omega \ k^2 \ \frac{4}{\pi} \ B(k) \qquad \text{where } \frac{\pi}{4} \leq B(k) \leq 1$$

$$\hat{J} = 2\Omega \cdot \frac{4}{\pi} \qquad \text{when } 0 \leq k \leq 1 \tag{5.2}$$

The notations for elliptic integrals are those of Jahnke and Emde (Ref. 12, p. 73). From (3.7) and (5.1) we get for the unperturbed motion

$$t = \int_{\phi_0}^{\phi} \frac{d\phi}{P} = \frac{1}{\Omega}(u-u_0) \qquad \text{where } u = \int_0^{\psi} \frac{d\psi}{\sqrt{1 - k^2 \sin^2 \psi}} \tag{5.3}$$

By definition,

$$sn(u,k) = \sin \psi \ , \quad cn(u,k) = \cos \psi \ , \quad dn(u,k) = \sqrt{1 - k^2 \sin^2 \psi} \tag{5.4}$$

To the period 2π in ψ corresponds the period $4K$ in u and $4K/\Omega$ in time; therefore, the angular frequency of synchrotron oscillations is

$$\omega_s = \Omega \cdot \frac{\pi}{2K(k)} \tag{5.5}$$

From (3.14), (5.3) and (5.5),

$$\theta - \theta_0 = \omega_s t = \frac{\pi}{2K}(u-u_0)$$

so that we may take

$$\theta = \frac{\pi u}{2K} \tag{5.6}$$

Finally, with (5.1) and (5.4), (3.16) becomes

$$\begin{cases} \sum_{m=-\infty}^{+\infty} b_m e^{im\theta} = P = 2\,\Omega\, k\, cn\, u \\ \sum_{m=-\infty}^{+\infty} c_m e^{im\theta} = U = \Omega^2 \cdot 2 \sin^2 \frac{\phi}{2} = 2\,\Omega^2 k^2 sn^2 u \end{cases} \tag{5.7}$$

These Fourier series are found in treatises on elliptic functions. For the first one, we have (Ref. 13, p. 256; Ref. 14, p. 911)

$$k \frac{2K}{\pi} cn\, u = \sum_{m\ odd=-\infty}^{+\infty} \frac{e^{im\theta}}{ch(mv)} \qquad \text{where } v = \frac{\pi}{2} \frac{K'(k)}{K(k)} \tag{5.8}$$

Therefore

$$b_m = 2\,\Omega \frac{\pi}{2K} \frac{1}{ch(mv)} = \frac{2\omega_s}{ch(mv)} \quad \text{for m odd} \ , \qquad b_m = 0 \quad \text{for m even} \tag{5.9}$$

The Fourier series of $k^2\, sn^2\, u$ may be obtained through Jacobi's Zeta function (Ref. 13, p. 292, 295):

$$Z(u) = \left(1 - \frac{E}{K}\right)u - \int_0^u k^2\, sn^2\, u \cdot du = \frac{\pi}{2K} \cdot \sum_{m\ even=-\infty}^{+\infty}{}' \frac{e^{im\theta}}{i\, sh(mv)} \qquad \text{(without m = 0)}$$

whence

$$\frac{\partial Z}{\partial u} = \left(1 - \frac{E}{K}\right) - k^2 \, sn^2 \, u = \left(\frac{\pi}{2K}\right)^2 \cdot \sum_{m \text{ even}=-\infty}^{+\infty}{}' \frac{m}{sh(mv)} e^{im\theta} \qquad \text{(without } m = 0\text{)} \tag{5.10}$$

Therefore,

$$c_0 = 2\,\Omega^2 \left(1 - \frac{E}{K}\right), \quad c_m = -2\,\Omega^2 \left(\frac{\pi}{2K}\right)^2 \frac{m}{sh(mv)} = -2\,\omega_s^2 \frac{m}{sh(mv)} \qquad \text{for } m \text{ even} \neq 0$$
$$c_m = 0 \qquad \text{for } m \text{ odd} \tag{5.11}$$

It remains to introduce (5.9) and (5.11) into (4.13) in order to get:

$$\frac{A_2}{2} = 4 \sum_{m=1,3,}^{\infty} \frac{(m\omega_s)^4}{ch^2(mv)} S_\varphi(m\omega_s) + 4 \sum_{m=2,4,}^{\infty} \frac{(m\omega_s)^4}{sh^2(mv)} S_a(m\omega_s) \tag{5.12}$$

This expression can be put in the form of a weighted average over the noise spectral density by using the relations [see Appendix A, Eq. (A.6) and (A.8)]

and

$$\begin{cases} 4 \displaystyle\sum_{m=1,3,}^{\infty} \frac{\left(m\frac{\pi}{2K}\right)^4}{ch^2(mv)} = \left[k^2 \frac{4}{\pi} B\right] \frac{2K}{\pi} (1 - k^2\alpha) \\ 4 \displaystyle\sum_{m=2,4,}^{\infty} \frac{\left(m\frac{\pi}{2K}\right)^4}{sh^2(mv)} = \left[k^2 \frac{4}{\pi} B\right]^2 \frac{2K}{\pi} \frac{\pi}{4B} \alpha \end{cases} \tag{5.13}$$

where, from eq. (A.4),

$$\alpha(k^2) = \frac{4}{15}\left[\frac{2E}{B} - (1-k^2)\frac{C}{B}\right]$$

We have

$$1 \geq \alpha \geq \frac{8}{15}$$

when

$$0 \leq k^2 \leq 1$$

If instead of J we use $x = J/\hat{J}$ as independent variable in the Fokker-Planck equation, the corresponding diffusion coefficient reads

$$D = \frac{A_2}{2\hat{J}^2} = x\,S_1 + x^2\,S_2 \tag{5.14}$$

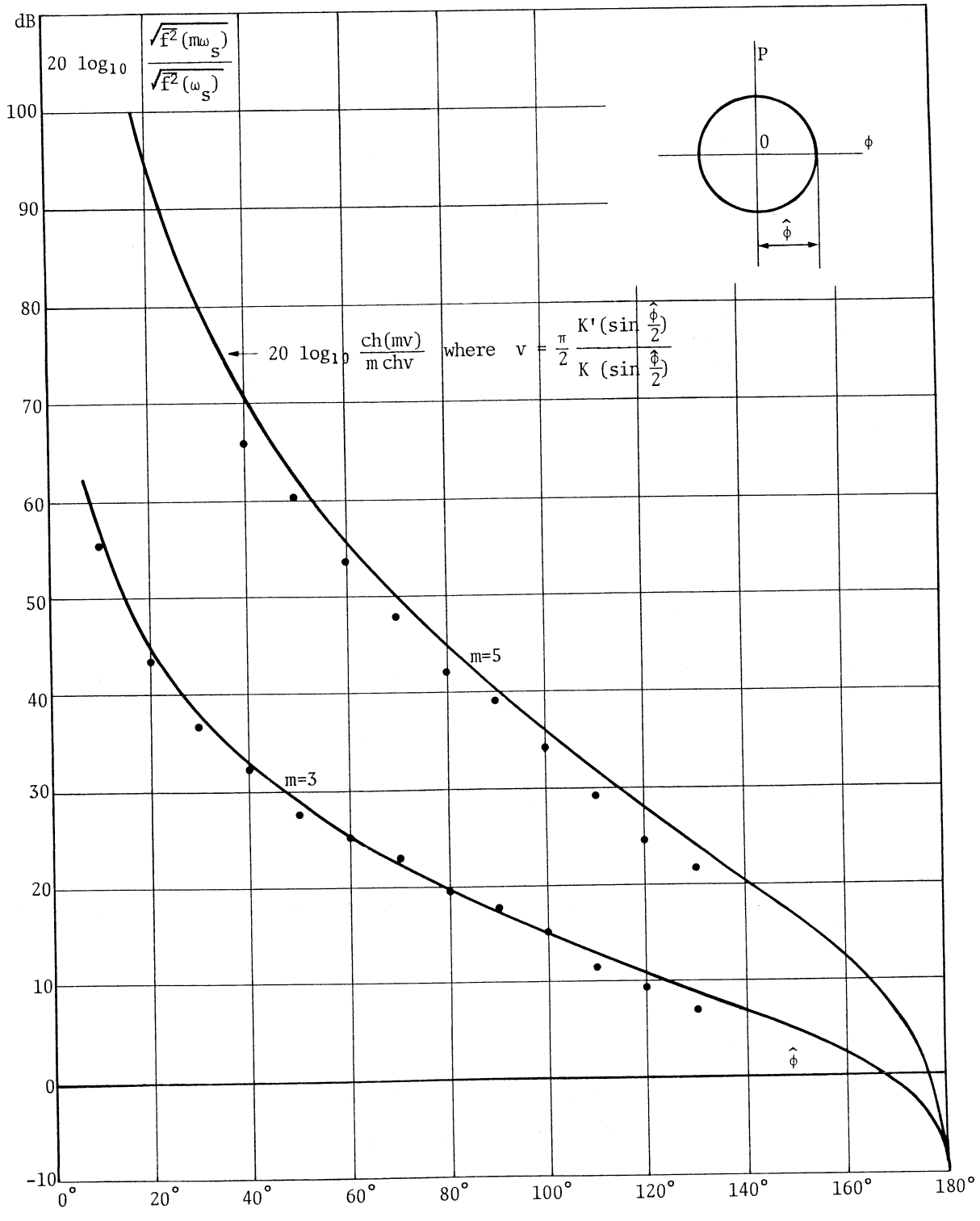


Fig. 5.1 Levels of frequency noise for equal rate of diffusion.
The dots are the results of numerical simulations;
the continuous curves are the theoretical estimates (5.17).

where

$$S_1 = \frac{\omega_s^2}{4}\left(\frac{2K}{\pi}\right)^3 \frac{\pi}{4}(1-k^2\alpha)\,\frac{\sum_{m=1,3,}^{\infty} \frac{m^4}{ch^2(mv)}\, S_\varphi(m\omega_s)}{\sum_{m=1,3,}^{\infty} \frac{m^4}{ch^2(mv)}}, \qquad S_2 = \frac{\omega_s^2}{4}\left(\frac{2K}{\pi}\right)^3 \frac{\pi}{4B}\,\alpha\,\frac{\sum_{m=2,4,}^{\infty} \frac{m^4}{sh^2(mv)}\, S_a(m\omega_s)}{\sum_{m=2,4,}^{\infty} \frac{m^4}{sh^2(mv)}} \tag{5.15}$$

This result has been given in Ref. 10, where the independent variable was taken as $J/2\Omega$; the only difference is that with $J/\hat{J}$ as independent variable, S_1 contains an extra factor $\pi/4$. Let us notice that the positive quantities S_1, S_2 are slowly varying functions of $k = \sin(\hat{\phi}/2)$.

