EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN – PS DIVISION

CERN/PS 99-010 (DI)

PROTON-ION MEDICAL MACHINE STUDY (PIMMS)
PART I

Accelerator Complex Study Group*
supported by the Med-AUSTRON, Onkologie-2000 and the TERA Foundation
and hosted by CERN

Abstract

The Proton-Ion Medical Machine Study (PIMMS) group was formed following an agreement between the Med-AUSTRON (Austria) and the TERA Foundation (Italy) to combine their efforts in the design of a cancer therapy synchrotron. CERN agreed to host this study in its PS Division and a close collaboration was set up with GSI (Germany). The study group was later joined by Onkologie-2000 (Czech Republic). Effort was first focused on the theoretical understanding of slow extraction and the techniques required to produce a smooth beam spill for the conformal treatment of complex-shaped tumours with a sub-millimetre accuracy by active scanning with proton and carbon ion beams. Considerations for passive scanning were also included. The more general and theoretical aspects of the study are recorded in Part I and the more specific technical design considerations are presented in a second volume Part II. The PIMMS team started their work in January 1996 in the PS Division and continued for a period of three years.

*) Full-time members: L. Badano[1)], M. Benedikt[2)], P.J. Bryant[2)] (Study Leader), M. Crescenti[1)],
P. Holy[3)], A. Maier[2)+4)], M. Pullia[1)], S. Rossi[1)],
Part-time member: P. Knaus[1)+2)]

1) TERA Foundation, via Puccini. 11, I-28100 Novara.
2) CERN, CH 1211 Geneva-23.
3) Oncology-2000 Foundation, Na Morani 4, CZ-12808 Prague 2.
4) Med-AUSTRON, c/o RIZ, Prof. Dr. Stephan Korenstr.10, A-2700 Wr. Neustadt.

Geneva, Switzerland
2 March 1999

CONTENTS PART I

Co-ordinate system v

Frequently used symbols and abbreviations v

Nomenclature v

I-1 Introduction 1

1.1 Background 1
1.2 Basic design considerations 2
1.3 Performance parameters 5

I-2 Third-order Resonance 9

2.1 Sextupole magnets 9
2.2 Sextupoles in normalised co-ordinates 10
2.3 Sign conventions for sextupole magnets 11
2.4 Basic theory of the third-order resonance 12
2.5 Kobayashi Hamiltonian 14
2.6 A more general Hamiltonian 16
2.7 Closed-orbit distortion at the sextupole 18
2.8 Phase-space maps along the machine 18
2.9 General equations for the separatrices 19
2.10 Resonance excitation by sextupoles 20
2.11 Chromatic effects of sextupoles 21
2.12 Planning sextupole families
22

I-3 Resonant slow extraction 25

3.1 Steinbach Diagram 25
3.2 Overview of extraction methods 26
3.3 The mechanism of extraction 29
3.4 Separatrix geometry at the electrostatic septum 32
3.5 Phase-space representation of beam and resonance 34
3.6 Hardt Condition for superposition of separatrices 37
3.7 1st and 3rd quadrant operation with the Hardt Condition 40
3.8 Choosing the dispersion function 42
3.9 Positioning the electrostatic and magnetic septa 43
3.10 Transfer of off-momentum particles between septa 44
3.11 Minimisation of chromatic effects 46
3.12 Achromatic transfer between extraction septa 47
3.13 Zero chromaticity and the Hardt condition 51

I-4 Spill characteristics 53

4.1 Translation of the Hamiltonian 53
4.2 Transit times under static conditions 55
4.3 Transit times under dynamic conditions 57
4.4 Beam models for the spill 61
4.5 Tune ripple 76
4.6 Emittance of the extracted beam 82

I-5 Ripple **87**
5.1 Conventions 87
5.2 Over-modulation of a spill 88
5.3 Poisson statistics in a spill 90
5.4 Duty factor and frequency 91
5.5 Effect of tune ripple at low frequencies 92
5.6 Effect of tune ripple at medium frequencies 94
5.7 Effect of tune ripple at high and ultra-high frequencies 96
5.8 Summary of the ripple régimes 96
5.9 Defence against tune ripple
97
5.10 Ramping power converters 98
5.11 Damping by eddy currents 100
5.12 Spill specification for voxel scanning 102

I-6 'Feeding' the resonance **105**
6.1 Betatron core 105
6.2 Stochastic noise 107
6.3 Phase-displacement acceleration 114
6.4 Unstacking 121
6.5 ‘Front-end’ acceleration by empty rf bucket channelling 123

I-7 Lattices **135**
7.1 Regular-cell lattice 135
7.2 ‘Square’ ring lattice 136
7.3 Customised ring lattice 137
7.4 Basic cell types 138
7.5 Comments 139

I-8 Matching to gantries **143**
8.1 Matching methods 143
8.2 Tools for describing the matching methods 144
8.3 Symmetric-beam method 148
8.4 Round-beam method 150
8.5 Rotator method 152
8.6 Rotator design examples 153
8.7 Length scaling 155
8.8 Beam sizes in rotators 156
8.9 Spreading and scanning systems 159

I-9 Extraction lines **161**
9.1 Design concept 161
9.2 ‘Telescopes’ 162
9.3 Example optics with a Riesenrad gantry 163

I-10 Multiple scattering **179**
10.1 Characteristic multiple scattering angle 179
10.2 Scattering seen through the Twiss functions 182
10.3 Multiple scattering and tracking 188

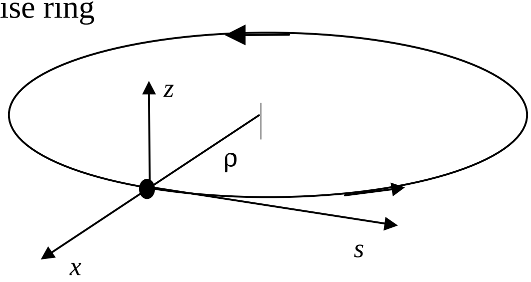
I-11 Raster scanning 195
11.1 Mini-voxel scanning 195
11.2 Raster scanning 205
11.3 Summary 214

Appendix I-A Poisson statistics 217

Index 221

CO-ORDINATE SYSTEM

Right-handed, curvilinear co-ordinate system (x, s, z) for the beam:
x is directed radially outwards in an anticlockwise ring
s is the direction along the beam
z is the vertical co-ordinate
ρ is the local radius of curvature



y is used as a general transverse co-ordinate that can replace both x and z.

FREQUENTLY-USED SYMBOLS AND ABBREVIATIONS

Subscript $_0$	denotes a reference value evaluated, e.g. at the origin, on central orbit, or at rest.
< >	Average over a distribution.
$^\bullet$, $'$	differentiation wrt time, differentiation wrt to s or $\Delta p/p$.
Superscript ^{T}	denotes transpose of a matrix.
Δ, δ, d	macroscopic, microscopic and infinitesimal steps.

FWHH	full width at half height.
RMS	root mean square.
F, D	focusing and defocusing lenses.
ES, MS	electrostatic and magnetic septa.
i	imaginary number.
rf	radio frequency.

NOMENCLATURE

A	atomic mass in atomic mass units.
$A = X^2+X'^2$	normalised betatron amplitude.
A_0, A_r	dc and ripple amplitudes of the particle spill intensity [particle/s].
A_m, B_m	harmonic coefficients in a 2-dimensional magnetic field
A_b, A_{sb}	rf bucket area, stationary rf bucket area [eV rad].
$\boldsymbol{B}$, B_x, B_s, B_z	magnetic induction vector and components [T].
C	machine circumference.
C_x, C_x', S_x, S_x' etc.	principal trajectories (cosine and sine-like).
$\boldsymbol{D}$, D_x, D_z	dispersion vector and components [m].
$\boldsymbol{D}_n$, $D_{n,x}$, $D_{n,z}$	normalised dispersion vector and components.
D	diffusion constant.
e	electronic charge.
E_x, E_z	emittance (phase-space area) quoted with π apparent e.g. 10π mm mrad. Unless otherwise stated the emittance of a distribution will the RMS value.
$\boldsymbol{E}$, E_x, E_s, E_z	electric field vector and components [V/m].
E, E_0	total energy, rest energy of a particle [eV].

F	duty factor for spill quality.
f	frequency [s^{-1}].
G	space-charge form factor.
h	perpendicular distance (normalised) from the separatrix to the origin.
h	rf harmonic number.
H_b, H_{sb}	rf bucket height, stationary bucket height [eV].
$\boldsymbol{H}$	hamiltonian.
$\boldsymbol{I}$	identity matrix.
I	current [A].
j	current density [A/m^2].
K	improvement factor for speed of crossing into the resonance.
$K(s)$	general focusing constant (equivalent to spring constant for simple harmonic motion) [m^{-2}].
$k = (1/B\rho)(dB_z/dx)_0$	Normalised (by momentum) quadrupole gradient [m^{-2}].
$k' = (1/B\rho)(d^2B_z/dx^2)_0$	Normalised (by momentum) sextupole gradient [m^{-3}].
L	thickness of a scatterer [units consistent with L_r].
L_r	radiation length [units consistent with L].
ℓ_s	effective magnetic length of a sextupole [m].
m, m_0	particle mass and rest mass[eV] also used as an integer.
m_e	mass of electron [MeV].
$\boldsymbol{M}$, m_{11}, m_{12}, etc.	transfer matrix and elements of transfer matrix.
N	number of particles.
N_A	Avogadro’s constant.
n	integer.
p	momentum [eV/c].
p, q	co-ordinates in a rotating system (rotator).
Q	spill quality.
$Q_{x,z}$	betatron tunes.
$q = ze$	charge [C].
r_e	classical radius of the electron [m].
$\boldsymbol{R}$	rotation matrix.
R	average radius of machine [m] also used for range of a particle in an absorber [m].
$S = (1/2)\beta_x^{3/2}\ell_s k'$	normalised (by β) sextupole strength [$m^{-1/2}$].
T, t_D	time [s] and dwell time on a mini-voxel [s].
T, T, T_{spill}	kinetic energy, transit time [s], spill length [s].
u, u', v, v'	co-ordinates in a rotating frame (gantry).
V	voltage [V] and volume [m^3].
v_{scan}	scanning velocity of beam spot [ms^{-1}].
W	FWHH of beam spot distribution equal to size of a voxel [m].
$\boldsymbol{W} = \boldsymbol{y}^T\boldsymbol{\sigma}^{-1}\boldsymbol{y}$	motion invariant.
x, s, z	local curvilinear co-ordinate system for the beam [m].
X, X', Z, Z'	normalised co-ordinates.
Y	used to replace X'.
z, z_{inc}	net number of electronic charges on a particle, net number of charges on a particle incident on a scatterer.
Z	atomic number and impedance [Ω].

α	adiabaticity factor.
$\beta = v/c$	relativistic β.
$\gamma\ = m/m_0$	relativistic γ.
γ_t	γ at transition.
$\alpha_{x,z}$, ,$\beta_{x,z}$, $\gamma_{x,z} = (1+\alpha_{x,z}^2)/\beta_{x,z}$	Courant and Snyder functions(β [m], γ[m^{-1}].
$\varepsilon = 6\pi.\delta Q$	modified tune distance.
$\phi = \mathrm{d}N/\mathrm{d}t$	particle flux [s^{-1}].
ϕ, ϕ_s	rf phase, synchronous rf phase [rad].
Φ	magnetic flux.
$\Gamma = \sin\phi_s$	
$\eta = \gamma^{-2}-\gamma_t^{-2}$	phase slip factor.
$\kappa(s_0) = \frac{1}{24\sqrt{\pi C}}\int_{s_0}^{s_0+C}\beta_x^{3/2}\left[\frac{-1}{\lvert B_0\rho\rvert}\left(\frac{\mathrm{d}^2B_z}{\mathrm{d}x^2}\right)_0\right]\exp(3\mathrm{i}\mu_x)\,\mathrm{d}s$	resonance driving term.
λ, Λ	normalised trajectory co-ordinates measured in units of h.
$\lambda = \mathrm{d}N/\mathrm{d}s$	linear particle density in a spill, or entering a resonance.
$\mu_{x,z}$	betatron phase advance [rad].
θ_0	projected RMS (or characteristic) scattering angle [rad].
θ_s	scattering angle [rad].
ρ	radius of curvature [m] and linear particle density along the side of an unstable triangle and in distributions.
σ	root mean square value.
$\boldsymbol{\sigma}$	sigma matrix.
τ	time constant [s].
Ω_s	synchrotron frequency [s^{-1}].
ω	angular frequency [s^{-1}].
$\Psi = \mathrm{d}N/\mathrm{d}(\Delta p/p)$	particle density in momentum space.

* * *

I-1 INTRODUCTION

1.1 BACKGROUND

High-energy, ionising radiation has proved to be effective in the treatment of cancerous tumours by causing double-strand breaks in the cell DNA. In particular, hadrons (that is protons and light ions) have the advantageous property of penetrating the body easily and then depositing their energy at a depth determined by their initial energy. This is often referred to as the *Bragg-peak behaviour*, see Figure 1.1 and, for example, Ref. 1. The abrupt cut-off of the beam at a controllable depth and the easy penetration compares extremely favourably with conventional radiation techniques using electrons or X-rays that deliver the highest dose at the surface diminishing with depth. The Bragg-peak behaviour offers the possibility of a *conformal treatment* of deep-seated tumours with minimum disturbance to the surrounding tissue.

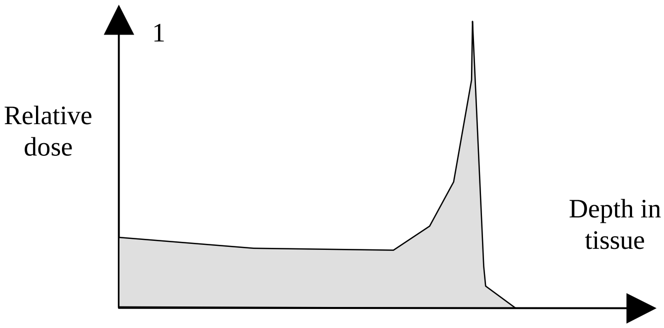


Figure 1.1 Bragg curve

Irradiation techniques fall into two broad categories:

- Passive spreading
- Active scanning.

The more usual technique is *passive spreading*, which uses a specially designed double scatterer to spread the beam by multiple Coulomb scattering uniformly over a large area that is sufficient to treat the whole tumour, or a large part of it. By treating the tumour in layers, defined by the depth of the Bragg peak, and applying collimators and shaped absorbers (bolus), a high degree of conformal treatment can be achieved [2]. This technique is well suited to large tumours and to ones that are difficult to immobilise. The second technique, which is of primary interest in this report, is *active scanning*, which uses a ‘pencil’ beam to ‘paint’ the tumour in three-dimensional space with sub-millimetre accuracy [3,4]. Longer times with smoother beam spills are required for this type of treatment to facilitate the on-line dosimetry and the accelerator has to produce a well-focused beam with a high spatial precision and an exact energy. However, the precision is only meaningful if a high positional stability can be assured for the tumour. Under the influence of tumour movements, active scanning can form ‘hot’ and ‘cold’ spots throughout the treatment volume, whereas passive spreading still irradiates uniformly and only the boundaries become uncertain.

1.2 BASIC DESIGN CONSIDERATIONS

The primary aim of this study is to design a machine that would allow the direct clinical comparison of protons and carbon ions for cancer therapy[*] [5] using high-precision active scanning. As a secondary aim, the machine should also be capable of delivering proton beams by passive scattering.

A synchrotron offers the flexibility needed for dual-species operation and the variable energy needed for active scanning. The higher rigidity of the ions determines the size and maximum power of the accelerator, while the protons for the passive spreading mode dominate the design of the injection system and the low-energy operation, due to their high space charge. The use of *slow resonant extraction* [6,7] extends the beam spill time sufficiently to perform on-line dosimetry at the patient and to switch the beam on and off according to the dose required. Either the half-integer or third-integer resonance can be used, but the current trend is towards the slower and more controllable spills from the third-integer resonance. Higher order resonances are not used because the angular separation of the separatrices becomes too small.