Remark: In terms of frequency noise, eq. (5.12) reads:

$$\frac{A_2}{2} = 4\sum_{m=1,3,}^{\infty} \frac{(m\omega_s)^2}{ch^2(mv)}\, S_f(m\omega_s) + 4\sum_{m=2,4,}^{\infty} \frac{(m\omega_s)^4}{sh^2(mv)}\, S_a(m\omega_s) \tag{5.16}$$

Using bandpass filters with the same bandwidth centred on the harmonics m = 1, 2, 3, 5 of ω_s, G. Gurov (Ref. 15) compared the levels of frequency noise $\sqrt{\overline{f^2}(m\omega_s)}$ which are necessary to produce the same diffusion of particles at a given synchrotron amplitude $\hat{\phi}$. Since $\overline{f^2}(m\omega_s)$ is then proportional to $S_f(m\omega_s)$, $\sqrt{\overline{f^2}(m\omega_s)}$ should be inversely proportional to its weight m/ch(mv) in (5.16), which means that one expects to have

$$\frac{\sqrt{\overline{f^2}(m\omega_s)}}{\sqrt{\overline{f^2}(\omega_s)}} = \frac{ch(mv)}{m\ ch\ v} \qquad \text{where } v = \frac{\pi}{2}\frac{K'(k)}{K(k)} \text{ is determined by } k = \sin\frac{\hat{\phi}}{2}\,. \tag{5.17}$$

In Fig. 5.1 the results of Gurov's numerical simulations are compared to the theoretical estimate (5.17) at m = 3 and 5; Gurov also verified that the frequency noise at m = 2 has no effect on the particles.

6. FINITE DIFFERENCE EQUATIONS

Even without noise, the system (3.6) of finite difference equations is integrable only in the case of a linear RF voltage. In this simple case it reduces to

$$\begin{cases} P_{n+1} - P_n = -K_0\,(1+a_n)\,\phi_n \\ \phi_{n+1} - \phi_n = P_{n+1} + \varphi_{n+1} - \varphi_n \end{cases} \tag{6.1}$$

Motion without noise

With μ defined as in (11.6) of Ref. 7, the unperturbed motion is given by (11.7) of Ref. 7:

$$\phi_n = \mathrm{Re}\left[a\, e^{in\mu}\right], \qquad P_n = \mathrm{Re}\left[i\, a\, 2\sin\frac{\mu}{2}\, e^{i(n-\frac{1}{2})\mu}\right]$$

where "a" is a complex amplitude. It is seen that

$$\phi_n - i\,\frac{P_{n+1} + P_n}{2 \sin \mu} = a\, e^{in\mu} \qquad \text{and} \qquad \frac{\phi_n + \phi_{n-1}}{2 \cos \frac{\mu}{2}} - i\,\frac{P_n}{2 \sin \frac{\mu}{2}} = a\, e^{i(n-\frac{1}{2})\mu} \tag{6.2}$$

From (6.2) we derive a constant for the unperturbed motion:

$$\phi_n^2 + \left(\frac{P_{n+1} + P_n}{2 \sin \mu}\right)^2 = \left(\frac{\phi_n + \phi_{n-1}}{2 \cos \frac{\mu}{2}}\right)^2 + \left(\frac{P_n}{2 \sin \frac{\mu}{2}}\right)^2 = |a|^2 \tag{6.3}$$

Using (11.6) of Ref. 7 and (6.1), both expressions are converted into a single one:

$$\phi_n^2 + \left(\frac{P_n}{\sin \mu} - \operatorname{tg} \frac{\mu}{2} \cdot \phi_n\right)^2 = \left(\frac{\phi_n}{\cos \frac{\mu}{2}} - \frac{P_n}{2 \cos \frac{\mu}{2}}\right)^2 + \left(\frac{P_n}{2 \sin \frac{\mu}{2}}\right)^2 = \frac{\phi_n^2 - P_n \phi_n}{\cos^2 \frac{\mu}{2}} + \frac{P_n^2}{4 \sin^2 \frac{\mu}{2} \cos^2 \frac{\mu}{2}} = |a|^2 \tag{6.4}$$

In the P, ϕ phase plane where P is defined as in (3.8), all successive points of a particle trajectory lie on an ellipse with area

$$2\pi J = \frac{N}{T_0}\, \pi |a|^2 \sin \mu$$

Therefore, if one takes $x_0 = |a/2|^2 = (\hat{\phi}/2)^2$ as generalized coordinate (as done in Ref. 8), x_0 is proportional to the action J and by (5.1), it is very close, for small amplitudes, to $k^2 = \sin^2 (\hat{\phi}/2)$ or to $\left[k^2 (4/\pi) B(k)\right]$. In terms of x_0, the action reads

$$J = x_0 \cdot \frac{2N}{T_0} \sin\left(\frac{\omega_s T_0}{N}\right) \qquad \text{with} \quad \frac{\omega_s T_0}{N} = \mu \tag{6.5}$$

It should be noticed that in the case of a nonlinear RF voltage, this relation only applies for small amplitudes.

Motion with noise

For finite difference equations with a linear RF voltage, it can be verified that the remarkable relation (4.9) still applies (Ref. 8). Therefore, with the action J as generalized coordinate, the Fokker-Planck equation still reduces to a diffusion equation; with (6.5) the diffusion coefficient reads

$$D = \frac{A_2}{2} = \left(\frac{2N}{T_0}\right)^2 \sin^2\left(\frac{\omega_s T_0}{N}\right) \cdot \left[x_0 S_1 + x_0^2 S_2\right]$$

where (Ref. 3)

$$S_1 = \left(\frac{N}{T_0}\right)^2 \operatorname{tg}^2\left(\frac{\omega_s T_0}{2N}\right) \cdot \sum_{n=-\infty}^{+\infty} S_\varphi(\omega_s + nN\,\omega_0) \;, \qquad S_2 = \left(\frac{N}{T_0}\right)^2 \operatorname{tg}^2\left(\frac{\omega_s T_0}{2N}\right) \cdot \sum_{n=-\infty}^{+\infty} S_a(2\omega_s + nN\omega_0) \tag{6.6}$$

Finally,

$$D = \frac{A_2}{2} = \left(\frac{2N}{T_0} \sin \frac{\omega_s T_0}{2N}\right)^4 \left[x_0 \sum_{n=-\infty}^{+\infty} S_\varphi(\omega_s + nN\,\omega_0) + x_0^2 \sum_{n=-\infty}^{+\infty} S_a(2\omega_s + nN\,\omega_0)\right] \tag{6.7}$$

In order to compare (6.7) to the corresponding expression (5.12) for differential equations with a sinusoidal RF voltage, we must take in (5.12) the limit of vanishingly small amplitudes, which means that the infinite series in m both reduce to their first term; by (5.13) these terms coincide with the n = 0 terms in (6.7), provided that in (5.12), ω_s^4 is replaced by the first factor of (6.7).

For small synchrotron tune Q_{s_0}/N, the coarse motion is still well represented by differential equations; therefore, a reasonable generalization of (5.12) for finite difference equations with a sinusoidal RF voltage would involve replacing

$$\omega_s^4 \text{ by } \left(\frac{2N}{T_0} \sin \frac{\omega_s T_0}{2N}\right)^4 , \quad S_\varphi(m\omega_s) \text{ by } \sum_{n=-\infty}^{+\infty} S_\varphi(m\omega_s + nN\,\omega_0) , \quad S_a(m\omega_s) \text{ by } \sum_{n=-\infty}^{+\infty} S_a(m\omega_s + nN\,\omega_0) \tag{6.8}$$

The main effect of a finite N is that the spectral density of noise has to be taken into account not only at frequencies $m\omega_s$, but at all frequencies $(m\omega_s + nN\,\omega_0)$.

7. DIFFUSION EQUATION

Taking $x = J/\hat{J}$ as independent variable in the Fokker-Planck equation, the diffusion equation (4.11) reads, with (5.14):

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left[(x\,S_1 + x^2\,S_2)\,\frac{\partial \rho}{\partial x}\right] \tag{7.1}$$

where x = 1 on the separatrix; S_1 and S_2 are slowly varying functions of x.

Boundary conditions

At x = 0 the flux of particles must vanish:

$$\lim_{x \to 0} (x\,S_1 + x^2\,S_2)\,\frac{\partial \rho}{\partial x} = 0 \tag{7.2}$$

At x = 1 particles are lost, which entails that

$$\rho\,(x = 1,\ t) = 0 \tag{7.3}$$

Normalization condition

At t = 0,

$$\int_0^1 \rho(x,0)\,dx = 1 \tag{7.4}$$

Let us restrict to phase noise. Since the action J is an area in phase space, we may put $x = r^2$ where r, φ are polar coordinates. The diffusion equation becomes

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(x\,S_1\,\frac{\partial \rho}{\partial x}\right) = \frac{1}{4}\,\frac{\partial}{\partial r}\left(r\,S_1\,\frac{\partial \rho}{\partial r}\right) \tag{7.5}$$

It is solved by separation of variables, putting $\rho(r,t) = R(r) \cdot T(t)$. Then

$$\frac{1}{T} \cdot \frac{dT}{dt} = \frac{1}{4R} \cdot \frac{d}{r\,dr}\left(r\,S_1\,\frac{dR}{dr}\right) = -\frac{\lambda^2}{4} \qquad \text{where } \lambda^2 \text{ is some constant.}$$

T varies according to

$$T = e^{-(\lambda^2/4)t} \tag{7.6}$$

whereas R must satisfy

$$\frac{d}{dx}\left(x\, S_1 \frac{dR}{dx}\right) + \frac{\lambda^2}{4} R = 0 \quad \text{with} \quad \lim_{x\to 0} x\, S_1 \frac{dR}{dx} = 0 \quad \text{and} \quad R(1) = 0$$

or (7.7)

$$\frac{d}{r\,dr}\left(r\, S_1 \frac{dR}{dr}\right) + \lambda^2 R = 0 \qquad \lim_{r\to 0} r\, S_1 \frac{dR}{dr} = 0 \qquad R(1) = 0$$

The conditions (7.7) define λ^2 and R(r) as eigenvalue and eigenfunction of the operator $-1/r.d/dr\ (r\, S_1\, d/dr)$; we anticipate here that all eigenvalues λ_n^2 are positive because $S_1 > 0$. The general solution of (7.1) then reads:

$$\rho(x,t) = \sum_{n=1}^{\infty} a_n\, e^{-(\lambda_n^2/4)t}\, R_n(x) \tag{7.8}$$

The smallest eigenvalue λ_1^2 determines the ultimate lifetime of the particles under noise diffusion; it can be computed easily in a few cases only, when $S_1(x)$ is simple enough.