The principal design requirement is that of a smooth spill. This directly determines the performance of the machine. A poor spill quality makes it necessary to slow down the treatment by lowering the spill intensity and the scanning speed, so that the spill imperfections can be corrected by the scanning system on-line. In the Proton-Ion Medical Machine Study (PIMMS) design, many of the features are chosen specifically to ensure a good spill quality [8].

Betatron core

The extraction is activated by accelerating the beam into the resonance with a betatron core [9]. The extraction time is the time needed to accelerate the beam by its own momentum spread. Thus, the favoured configuration for the beam is one of small emittance and large momentum spread, so that extraction can be extended smoothly in time. This technique has the great advantage that it maintains all transverse optical parameters (and hence all power converters) in the machine constant. The only system that changes is the power converter for the betatron core. Since this is a single unit, special care can be taken with its design and that of its DAC (Digital to Analogue Converter). An 18-bit DAC, or a 16-bit DAC with smoothing of the DAC steps [10] is needed. The betatron core is a high inductance device and is intrinsically smooth in its operation.

Hardt condition

The Hardt condition [11] imposes a special configuration on the resonance that aligns the extraction separatrices for all momenta and thus minimises the beam losses on the extraction septum. An acceleration-driven extraction, such as that obtained with a betatron core, is required to implement this feature and certain optical conditions have to be carefully integrated into the lattice design [12].

---

[*] In 1946, R.R. Wilson proposed the use of protons for cancer therapy and predicted the future use of heavier ions, specifically mentioning carbon.

Intrinsic smoothing

An acceleration-driven extraction with the Hardt Condition has an additional advantage for the spill quality. When a batch of particles enters the resonance, they become trapped for many hundreds of turns in the machine before being released into the spill [13]. When this occurs, half of the particles are concentrated in a spike and the other half are spread out in a long tail [14]. The delay between entering the resonance and emerging in the spike depends upon the initial betatron amplitude of the particles. The Hardt condition configures the resonance and the beam in such a way that a mixture of all amplitudes enter the resonance at all times. This has the effect of spreading the spikes and reducing the sensitivity to power converter ripple.

Channelling rf bucket

Ripple in power converters affects the spill uniformity by causing a relative motion in tune between the beam and the resonance. One way of making the spill less sensitive to this motion is to cause the particles to enter the resonance with a velocity that is much in excess of the ripple velocity. Changing the betatron core more quickly is a too small an effect and is, in any case, counter productive, because it shortens the spill time. Instead, the technique is to create a region between the beam and the resonance where the particle velocity is higher, but the density is lower (so that the particle flux is constant). This can be done in different ways, for example by stochastic noise [15], but the method chosen for PIMMS is the channelling rf bucket [16,17]. The choice is partly justified by the fact that no new equipment is needed. The main rf cavity that is used for the acceleration of the beam can also be used for this task. The action of the cavity is based on a technique known as phase-displacement acceleration [18]. All particles in the beam are accelerated by the betatron core and, at the same time, the rf cavity is set so that it would decelerate particles by the same amount if they were trapped inside the bucket. The beam, however, is outside the rf bucket and the influence of the cavity is only felt as the revolution frequency of the particles approaches that of the cavity. Close to the cavity frequency, the particles are compressed into a narrower and narrower region of phase space and have to move rapidly around the bucket, which remains empty. This can be visualised by thinking of a river flowing past the piers of a bridge. The narrower the space allowed between the piers, the greater the river’s velocity.

Rotator

The choice of resonant slow extraction has the consequence that the extracted beam has unequal transverse emittances. This difference makes it impossible to rely on the conventional method used for matching gantries to cyclotrons where the beam emittances and the optics functions are all assumed to be equal and the dispersion to be zero at the interface between the fixed line and the rotating gantry. There is, in fact, a marked difference in the emittances (a factor of 50 would be a practical case with low coupling from the orthogonal plane) and a more advanced technique is required for matching to the rotating gantry. This technique uses a device known as a *rotator* [19]. This maps, one-to-one, the optics of the fixed machine directly to those of the gantry and makes this liaison completely independent of the gantry rotation angle. The mathematics of the rotator is simple and rigorous, but the practical design has to be approached with some care [20]. The rotator not only allows unequal emittances and optics functions, it also allows finite dispersion functions.

'Bar' of charge

The slow-extracted beam from a synchrotron is not only asymmetric in terms of emittance, it is also asymmetric in the shape of its emittance (or 'footprint') in phase space. In the vertical plane, the beam occupies the usual elliptical area, whereas, in the horizontal plane, it is a narrow bar. This bar must arrive at the patient with a known and controllable orientation, since this determines the spot size in that plane. The positive aspect of this behaviour is that it provides an independent handle on the control of the beam size in the horizontal plane. This opens the way to a new concept for controlling the beam size, not from the gantry, but from an optics module set upstream in the transfer line [21]. In the next section, the vertical beam size will be similarly treated upstream of the gantry, but by a different technique. Moving the controls of the horizontal and vertical beam sizes upstream is a new philosophy that makes it possible to control the spot sizes in all gantries and fixed beam lines in the complex with just two optics modules. At the same time, it reduces the number of optical constraints placed on the gantry design.

Vertical beam size control

The vertical beam-size control is also moved out of the gantry and to a point closer to the accelerator. The technique that allows this to be done is the use of one-to-one and telescopic modules for the extraction line optics [21]. Once the modules are all of the one-to-one or telescopic type, the vertical betatron amplitude function is simply handed from one module to the next with a constant magnification (usually unity) until it arrives at the patient. One-to-one modules are also very convenient structures in which to embed closed dispersion bends such as those needed when turning away from the main extraction line towards a gantry and treatment room.

Extracted beam intensity

The extracted beam intensity can be varied in two ways:

- Varying the rate of field change in the betatron core
- Varying the intensity of the injected beam.

In theory, the rate of field change in the betatron core can be modulated from approximately four times the nominal rate needed for active scanning (i.e. the rate foreseen for passive spreading) down to zero. However, very low rates are not advisable because the 'granularity' of the DAC will become apparent. For this reason, a maximum intensity variation of one to ten is foreseen during a spill using the betatron core. However, this ratio can be extended from spill to spill by varying the injected beam current. Ideally, the accelerator should always be filled in the same way in order to have reproducible operation, but this ideal is already lost because the low space-charge carbon-ion beam and the high space-charge proton beam must have different operational cycles (working line corrections etc.). Therefore, some additional cycles for intermediate proton beam intensities will not greatly add to the existing complexity of the operation. The delivered dose intensity can be further regulated at the patient by varying the scanning speed.

Active scanning

The aim is to provide a maximum speed of 10 m/s for the scanning beam spot. The boundary conditions for carbon ions and protons differ strongly and a successful scanning system for the carbon ions is not trivial [22]. The phase-space asymmetry of the slow-extracted beam makes one important difference with respect to scanning systems that deploy cyclotron beams. Since the beam distribution is near-rectangular in one plane and near-gaussian in the other, it is necessary to scan in the direction of the rectangular distribution and to keep the gaussian distribution at right-angles to the spot motion. In some cases, it is possible to exploit this rectangular shape to bestow a sharp edge on the scan. Perpendicular to the scanning direction, the overlap of the gaussian tails makes the alignment of the adjacent scans lines insensitive to small errors.

Alternative gantry

Finally, the high magnetic rigidity of the carbon ions led to an investigation of an alternative gantry design that has become known as the Riesenrad gantry [23,24]. In this variant, the heavy accelerator equipment is kept on the axis while the treatment room with the patient couch is positioned off-axis. The treatment room would be positioned before the entry of the patient and would be accessed from the rear by a ramp or lift to take into account changes in level. The reason for inverting the conventional patient-gantry geometry is one of engineering. For protons, it is feasible to build a gantry structure that can carry dipoles and quadrupoles with a total bending of close to three-quarters of the main synchrotron and still maintain sub-millimetre accuracy along the whole line while the gantry assumes different orientations, but for light ions, this is a more difficult task and has yet to be demonstrated. The point at which one system becomes more attractive than the other has not been established, but a first guess is that for a gantry of more than 100 tonnes it is more convenient to move the lighter treatment room than the heavier magnets. A major optics constraint for the Riesenrad gantry, when it was first proposed, was the closing of the dispersion bump, but this problem is now rendered void by the use of a *rotator*.

1.3 PERFORMANCE PARAMETERS

The general performance specifications for the machine are based on the following premises for the clinical requirements [25],

- A treatment will on average be 30 fractions with 2 Gray per fraction.
- An acceptable treatment time is 2 to 2.5 minutes per fraction.
- The maximum size of the treatment volume is assumed to be 2 liter when using active scanning and 7.5 liter when using passive scanning.
- An acceptable maximum depth is assumed to be ~27.5 cm
- An acceptable minimum depth is assumed to be ~3.5 cm.

The above guide lines lead to the more technical specifications listed in Table 1.1.

The theoretical aspects of this study and, in particular, the considerations concerned with how to generate a smooth beam spill are presented in this report and the technical design of a generic machine is presented in a second volume.

PIMMS performance parameters

	Active scanning (Pencil beam)	**Passive scanning (large area beam)**
Extraction energies for carbon ions	120-400 MeV/u	-
Extraction energies for protons*	60-250 MeV	60-250 MeV
Beam distributions	Spot is gaussian in direction perpendicular to scan and near-rectangular in scan direction. Scanning is parallel over a $20 \times 20\ \text{cm}^2$ rectangular area .	Beam is scattered up to an elliptical field of $20 \times 15\text{cm}^2$ with a ±2% uniformity. Scatterer has 40-50% efficiency.
Nominal treatments	60 spills of 1 s +1s to ramp up and down = 2 minute	120 spills of 0.25s +1s to ramp up and down = 2.5 minute
Nominal dose delivered	2 Gray in 2 liter	2 Gray in 7.5 liter
Number of protons in one spill	10^{10}	2×10^{10}
Number of carbon ions in one spill	4×10^{8}	-
Start of spill can be synchronised to breathing	Yes	Yes
Spot sizes variation at all energies (FWHH).	4-10 mm	-
Intensity levels	The spill rate within a spill can be adjusted by the rate of change of the betatron core. The maximum rate is for passive scanning, which will be unsuitable for active scanning. A minimum variation of 1:10 is expected within a spill for active scanning. The number of intermediate levels is more a function of the control system than a fundamental limit. Wider variations from spill to spill can be obtained by changing the beam intensity at injection	
Energy levels	The number of energy steps is limited only by the control system	

* The top extraction energy for protons is purely nominal. The accelerator can deliver protons up to 1.2 GeV.

Table 1.1 PIMMS performance parameters

References:

[1] G.F. Knoll, *Radiation detection and measurement,* (Wiley 1989, ISBN 0-471-81504-7,p33.

[2] E. Pedroni, *Beam delivery*, Proc. 1st Int. Symposium. On Hadrontherapy, Como, 1993, (Elsevier, 1994), p438-41.

[3] E. Pedroni, H. Blattmann, T. Böhringer, A. Coray, S. Lin, S. Scheib, U. Schneider, *Voxel scanning for proton therapy,* Proc. of the NIRS International Workshop on Heavy Charged Particle Therapy and Related Subjects, Chiba, Japan, (July 1991), 94-109.

[4] Th. Haberer, W. Becher, D. Schardt, G. Kraft, *Magnetic scanning system for heavy ion therapy*, Nucl. Instr. Methods in Physics Research, A330 (1993) p296-305,

[5] R.R. Wilson, *Radiobiological use of fast protons,* Radiobiology **47**, p487-491 (1946).

[6] H.G. Hereward, *The possibility of resonant extraction from the CPS*, AR/Int. GS/61-5 (1961).

[7] H.G. Hereward, Proc. Vth Int. Conf. on High Energy Accelerators, Dubna, (1963).

[8] L. Badano,, M. Benedikt, P. Bryant, M. Crescenti, P. Holy, P. Knaus, A. Maier, M. Pullia, S. Rossi, *Synchrotrons for hadron therapy,* (submitted to Nucl. Instru. & Methods,. December 1998)

[9] L. Badano, S. Rossi, *Characteristics of a betatron core for extraction in a proton-ion medical synchrotron*, CERN/PS 97-19 (DI).

[10] J. Bosser, *Rappel sur les systemes lineaires exchantillonnes*, CERN Internal Note SPS/ABM/Note /84-10.

[11] W. Hardt, *Ultraslow extraction out of LEAR (transverse aspects),* CERN Internal Note PS/DL/LEAR Note 81-6, (1981).

[12] M. Benedikt, *Optical design of a synchrotron with optimisation of the slow extraction for hadron therapy*, University of Technology Vienna, Ph. D. thesis (1997).

[13] M. Pullia, *Transit time for third order resonance extraction*, CERN/PS 96-36 (DI).

[14] M. Pullia, *Time profile of the slowly extracted beam*, CERN/PS 97-50 (DI).

[15] S. van der Meer, *Stochastic extraction, a low-ripple version of resonant extraction*, CERN/PS/AA 78-6, (1978).

[16] R. Cappi, C. Steinbach, *Low frequency duty factor improvement for the CERN PS slow extraction using rf phase displacement techniques*, Accel. Conf., Washington, IEEE Trans. Nucl. Sci., Vol. **NS-28**, No. 3, (1981).

[17] M. Crescenti, *RF empty bucket channelling with a betatron core to improve slow extraction in medical synchrotrons*, CERN/PS 97-68 (DI), (Jan. 1998).

[18] E. Ciapala, *Stacking and phase displacement acceleration*, CERN 85-19, (Nov. 1985), p195-225.

[19] L.C. Teng, Private communication, Laboratory notebook (Jan. 1970) and Internal Report LL-134 (Oct.1986).

[20] M. Benedikt, C. Carli, *Matching to gantries for medical synchrotrons*, Part. Accel. Conf. 1997, Vancouver, (1997) and M. Benedikt, C. Carli, *Optical design of a beam delivery system using a rotator*, CERN/PS 96-041 (1996).

[21] M. Benedikt, P. Bryant, M. Pullia, *A new concept for the control of a slow-extracted beam in a line with rotational optics*, (submitted to Nucl. Instru. & Methods,. December 1998).

[22] P. Holy, *Analysis of scanning techniques for hadron therapy*, CERN/PS 98-064 (DI) and P. Holy, *Ion gantry design and scanning techniques in hadron therapy*, Ph.D. thesis, (to be published).

[23] U. Amaldi, Private communication, The gantry design referred to here has been named the 'Riesenrad' and was first suggested during the preparation of the TERA Blue Book (*The TERA Project and Centre for Oncological Hadrontherapy*, ed. U. Amaldi and M. Silari, INFN, Frascati, 1995, II Ed.) and was later mentioned by E. Pedroni in the Grey Book (*Hadrontherapy in Oncology*, Como Proceedings, Excerpta Medica, International Congress Series, 1077, Elsevier, Amsterdam (1994), p434).

[24] M. Benedikt, P. Bryant, P. Holy, M. Pullia, *'Riesenrad' ion gantry for hadron therapy*, (submitted to Nucl. Instru. & Methods, December 1998).

[25] U. Amaldi, M. Silari (editors), *The TERA project and the centre for oncological hadrontherapy*, The TERA Foundation, (INFN, Frascati, 1995, II Ed.)

* * *

I-2 THIRD-ORDER RESONANCE

The third-order resonance can be used to extract particles from a synchrotron over a large number of turns. The slowly extracted beam is known as the *spill*. In a medical machine, the spill time is extended to about one second, (~10^6 turns), in order to facilitate the measurement and control of the radiation dose delivered to the patient. A simplified Hamiltonian due to Kobayashi [1,2] describes the important characteristics of the extraction process and gives a useful insight into the physics. This method describes the effect of a *sextupole* magnet as a perturbation to the linear machine. Once the basic physics is understood the analytic results can be used to guide a practical machine design based on numerical simulations.

2.1 SEXTUPOLE MAGNETS

In the current-free region of a magnet gap, the field can be derived from a scalar potential ϕ [3]. Assuming that the magnetic field has only transverse components then the scalar potential of a magnet with 2*m* poles is given by:

$$\phi = \underbrace{A_{\mathrm{m}}\,\mathrm{Re}(x+iz)^m}_{\text{Skew}} + \underbrace{B_{\mathrm{m}}\,\mathrm{Im}(x+iz)^m}_{\text{Normal}} \tag{2.1}$$

and the field components are obtained by differentiation according to

$$B_x = -\frac{\partial\phi}{\partial x} \quad \text{and} \quad B_z = -\frac{\partial\phi}{\partial z}. \tag{2.2}$$

The transverse fields in a *normal sextupole* magnet ($m = 3$) are given by the differentiation of the imaginary terms in (2.1)[*],

$$B_x(x,z) = -6B_3xz \quad \text{and} \quad B_z(x,z) = -3B_3(x^2 - z^2). \tag{2.3}$$

The relationship of the coefficient B_3 to the gradient is found by comparing (2.3) to the Taylor expansion of the magnetic field in the horizontal plane,

$$B_z(x, z=0) = B_0 + \frac{1}{1!}\left(\frac{\mathrm{d}B_z}{\mathrm{d}x}\right)_0 x + \frac{1}{2!}\left(\frac{\mathrm{d}^2B_z}{\mathrm{d}x^2}\right)_0 x^2 + \ldots \tag{2.4}$$

so that

$$B_3 = -\frac{1}{6}\left(\frac{d^2B_z}{dx^2}\right)_0 . \tag{2.5}$$

The horizontal and vertical fields in a sextupole can then be written as

$$B_x = \left(\frac{\mathrm{d}^2B_z}{\mathrm{d}x^2}\right)_0 xz \quad \text{and} \quad B_z = \frac{1}{2}\left(\frac{\mathrm{d}^2B_z}{\mathrm{d}x^2}\right)_0 (x^2 - z^2). \tag{2.6}$$

---

[*] * The *skew sextupole* fields are obtained by the differentiation of the real terms.

The effect of a sextupole on a particle trajectory can be described in a simple way by considering the magnet as a thin lens. For positively charged particles in an anticlockwise ring (see Co-ordinate System),

$$\Delta x' = \frac{B_z \ell_{\mathrm{S}}}{|B\rho|} = \frac{1}{2}\frac{\ell_{\mathrm{S}}}{|B\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}x^2}\right)_0 \left(x^2 - z^2\right) = \frac{1}{2}\ell_{\mathrm{S}} k' \left(x^2 - z^2\right) \tag{2.7}$$

$$\Delta z' = -\frac{\ell_{\mathrm{S}}}{|B\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}\,x^2}\right)_0 xz = -\ell_{\mathrm{S}} k' xz\,, \tag{2.8}$$

where k' is the *normalised sextupole gradient,*

$$k' = \frac{1}{|B\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}x^2}\right)_0 . \tag{2.9}$$

2.2 SEXTUPOLES IN NORMALISED COORDINATES

The effect of a thin-lens sextupole in *normalised co-ordinates* can be found by applying the transformations for normalised co-ordinates to equations (2.7) and (2.8). The transformation is simplified since in a thin lens $\Delta x = \Delta z = 0$, so that

$$x \Rightarrow \sqrt{\beta_x}\,X, \quad \Delta x' \Rightarrow \frac{1}{\sqrt{\beta_x}}\Delta X', \quad z \Rightarrow \sqrt{\beta_z}\,Z \quad \text{and} \quad \Delta z' \Rightarrow \frac{1}{\sqrt{\beta_z}}\Delta Z'.$$

Thus the effect of a thin-lens sextupole in normalised co-ordinates appears as,

$$\Delta X = 0 \text{ and } \Delta X' = \left[\frac{1}{2}\beta_{\mathrm{x}}^{\;3/2}\frac{\ell_{\mathrm{S}}}{|\mathrm{B}\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}x^2}\right)_0\right]\left(X^2 - \frac{\beta_z}{\beta_x}Z^2\right) = S\left(X^2 - \frac{\beta_z}{\beta_x}Z^2\right) \tag{2.10}$$

$$\Delta Z = 0 \text{ and } \Delta Z' = -2\left[\frac{1}{2}\beta_{\mathrm{x}}^{\;3/2}\frac{\ell_{\mathrm{S}}}{|\mathrm{B}\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}x^2}\right)_0\right]\frac{\beta_z}{\beta_x}XZ = -2S\frac{\beta_z}{\beta_x}XZ\,, \tag{2.11}$$

where S is the *normalised sextupole strength*

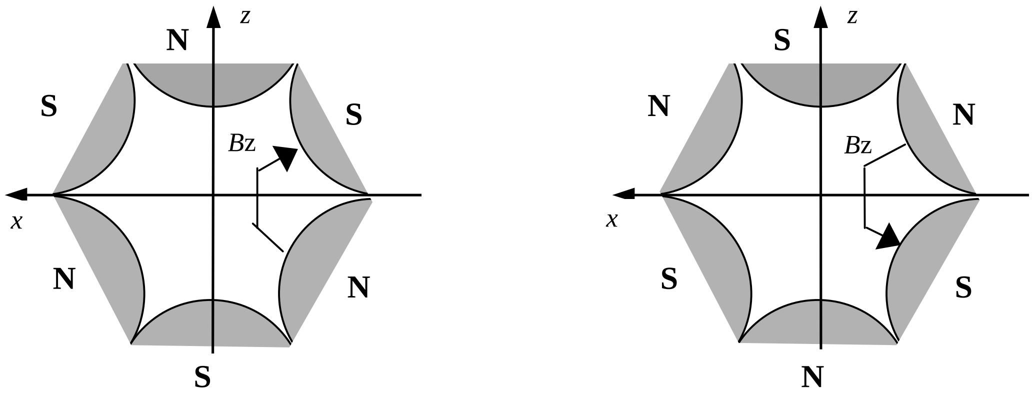
$$S = \frac{1}{2}\beta_x^{\;3/2}\frac{\ell_{\mathrm{S}}}{|\mathrm{B}\rho|}\left(\frac{\mathrm{d}^2 B_z}{\mathrm{d}x^2}\right)_0 = \frac{1}{2}\beta_x^{\;3/2}\ell_{\mathrm{S}} k'. \tag{2.12}$$

Unless $Z = 0$, a sextupole couples the horizontal and vertical motions. The strength of the coupling is proportional to the ratio of the vertical and horizontal betatron amplitude functions (β_z/β_x) at the sextupole. For a horizontal extraction, Z is generally much smaller than X and, provided the vertical tune does not satisfy a resonance condition, the influence of the vertical motion can be neglected to first order. For this reason, only the horizontal motion is considered in the resonance analysis in Section 2.4 and the equations (2.10) and (2.11) are replaced by,

Simplified form: $\quad \Delta X = \Delta Z = \Delta Z' = 0 \quad \text{and} \quad \Delta X' = SX^2 \qquad (2.13)$

2.3 SIGN CONVENTIONS FOR SEXTUPOLE MAGNETS

The sign conventions and nomenclature that have been used in the earlier sections are summarised in Figure 2.1.



Normal sextupole
(no pole on the median plane)
Positive field configuration

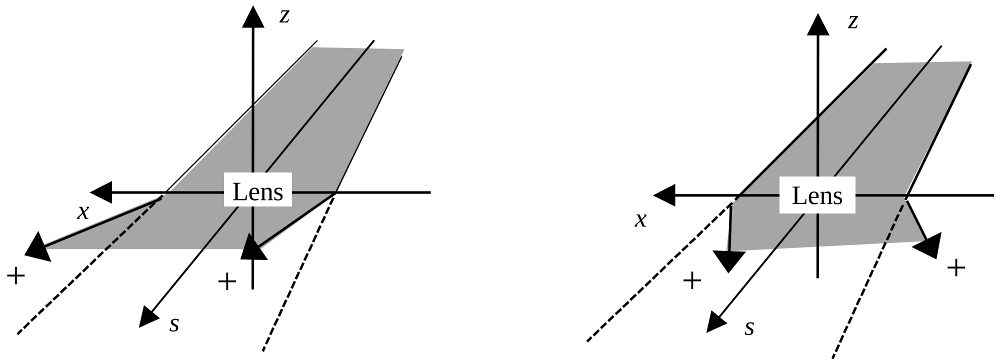
B_z positive on median plane
Sextupole gradient (d^2B_z/dx^2) positive
Normalised gradient, k' positive
Normalised strength, S positive

Normal sextupole
(no pole on the median plane)
Negative field configuration

B_z negative on median plane
Sextupole gradient (d^2B_z/dx^2) negative
Normalised gradient, k' negative
Normalised strength, S negative

Figure 2.1 Sign definitions for sextupole magnets

In general, lattice programs suppress the absolute signs of the fields and particles and replace them by normalised quantities that are defined according to the geometry of the beam. Thus, when a lens deflects the beam in the outer part of the aperture ($x > 0$) towards the axis, the lens is defined as focusing (F-type) and in the MAD program [4], for example, the normalised gradient k' is defined as positive (see Figure 2.2 for sextupoles). For positive ions, this unfortunately leads to a reversal of sign with respect to the above when using MAD (i.e. $k' = - k'_{\mathrm{MAD}}$).



Sextupole D-type
Normal type (no pole on median plane)

Positively charged ions
Trajectories are deflected outwards
B_z positive on median plane
$k' = - k'_{\mathrm{MAD}}$

Sextupole F-type
Normal type (no pole on median plane)

Positively charged ions
Trajectories are deflected inwards
B_z negative on median plane
$k' = - k'_{\mathrm{MAD}}$

Figure 2.2 Conventions for lattice programs

2.4 BASIC THEORY OF THE THIRD-ORDER RESONANCE

The general transfer matrix $\boldsymbol{M}_\mathrm{n}$ for normalised co-ordinates, describing *n* turns in the machine is given by:

$$\boldsymbol{M}_\mathrm{n} = \begin{pmatrix} \cos 2\pi(nQ_\mathrm{x}) & \sin 2\pi(nQ_\mathrm{x}) \\ -\sin 2\pi(nQ_\mathrm{x}) & \cos 2\pi(nQ_\mathrm{x}) \end{pmatrix}. \tag{2-14}$$

Consider a particle with a horizontal betatron tune close to a third-integer, i.e. $Q_\mathrm{x} = m \pm 1/3 + \delta Q$,, where m is integer and $|\delta Q| << 1/3$). The tune increment δQ is defined as the *tune distance* of the particle from the resonance,

$$\delta Q = Q_\mathrm{particle} - Q_\mathrm{resonance}\,. \tag{2-15}$$

The explicit transfer matrices for *n* turns in the unperturbed machine can then be written as:

$$\boldsymbol{M}_1 = \begin{pmatrix} \cos[2n\pi(m \pm 1/3 + \delta Q)] & \sin[2n\pi(m \pm 1/3 + \delta Q)] \\ -\sin[2n\pi(m \pm 1/3 + \delta Q)] & \cos[2n\pi(m \pm 1/3 + \delta Q)] \end{pmatrix}. \tag{2-16}$$

Thus, the co-ordinates of the particle after one, two and three turns become

1st turn neglecting δ*Q*
$$\begin{pmatrix} X \\ X' \end{pmatrix}_1 \cong \begin{pmatrix} -1/2 & \pm\sqrt{3}/2 \\ \mp\sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} X \\ X' \end{pmatrix}_0 \tag{2-17}$$

2nd turn neglecting δ*Q*
$$\begin{pmatrix} X \\ X' \end{pmatrix}_2 \cong \begin{pmatrix} -1/2 & \mp\sqrt{3}/2 \\ \pm\sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} X \\ X' \end{pmatrix}_0 . \tag{2-18}$$

3rd turn with δ*Q*
$$\begin{pmatrix} X \\ X' \end{pmatrix}_3 \cong \begin{pmatrix} 1 & \varepsilon \\ -\varepsilon & 1 \end{pmatrix} \begin{pmatrix} X \\ X' \end{pmatrix}_0 , \tag{2-19}$$

where ε replaces $6\pi\delta Q$ for brevity. The small quantity ε will be called the *modified tune distance*. It can be seen that a particle with exactly the resonant tune (i.e. $\varepsilon = 0$) will return to its initial position every three turns. The effect of the sextupole during three turns in the machine is now calculated as a perturbation by the linear addition of:

(A) The effect of 3 turns with a sextupole placed after the 3rd turn,
$\boldsymbol{M}_3$ + *Sextupole*

(B) The effect of 3 turns with a sextupole placed after the 2nd turn,
$\boldsymbol{M}_2$ + *Sextupole* + $\boldsymbol{M}_1$

(C) The effect of 3 turns with a sextupole placed after the 1st turn
$\boldsymbol{M}_1$ + *Sextupole* + $\boldsymbol{M}_2$

where the sextupole is represented by the simplified, thin-lens expression (2.13).

$$\Delta X = \Delta Z = \Delta Z' = 0 \quad \text{and} \quad \Delta X' = SX^2 . \tag{2.13}$$

The full evaluation of the three terms (A), (B) and (C) gives:

(A) 3 turns + sextupole

$$\begin{aligned} X_3 &= \left(X_0 + \varepsilon X_0'\right) \\ X_3' &= -\varepsilon X_0 + X_0' + S\left(X_0 + \varepsilon X'\right)^2. \end{aligned} \tag{2.20}$$

(B) 2 turns + sextupole+1 turn

$$\begin{aligned} X_3 &= -\frac{1}{2}\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X'\right) \pm \frac{\sqrt{3}}{2}\left[\pm\frac{\sqrt{3}}{2}X_0 - \frac{1}{2}X_0' + S\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X_0'\right)^2\right] \\ X_3' &= \mp\frac{\sqrt{3}}{2}\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X'\right) - \frac{1}{2}\left[\pm\frac{\sqrt{3}}{2}X_0 - \frac{1}{2}X_0' + S\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X_0'\right)^2\right]. \end{aligned} \tag{2.21}$$

(C) 1 turn + sextupole+2 turns

$$\begin{aligned} X_3 &= -\frac{1}{2}\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X'\right) \mp \frac{\sqrt{3}}{2}\left[\mp\frac{\sqrt{3}}{2}X_0 - \frac{1}{2}X_0' + S\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X_0'\right)^2\right] \\ X_3' &= \pm\frac{\sqrt{3}}{2}\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X'\right) - \frac{1}{2}\left[\mp\frac{\sqrt{3}}{2}X_0 - \frac{1}{2}X_0' + S\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X_0'\right)^2\right]. \end{aligned} \tag{2.22}$$

After the addition of the three terms (A), (B) and (C), only first-order correction terms in ε are retained to give

$$\begin{aligned} \Delta X_3 &= \varepsilon X_0' \pm \frac{\sqrt{3}}{2}S\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X_0'\right)^2 \mp \frac{\sqrt{3}}{2}S\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X_0'\right)^2 \\ \Delta X_3' &= -\varepsilon X_0 + SX_0^2 - \frac{1}{2}S\left(-\frac{1}{2}X_0 \mp \frac{\sqrt{3}}{2}X_0'\right)^2 - \frac{1}{2}S\left(-\frac{1}{2}X_0 \pm \frac{\sqrt{3}}{2}X_0'\right)^2 . \end{aligned} \tag{2.23}$$

The cancellation of signs shows that there is no fundamental difference between the 1/3rd and 2/3rd resonances.

$$\begin{aligned} \Delta X_3 &= \varepsilon X_0' + \frac{\sqrt{3}}{2}S\left(-\frac{1}{2}X_0 - \frac{\sqrt{3}}{2}X_0'\right)^2 - \frac{\sqrt{3}}{2}S\left(-\frac{1}{2}X_0 + \frac{\sqrt{3}}{2}X_0'\right)^2 \\ \Delta X_3' &= -\varepsilon X_0 + SX_0^2 - \frac{1}{2}S\left(-\frac{1}{2}X_0 - \frac{\sqrt{3}}{2}X_0'\right)^2 + \frac{1}{2}S\left(-\frac{1}{2}X_0 + \frac{\sqrt{3}}{2}X_0'\right)^2 . \end{aligned} \tag{2.24}$$

The final expressions for the change of position and divergence of the particle over three revolutions, known as the *spiral step* and *spiral kick* are obtained as:

Spiral step and kick:

$$\begin{aligned} \Delta X_3 &= \varepsilon X_0' + \frac{3}{2}SX_0X_0' \\ \Delta X_3' &= -\varepsilon X_0 + \frac{3}{4}S\left(X_0^{\,2} - X_0'^{\,2}\right) \end{aligned} \tag{2.25}$$

2.5 KOBAYASHI HAMILTONIAN

The time needed for three revolutions in the machine is short compared to the spill time and can be safely used as the basic time unit. The elementary changes occurring in this time are also the smallest that need to be resolved to understand the physics of the extraction. Thus the subscripts are no longer needed and (2.25) can be treated as a continuous function that is derived from a Hamiltonian ***H***, such that

$$\begin{aligned}\Delta X_3 &\Rightarrow \left(\frac{\Delta X}{\Delta t}\right)_{\Delta t=1(3\,\text{turn})} = \frac{\partial \boldsymbol{H}}{\partial X'} = \varepsilon X' + \frac{3}{2}SXX' \\ \Delta X'_3 &\Rightarrow \left(\frac{\Delta X'}{\Delta t}\right)_{\Delta t=1(3\,\text{turn})} = -\frac{\partial \boldsymbol{H}}{\partial X} = -\varepsilon X + \frac{3}{4}S\left(X^2 - X'^2\right)\end{aligned}. \tag{2.26}$$

The Kobayashi Hamiltonian is found by integrating the above partial differentials:

$$\boldsymbol{H} = \frac{\varepsilon}{2}\left(X^2 + X'^2\right) + \frac{S}{4}\left(3XX'^2 - X^3\right). \tag{2.27}$$

It should be noted that in this formulation time is dimensionless.