Case 1: $S_1(x) = d_0^2$ constant

$$R(r) = J_0\left(\frac{\lambda_n}{d_0} r\right)$$

with

$$\frac{\lambda_n}{d_0} = j_{0n}, \qquad n^{th} \text{ root of Bessel function } J_0.$$

The smallest eigenvalue is thus given by

$$\frac{\lambda_1}{d_0} = j_{01} = 2.404\ 826$$

Case 2: $S_1(x) = d_0^2(1-x)^{-\nu}$. The solution satisfying the boundary condition (7.7) when $x \to 0$ is of the form

$$R(x) = \sum_{m=0}^{\infty} \alpha_m\, x^m$$

where the coefficients α_m can be computed numerically by a recurrence relation derived from the differential equation (7.7); λ^2 is then adjusted until R(1) = 0.

If $\nu \leq -1$, the condition R(1) = 0 cannot be met. For $-2 < \nu \leq -1$, λ^2 is determined by requiring that R(1) remains finite; for $\nu \leq -2$ even the latter condition is only met when $\lambda^2 = 0$, and the λ^2 spectrum is continuous. In all these cases the flux of particles at the separatrix is zero, which means that no particles are lost; when $t \to \infty$, $\rho(x,t)$ approaches an equilibrium value (corresponding to $\lambda^2 = 0$) which is independent of x.

Unfortunately, the condition $\nu \leq -1$ implies that the RF noise goes fast enough to an absolute zero (i.e. $-\infty$ dB) close to the separatrix; this cannot be realized in practice.

If $\nu = -1$, the eigenvalues are given by

$$\frac{\lambda_{n+1}^2}{4d_0^2} = n(n+1)$$

where $n = 0, 1, 2, \ldots$; the corresponding eigenfunctions are the Legendre polynomials $P_n(1-2x)$, and $R_n(1) = (-1)^n$. The smallest eigenvalue λ_1^2 is zero, with eigenfunction $P_0(1-2x) = 1$ independent of x.

If $\nu > -1$, numerical computations yield the following results:

ν	$\frac{\lambda_1^2}{4d_0^2}$	
$-1+\varepsilon$	$< (\nu+1)\,\frac{1+2(\nu+1)}{1+(\nu+1)}$	obtained by a Rayleigh-Ritz variational method (see Appendix B)
-0.5	0.658 424	
0	1.445 796	
0.5	2.264 778	
1	3.094 169	
2	4.763 840	
∞	$< \frac{27}{16}\,(\nu + \frac{5}{6})$	obtained by a Rayleigh-Ritz variational method (see Appendix B)

Case 3: $S_1(x) = d_0^2 \cdot x^\nu$. The solution satisfying (7.7) is

$$R = z^{-\mu} J_\mu(z) \qquad \text{where } \mu = \frac{\nu}{1-\nu}, \quad z = \frac{\lambda}{d_0}\,\frac{r^{1-\nu}}{1-\nu} \tag{7.9}$$

The condition $R(1) = 0$ reads

$$J_\mu\left(\frac{\lambda}{d_0}\,\frac{1}{1-\nu}\right) = 0 \tag{7.10}$$

From properties of zeros of Bessel functions, one gets the following results:

ν	μ	$\frac{\lambda_1^2}{4d_0^2}$	
$-\infty$	-1	$1.5 - \nu - \ldots$	$\frac{\lambda_1}{d_0} < \sqrt{2-\nu} + \sqrt{1-\nu}$ (7.11)
-1	$-\frac{1}{2}$	2.467 401	obtained by a Rayleigh-Ritz variational method (see Appendix B)
0	0	1.445 796	
0.5	1	0.917 623	
$1-\varepsilon$	∞	$0.25 + 0.927\,879\,(1-\nu)^{2/3} + \ldots$	
≥ 1	< -1	$0 < \frac{\lambda^2}{4d_0^2} < \infty$	continuous spectrum

In the case of a continuous spectrum, the diffusion equation can be solved by using Laplace transforms.

Let us remark in passing that (7.11) provides a very good upper estimate for the first zero $j_{\mu 1}$ of the Bessel function $J_\mu(x)$. From (7.9) we have

$$\mu + 1 = \frac{1}{1-\nu} \geq 0$$

Comparing (7.10) and (7.11) we obtain

$$j_{\mu 1} = \frac{\lambda_1}{d_o} \cdot \frac{1}{1-\nu} < \frac{1}{\sqrt{1-\nu}} \left[\sqrt{\frac{1}{1-\nu} + 1} + 1 \right]$$

or

$$\boxed{j_{\mu 1} < \sqrt{\mu + 1} \left[\sqrt{\mu + 2} + 1 \right] \qquad \mu \geq -1} \tag{7-12}$$

In practice, the noise level increases from the centre of the bucket towards the separatrix, which corresponds to ν being positive in cases 2 and 3.

Bunch current $I(\phi)$

With the normalization (7.4) we have, if q is the total charge in the bunch at t = 0:

$$I(\phi)\, dt = q \int \rho \frac{dJ}{\hat{J}} \frac{d\theta}{2\pi} \tag{7.13}$$

But $dJ \cdot d\theta = dP \cdot d\phi$ since it is the elementary area in phase space. From (5.1) we may write

$$k^2 = \sin^2 \frac{\phi}{2} + y^2 \qquad \text{where } y = \frac{P}{2\Omega} \tag{7.14}$$

so that (7.13) reads

$$I(\phi)\, dt = \frac{q}{2\pi \hat{J}} \int \rho\, dP\, d\phi = q \frac{2\Omega}{2\pi \hat{J}} \int \rho\, dy\, \omega_{RF}\, dt$$

With (5.2) and (7.14):

$$I(\phi) = q \frac{\omega_{RF}}{4} \int_0^{\cos \frac{\phi}{2}} \rho(x,t)\, dy$$

Using (7.8) we finally get

$$I(\phi) = q \frac{\omega_{RF}}{4} \sum_{n=1}^{\infty} a_n\, e^{-(\lambda_n^2/4)t} \underbrace{\int_0^{\cos \frac{\phi}{2}} R_n(x)\, dy}_{\cos \frac{\phi}{2} \cdot (\text{Function of } \sin^2 \frac{\phi}{2})} \qquad \begin{array}{l} \text{where } x = k^2 \cdot B(k) \\ \text{and } k^2 = \sin^2 \frac{\phi}{2} + y^2 \end{array} \tag{7.15}$$

This relation yields the shape of the instantaneous bunch current as a function of time.

8. CONTRIBUTION OF RF NOISE TO THE FINITE BEAM LIFE-TIME IN THE SPS COLLIDER

This is a short historical survey of the improvements achieved by D. Boussard and the RF group on the life time of proton (and later anti-proton) bunches in the SPS when operated as a storage ring for p $\bar{p}$ collisions.
In the following, τ represents that contribution to the finite beam life-time which is due to RF noise only.

- With original RF system (June 1978): $\tau \approx 1$ minute.
- Insertion of a low pass filter (cut-off 10 Hz) in radial loop (December 1978): $\tau \approx 10$ minutes.
- Sampling of the beam-phase pick-up during the passage of the reference (proton) bunch (October 1979): $\tau \approx 10$ hours.
  This sampling is the change which produced the most dramatic improvement in τ. With an ideal phase loop having infinite gain and no phase noise, the particles in the reference bunch would see no phase noise.
- Replacement of the radial loop by a frequency programme (January 1980): $\tau \approx 40$ hours.
- New low-noise master oscillator (August 1981): $\tau = 100$ hours for the reference bunch, and almost the same for the other bunches.

REFERENCES

(1) G.M. Jenkins and D.G. Watts, Spectral analysis and its applications, Holden-Day, San Francisco, 1968.

(2) A. Papoulis, Signal analysis, McGraw-Hill International Student Edition, Tokyo, 1977.

(3) C.K. Yuen and D. Frazer, Digital spectral analysis, Pitman, London, 1979.

(4) G.E. Uhlenbeck and L.S. Ornstein, On the theory of the Brownian motion, Physical Review, Vol. 36 (1930), p. 823-841 (see §6, p. 828); reprinted in "Selected papers on noise and stochastic processes", edited by N. Wax, Dover, New York, 1954, p. 93-111.

(5) D. Möhl, G. Petrucci, L. Thorndahl and S. van der Meer, Physics and technique of stochastic cooling, CERN-PS-AA/79-23 (p. 24), or Physics Reports, Vol. 58 (1980), p. 73-119.

(6) E. Peschardt, F. Sacherer, L. Thorndahl, S. van der Meer, A. Vaughan, Diffusion in momentum caused by filtered noise, ISR Performance Report, 19th August 1977 (unpublished).

(7) G. Dôme: Theory of RF acceleration, these proceedings, p. 110-158.

(8) G. Dôme, Synchrotron motion in a stationary bucket with RF noise, Workshop on $p\bar{p}$ in the SPS, SPS-$p\bar{p}$-1, 9 May 1980, p. 102-119 (see References p. 91).

(9) H. Goldstein, Classical Mechanics, Addison-Wesley, Reading, Mass. (1965).

(10) D. Boussard, G. Dôme, C. Graziani, The influence of RF noise on the lifetime of bunched proton beams, SPS/ARF/80-42; or 11th International Conference on High-Energy Accelerators, CERN, 1980, p. 620-626.

(11) S. Krinsky and J.M. Wang, Bunch diffusion due to RF noise, BNL 30877, 1982; or Particle Accelerators, Vol. 12 (1982), p. 107-117.

(12) E. Jahnke and F. Emde, Tables of functions, Dover Publications, New York, 1945.

(13) H. Hancock, Theory of elliptic functions, Dover Publications, New York, 1958.

(14) I.S. Gradshteyn and I.M. Ryzhik, Table of Integrals, series and products, Academic Press, New York, 1980.

(15) G. Gurov, private communication, November 1979.