The Hamiltonian is time independent and a constant of the motion. Contours of constant ***H*** show the particle trajectories in normalised phase space at the sextupole. This presentation of the motion is known as a *phase-space map*. The first term in (2.27) describes the unperturbed particle motion in the linear, machine (i.e. $S = 0$). These trajectories are circles of radius $\sqrt{(2\boldsymbol{H}/\varepsilon)}$ in normalised phase space. The second term contains the perturbation that distorts the circular phase-space trajectories into a triangular form as illustrated in Figure 2.3. At a certain level of excitation, the triangle 'breaks' into open phase-space trajectories. A change in sign of either the modified tune distance ε or the normalised sextupole strength S is equivalent to a rotation of the phase-space trajectories by 180°.

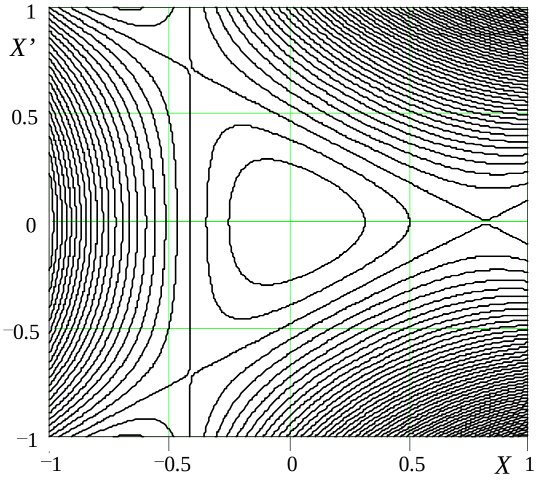
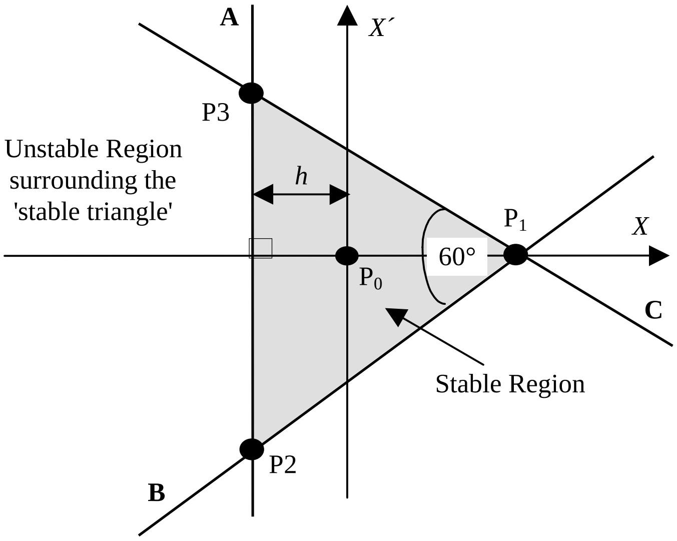


Figure 2.3 Phase-space map calculated from the Kobayashi Hamiltonian

All the properties of the system can be derived from the Hamiltonian. In particular, when ***H*** has the value $[(2\varepsilon/3)^3/S^2]$, it factorises into three straight lines

$$\left(\frac{S}{4}X+\frac{\varepsilon}{6}\right)\left(\sqrt{3}X'+X-\frac{4\varepsilon}{3S}\right)\left(\sqrt{3}X'-X+\frac{4\varepsilon}{3S}\right)=0\,, \qquad (2.28)$$

called the *separatrices,* that define the boundaries between the *stable triangle* and unstable regions in phase space. This situation is exactly analogous to that of an rf bucket in longitudinal phase space, where the separatrices define the well-known fish-shaped stability region. The phase space area of the stable region is the *acceptance* of the system at the given momentum. It could also be regarded as the *dynamic aperture*. The size of the stable region is determined by the ratio $|\varepsilon/S|$. For a particle that has exactly the resonance tune the stable region shrinks to zero. Figure 2.4 shows the geometry of the separatrices at the sextupole and the four stable fixed points, P_0 to P_3



$$P_1=\left(\frac{4}{3}\frac{\varepsilon}{S},0\right)$$

$$P_2=\left(-\frac{2}{3}\frac{\varepsilon}{S},-\frac{2\varepsilon}{\sqrt{3}S}\right)$$

$$P_3=\left(-\frac{2}{3}\frac{\varepsilon}{S},\frac{2\varepsilon}{\sqrt{3}S}\right)$$

$$h=\frac{2}{3}\frac{\varepsilon}{S}=\frac{4\pi}{S}\delta Q$$

Figure 2.4 Geometry of the separatrices and stable triangle at the sextupole

The geometry of the stable triangle is conveniently described by introducing the distance *h* between the upright separatrix **A** and the *X′*-axis. A change in the sign of *h* is equivalent to a 180° rotation of the stable triangle around the origin.

$$h=\frac{2}{3}\frac{\varepsilon}{S}=\frac{4\pi}{S}\delta Q\,. \qquad (2.29)$$

The area (i.e. the acceptance) of the stable triangle can also be expressed in terms of *h*:

$$\text{Area (acceptance) of stable triangle} = 3\sqrt{3}h^2=\frac{48\sqrt{3}\pi}{S^2}(\delta Q)^2\,\pi\,. \qquad (2.30)$$

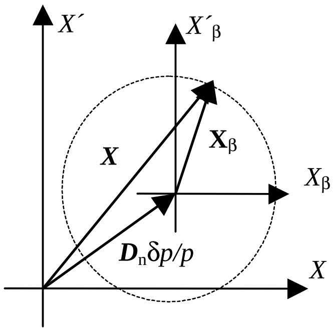
Providing the sextupole is powered in an adiabatic way*, the emittance will be conserved and (2.30) gives the fraction of the original beam emittance that remains stable.

2.6 A MORE GENERAL HAMILTONIAN

For the purposes of the PIMMS design, it is useful to extend the basic theory to off-momentum particles and to machines with closed-orbit distortions [5]. To first order in the momentum deviation $\delta p/p$, the equilibrium orbit in normalised co-ordinates is given by:

$$X_{\mathrm{EQ.O}} = D_n \frac{\delta p}{p}, X'_{\mathrm{EQ.O}} = D'_n \frac{\delta p}{p}, \qquad (2.31)$$

where $(D_{\mathrm{NA}}, D'_{\mathrm{n}})$ is the normalised dispersion function. A second co-ordinate system (X_β, X'_β) is introduced with its origin on the off-momentum equilibrium orbit. The particle co-ordinates are then split into a constant term, given by the dispersion function, and a betatron term as shown in Figure 2.5.



$$X = X_\beta + D_{\mathrm{n}} \frac{\delta p}{p} \Leftrightarrow X_\beta = X - D_{\mathrm{n}} \frac{\delta p}{p} \qquad (2.32)$$

$$X' = X'_\beta + D'_{\mathrm{n}} \frac{\delta p}{p} \Leftrightarrow X'_\beta = X' - D'_{\mathrm{n}} \frac{\delta p}{p} \qquad (2.33)$$

Figure 2.5 Co-ordinates of the betatron motion

The general Hamiltonian is derived following the same steps as for the Kobayashi Hamiltonian except that the betatron motion and the total motion now have to be separated. The analysis starts with the betatron motion for n turns in the unperturbed machine,

$$\begin{pmatrix} X_\beta \\ X'_\beta \end{pmatrix}_{\mathrm{k}} = \boldsymbol{M}_{\mathrm{n}} \begin{pmatrix} X_\beta \\ X'_\beta \end{pmatrix}_{0} . \qquad (2.34)$$

The main difference appears when calculating the kick of the sextupole that depends on the total motion of the ion:

Dispersion region: $\Delta X_\beta = 0, \quad \Delta X'_\beta = SX^2 = S\left(X_\beta + D_n \frac{\delta p}{p} \right)^2$ (2.35)

---

* Strictly, the emittance is conserved under all conditions, but if the excitation rate is too great the phase-space ellipse becomes distorted. The distorted ellipse has the same area, but later it filaments to cover a larger phase-space area.

This yields a more general Hamiltonian in the betatron co-ordinates (X_β, X'_β):

$$\boldsymbol{H} = \frac{\varepsilon}{2}\left(X_\beta^2 + X_\beta'^2\right) + \frac{S}{4}\left(3X_\beta X_\beta'^2 - X_\beta^3\right) - \frac{3}{2}S\left(D_\mathrm{n}\frac{\delta p}{p}\right)\left(X_\beta^2 + X_\beta'^2\right). \qquad (2.36)$$

Re-ordering the terms in (2.36) shows that the dispersion-dependent term affects only the 'circular' trajectories and this can be considered as a change in the tune distance, ε.

General form: $\boldsymbol{H} = \frac{1}{2}\left[\varepsilon - 3S\left(D_\mathrm{n}\frac{\delta p}{p}\right)\right]\left(X_\beta^2 + X_\beta'^2\right) + \frac{S}{4}\left(3X_\beta X_\beta'^2 - X_\beta^3\right).$ (2.37)

Let $\delta\tilde{Q}_\mathrm{x}$ be the change in tune implied by the dispersion-dependent term, so that

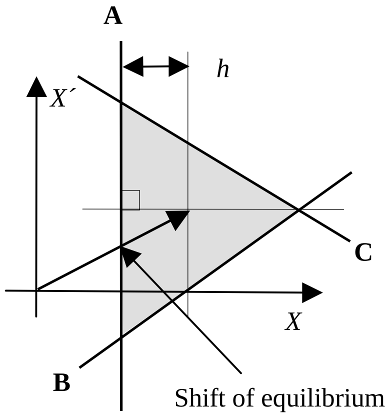
$$\delta\tilde{Q}_x = -\frac{1}{6\pi}3S\left(D_\mathrm{n}\frac{\delta p}{p}\right) = -\frac{1}{4\pi}\beta_\mathrm{x}\ell_\mathrm{S}k'D_\mathrm{x}\frac{\delta p}{p}. \qquad (2.38)$$

By introducing the *chromaticity* Q_x' as the linear change of the betatron tune with momentum, the chromatic effect of a sextupole can be expressed as

$$Q_\mathrm{x}' = \left(\frac{d\tilde{Q}_\mathrm{x}}{dp/p}\right) = -\frac{1}{4\pi}\beta\ell_\mathrm{S}k'D_\mathrm{x}, \qquad (2.39)$$

which corresponds exactly to the well-known linear form for the tune shift introduced by a sextupole. Although the above derivation is restricted to the region in tune close to the third-order resonance the result is in fact more general and the 'tilda' on the tune can be omitted.

The general Hamiltonian (2.37) describes correctly the physics of a third-integer resonance to first order in a perfect lattice without any further restrictions on sextupole locations or particle momenta. The phase space at the sextupole has the same qualitative shape as the earlier simple theory, but it is scaled by the tune shift introduced by the sextupole. For completeness the geometry and general equations of the separatrices are given in Figure 2.6. From these results, it is clear that the resonance sextupole is best put in a dispersion-free region, otherwise any change in the sextupole strength will also change the tune distance of the particles and therefore the scale of the extraction phase space.



Shift of equilibrium orbit:
$\Delta X = D_n\,\delta p/p\Delta X' = D'_n\,\delta p/p$

A $$X - D_n \frac{\delta p}{p} = -h \tag{2.40}$$

B $$-\frac{\sqrt{3}}{2}\left(X' - D'_n \frac{\delta p}{p}\right) + \frac{1}{2}\left(X - D_n \frac{\delta p}{p}\right) = h \tag{2.41}$$

C $$\frac{\sqrt{3}}{2}\left(X' - D'_n \frac{\delta p}{p}\right) + \frac{1}{2}\left(X - D_n \frac{\delta p}{p}\right) = h \tag{2.42}$$

Figure 2.6 Geometry of the separatrices at the sextupole for an off-momentum beam

2.7 CLOSED-ORBIT DISTORTION AT THE SEXTUPOLE

In practice, the equilibrium orbit will be distorted by magnet imperfections and misalignments. The effect will be similar to that caused by the dispersion function, except that the orbit distortion will be independent of particle momentum to first order. It is straight forward to include the orbit distortions at the sextupole in the Hamiltonian.

$$\boldsymbol{H} = \frac{1}{2}\left[\varepsilon - 3S\left(D_n \frac{\delta p}{p} + \Delta X_{\mathrm{CO}}\right)\right]\left(X_\beta^2 + X_\beta'^2\right) + \frac{S}{4}\left(3X_\beta X_\beta'^2 - X_\beta^3\right). \tag{2.43}$$

The geometry and dynamics of the phase space can be analysed as in the previous section. The shift in tune consists of the momentum-dependent dispersion part and the constant offset given by the closed-orbit distortion:

$$\delta\tilde{Q}_{\mathrm{x}} = \left(\delta\tilde{Q}_{\mathrm{x}}\right)_{\mathrm{D}} + \left(\delta\tilde{Q}_{\mathrm{x}}\right)_{\mathrm{CO}} = -\frac{1}{4\pi}\beta_{\mathrm{x}}\ell_{\mathrm{S}}k'\left(D\frac{\delta p}{p}\right) - \frac{1}{4\pi}\beta_{\mathrm{x}}\ell_{\mathrm{S}}k'(\Delta x)_{\mathrm{CO}}. \tag{2.44}$$

2.8 PHASE-SPACE MAPS ALONG THE MACHINE

When designing an extraction scheme, it is essential to know how the phase-space map changes with longitudinal position in the machine. All the considerations are based on the general Hamiltonian (2.37) derived earlier. It is assumed that there is only one sextupole[*], defining the reference position ($\mu = 0$) in the machine. In general, it is sufficient to describe the evolution of the separatrices around the machine as all the relevant physics of the slow-extraction process can be obtained from this.

---

[*] This is not a restriction on the generality of the theory at the level of approximation being used. It will be shown later that from the standpoint of resonance excitation (described on a time scale of 3 turns) many sextupoles can be replaced by a single virtual sextupole.