APPENDIX A Computation of $\sum_{m=1,3,}^{\infty} \frac{m^4}{ch^2(mv)}$ and $\sum_{m=2,4,}^{\infty} \frac{m^4}{sh^2(mv)}$

If $h(\theta) = \sum_{m=-\infty}^{+\infty} a_m e^{im\theta}$, then

$$\sum_{m=-\infty}^{+\infty} |a_m|^2 = \frac{1}{2\pi} \int_0^{2\pi} |h(\theta)|^2 \, d\theta \tag{A.1}$$

From (5.8) and (5.6),

$$\sum_{m\ odd=-\infty}^{+\infty} \frac{e^{im\theta}}{ch(mv)} = k\left(\frac{2K}{\pi}\right) cn(u,k) \qquad \text{where } \theta = \frac{\pi u}{2K}$$

Therefore,

$$\sum_{m\ odd=-\infty}^{+\infty} \frac{(im)^2}{ch(mv)} e^{im\theta} = k\left(\frac{2K}{\pi}\right)^3 \cdot \frac{\partial^2}{\partial u^2} cn(u,k)$$

and

$$\sum_{m\ odd=-\infty}^{+\infty} \frac{m^4}{ch^2(mv)} = \frac{1}{2\pi}\frac{\pi}{2K} \int_0^K \left| k\left(\frac{2K}{\pi}\right)^3 \cdot \frac{\partial^2}{\partial u^2} cn\ u \right|^2 du = \frac{2}{\pi} k^2 \left(\frac{2K}{\pi}\right)^5 \int_0^K \left| \frac{\partial^2}{\partial u^2} cn\ u \right|^2 du \tag{A.2}$$

Now

$$\frac{\partial^2}{\partial u^2} cn\ u = -\frac{\partial}{\partial u}\left[sn\ u \cdot dn\ u\right] = -cn\ u \cdot dn^2\ u + k^2\ sn^2\ u \cdot cn\ u$$

$$= -\ cn\ u\ (dn^2\ u - k^2\ sn^2\ u) = -\ cn\ u\ (2dn^2\ u - 1)$$

hence

$$\left|\frac{\partial^2}{\partial u^2} cn\ u\right|^2 = cn^2\ u\ (1 - 4\ dn^2\ u + 4\ dn^4\ u) = cn^2\ u\ + 4\ cn^2\ u\ dn^2\ u\ (dn^2\ u - 1)$$

$$= cn^2\ u - 4\ k^2\ sn^2\ u\ cn^2\ u\ dn^2\ u$$

The integral with respect to u is most easily computed by reverting to (5.3), (5.4):

$$\int_0^K \left|\frac{\partial^2}{\partial u^2} cn\ u\right|^2 du = \int_0^K cn^2\ u \cdot du - 4k^2 \int_0^K sn^2\ u\ cn^2\ u\ dn^2\ u \cdot du$$

$$= \int_0^{\pi/2} \frac{\cos^2\psi}{\sqrt{1 - k^2 \sin^2\psi}}\, d\psi - 4k^2 \int_0^{\pi/2} \sin^2\psi \cdot \cos^2\psi \sqrt{1 - k^2 \sin^2\psi}\, d\psi \tag{A.3}$$

The first integral in (A.3) is B(k) (Ref. 12). The second integral is given by (Ref. 23, p. 386)

$$\frac{2}{\pi}\int_0^{\pi/2} \sin^2\psi \cdot \cos^2\psi \cdot \sqrt{1 - k^2\sin^2\psi}\, d\psi = \frac{1}{\pi} B\left(\frac{3}{2}, \frac{3}{2}\right) F\left[-\frac{1}{2}, \frac{3}{2}; 3; k^2\right] = \frac{1}{8} F\left[-\frac{1}{2}, \frac{3}{2}; 3; k^2\right]$$

By repeated application of Gauss' relations for contiguous hypergeometric functions, the above hypergeometric function can be expressed as a linear combination of

$$F\left[\frac{1}{2}, \frac{1}{2}; 1; k^2\right] = \frac{2K}{\pi} \quad \text{and} \quad F\left[-\frac{1}{2}, \frac{1}{2}; 1; k^2\right] = \frac{2E}{\pi} ;$$

there results

$$\int_0^K sn^2\, u\, cn^2\, u\, dn^2\, u\, du$$

$$= \int_0^{\pi/2} \sin^2\psi \cdot \cos^2\psi \cdot \sqrt{1 - k^2\sin^2\psi}\, d\psi = \frac{1}{15\, k^4}\left[(k^2-2)\, k'^2 K + 2(1 - k^2 + k^4)E\right]$$

With the notations of Ref. 12, this is transformed into

$$\frac{1}{15\, k^4}\left[(k^2-2)(E-k^2B) + 2(1-k^2+k^4)E\right] = \frac{1}{15\, k^2}\left[(1+k'^2)B + E(2k^2-1)\right]$$

$$= \frac{1}{15\, k^2}\left[(E-k^2k'^2C) + E(2k^2-1)\right] = \frac{1}{15}\left[2E - k'^2C\right] = \frac{B}{4}\alpha(k^2) \qquad (A.4)$$

where $\alpha(k^2)$ is defined by the last relation; it decreases from 1 to 8/15 when k^2 increases from 0 to 1.

Finally,

$$4\int_0^K sn^2\, u\, cn^2\, u\, dn^2\, u\, du = B \cdot \alpha(k^2) \qquad (A.5)$$

and (A.3) becomes

$$\int_0^K \left|\frac{\partial^2}{\partial u^2} cn\, u\right|^2 du = B(1-k^2\alpha)$$

hence, from (A.2):

$$2\sum_{m=1,3,}^{\infty} \frac{m^4}{ch^2(m\nu)} = k^2\, \frac{2}{\pi}\, B\left(\frac{2K}{\pi}\right)^5 (1-k^2\alpha) \qquad (A.6)$$

For the second series, we start from (5.10):

$$\text{(without } m = 0) \qquad \sum_{m \text{ even}=-\infty}^{+\infty}{}' \frac{m}{\mathrm{sh}(mv)} e^{im\theta} = \left(\frac{2K}{\pi}\right)^2 \left[\left(1 - \frac{E}{K}\right) - k^2 \, \mathrm{sn}^2 \, u\right]$$

Therefore,

$$\sum_{m \text{ even}=-\infty}^{+\infty}{}' \frac{im^2}{\mathrm{sh}(mv)} e^{im\theta} = -k^2 \left(\frac{2K}{\pi}\right)^3 \frac{\partial}{\partial u} \mathrm{sn}^2 \, u$$

and, with (A.1):

$$\sum_{m \text{ even}=-\infty}^{+\infty}{}' \frac{m^4}{\mathrm{sh}^2(mv)} = \frac{1}{2\pi} \frac{\pi}{2K} \int_0^{4K} \left| k^2 \left(\frac{2K}{\pi}\right)^3 \frac{\partial}{\partial u} \mathrm{sn}^2 \, u \right|^2 du = \frac{2}{\pi} k^4 \left(\frac{2K}{\pi}\right)^5 \int_0^{K} \left| \frac{\partial}{\partial u} \mathrm{sn}^2 \, u \right|^2 du \tag{A.7}$$

Now $\partial/\partial u(\mathrm{sn}^2 u) = 2 \, \mathrm{sn} \, u \, \mathrm{cn} \, u \, \mathrm{dn} \, u$ and the integral in (A.7) is the same as (A.5); thus

$$2 \sum_{m=2,4,}^{\infty} \frac{m^4}{\mathrm{sh}^2(mv)} = k^4 \frac{2}{\pi} B \left(\frac{2K}{\pi}\right)^5 \alpha \tag{A.8}$$

APPENDIX B Rayleigh-Ritz method and diffusion equation

Rayleigh-Ritz method for the smallest eigenvalue λ^2 of a second order differential equation

$$\frac{d}{dx}\left[p(x)\frac{dR}{dx}\right] + \lambda^2 q(x)\, R = 0 \qquad \text{where} \qquad p(x) \geq 0\ ,\ q(x) \geq 0 \tag{B.1}$$

with boundary conditions

$$\lim_{x\to 0}\left[p(x)\,\frac{dR}{dx}\right] = 0 \qquad\qquad R(x=1) = 0 \tag{B.2}$$

From (B.1), for any function $Z(x)$ we have

$$Z\,\frac{d}{dx}\left[p(x)\,\frac{dR}{dx}\right] + \lambda^2 q(x)\, R\, Z = 0$$

hence

$$\left[Z\,p(x)\,\frac{dR}{dx}\right]_0^1 - \int_0^1 p(x)\,\frac{dR}{dx}\,\frac{dZ}{dx}\,dx + \lambda^2 \int_0^1 q(x)\, R\, Z\, dx = 0 \tag{B.3}$$

If $\qquad |Z(x=0)| < \infty \qquad$ and $\qquad Z(x=1) = 0 \tag{B.4}$

(in particular one may take $Z = R$), then from (B.3)

$$\lambda^2 = \frac{\displaystyle\int_0^1 p(x)\,\frac{dR}{dx}\,\frac{dZ}{dx}\,dx}{\displaystyle\int_0^1 q(x)\, R\, Z\, dx} = \frac{\displaystyle\int_0^1 p(x)\,\left(\frac{dR}{dx}\right)^2 dx}{\displaystyle\int_0^1 q(x)\, R^2\, dx}$$

By expanding Z in a series of eigenfunctions R_n of (B.1), it can be shown that

$$\lambda_1^2 \leq \frac{\displaystyle\int_0^1 p(x)\,\left(\frac{dZ}{dx}\right)^2 dx}{\displaystyle\int_0^1 q(x)\, Z^2\, dx} \qquad \text{for the smallest eigenvalue} \tag{B.5}$$

Diffusion equation

The equation (7.7) for the eigenvalue λ^2 reads

$$\frac{d}{dx}\left[x\, S_1\,\frac{dR}{dx}\right] + \frac{\lambda^2}{4} R = 0 \tag{B.6}$$

with

$$\lim_{x\to 0} x\, S_1\,\frac{dR}{dx} = 0 \qquad\qquad R(x=1) = 0$$

Case 2: $S_1(x) = d_o^2 (1 - x)^{-\nu}$

In order to meet the boundary condition $R(1) = 0$, we must have

$$\nu + 1 = \frac{1}{m} > 0 \tag{B.7}$$

Let $z = 1 - x$ in (B.6):

$$\frac{d}{dz}\left[(1 - z)\, z^{-\nu} \frac{dR}{dz}\right] + \frac{\lambda^2}{4 d_o^2} R = 0 \tag{B.8}$$

with boundary conditions

$$\lim_{z \to 1} (1 - z)\, z^{-\nu} \frac{dR}{dz} = 0 \qquad R(z = 0) = 0$$

A simple function for $Z(z)$ is $Z(z) = z^{\ell}$ where $\ell > 0$ in order to satisfy (B.4). From (B.5),

$$\frac{\lambda_1^2}{4 d_o^2} \leq \frac{\int_0^1 (1 - z)\, z^{-\nu} \left(\ell z^{\ell-1}\right)^2 dz}{\int_0^1 \left(z^{\ell}\right)^2 dz} = \ell^2 \frac{\int_0^1 (1 - z)\, z^{2\ell-\nu-2} dz}{\int_0^1 z^{2\ell} dz}$$

where, for convergence of the numerator:

$$2\ell - (\nu + 1) > 0 \tag{B.9}$$

Therefore

$$\frac{\lambda_1^2}{4 d_o^2} \leq \ell^2 (2\ell + 1) \left[\frac{1}{2\ell - \nu - 1} - \frac{1}{2\ell - \nu}\right]$$

or, with (B.7):

$$\frac{\lambda_1^2}{d_o^2} \leq (2\ell)^2 \frac{(2\ell + 1)}{(2\ell - \nu - 1)(2\ell - \nu)} = \frac{1}{m} (2\ell m)^2 \frac{(2\ell m + m)}{(2\ell m - 1)(2\ell m + m - 1)}$$

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{m} \frac{(L + 1)^2}{L} \frac{L + m + 1}{L + m} \tag{B.10}$$

where from (B.7) and (B.9) $\qquad L = 2\ell m - 1 > 0 \tag{B.11}$

The best estimate for λ_1^2 is obtained by minimizing this expression with respect to L ; this entails

$$(L - 1)\left[(m + L)^2 + m - 2\right] = 2 \qquad \text{(B.12)}$$

When $m = \infty$, $L = 1$

When $m = 0$, $(L - 1)(L - 2) = 2$ i.e. $L = \begin{cases} 2 \\ 0 \\ -1 \end{cases}$

With (B.11), the only valid solution is $L = 2 > 0$ when $m = 0$.