There are two contributions to the evolution of the separatrices and the stable triangle when tracking around the machine (see Figure 2.6). The momentum-dependent equilibrium orbit defines the centre of the stable triangle at any position *s* in the machine according to,

$$X_{\mathrm{EQ.O}}(s) = D_n(s)\frac{\delta p}{p}, X'_{\mathrm{EQ.O}}(s) = D'_n(s)\frac{\delta p}{p} \tag{2.45}$$

and the phase advance $\Delta\mu$ from the sextupole to a position *s* determines the orientation of the stable triangle according to simple rotation in the normalised phase space,

$$\boldsymbol{M}_{\Delta\mu} = \begin{pmatrix} \cos(\Delta\mu) & \sin(\Delta\mu) \\ -\sin(\Delta\mu) & \cos(\Delta\mu) \end{pmatrix} = \boldsymbol{R}\,(\Delta\mu). \tag{2.46}$$

The size of the stable triangle is determined by the normalised sextupole strength *S* and the modified tune distance ε (including chromatic effects) and remains unchanged in normalised phase space at all positions around the machine. A schematic example is shown in Figure 2.7 of the evolution of the map over a 90° phase advance from the sextupole with a typical change in the dispersion vector.

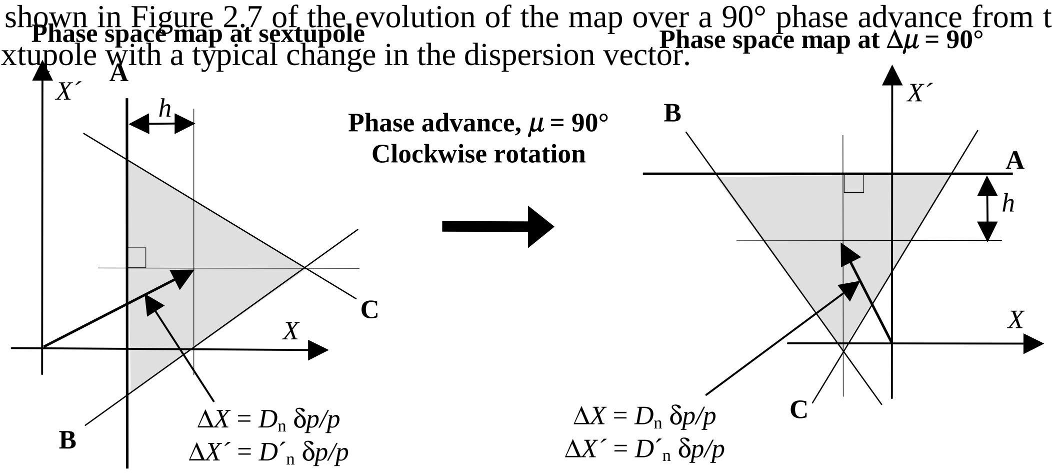
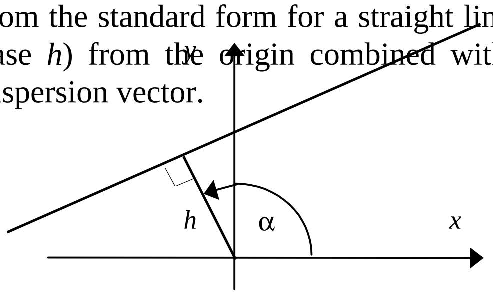


Figure 2.7 Maps separated by $\Delta\mu$ = 90°

2.9 GENERAL EQUATIONS FOR THE SEPARATRICES

A general equation of a separatrix in the phase-space map can be constructed from the standard form for a straight line that uses the perpendicular distance (in this case *h*) from the origin combined with a shift of the origin corresponding to the dispersion vector.



[Angle α is measured anticlockwise from *x*-axis to perpendicular *h*]

$$x\cos\alpha + y\sin\alpha = h \tag{2.47}$$

Figure 2.8 Perpendicular form for a straight line

Thus, the general equation for a separatrix has the form:

$$\left(X - D_n(s)\frac{\delta p}{p}\right)\cos\alpha + \left(X' - D_n'(s)\frac{\delta p}{p}\right)\sin\alpha = h. \tag{2.48}$$

By inspection of Figure 2.4, the values for α_0 at the sextupole are:

(**A**) α_0 = 180° (anticlockwise)
(**B**) α_0 = 300° (anticlockwise)
(**C**) α_0 = 420° (anticlockwise)

The separatrices rotate clockwise with the betatron phase advance $\Delta\mu$ measured from the sextupole. At a given position $\alpha = (\alpha_0\text{-}\Delta\mu)$, but for convenience, it is easier to think in terms of $\Delta\mu$ appearing as a positive term, which gives:

General equations of the separatrices at any position *s* with phase advance from the sextupole of $\Delta\mu$:

(**A**)

$$-\left(X - D_n(s)\frac{\Delta p}{p}\right)\cos(\Delta\mu) + \left(X' - D_n'(s)\frac{\Delta p}{p}\right)\sin(\Delta\mu) = h \tag{2.49}$$

(**B**)

$$-\left(X - D_n(s)\frac{\Delta p}{p}\right)\cos(\Delta\mu - 120°) + \left(X' - D_n'(s)\frac{\Delta p}{p}\right)\sin(\Delta\mu - 120°) = h \tag{2.50}$$

(**C**)

$$-\left(X - D_n(s)\frac{\Delta p}{p}\right)\cos(\Delta\mu + 120°) + \left(X' - D_n'(s)\frac{\Delta p}{p}\right)\sin(\Delta\mu + 120°) = h \tag{2.51}$$

2.10 RESONANCE EXCITATION BY SEXTUPOLES

In a machine, there are likely to be several sextupoles as well as distributed sextupole errors (mainly in the dipoles), whereas the theory presented so far is based on a single sextupole. The combined effect of many sextupole fields on the resonance is described by the so-called *driving term* κ [6]. For the third-order resonance $3Q_x = n$ the driving term is,

$$\kappa(s_0) = \frac{1}{24\sqrt{\pi C}}\int_{s_0}^{s_0+C}\beta_x^{3/2}\left[\frac{-1}{|B_0\rho|}\left(\frac{\mathrm{d}^2B_z}{\mathrm{d}x^2}\right)_0\right]\exp(3\mathrm{i}\mu_x)\,\mathrm{d}s. \tag{2.52}$$

The integral is made around the full machine circumference and includes all sextupole fields. For short sextupoles, the above can be rewritten as a sum, using the normalised sextupole strength, *S*, as defined earlier,

$$\kappa = \frac{-1}{12\sqrt{\pi C}}\sum_{\mathrm{n}} S_{\mathrm{n}}\exp(3\mathrm{i}\mu_{\mathrm{x,n}}). \tag{2.53}$$

An equivalent sextupole can be found by evaluating the above driving term and equating to a single *virtual sextupole*, so that

$$S_{\mathrm{virt}}\exp\left(3\mathrm{i}\,\mu_{\mathrm{x,virt}}\right)=\sum_{n}S_{\mathrm{n}}\exp\left(3\mathrm{i}\mu_{\mathrm{x,n}}\right).$$

By separating the real and imaginary parts,

$$S_{\mathrm{virt}}\cos\left(3\mu_{\mathrm{x,virt}}\right)=\sum_{\mathrm{n}}S_{\mathrm{n}}\cos\left(3\mu_{\mathrm{x,n}}\right),S_{\mathrm{virt}}\sin\left(3\mu_{\mathrm{x,virt}}\right)=\sum_{\mathrm{n}}S_{\mathrm{n}}\sin\left(3\mu_{\mathrm{x,n}}\right)$$

the betatron phase and strength of the virtual sextupole can be found as

Virtual sextupole:

$$\tan\left(3\mu_{\mathrm{x,virt}}\right)=\frac{\sum_{\mathrm{n}}S_{\mathrm{n}}\sin\left(3\mu_{\mathrm{x,n}}\right)}{\sum_{n}S_{\mathrm{n}}\cos\left(3\mu_{\mathrm{x,n}}\right)},\qquad(2.54)$$

$$S_{\mathrm{virt}}^{2}=\left[\sum_{\mathrm{n}}S_{\mathrm{n}}\cos\left(3\mu_{\mathrm{x,n}}\right)\right]^{2}+\left[\sum_{\mathrm{n}}S_{\mathrm{n}}\sin\left(3\mu_{\mathrm{x,n}}\right)\right]^{2}.\qquad(2.55)$$

Two standard configurations should be noted. Arranging an even number of sextupoles with the same normalised strength, S_0 and a regular spacing in phase of $\Delta\mu = \pi/3$ leads to cancellation of the resonance driving term in equation (2.55). Similarly, a spacing of $\Delta\mu = 2\pi/3$ leads to a reinforcement of the driving term,

Cancellation:

$$\begin{aligned}S_{\mathrm{Virtual}}^{2}&=\left\{S_{0}\cos\left(3\mu_{\mathrm{x,1}}\right)+S_{0}\cos\left[3\left(\mu_{\mathrm{x,1}}+60^{\circ}\right)\right]+\ldots\right\}^{2}+\\&\left\{S_{0}\sin\left(3\mu_{\mathrm{x,1}}\right)+S_{0}\sin\left[3\left(\mu_{\mathrm{x,1}}+60^{\circ}\right)\right]+\ldots\right\}^{2}.\end{aligned}\qquad(2.56)$$

Addition:

$$\begin{aligned}S_{\mathrm{Virtual}}^{2}&=\left\{S_{1}\cos\left(3\mu_{\mathrm{x,1}}\right)+S_{2}\cos\left[3\left(\mu_{\mathrm{x,1}}+120^{\circ}\right)\right]+\ldots\right\}^{2}+\\&\left\{S_{1}\sin\left(3\mu_{\mathrm{x,1}}\right)+S_{2}\sin\left[3\left(\mu_{\mathrm{x,1}}+120^{\circ}\right)\right]+\ldots\right\}^{2}.\end{aligned}\qquad(2.57)$$

2.11 CHROMATIC EFFECTS OF SEXTUPOLES

The tune shifts on an off-momentum orbit due to sextupoles are given by,

$$\Delta Q_{x}'=-\frac{1}{4\pi}\int_{0}^{C}k'(s)\beta_{x}(s)D_{x}(s)\mathrm{d}s,\Delta Q_{z}'=\frac{1}{4\pi}\int_{0}^{C}k'(s)\beta_{z}(s)D_{x}(s)\mathrm{d}s.\qquad(2.58)$$

For short lenses, the integrals can be replaced by summations to give

$$\Delta Q_{\mathrm{x}}'=-\frac{1}{4\pi}\sum_{\mathrm{i}}\left(k'\ell\right)_{i}\left(\beta_{\mathrm{x}}D_{\mathrm{x}}\right)_{\mathrm{i}},\Delta Q_{\mathrm{x}}'=\frac{1}{4\pi}\sum_{\mathrm{i}}\left(k'\ell\right)_{i}\left(\beta_{\mathrm{z}}D_{\mathrm{x}}\right)_{\mathrm{i}}.\qquad(2.59)$$

Expressions (2.58) and (2.59) describe the accumulated chromatic effect of all the sextupoles in the machine and the modified tune distance ε of a particle with a momentum deviation $\delta p/p$ becomes

$$\varepsilon = \varepsilon_0 + 6\pi\Delta Q'_x \frac{\delta p}{p}\,. \tag{2.60}$$

For an independent adjustment of the horizontal and vertical chromaticity in a machine at least two independently powered sextupoles are needed. It can be seen from (2.59) that $\Delta Q'_\mathrm{x}$ depends on β_x while $\Delta Q'_\mathrm{z}$ depends on β_z. Since strong focusing lattices often have positions with large ratios of $\beta_\mathrm{x}/\beta_\mathrm{z}$ and $\beta_\mathrm{z}/\beta_\mathrm{x}$, it is possible to combine two or more lenses to obtain a desired combination of $\Delta Q'_\mathrm{x}$ and $\Delta Q'_\mathrm{z}$ and to have a reasonable degree of orthogonality between the series or *families*. In most cases, a family of F-type sextupoles is used with a family of D-type sextupoles. Assuming that all the magnets are of the same construction, the expressions for controlling the chromaticities are,

$$\Delta Q'_\mathrm{x} = -\frac{\ell_\mathrm{S}}{4\pi}\left[k'_\mathrm{SF}\sum_{\mathrm{i}=1}^{\mathrm{NF}}\left(\beta_\mathrm{x} D_\mathrm{x}\right)_\mathrm{i} + k'_\mathrm{SD}\sum_{\mathrm{i}=1}^{\mathrm{ND}}\left(\beta_\mathrm{x} D_\mathrm{x}\right)_\mathrm{i}\right] \tag{2.61}$$

$$\Delta Q'_\mathrm{z} = \frac{\ell_\mathrm{S}}{4\pi}\left[k'_\mathrm{SF}\sum_{\mathrm{i}=1}^{\mathrm{NF}}\left(\beta_\mathrm{z} D_\mathrm{x}\right)_\mathrm{i} + k'_\mathrm{SD}\sum_{\mathrm{i}=1}^{\mathrm{ND}}\left(\beta_\mathrm{z} D_\mathrm{x}\right)_\mathrm{i}\right], \tag{2.62}$$

where *NF* and *ND* are the numbers of lenses in each family.

2.12 PLANNING SEXTUPOLE FAMILIES

Medium and large machines

When planning a machine lattice, it is desirable to keep the resonance excitation and the chromaticity correction separate. The first choice is to place resonance excitation sextupoles in dispersion-free regions. In this way, they do not affect the chromaticity and the phase-space map of the resonance. The two families for the chromaticity control should contain even numbers of lenses separated in phase by $\pi/3$ so that the resonance driving term vanishes. Furthermore the F-type family should have large and equal ratios of $\beta_\mathrm{x}/\beta_\mathrm{z}$ and the D-type family large and equal values of $\beta_\mathrm{z}/\beta_\mathrm{x}$ to give some degree of orthogonality.

Small machines

In small machines, it is not possible to have several sextupoles in a family, or even several families, and some specially tailored combination is required. One example is given below for the case of slow-extraction on a third-order resonance.

The following is based on a lattice comprising two equal 180° arcs joined by two equal dispersion-free straight sections. The sextupole(s) controlling the resonance excitation are placed in one, or both, dispersion-free straight sections so that they can be used without affecting the chromaticity. One chromaticity sextupole is placed in each arc at symmetrically opposed positions (with equal lattice functions). Since the lattice is adjusted for slow extraction the tune will be close to $Q_x \approx n \pm 1/3$ and the two sextupoles will be mutually separated in betatron phase by $\Delta\mu_x = Q\pi \approx (n \pm 1/3)\pi$. The effect on the resonance and the chromaticities can now be evaluated from (2.55), (2.61) and (2.62). For resonance excitation:

$$S_{\mathrm{virt}}^2 = \{S_1 \cos(0) + S_2 \cos(3Q\pi)\}^2 + \{S_1 \sin(0) + S_2 \sin(3Q\pi)\}^2$$

$$S_{\mathrm{virt}} = S_1 \pm S_2 \,. \tag{2.65}$$

The sine terms are close to zero and the cosine terms are close to unity. The upper sign is for n odd and the lower sign for n even.
For chromaticity control:

$$\Delta Q_x' = -\frac{\ell_S}{4\pi}\beta_x D_x (k_1' + k_2'); \qquad \Delta Q_z' = \frac{\ell_S}{4\pi}\beta_z D_x (k_1' + k_2'). \tag{2.66}$$

For n even, $(S_1 - S_2)$ excites the resonance, but not the chromaticity and $(S_1 + S_2)$ does the converse. The control on the chromaticity is not universal, but since the horizontal chromaticity is the more important parameter for extraction, a workable scheme could be built. For n odd, the separation of the functions no longer occurs.

References

[1] Y. Kobayashi, H. Takahashi, *Improvement of the emittance in the resonant ejection*, Proc. VIth Int. Conf. on High Energy Accelerators, Cambridge, Massachusetts, (1967), p347-51.

[2] M.Q. Barton, *Beam extraction from synchrotrons*, Proc. VIIIth Int. Conf. on High Energy Accelerators, CERN, Geneva, (1971), p85-8.

[3] W.C. Elmore, M.W. Garrett, *Measurement of two-dimensional fields. Part I*:Theory, Rev. Sci. Instr., **25**, No. 5 (1954), p480-5.

[4] H. Grote, F.C. Iselin, *The MAD program version 8.4 users' reference manual,* CERN/SL90-13 (AP).