Therefore, using (B.12):

$$L - 1 = \frac{2}{(m + L)^2 + m - 2} \rightarrow \begin{cases} \dfrac{2}{(m + 1)^2 + m - 2} = \dfrac{2}{m^2 + 3m - 1} & \text{when } m \rightarrow \infty \\ \dfrac{2}{(m + 2)^2 + m - 2} = \dfrac{2}{m^2 + 5m + 2} = 1 - \dfrac{5}{2} m + \ldots & \text{when } m \rightarrow 0 \end{cases} \qquad \text{(B.13)}$$

Now (B.10) may be rewritten as

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{m}\left[L + 2 + \frac{1}{L}\right]\left[1 + \frac{1}{m + L}\right]$$

With (B.13) and (B.7), this becomes

$$\boxed{\frac{\lambda_1^2}{d_o^2} \leq \frac{4}{m}\left[1 + \frac{1}{m + 1 + 2/m^2}\right] < \frac{4}{m}\frac{m+2}{m+1} = 4(\nu + 1) \cdot \frac{1 + 2(\nu + 1)}{1 + (\nu + 1)}}$$

when $m \rightarrow \infty$ i.e. $\nu + 1 \rightarrow 0$ (B.14)

and, when $m \rightarrow 0$:

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{m}\left[4 - \frac{5}{2} m + \frac{1}{2 - \frac{5}{2} m}\right]\left[1 + \frac{1}{2 - \frac{3}{2} m}\right]$$

$$= \frac{1}{m}\left[\frac{9}{2} - \frac{3}{4}\frac{5}{2} m + \ldots\right]\left[\frac{3}{2} + \frac{3}{8} m + \ldots\right] = \frac{1}{m}\frac{27}{4}\left[1 - \frac{5}{3.4} m + \ldots\right]\left[1 + \frac{1}{4} m + \ldots\right]$$

$$\boxed{\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{m}\frac{27}{4}\left[1 - \frac{2}{3.4} m\right] = \frac{27}{4}\left[\frac{1}{m} - \frac{1}{6}\right] = \frac{27}{4}\left[\nu + \frac{5}{6}\right]}$$

when $m \rightarrow 0$ i.e. $\nu + 1 \rightarrow \infty$ (B.15)

Case 3: $S_1(x) = d_o^2\, x^\nu$

Then (B.6) becomes

$$\frac{d}{dx}\left(x^{\nu+1}\frac{dR}{dx}\right) + \frac{\lambda^2}{4d_o^2}R = 0 \qquad \text{(B.16)}$$

with boundary conditions

$$\lim_{x\to 0} x^{\nu+1}\frac{dR}{dx} = 0 \qquad R(x=1) = 0$$

A simple function for $Z(x)$ is $Z(x) = \left(1 - x^{\frac{1}{\sigma}}\right)^\ell$ where from (B.4) $\sigma > 0$, $\ell > 0$.

Applying (B.5):

$$\frac{\lambda_1^2}{4d_o^2} \leq \frac{\int_0^1 x^{\nu+1}\left(\frac{dZ}{dx}\right)^2 dx}{\int_0^1 Z^2\, dx}$$

$$= \frac{\int_0^1 \left[-\ell(1-z)^{\ell-1}\frac{1}{\sigma z^{\sigma-1}}\right]^2 \sigma z^{\sigma(\nu+2)-1}\, dz}{\int_0^1 \left[(1-z)^\ell\right]^2 \sigma z^{\sigma-1}\, dz}$$

where we have put $x = z^\sigma$

This becomes

$$\frac{\lambda_1^2}{4d_o^2} \leq \left(\frac{\ell}{\sigma}\right)^2 \frac{\int_0^1 (1-z)^{2\ell-2} z^{\sigma\nu+1}\, dz}{\int_0^1 (1-z)^{2\ell} z^{\sigma-1}\, dz}$$

where, for convergence of the numerator:

$$L = 2\ell - 1 > 0 \qquad \sigma\nu + 2 > 0 \qquad \text{(B.17)}$$

Therefore

$$\frac{\lambda_1^2}{d_o^2} \leq \left(\frac{2\ell}{\sigma}\right)^2 \frac{\frac{\Gamma(2\ell - 1)\ \Gamma(\sigma\nu + 2)}{\Gamma(2\ell + 1 + \sigma\nu)}}{\frac{\Gamma(2\ell + 1)\ \Gamma(\sigma)}{\Gamma(2\ell + 1 + \sigma)}}$$

or

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{\sigma} \frac{2\ell}{2\ell-1} \frac{\Gamma(2\ell + 1 + \sigma)\ \Gamma(\sigma\nu + 2)}{\Gamma(2\ell + 1 + \sigma\nu)\ \Gamma(\sigma + 1)} \tag{B.18}$$

If $\nu > 1$,

$$\frac{\lambda_1^2}{d_o^2} \leq \lim_{2\ell \to \infty} \frac{1}{\sigma} \cdot \left[2\ell + \frac{1}{2} + \sigma \frac{\nu + 1}{2}\right]^{\sigma(1-\nu)} \cdot \frac{\Gamma(\sigma\nu + 2)}{\Gamma(\sigma + 1)} = 0 \quad ;$$

the λ^2 spectrum is continuous.

If $\nu < 1$, choose σ such that $\sigma\nu = \sigma - 1$ i.e. $\sigma = \frac{1}{1 - \nu} > 0$ (B.19)

This value of σ satisfies also (B.17) since $\sigma\nu + 2 = \sigma + 1 > 0$.

Then (B.18) reduces to

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{\sigma} \frac{2\ell}{2\ell-1} (2\ell + \sigma) = \frac{1}{\sigma} \frac{L + 1}{L} (L + \sigma + 1) = \frac{1}{\sigma} (L + 1) \left(\frac{\sigma+1}{L} + 1\right)$$

Minimizing with respect to L yields $L^2 = \sigma + 1$, $L = \sqrt{\sigma + 1}$

Then

$$\frac{\lambda_1^2}{d_o^2} \leq \frac{1}{\sigma} (L + 1)^2 \qquad \text{or} \qquad \frac{\lambda_1}{d_o} \leq \frac{L + 1}{\sqrt{\sigma}} = \frac{\sqrt{\sigma + 1} + 1}{\sqrt{\sigma}}$$

Finally, with (B.19):

$$\boxed{\frac{\lambda_1}{d_o} \leq \sqrt{2 - \nu} + \sqrt{1 - \nu} \qquad \text{when} \qquad \nu < 1} \tag{B.20}$$

INTRA-BEAM SCATTERING

A. Piwinski*)
CERN, Geneva, Switzerland

ABSTRACT

Intra-beam scattering is analysed and the rise times or damping times of the beam dimensions are derived. The theoretical results are compared with experimental values obtained on the CERN AA and SPS machines.

1. INTRODUCTION

Intra-beam scattering or multiple scattering is a Coulomb scattering between the particles within a bunch or within an unbunched beam. It can be compared with the scattering of gas molecules in a closed box. The focusing forces and the rf accelerating voltage of the storage ring play the same role as the walls of the box since they keep the particles together.

It is well known that the scattering of the molecules in a closed box leads to a Gaussian distribution of the three components of the velocities:

$$\phi(\nu_s,\nu_x,\nu_z) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left\{-\frac{m}{kT}\left(\nu_s^2 + \nu_x^2 + \nu_z^2\right)\right\}$$

where m is the mass of the molecules, T is the absolute temperature and k is Boltzmann's constant. This can be proved by assuming that the distributions of the three velocity components are independent.

The situation is different in the case of intra-beam scattering. Due to the dispersion a change in energy always causes a change in the betatron amplitude, and a coupling arises between the synchrotron oscillation and the betatron oscillation. Furthermore, above transition energy the particles behave as if they had a negative mass, i.e. an increase of energy reduces the revolution frequency. We will see that due to this behaviour an equilibrium distribution of the particles cannot exist above transition energy, and the intra-beam scattering will increase all three dimensions of the bunch in so far as they do not hit other limitations.

But even when an equilibrium distribution exists (below transition energy) the initial distribution will, in general, be different from the equilibrium distribution and the change of the distribution due to the intra-beam scattering can reduce the lifetime or the luminosity of the storage ring. The main purpose of the following investigation is to determine the rise times or damping times of the bunch dimensions. This is done as follows:

---

* On leave from DESY, Hamburg, Fed. Rep. Germany

a) The transformation of the momenta of two colliding particles into their centre-of-mass system.

b) The changes of the momenta due to the collision are calculated.

c) The changed momenta are then transformed back into the laboratory system.

d) The changes of the momenta give the changes of the emittances of the betatron oscillations and the change of the amplitude of the synchrotron oscillation.

e) At first the average is taken over all scattering angles using the Rutherford cross section.

f) Then it is averaged over all momenta and positions of the colliding particles assuming Gaussian distributions.

g) Now one can see that the result of the intra-beam scattering is different below and above transition energy.

h) The average values also give the rise times or damping times of the bunch dimensions.

i) Finally experimental results of accelerators at CERN are discussed.

2. CALCULATION OF RISE TIMES AND DAMPING TIMES

a) Lorentz transformation

The three momenta of the two particles before the collision are given by

$$\vec{P}_{1,2} = P_{1,2}\, s\, \{1,\ x'_{1,2},\ z'_{1,2}\}_{s,x,z}$$

in the coordinate system {s,x,z} (see Fig. 1). For the Lorentz-transformation a new coordinate system {u,v,w} is defined by

$$\vec{e}_u = (\vec{p}_1 + \vec{p}_2)/|\vec{p}_1 + \vec{p}_2|$$

$$\vec{e}_v = \vec{p}_1 \times \vec{p}_2/|\vec{p}_1 \times \vec{p}_2|$$

$$\vec{e}_w = \vec{e}_u \times \vec{e}_w$$

where $\vec{e}_u$, $\vec{e}_v$ and $\vec{e}_w$ are the unit vectors parallel to the coordinate axes. In this coordinate system the two momenta can be written as

$$\vec{p}_{1,2} = p_{1,2}\, \{\cos\alpha_{1,2},\ 0,\ \pm\sin\alpha_{1,2}\}_{u,v,w}$$

where α_1 and α_2 are shown in Fig. 1.

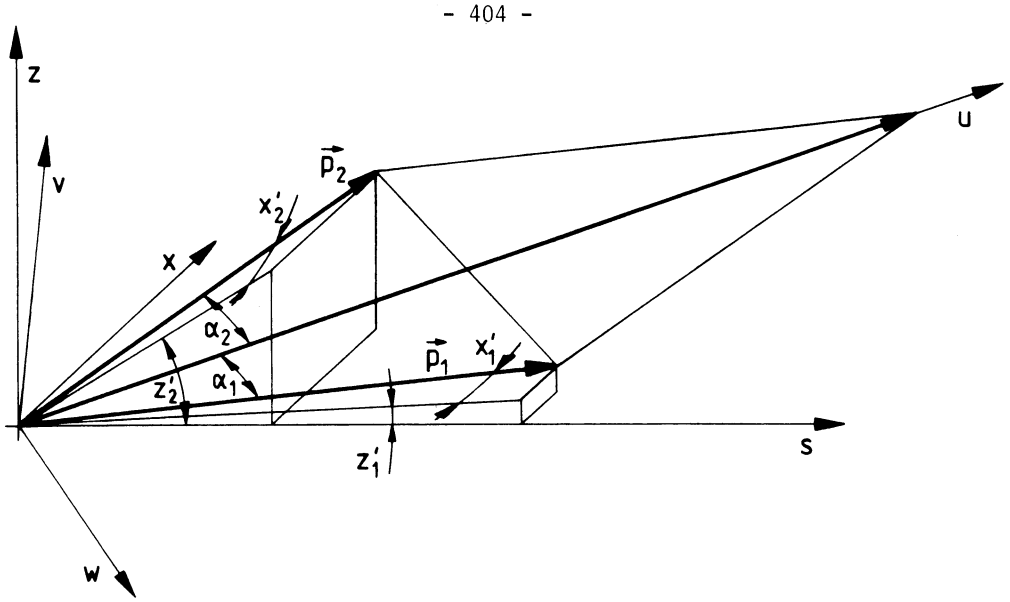


Fig. 1 Relationship of the (u,v,ω) coordinate system aligned on the centre-of-mass motion and the local coordinate system (s,x,z)

A Lorentz-transformation parallel to the u-axis can now be made and gives for the momenta in the centre of mass system

$$\vec{p}_{1,2} = \pm \{\bar{p}_{\bar{u}}, 0, \bar{p}_{\bar{w}}\}_{\bar{u},\bar{v},\bar{w}}$$

with

$$\bar{p}_{\bar{u}} = \pm p_{1,2}\, \gamma_{\ell t} \left(\cos \alpha_{1,2} - \frac{\vec{\beta}_{\ell t}}{\beta_{1,2}}\right)$$

$$\bar{p}_{\bar{w}} = p_{1,2} \sin \alpha_{1,2}$$

where the bars denote all quantities in the centre of mass system. $\beta_{\ell t}$ is given by the condition

$$\vec{\bar{p}}_1 + \vec{\bar{p}}_2 = 0$$

or

$$p_1 \left(\cos \alpha_1 - \frac{\beta_{\ell t}}{\beta_1}\right) + p_2 \left(\cos \alpha_2 - \frac{\beta_{\ell t}}{\beta_2}\right) = 0$$

and

$$\beta_{\ell t} = \frac{\gamma_1 \beta_1 \cos \alpha_1 + \gamma_2 \beta_2 \cos \alpha_2}{\gamma_1 + \gamma_2} .$$

b) Change of the momenta

The changes of the momenta of the two colliding particles have, for symmetry reasons, in the centre of mass system the same absolute values but opposite signs. They can be described by the polar angle ψ and the azimuthal angle ϕ (see Fig. 2).

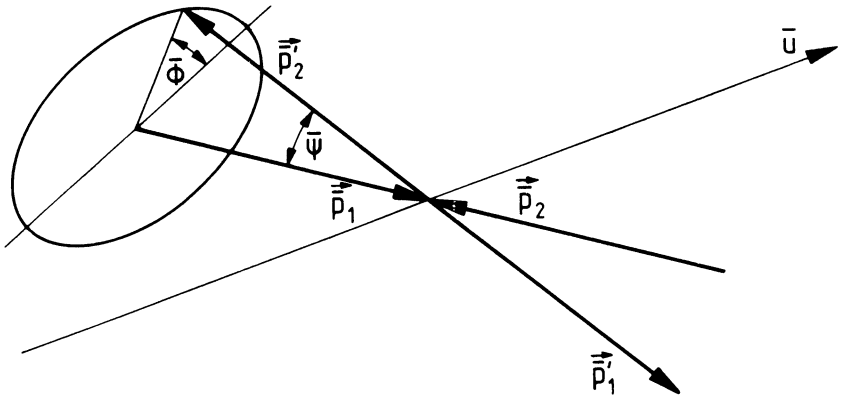


Fig. 2 Change of momenta in a two-particle collision in the centre-of-mass coordinate system

After the collision the momenta can be written in the form

$$\vec{\bar{p}}'_{1,2} = \pm \{\bar{p}_{\bar{w}} \sin\bar{\psi}\cos\bar{\phi} + \bar{p}_{\bar{u}} \cos\bar{\psi}\,,\ \bar{p}\sin\bar{\psi}\sin\bar{\phi}\,,\ \bar{p}_{\bar{w}}\cos\bar{\psi}$$

$$- \bar{p}_{\bar{u}} \sin\bar{\psi}\cos\bar{\phi}\}_{\bar{u},\bar{v},\bar{w}} \,.$$

We now assume that the particle velocities are non-relativistic in the centre of mass system:

$$\bar{\beta}^2 = \frac{\beta^2}{4}\left[\frac{(p_1-p_2)^2}{p^2} + \gamma^2\,(\alpha_1+\alpha_2)^2\right] \ll 1$$

where β is the mean velocity of the particles divided by the velocity of light and p is the mean momentum. This condition is satisfied very well in the SPS with $\alpha_{1,2} \simeq 3\cdot10^{-5} - 3\cdot10^{-4}$ at $\gamma = 300$. In this case one gets

$$\gamma_{\ell t} = \gamma$$

and

$$\cos\alpha_{1,2} - \frac{\beta_{\ell t}}{\beta_{1,2}} = \pm\frac{p_1 - p_2}{2\gamma^2 p} \,.$$

With the abbreviations

$$\frac{p_1 - p_2}{\gamma p} = \xi \,, \quad x'_1 - x'_2 = \Theta \,, \quad z'_1 - z'_2 = \zeta$$

$$\alpha_1 + \alpha_2 = 2\alpha = \sqrt{(x'_1 - x'_2)^2 + (z'_1 - z'_2)^2} = \sqrt{\Theta^2 + \zeta^2}$$

one gets for the changes of the momenta in the storage ring system the expression

$$\delta\vec{p}_{1,2} = \vec{p}'_{1,2} - \vec{p}_{1,2}$$

$$= \pm \frac{p}{2} \{2\alpha\gamma \sin\bar{\psi} \cos\bar{\phi} + \gamma\xi(\cos\bar{\psi} - 1) \,,$$

$$[\zeta\sqrt{1 + \frac{\xi^2}{4\alpha^2}} \sin\bar{\phi} - \frac{\xi\Theta}{2\alpha} \cos\bar{\phi}] \sin\bar{\psi} + \Theta(\cos\bar{\psi} - 1) \,,$$

$$[\Theta\sqrt{1 + \frac{\xi^2}{4\alpha^2}} \sin\bar{\phi} - \frac{\xi\zeta}{2\alpha} \cos\bar{\phi}] \sin\bar{\psi} + \zeta(\cos\bar{\psi} - 1)\}_{s,x,z} \,.$$

d) Change of the oscillation amplitudes

The betatron amplitudes are given by the emittances. Neglecting β'_x one gets for the change of the horizontal emittance:

$$\beta_x \delta\varepsilon_x = 2\,\delta x_\beta + \delta^2 x_\beta + \beta_x^2 (2x'\delta x' + \delta^2 x')$$

δx_β is given, as in the case of quantum fluctuation, by the change of the absolute value of the momentum, and $\delta x'$ is given by the change of the direction of the momentum:

$$\beta_x \delta\varepsilon_x = -2x_\beta D \frac{\delta p}{p} + D^2 \frac{\delta^2 p}{p^2} + \beta_x^2 (2x' \frac{\delta p_x}{p} + \frac{\delta^2 p_x}{p^2}) \,.$$

Here we have neglected also the derivative of the dispersion D'. For the change of the vertical emittance one obtains

$$\delta\varepsilon_z = \beta_z (2z' \frac{\delta p_z}{p} + \frac{\delta^2 p_z}{p^2})$$

since the vertical dispersion is usually zero.

For the invariants of the linearized longitudinal motion one gets

$$H = \begin{cases} \left(\frac{\Delta p}{p}\right)^2 + \frac{1}{\Omega_s^2}\left(\frac{d}{dt}\frac{\Delta p}{p}\right)^2 & \text{for bunched beams} \\ \left(\frac{\Delta p}{p}\right)^2 & \text{for unbunched beams} \end{cases}$$

where $\Omega_s/2\pi$ is the synchrotron frequency. The change of H due to a scattering event is in both cases

$$\delta H = 2\,\frac{\Delta p}{p}\,\frac{\delta p}{p} + \frac{\delta^2 p}{p^2}\,.$$

e) Averaging over all scattering angles

The distribution of the scattering angles ψ and ϕ is given by the Rutherford cross-section:

$$d\bar{\sigma} = \left(\frac{r_p}{4\bar{\beta}^2 \sin^2\frac{\bar{\psi}}{2}}\right)^2 \sin\bar{\psi}\, d\bar{\psi}\, d\bar{\phi}$$

where r_p is the classical proton radius. Integration over all scattering angles gives

$$\int_{\bar{\psi}_m}^{\pi}\int_0^{2\pi} \delta\vec{p}_1\, d\bar{\sigma} = -\frac{\pi}{2}\frac{p r_p^2}{\bar{\beta}^4}\ln\frac{2\bar{\beta}^2\bar{d}}{r_p}\,\{\gamma\xi,\Theta,\zeta\}_{s,x,z}$$

$$\int_{\bar{\psi}_m}^{\pi}\int_0^{2\pi} \delta^2 p\, d\bar{\sigma} = \frac{\pi}{4}\frac{p^2 r_p^2}{\bar{\beta}^4}\ln\frac{2\bar{\beta}^2\bar{d}}{r_p}\left[\gamma^2(\Theta^2+\zeta^2) + (\xi^2+\zeta^2) + (\xi^2+\Theta^2)\right].$$

The minimum scattering angle ψ_m is determined by the impact parameter $\bar{d}$ by

$$\tan\frac{\bar{\psi}_m}{2} = \frac{r_p}{2\bar{\beta}^2\bar{d}}$$

with

$$\bar{d} \approx \frac{1}{2}\,\text{beam height}\,.$$

Furthermore,

$$2\bar{\beta}^2\bar{d} \gg r_p$$

is assumed for the calculation of the integrals.