[5] M. Benedikt, *Optimisation of the slow extraction and optical design of a synchrotron for hadron therapy*, Doctoral thesis, Technical University of Vienna (1997).

[6] G. Guignard, *A general treatment of resonances in accelerators*, CERN 78-11 (Nov. 1978).

* * *

I-3 RESONANT SLOW EXTRACTION

A slow extraction scheme requires a mechanism for moving the beam into the resonance. The separatrices and extraction equipment have to be configured for low loss, reasonable apertures and practical septum strengths. The possibilities are numerous and there are definite preferences for choosing different schemes for different applications. Once chosen, the extraction scheme imposes requirements on the momentum spread and emittance of the beam that is waiting to be 'fed' to the resonance and this in turn has consequences for the method of injection and the rf equipment. The characteristics of the extracted beam will only be partially treated in this Chapter and will be continued in Chapter 4. For simplicity, it is assumed that one sextupole is used to excite the resonance and that it is located in a zero-dispersion region. It is also assumed that the resonant tune is located on the central orbit.

3.1 STEINBACH DIAGRAM

Frequent use will be made of the *Steinbach diagram*, which shows the beam and resonance in amplitude-momentum space (see Figure 3.1). The abscissa is the momentum deviation with respect to central orbit, which can equally be the position across the aperture in dispersive regions and the tune in machines with finite chromaticity. The ordinate is the effective, normalised amplitude of the ion's betatron motion, $A = \sqrt{(E/\pi)}$, in normalised co-ordinates calculated from its single-particle emittance. Assuming that the sextupole has been applied adiabatically, then A is equal to the unperturbed circular motion in normalised phase space. In Figure 3.1, the particle density, expressed as dN/dA, has been added on the left-hand side for a typical beam. The large grey arrows indicate the motion of the beam into the resonance and the outward movement of the unstable particles that form the spill. The spill slopes slightly as these particles are still accelerated at the same rate as the main beam.

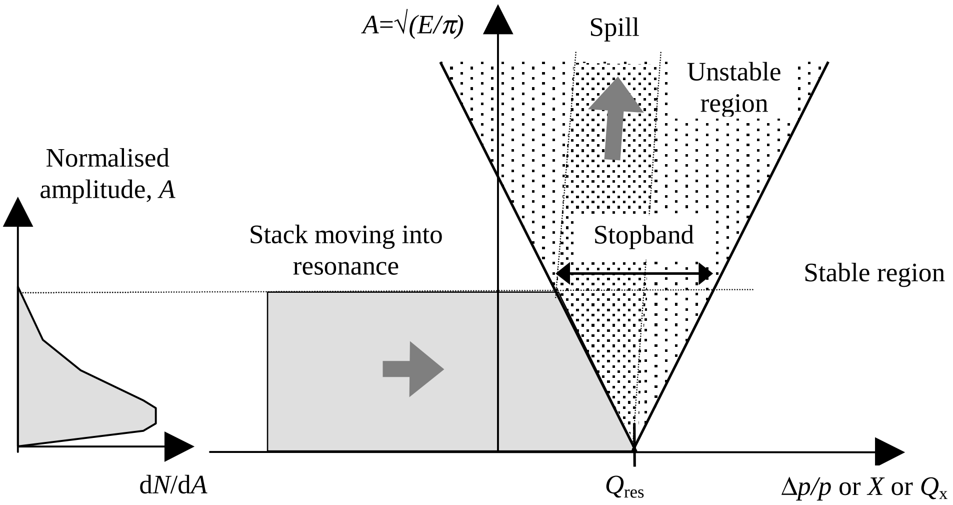

$A=\sqrt{(E/\pi)}$
Spill
Unstable region
Normalised amplitude, A
Stack moving into resonance
Stopband
Stable region
dN/dA
Q_{res}
$\Delta p/p$ or X or Q_x

Figure 3.1 Steinbach diagram showing a beam entering a resonance

In Figure 3.1, the resonance appears as a 'V'-shaped region centred on the resonant tune and the beam is cut off by the sloping side of the 'V'. A particle will be stable if its emittance is smaller, or equal to, the acceptance of the stable triangle defined in Section 2.5 and equation 2.30 and it is this equation that defines the sloping lines of the 'V'. The width of the 'V' at a given emittance defines an interval in tune known as the *stopband.*

From (2.30)
$$\hat{E}_{\mathrm{stable}} = \hat{A}^2_{\mathrm{stable}}\pi \le \frac{48\pi\sqrt{3}}{S^2}\left(\delta Q_{\mathrm{stopband}}\right)^2\pi \tag{3.1}$$

and in absolute tune the stopband is defined by,

$$Q_{\mathrm{res}} - \underbrace{\sqrt{\frac{1}{48\pi\sqrt{3}}\frac{\hat{E}_{\mathrm{stable}}}{\pi}}\left|S\right|}_{\delta Q_{\mathrm{stopband}}} < Q_{\mathrm{particle}} < Q_{\mathrm{res}} + \underbrace{\sqrt{\frac{1}{48\pi\sqrt{3}}\frac{\hat{E}_{\mathrm{stable}}}{\pi}}\left|S\right|}_{\delta Q_{\mathrm{stopband}}}. \tag{3.2}$$

3.1.1 Distribution in the extracted beam

As the beam is moved into the resonance in Figure 3.1, the large emittance ions become unstable first. Due to the chromaticity, there is a one-to-one correspondence between the emittance at which each ion becomes unstable and its momentum. In this way, the particle density distribution in betatron amplitude (dN/dA) is converted into the particle density distribution in momentum space (dN/dp/p) in the extracted beam.

3.1.2 Special case of zero chromaticity

As the chromaticity approaches zero, the 'V' in Figure 3.1 becomes wider and flatter. If the tune of the machine is then set just off-resonance, the resonant tune moves away to infinity and the resonance appears as a straight line quasi-parallel to the horizontal axis (momentum). The height of the line above the axis defines the boundary of the unstable region.

3.2 OVERVIEW OF EXTRACTION METHODS

A beam is first accumulated and positioned on one side of the resonance. To extract, either the resonance has to be moved towards the beam, or vice versa. This distinction provides a first way of classifying extraction methods. An alternative classification is based on how the ions are selected by the resonance. The selection can either be according to amplitude, or according to tune (momentum). Thus ions with large amplitudes would enter the resonance before those with smaller amplitudes, or those with a higher (or lower) momentum before those with a lower (or higher) momentum. There will also be hybrid situations, especially for the amplitude and momentum selections, which will nearly always be mixed to some degree.

3.2.1 First classification

Moving the beam

This method has the great advantage of leaving the optical parameters of the machine constant, and hence also those of the resonance, during the extraction. How the beam is moved can be further sub-divided into the two most promising categories for future machines:

- ***Acceleration-Driven*** The beam is accelerated towards the stationary resonance by a betatron core, or by stochastic noise, or possibly by a phase displacement or a rf micro-bucket acceleration system.
- ***RF 'knockout'*** The beam is excited by transverse stochastic noise or rf excitation at the revolution frequency, so that its betatron amplitudes grow. The chromaticity is set to zero, or a low value, so that the resonance line acts as a threshold in amplitude above which the ions become unstable [1].

Moving the resonance

There are two ways in which the resonance can be moved:

- ***Quadrupole-driven*** The tune of the machine is changed so that the resonance region moves towards the beam. This is the conventional way to operate a slow extraction scheme. It requires no additional equipment and its basic idea is simple.
- ***Sextupole-driven*** The resonance excitation is changed by increasing the sextupole strength. This method is included only for academic completeness and is not suitable for a medical application.

3.2.2 Alternative classification

An alternative method of classification is to look at how the particles are selected by the resonance. This divides the extraction systems into two different broad categories:

- ***Amplitude selection*** In this case, the large betatron amplitude ions enter the resonance first followed progressively by the smaller and smaller amplitudes. This implies either a rather flat resonance line in the Steinbach diagram or a narrow momentum spread in the beam.
- ***Momentum selection*** In this case, the high (or low) momentum ions enter the resonance first and are progressively followed by lower and lower (or higher and higher) momentum ions.

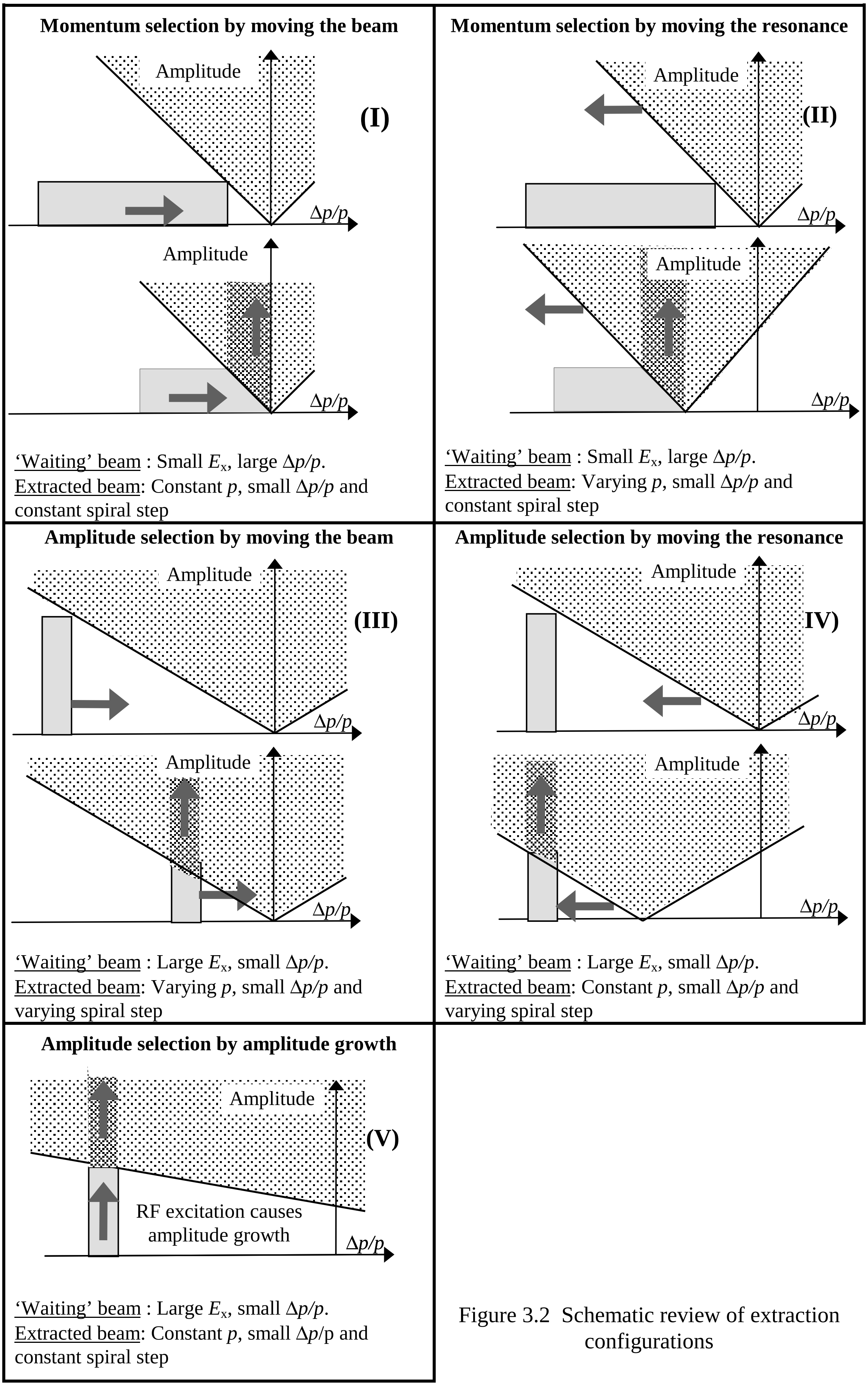
The most interesting possibility is a hybrid method, which correlates the amplitude to the momentum in a precise way* that imposes the same extraction trajectory on ions of all momenta. This is the proposed extraction technique for PIMMS and will be referred to as the *amplitude-momentum selection method* that is *acceleration-driven*.

3.2.3 Extraction schemes

The various possibilities are schematically presented in Figure 3.2. From the five cases, an initial choice can be made. Method **(II)** *'Momentum selection by moving the resonance'* and method **(III)** *'Amplitude selection by moving the beam'* are less attractive because the momentum of the extracted beam varies during the extraction. Method **(IV)** *'Amplitude selection by moving the resonance'* is also less attractive because the spiral step varies during the extraction. Finally, the choices with the best extracted beam characteristics are **(I)** *‘Momentum selection by moving the beam’* and **(V)** *‘Amplitude selection by amplitude growth’*.

---

* The correlation is known as the Hardt Condition and will be discussed in Sections 3.6 to 3.8


Amplitude
(I)
Δp/p
Amplitude
Δp/p
Amplitude
(II)
Δp/p
Amplitude
Δp/p
Amplitude
(III)
Δp/p
Amplitude
Δp/p
Amplitude
IV)
Δp/p
Amplitude
Δp/p
Amplitude
(V)
RF excitation causes amplitude growth
Δp/p

Momentum selection by moving the beam

‘Waiting’ beam : Small E_x, large $\Delta p/p$.
Extracted beam: Constant p, small $\Delta p/p$ and constant spiral step

Momentum selection by moving the resonance

‘Waiting’ beam : Small E_x, large $\Delta p/p$.
Extracted beam: Varying p, small $\Delta p/p$ and constant spiral step

Amplitude selection by moving the beam

‘Waiting’ beam : Large E_x, small $\Delta p/p$.
Extracted beam: Varying p, small $\Delta p/p$ and varying spiral step

Amplitude selection by moving the resonance

‘Waiting’ beam : Large E_x, small $\Delta p/p$.
Extracted beam: Constant p, small $\Delta p/p$ and varying spiral step

Amplitude selection by amplitude growth

‘Waiting’ beam : Large E_x, small $\Delta p/p$.
Extracted beam: Constant p, small Δp/p and constant spiral step

Figure 3.2 Schematic review of extraction configurations

Choosing between the last two possibilities depends on subjects that will be treated later and upon the type of treatment that is envisaged. In anticipation of a fuller explanation, the following points can be noted:

(I) ***Acceleration-driven momentum selection (moving the beam)***

Advantages	Disadvantages
• Lattice parameters and spiral step are constant during the spill. • Hardt condition can be applied (Section 3.6). • A front-end acceleration mechanism can be added (Chapter 6). • Extraction takes place from a range of emittances and this smoothes the spill.	• Less convenient for starting and stopping the spill at the level of the resonance. • Additional constraints on the optics.

(V) ***RF knockout amplitude selection (amplitude growth)***

Advantages	Disadvantages
• Lattice parameters and spiral step are constant during spill. • Spill can be turned on and off very cleanly and easily by the rf kicker, which is convenient for breathing synchronisation	• The near-zero chromaticity may cause stability problems in the ‘waiting’ beam. • Not possible to add a front-end acceleration mechanism.

Method **(I)** is that proposed for PIMMS and Method **(V)** is used at HIMAC for treatments with respiration gating [2].

3.3 THE MECHANISM OF EXTRACTION

3.3.1 Jumping into the electrostatic septum

In the theoretical model built in Chapter 2, all the separatrices are essentially equal and they rotate clockwise, simply exchanging their positions from turn to turn. However, when analysing the extraction process, it is more convenient to consider the separatrices as geometrically fixed in phase space. An unstable particle then moves steadily outwards while jumping from one separatrix to the next at every turn. After every third turn the particle returns to its initial separatrix, but at a position more distant from the centre. The change of the particle co-ordinates after three revolutions in the machine was derived as,

$$\begin{aligned}\Delta X_3 &= \varepsilon X_0' + \frac{3}{2} S X_0 X_0' \\ \Delta X_3' &= -\varepsilon X_0 + \frac{3}{4} S\left(X_0^{\,2} - X_0'^{\,2}\right)\end{aligned} \tag{2.25}$$

and these changes were called the *spiral step* and the *spiral kick*. The change in amplitude of the particle can be written as,

$$(\Delta A)^2 = (\Delta X)^2 + (\Delta X')^2 \tag{3.3}$$

where Δ indicates the change over three turns. To find the maximum change in amplitude, an on-resonance particle is considered (i.e. ε = 0). This simplifies (2.25) and once substituted into (3.3) yields, Maximum change: $$\Delta A = \frac{3}{4} S A^2 . \tag{3.4}$$

Equation (3.4) shows that the growth increases rapidly as the particle progresses along the unstable separatrix. After a certain number of turns the particle amplitude has increased so much that the particle can ‘jump’ by ΔA into an electrostatic septum and be extracted.