The change of the three invariants is then given by

$$\int_{\bar{\psi}_m}^{\pi} \int_0^{2\pi} \begin{bmatrix} \delta H_1/\gamma^2 \\ \delta\varepsilon_{x_1}/\beta_x \\ \delta\varepsilon_{z_1}/\beta_z \end{bmatrix} d\bar{\sigma} =$$

$$\frac{\pi r_p^2}{4\bar{\beta}^4} \ln \frac{2\bar{d}\bar{\beta}^2}{r_p} \begin{bmatrix} -4\frac{\Delta p_1}{\gamma p}\xi + \theta^2 + \zeta^2 \\ -4\, x'_1\theta + \xi^2 + \zeta^2 - 4\frac{x_{\beta_1}}{\beta_x^2}\gamma\zeta + \frac{D^2\gamma^2}{\beta_x^2}(\theta^2+\zeta^2) \\ -4\, z'_1\zeta + \xi^2 + \theta^2 \end{bmatrix}$$

f) Averaging over all particles

The relative velocity between two colliding particles in the centre of mass system is $2c\bar{\beta}$. The probability for a scattering into $\bar{\psi}$ and $\bar{\phi}$ per unit time in the centre of mass system is $2c\bar{\beta}\bar{\rho}\, d\bar{\sigma}$ where $\bar{\rho} = \rho/\gamma$ is the particle density. The probability for scattering into ψ and ϕ per unit time in the storage ring system is

$$P_{scat} = \frac{2c\bar{\beta}\rho\, d\bar{\sigma}}{\gamma^2}$$

since $dt = \gamma d\bar{t}$. In order to calculate the mean change of the invariants of all particles we have to average with respect to the 12 variables of the two colliding particles:

$$s_1, s_2, x_{\beta_1}, x_{\beta_2}, z_{\beta_1}, z_{\beta_2}, \Delta p_1/p, \Delta p_2/p, x'_1, x'_2, z'_1, z'_2$$

where three of the twelve variables are dependent:

$$s_1 = s_2 = s$$

$$x_{\beta_1} + D\frac{\Delta p_1}{p} = x_{\beta_2} + D\frac{\Delta p_2}{p}$$

$$z_{\beta_1} = z_{\beta_2} = z_\beta$$

which follows from the assumption that the two colliding particles have the same position. It is assumed that all variables have Gaussian density distribution except for the longitudinal distribution w_s which can have a Gaussian distribution or a continuous distribution (unbunched beams). The total density function is then

$$P = w_s^2(s) w_{x_\beta}(x_{\beta_1}) w_{x_\beta}(x_{\beta_2}) w_{z_\beta}^2(z_\beta) w_p(\frac{\Delta p_1}{p}) w_p(\frac{\Delta p_2}{p}) w_{x'}(x'_1) w_{x'}(x'_2) w_{z'}(z'_1) w_{z'}(z'_2) \ .$$

Before integrating over the three variables we make the following substitutions

$$x_{\beta 1,2} = x_\beta \mp D\gamma\xi/2 \qquad \frac{\Delta p_{1,2}}{p} = \eta \pm \gamma\xi/2$$

$$x'_{1,2} = x' \pm \Theta/2 \qquad z'_{1,2} = z' \pm \zeta/2$$

and the differential becomes

$$dV = \gamma ds dx_\beta dz_\beta d\tau d\xi dx' d\Theta dz' d\zeta \ .$$

Now six of the nine integrals can be solved immediately and one obtains

$$\frac{d}{dt}\begin{bmatrix} \langle \delta H\rangle/\gamma^2 \\ \langle d\varepsilon_x\rangle/\beta_x \\ \langle \delta\varepsilon_2\rangle/\beta_z \end{bmatrix} = \int_V \frac{2c\bar{\beta}}{\gamma^2} \int_{\psi_m}^{\pi} \int_0^{2\pi} \begin{bmatrix} \delta H_1/\gamma^2 \\ \delta\varepsilon_{x_1}/\beta_x \\ \delta\varepsilon_{z_1}/\beta_z \end{bmatrix} d\bar{\sigma}\, dV =$$

$$2A \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{-\frac{\gamma^2\xi^2}{4}(\frac{1}{\sigma_\eta^2} + \frac{D^2}{\beta_1^2}) - \frac{\Theta^2}{4\sigma_{x'}^2} - \frac{\zeta^2}{4\sigma_x^2}\}$$

$$\begin{bmatrix} \Theta^2 + \zeta^2 - 2\xi^2 \\ \xi^2 + \zeta^2 - 2\Theta^2 + \frac{\gamma^2 D^2}{\beta_x^2}(\Theta^2 + \zeta^2 - 2\xi^2) \\ \xi^2 + \Theta^2 - 2\zeta^2 \end{bmatrix} \ln(2\frac{\bar{d}}{r_p}\bar{\beta}^2) \frac{d\xi d\Theta d\zeta}{(\xi^2 + \Theta^2 + \zeta^2)^{3/2}}$$

with

$$A = \frac{r_p^2 c N_b}{64\pi^2 \sigma_s \sigma_p \sigma_{x_\beta} \sigma_{z_\beta} \sigma_{x'} \sigma_{z'} \beta^3 \gamma^4} \qquad \text{(bunched)}$$

where N_b is the number of particles in the bunch. For an unbunched beam N_b/σ_s must be replaced by $2\sqrt{\pi}N/C$ where N is the number of particles in the beam and C is the circumference.

g) Invariants

Multiplying the equation for $\langle\delta H\rangle$ by $1-\gamma^2D^2/\beta^2_x$ and adding it to the equations for $\langle\delta\varepsilon_x\rangle$ and $\langle\delta\varepsilon_z\rangle$ gives

$$\frac{d}{dt}\left[\langle\delta H\rangle\left(\frac{1}{\gamma^2}-\frac{D^2}{\beta_z^2}\right)+\frac{\langle\delta\varepsilon_x\rangle}{\beta_x}+\frac{\langle\delta\varepsilon_z\rangle}{\beta_z}\right]=0\ .$$

In a weak focusing machine the momentum compaction factor α_M is given by

$$\alpha_M=\frac{D^2}{\beta_x^2}\ .$$

After integration one obtains

$$\langle H\rangle\left(\frac{1}{\gamma^2}-\alpha_M\right)+\frac{\langle\varepsilon_x\rangle}{\beta_x}+\frac{\langle\varepsilon_z\rangle}{\beta_z}=\text{const}\ .$$

This equation shows that the behaviour of the particles is different below transition energy ($\gamma^2 < 1/\alpha_M$) and above transition energy ($\gamma^2 > 1/\alpha_M$). Below transition energy the sum of the three positive invariants is limited, i.e. the three oscillation amplitudes are limited. The particles behave like the molecules of a gas in a closed box (focusing corresponds to the walls of the box). They can only exchange their oscillation energy, but the total oscillation energy or the temperature is limited. In this case an equilibrium distribution must exist where the intra beam scattering does not change the beam dimensions.

Above transition energy the coefficient of $\langle H\rangle$ is negative and the total oscillation energy can increase as long as it does not exceed other limitations. In this case an equilibrium distribution does not exist.

h) Rise times

The rise times for the mean oscillation amplitudes, which determine the bunch dimensions, are given by[1,2]

$$\frac{1}{\tau_p}=\frac{1}{2\sigma_p^2}\frac{d\sigma_p^2}{dt}=A\frac{\sigma_h^2}{\sigma_p^2}f(a,b,c)$$

$$\frac{1}{\tau_x}=\frac{1}{2\sigma_x^2}\frac{d\sigma_x^2}{dt}=A\left[f\left(\frac{1}{a},\frac{b}{a},\frac{c}{a}\right)+\frac{D^2\sigma_p^2}{\sigma_{x\beta}^2}f(a,b,c)\right]$$

$$\frac{1}{\tau}=\frac{1}{2\sigma_z^2}\frac{d\sigma_z^2}{dt}=A\,f\left(\frac{1}{b},\frac{a}{b},\frac{c}{b}\right)$$

with

$$a = \frac{\sigma_h}{\gamma\sigma_{x'}}, \quad b = \frac{\sigma_h}{\gamma\sigma_{z'}}, \quad c = \beta\sigma_h \sqrt{2\frac{d}{r_p}}$$

$$\frac{1}{\sigma_h^2} = \frac{1}{\sigma_p^2} + \frac{D^2}{\sigma_{x\beta}^2}$$

and

$$f(a,b,c) = 8\pi \int_0^1 [\ln (\frac{c^2}{2} (\frac{1}{\sqrt{p}} + \frac{1}{\sqrt{q}})) - 0.577 \ldots](1 - 3x^2) \frac{dx}{\sqrt{pq}}$$

$$p = a^2 + x^2 (1-a^2), \quad q = b^2 + x^2 (1-b^2) .$$

The function f(a,b,c) is plotted in Fig. 3. For

$$a = b = 1$$

one obtains

$$f(1,1,c) = 0$$

i.e. for

$$\sigma_p \sqrt{\frac{1}{\gamma^2} - \frac{D^2}{\beta^2}} = \sigma_{x'} = \sigma_{z'}$$

an equilibrium is reached, and the Gaussian function remains stable. This is possible only below transition energy, i.e. $\gamma < \beta/D$.

3. EXPERIMENTAL RESULTS

The last section shows some experimental results obtained in the two CERN storage rings, the Super Proton Synchrotron (SPS) and the Antiproton Accumulator (AA). Investigations in the SPS have shown that the decay of the luminosity consists of three parts[3]. The first part is the increase of the dimensions of the proton bunches due to the intra-beam scattering, the second is the loss of protons due to the intra-beam scattering and the third is the loss of antiprotons due to the beam-beam interaction. The total decay time of the luminosity is then given by:

$$\frac{1}{\tau_{lum}} = \frac{1}{\tau_{x'}^{+}} + \frac{1}{\tau_{life}^{+}} + \frac{1}{\tau_{life}^{-}} .$$

The times $\tau_{x'}^{+}$, τ_{life}^{+} and τ_{life}^{-} are, very roughly, in the order of 60 h so that the decay time of the luminosity is in the order of 20 h and mainly determined by the intra-beam scattering.