Figure 3.3 gives a schematic view of the normalised phase space at the electrostatic septum. A particle at point **0** is just passing the electrostatic septum on the machine side. One turn later, at the electrostatic septum, this particle has jumped to point **1** on the next separatrix that is 120° ahead. One more turn later, still at the electrostatic septum, the particle re-appears at point **2**. Finally, after three turns, the particle returns to the initial separatrix, but at point **3**. The growth in amplitude that has taken place during the three turns brings the particle inside the aperture of the electrostatic septum where it is deflected and extracted.

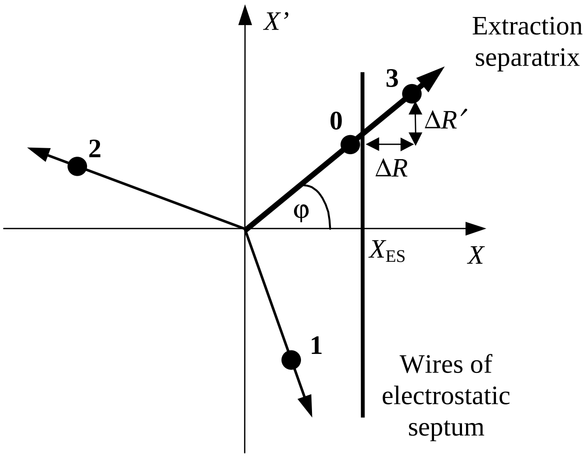


Figure 3.3 Amplitude increase during the last three turns before extraction

The maximum possible amplitude, A_{last}, of a particle, which is not extracted, can be calculated from the normalised position of the electrostatic septum, X_{ES}, and the angle φ between the extraction separatrix and the horizontal axis,

$$A_{last} = \frac{1}{\cos(\varphi)} X_{ES} . \tag{3.5}$$

Three turns later the particle will have increased in amplitude by, ΔA found from (3.4) and (3.5).

$$\Delta A = \frac{3}{4} S A_{\mathrm{last}}^2 = \frac{3}{4} S \left(\frac{X_{\mathrm{ES}}}{\cos\varphi} \right)^2 . \tag{3.6}$$

From (3.6), it can be calculated how far the particle jumps into the septum, the so-called spiral step Δ*R*, and the spiral kick Δ*R*′,

Maximum spiral step:
$$\Delta R = \Delta A_{\text{last}} \cos\varphi = \frac{3}{4} S \frac{1}{\cos\varphi} X_{\text{ES}}^{2}, \tag{3.7}$$

Maximum spiral kick
$$\Delta R' = \Delta A_{\text{last}} \sin\varphi = \frac{3}{4} S \frac{\tan\varphi}{\cos\varphi} X_{\text{ES}}^{2}. \tag{3.8}$$

The spiral step and kick define the size (emittance) of the extracted beam. It should be noted that (3.7) and (3.8) are only valid for particles that are exactly on resonance ($\delta Q = \varepsilon = 0$). Particles with a finite tune distance already have a finite amplitude when they become unstable, they spend fewer turns on the separatrix before they reach the electrostatic septum and their spiral step and kick are smaller. Thus the extraction separatrix will be continuously populated along its length by particles with different starting conditions. Some of the particles will hit the septum wall and will be lost. Thus the extraction efficiency depends critically on the thickness of the septum wall compared to the average spiral step. This is the main reason for the use of an electrostatic septum, which is usually built as a wire septum with a wire thickness of the order of 0.1 mm. Electric fields of up to 100 kV/cm can be obtained in such septa, but the kick obtained is rarely sufficient to extract the beam directly out of the machine, so a further step is required in the extraction process.

3.3.2 Introducing the magnetic septum

The small deflection provided by the electrostatic septum translates with phase advance into a physical gap between the circulating beam and the extracted particles. This gap can be used to introduce the wall of a magnetic septum, which can definitively extract the particles. The action of translating the deflection ϕ of the electrostatic septum into a gap for the magnetic septum wall is illustrated in Figure 3.4.

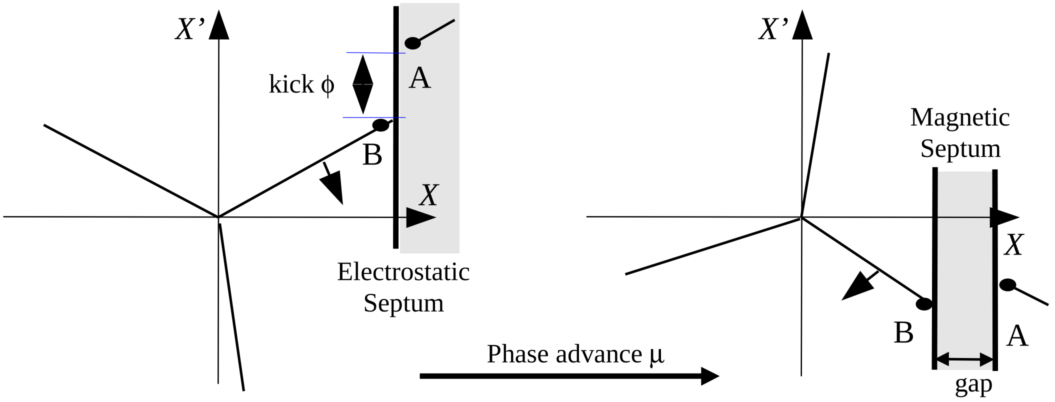


Figure 3.4. Effect of the kick of the electrostatic septum in normalised phase space

The gap size is calculated by comparing the trajectories of two on-momentum ($\delta p = 0$) particles from the electrostatic to the magnetic septum. Particle A starts just inside the electrostatic septum and particle B starts just outside, as shown in Figure 3.4. The thickness of the electrostatic septum (typically 0.1 mm) is neglected, both particles are assumed to start from the radial position of the septum x_{ES}, and with the same angle x'_{ES}, but only particle A receives the kick ϕ. The positions and angles of the particles at the magnetic septum are obtained as,

Particle A:

$$\begin{aligned} x_{\mathrm{MS}} &= m_{11} \cdot x_{\mathrm{ES}} + m_{12} \cdot x'_{\mathrm{ES}} + m_{12} \cdot \phi \\ x'_{\mathrm{MS}} &= m_{21} \cdot x_{\mathrm{ES}} + m_{22} \cdot x'_{\mathrm{ES}} + m_{22} \cdot \phi \end{aligned} \tag{3.9}$$

and

Particle B:

$$\begin{aligned} x_{\mathrm{MS}} &= m_{11} \cdot x_{\mathrm{ES}} + m_{12} \cdot x'_{\mathrm{ES}} \\ x'_{\mathrm{MS}} &= m_{21} \cdot x_{\mathrm{ES}} + m_{22} \cdot x'_{\mathrm{ES}} \end{aligned} \tag{3.10}$$

where the coefficients m_{11}, m_{12} etc. are elements of the transfer matrix between the electrostatic and magnetic septa. Thus, the effect of the kick appears at the magnetic septum as a difference in position and angle between the two particles of,

$$\begin{aligned} \Delta x_{\mathrm{MS}} &= m_{12} \cdot \phi \\ \Delta x'_{\mathrm{MS}} &= m_{22} \cdot \phi \end{aligned} \tag{3.11}$$

where Δx_{MS} is the gap, available for the wall of the thicker magnetic septum. This gap is explicitly given by,

$$\Delta x_{\mathrm{MS}} = \phi \cdot \sqrt{\beta_{\mathrm{ES}} \cdot \beta_{\mathrm{MS}}} \cdot \sin\mu \tag{3.12}$$

where μ is the phase advance between the two septa.

It follows from (3.12) that in order to create the space for the magnetic septum efficiently, the lattice functions of the machine and the positions for the septa have to be chosen such that:

- The phase advance between the septa is close to 90° + $n \cdot 180°$.
- The beta-functions at the septa have reasonably large values.

3.4 SEPARATRIX GEOMETRY AT ELECTROSTATIC SEPTUM

At first sight, the principle of transforming the angular kick of the electrostatic septum to a physical displacement looks to be equally applicable starting in any of the four quadrants. In fact, only the first and third quadrants are usable. This follows from purely geometrical considerations as shown in Figure 3.5. First, the septum kick must always be directed away from the *X*-axis, otherwise it tends to overlap the two sections of the separatrix rather than separating them. Secondly, when *X* is positive, the kick must also be positive and, when *X* is negative, the kick must be negative, otherwise the kick drives the separated section of the separatrix towards the axis of the machine and back into the wires of the septum. Once these restrictions are taken into account only the 1st and 3rd quadrants remain useful.

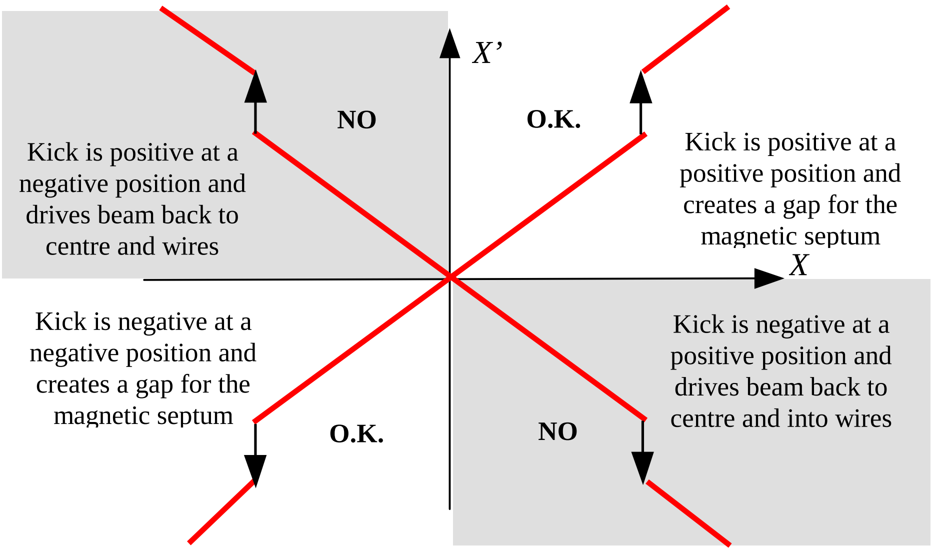

X'
X
NO
O.K.
O.K.
NO
Kick is positive at a negative position and drives beam back to centre and wires
Kick is positive at a positive position and creates a gap for the magnetic septum
Kick is negative at a negative position and creates a gap for the magnetic septum
Kick is negative at a positive position and drives beam back to centre and into wires

Figure 3.5 Possible working quadrants for the electrostatic septum

The next parameter to be considered is the angle the separatrix in phase at the electrostatic septum. From equation (3.12), it was found that for the optimum usage of the electrostatic septum's kick, the phase advance to the magnetic septum should be 90° + $n\cdot180°$. Figure 3.6 shows a logical layout for first quadrant operation with the extraction separatrix at 45° at the electrostatic septum and a phase advance of 90° to the magnetic septum.

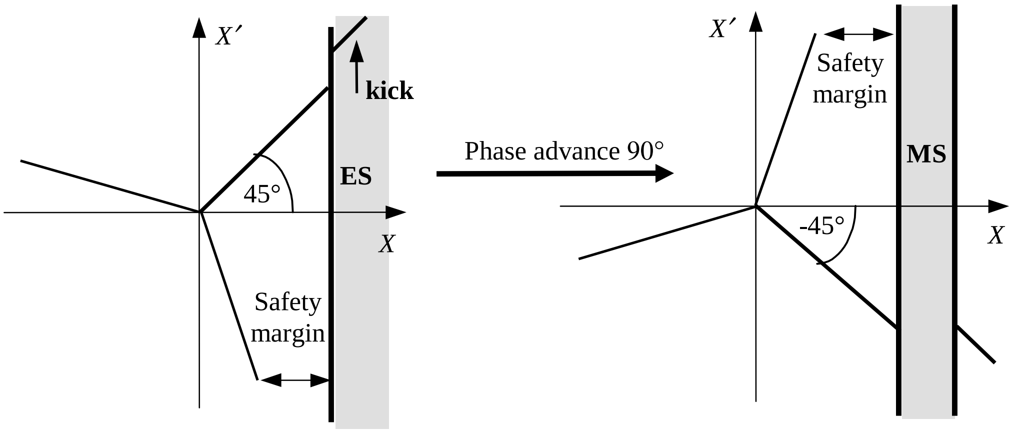

X′
kick
45°
ES
X
Safety margin
Phase advance 90°
X′
Safety margin
MS
-45°
X

Figure 3.6 Ideal separatrix geometry for first quadrant operation

Figure 3.6 shows that the 120° separation of the separatrices leaves a margin of security first at the electrostatic septum and then later at the magnetic septum. The extraction separatrix could be rotated by a maximum of 15° anticlockwise before the preceding separatrix hits the electrostatic septum as shown in Figure 3.7 (a), or by 15° clockwise before the following separatrix hits the magnetic septum, Figure 3.7 (b). This limits the possible angles for the extraction separatrix at the electrostatic septum to 45° ± 15° for first quadrant operation and to 225° ± 15° for third quadrant operation.

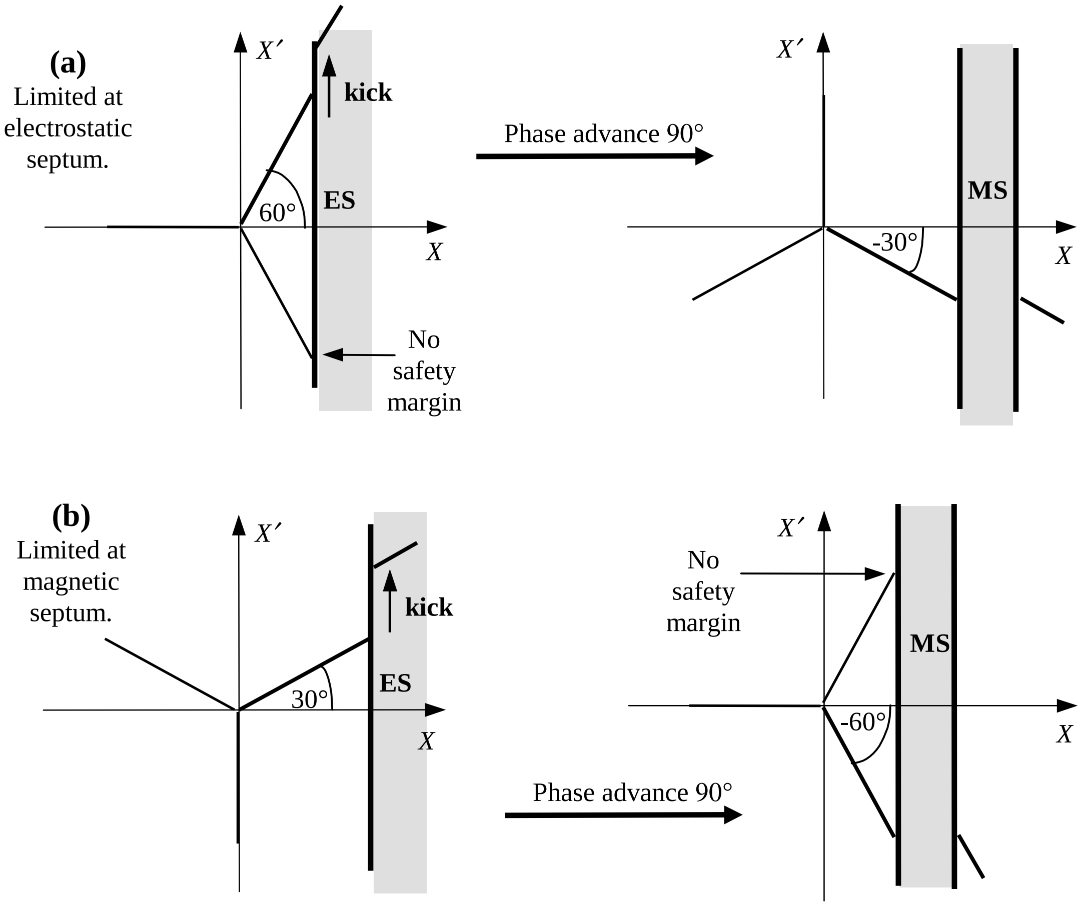

(a)
Limited at electrostatic septum.
X′
kick
60°
ES
X
No safety margin
Phase advance 90°
X′
MS
-30°
X
(b)
Limited at magnetic septum.
X′
kick
30°
ES
X
Phase advance 90°
X′
No safety margin
MS
-60°
X

Figure 3.7 Limitations of separatrix geometry by electrostatic and magnetic septa

Alternative layouts to the above are limited. They include:

- ***To accept a much smaller phase advance***. This may be imposed by lack of space, but it requires a stronger electrostatic septum, which may have implications for the reliability. This solution fits the lower energy proton machines better, where the extraction could then be easily made in a single straight section. However, it was proposed for the EULIMA light-ion medical machine [3].
- ***To use a 270° phase advance.*** This has two unfavourable aspects. Firstly, the extracted beam has to be transported for a longer distance in the machine, which for a small synchrotron means that the crossing of non-linear magnetic elements (e.g. chromaticity or resonance sextupoles) is almost unavoidable. Any change of these elements would result in a change of the extraction geometry. Secondly, a phase advance of 270° (more general 270° + $n \cdot 360°$) means that the electrostatic and magnetic septa are on opposite sides of the vacuum chamber. This may have the drawback that aperture is lost for the circulating beam by ‘encasing’ it between the septa.