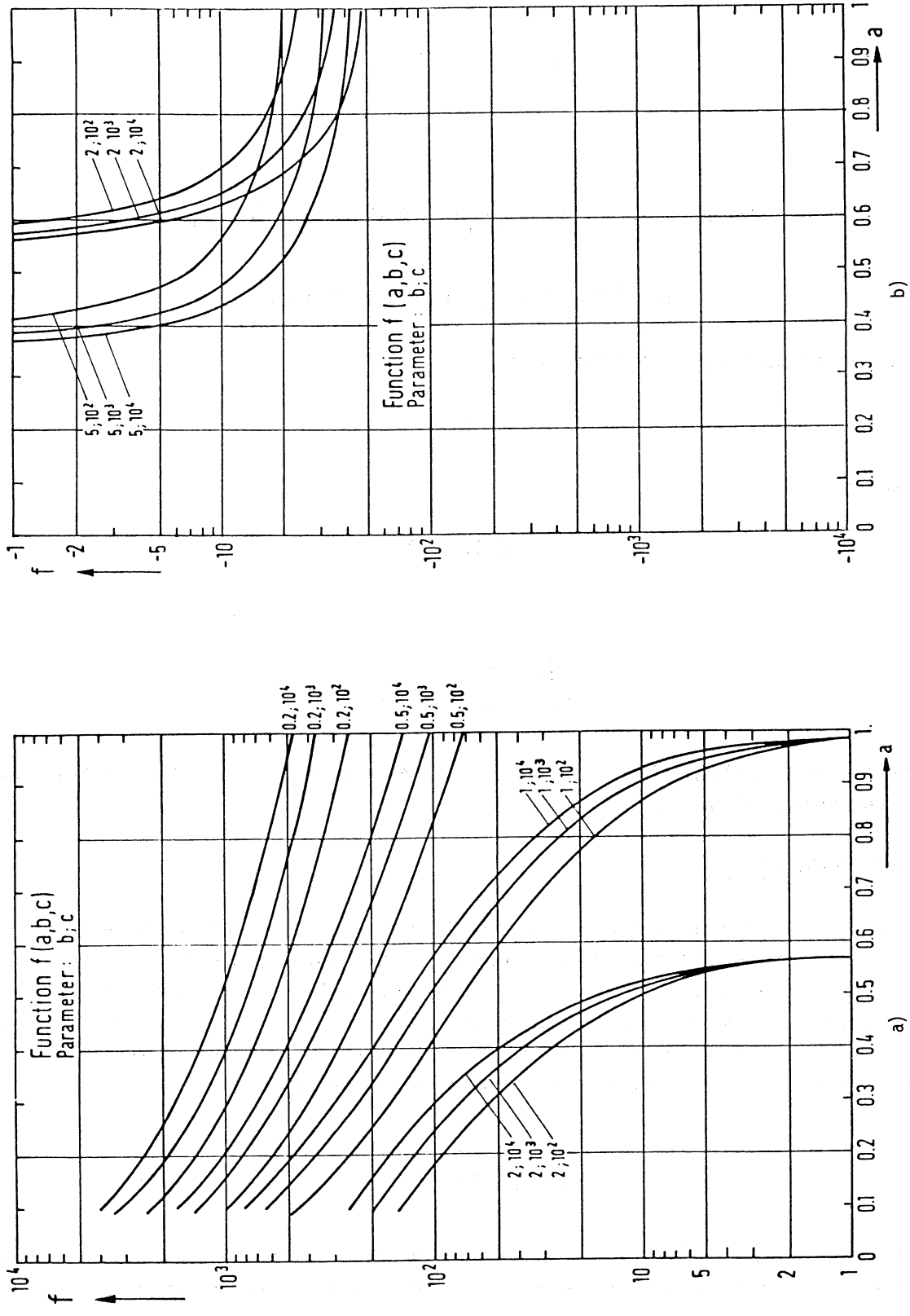
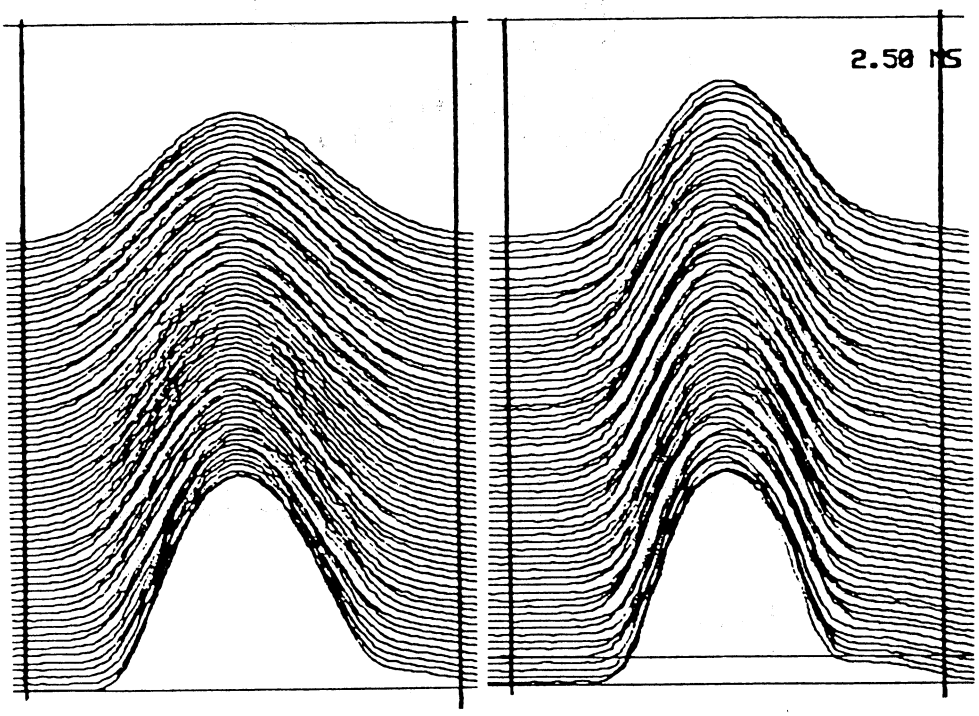


Fig. 3 Graphical representation of the function f(a,b,c)

Figure 4 shows the longitudinal distribution of proton and antiproton bunches measured at time intervals of 15 minutes[4]. It can be seen that in the case of antiprotons (b) the distribution remains almost constant whereas in the case of protons, which have a much larger density ($N^+ = 1.5 \times 10^{11}$, $N^- = 1.2 \times 10^{10}$), the bunch becomes longer and the density decreases. The rise times agree with the calculation[4].



a) b)

Fig. 4 Measurements made in the CERN SPS of the longitudinal distribution of (a) proton and (b) antiproton bunches

Figure 5 shows a direct comparison of the measured and calculated rise times for the horizontal emittance as a function of time[5]. The measured growth rate, which is the inverse of the rise time, is equal to the calculated growth rate plus a small constant. This constant is caused by the scattering of the protons by the residual gas[5].

In the AA, where the antiprotons are cooled, intra-beam scattering limits the maximum antiproton density which can be achieved by stochastic cooling[6,7]. In an experiment the cooling system was switched off after the antiprotons were cooled down and the growth times due to the intra-beam scattering were measured. Figure 6 shows the growth times for the momentum spread and the horizontal emittance[6].

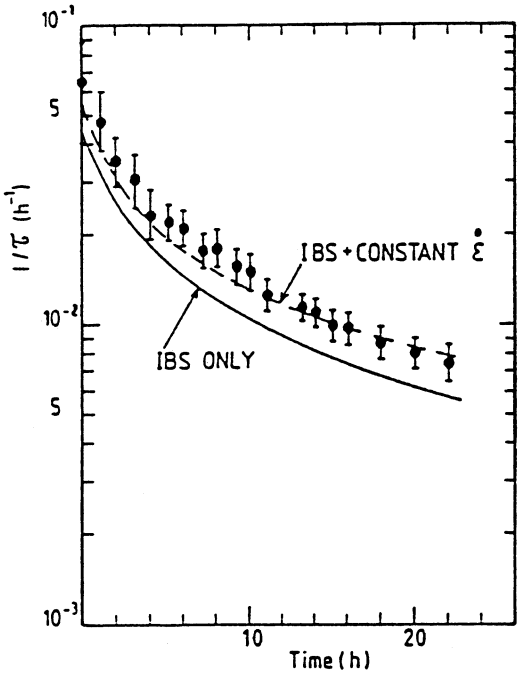


Fig. 5 Measured radial emittance growth rate in CERN SPS $\tau^{-1} = \dot{\varepsilon}/\varepsilon$ compared with the theoretical intra-beam scattering rate. The dotted line includes a correction for gas scattering.

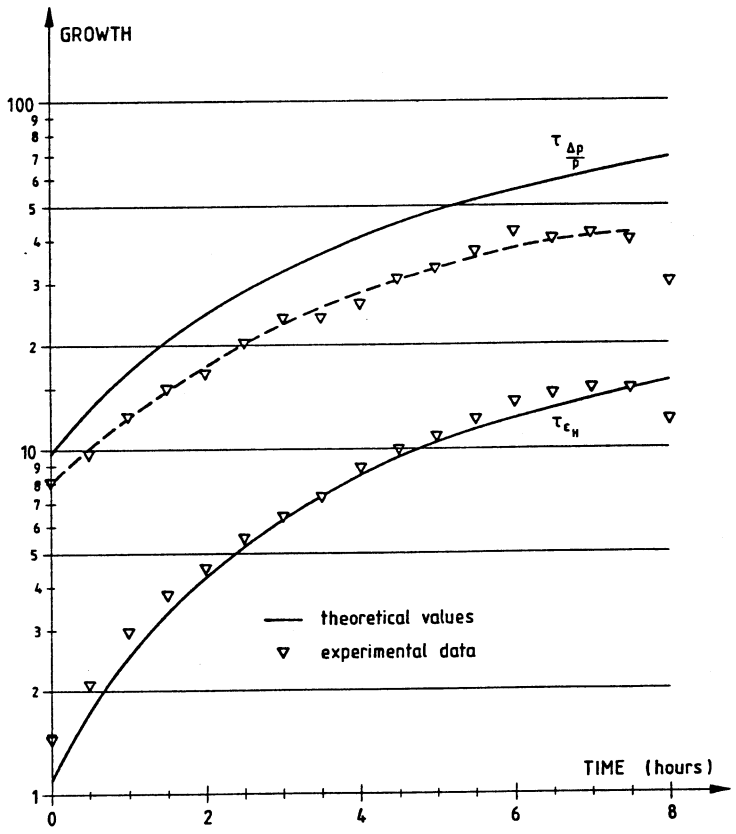


Fig. 6 Comparison of the measured and theoretical time constants for the growth of the momentum spread and horizontal emittance in the CERN AA with the cooling switched off

REFERENCES

1) A. Piwinski, Proc. 9th Int. Conf. on High Energy Accelerators, Stanford, 1974 (SLAC, Stanford, 1974) p. 405.

2) J.D. Bjorken and S.K. Mtingwa, Particle Accelerators, Vol. 13, p. 115 (1983).

3) L.R. Evans, CERN/SPS/83-30 (1983).

4) L.R. Evans, Private communication.

5) L.R. Evans and J. Gareyte, 1985 Particle Accelerator Conference, IEEE Trans. Nucl. Sci. NS-32, (1985).

6) M. Martini and S. van der Meer, PS/AA/ME/Note 75 (1984).

7) M. Martini, CERN PS/84-9 (AA) (1984).