3.5 PHASE-SPACE REPRESENTATION OF BEAM AND RESONANCE

The beam that is waiting to enter the resonance will have a certain momentum spread $\Delta p/p$ and, due to the chromaticity in the machine, this momentum spread will translate into a tune spread according to

$$\Delta Q = Q' \frac{\Delta p}{p}\,. \tag{3.13}$$

The ‘size’ of the beam in an accelerator is characterised by the *emittance*, which is defined as the area in phase space that contains a certain percentage of the beam particles of a given momentum. In general, the beam will contain particles with amplitudes between zero and a maximum amplitude, which corresponds to the *total emittance* of the beam. At a given azimuthal position s in the machine, the beam can be represented in phase space by a series of ellipses (circles in normalised phase space), centred on the dispersion vector $\boldsymbol{D}(s)\cdot\Delta p/p$, as shown in Figure 3.8.

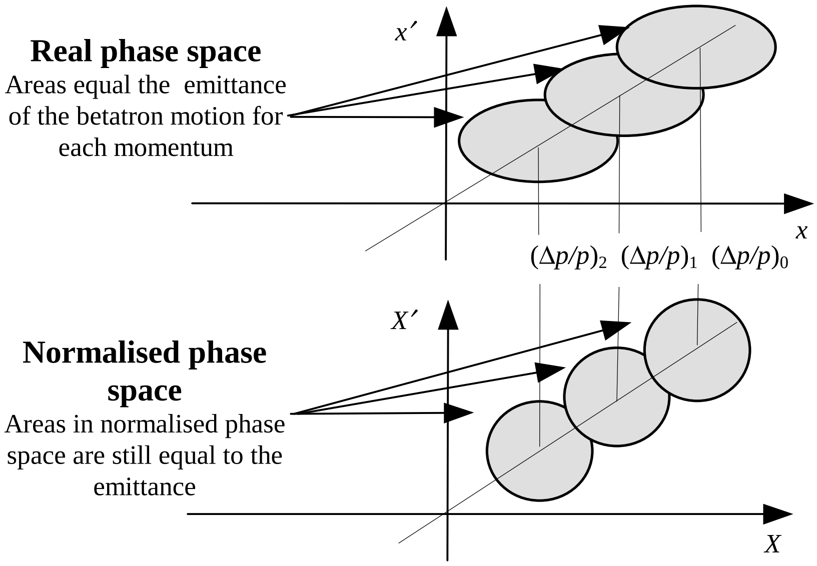

Real phase space
Areas equal the emittance of the betatron motion for each momentum
x′
x
(Δp/p)₂
(Δp/p)₁
(Δp/p)₀
Normalised phase space
Areas in normalised phase space are still equal to the emittance
X′
X

Figure 3.8 Representation of a beam in phase space and normalised phase space

The circles, that represent the beam emittance in normalised phase space, become triangles under the influence of the resonance as shown in Figure 3.9. The size of a triangle corresponds to the last stable orbit and is graded according the its distance in tune from the resonance. Any particles outside the last stable triangle are lost along the outward separatrices. Equation (3.13) gives the tune for each momentum and hence also the tune distance from the resonance.

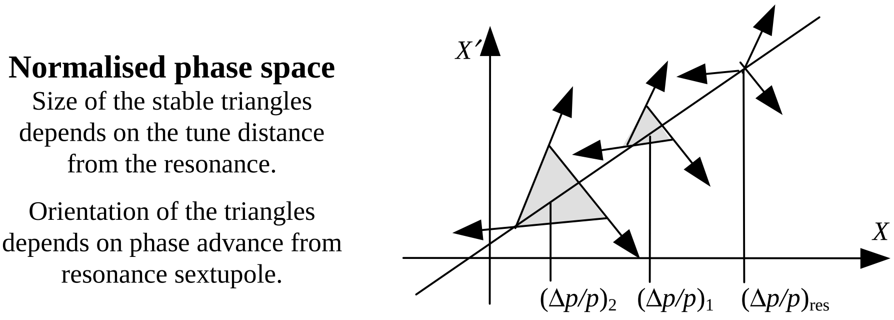

Normalised phase space
Size of the stable triangles depends on the tune distance from the resonance.
Orientation of the triangles depends on phase advance from resonance sextupole.
X′
X
(Δp/p)₂
(Δp/p)₁
(Δp/p)res

Figure 3.9 A beam under the influence of the resonance in normalised phase space

When considering the interactions of the beam and the resonance, it is convenient to use the Steinbach diagram that represents the stopband of the third-integer resonance (see Section 3.1). With (3.13) the expression for the stopband (3.1) can be rewritten with the momentum spread and the chromaticity as

$$A=\sqrt{48\pi\sqrt{3}}\left|\frac{Q'}{S}\right|\cdot\left|\frac{\Delta p}{p}\right|. \tag{3.14}$$

It is sometimes more convenient to use the momentum spread rather than the tune for abscissa, as the momentum spread is an independent beam parameter (changing the chromaticity will change the tune but not the momentum). Figure 3.10 (a) shows the circulating beam before extraction, corresponding to the phase-space representation in Figure 3.8. Figure 3.10 (b) shows the beam during the extraction process, corresponding to the phase-space diagram in Figure 3.9.

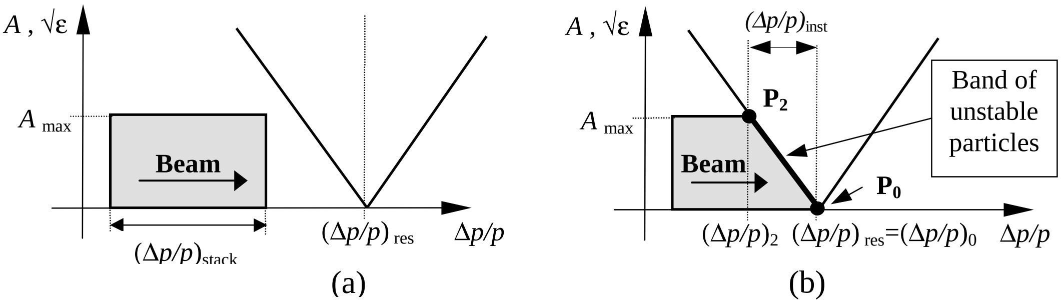

A , √ε
A max
Beam
(Δp/p)stack
(Δp/p) res
Δp/p
(a)
A , √ε
(Δp/p)inst
P2
Band of unstable particles
A max
Beam
P0
(Δp/p)2
(Δp/p) res=(Δp/p)0
Δp/p
(b)

Figure 3.10. Steinbach diagram of the beam before and during extraction

Providing the beam is wide in momentum and small in amplitude a quasi-static situation is reached in which a band of particles of all amplitudes continuously enters the resonance (see *Momentum selection by moving the beam* in Figure 3.2). These particles have different momenta and, due to the chromaticity different tunes, which means the corresponding stable triangles are of different sizes. Figure 3.11 shows the separatrix geometry for the zero and maximum amplitude particles that define the unstable band at the resonance sextupole and the electrostatic septum in normalised phase space.

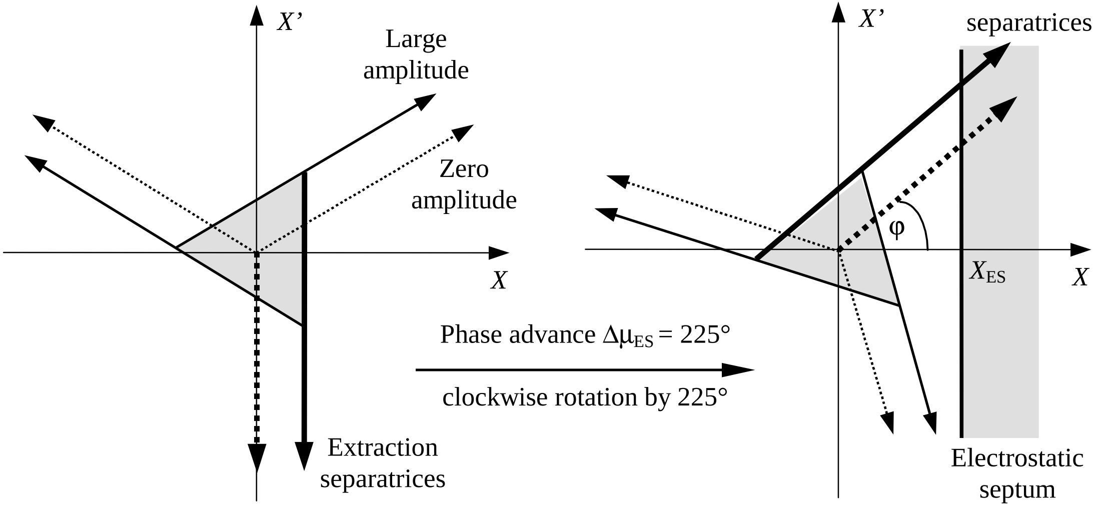

X'
Large amplitude
Zero amplitude
X
Extraction separatrices
Phase advance Δμ_ES = 225°
clockwise rotation by 225°
X'
Extraction separatrices
φ
X_ES
X
Electrostatic septum

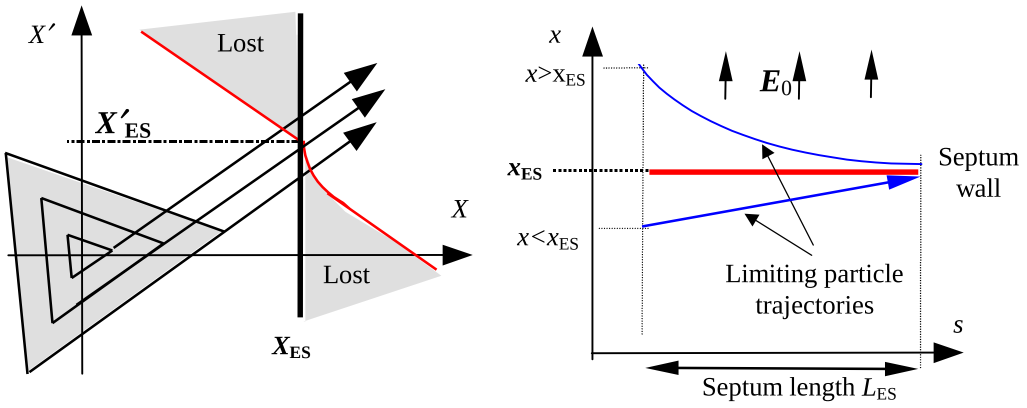
Figure 3.11 Separatrices at the resonance sextupole and electrostatic septum

The dotted lines are the separatrices corresponding to the zero amplitude particles P_0 in Figure 3.10 (b), the full lines are the separatrices of the maximum amplitude particles, P_2. The separatrices of all other particles are found in between these two extreme cases. The instantaneous momentum spread of the extracted beam can be calculated from (3.14) with the maximum amplitude $A_{\max}$ as

$$\left|\frac{\Delta p}{p}\right|_{\text{inst.}} = A_{\max}\sqrt{\frac{1}{48\sqrt{3}\pi}}\left|\frac{S}{Q'}\right|. \tag{3.15}$$

3.6 HARDT CONDITION FOR SUPERPOSITION OF SEPARATRICES

In Figure 3.10 of the previous section, particles of all amplitudes become unstable at the same time and are extracted along separatrices from stable triangles of different sizes. Figure 3.12 shows this general case when the separatrices at the electrostatic septum are not superimposed and the particles move outwards along different trajectories and therefore reach the electrostatic septum with different angles. This angular spread increases the effective thickness of the septum wall and increases the beam losses.



(a) **(b)**

Figure 3.12 Particle losses on the electrostatic septum due to angular spread of separatrices

Figure 3.12 (a) shows phase space map at the septum. This is like a cross-section of the beam with the particles moving perpendicular to the paper. Those particles arriving exactly on the thick line denoting the electrostatic septum are lost on the entry face of the septum wall (wire). Those arriving in the grey areas (shadow regions) hit the septum wall (wires) somewhere along the length of the septum. The boundary between survival and loss is curved for the those particles that arrive with a radial position inside the septum aperture ($x > x_{ES}$) and directed towards the axis with respect to the septum ($x' < x'_{ES}$), because the electric field of the septum deflects these particles away from the wires and in some case this is sufficient to prevent them from being lost. Figure 3.12 (b) shows a longitudinal view with the limiting trajectories that separate the survival and loss regions.

Minimum losses would be obtained if all the separatrices could be superpositioned and directed through the point between the two shadow regions. This is achieved by the Hardt Condition. It was mentioned earlier that stable triangles of different sizes correspond to different momenta. At positions in the machine where the dispersion is non-zero, the triangles will be shifted according to their momentum. With a suitable dispersion vector at the electrostatic septum, this effect can be used to superimpose the extraction separatrices and to avoid the shadow regions. It should be noted that in Figures 3.11 and 3.12 (a) all the triangles are centred which means that the dispersion was assumed to be zero at the electrostatic septum.

The restrictions on the lattice functions to satisfy the Hardt Condition can be derived with a purely mathematical approach from the general expression for the separatrix (2.48) in Section 2.9.

$$\left(X - D_n(s)\frac{\delta p}{p}\right)\cos(\alpha - \Delta\mu) + \left(X' - D_n'(s)\frac{\delta p}{p}\right)\sin(\alpha - \Delta\mu) = h\,. \qquad (2.48)'$$

The angle α describes the orientation of the separatrices at the sextupole, $\Delta\mu$ is the phase advance from the sextupole to the electrostatic septum and h is the distance from the side of the stable triangle to its centre given in (2.29).

$$h = \frac{2}{3}\frac{\varepsilon}{S} = \frac{4\pi}{S}\delta Q\,. \qquad (2.29)'$$

With (3.13), the above expression can be rewritten as

$$h = \frac{4\pi}{S}Q'\frac{\delta p}{p}\,. \qquad (3.16)$$

The substitution of (3.16) into (2.48)′ gives the general expression for a separatrix as a function of particle momentum and the chromaticity of the machine,

$$\left(X - D_n(s)\frac{\delta p}{p}\right)\cos(\alpha - \Delta\mu) + \left(X' - D_n'(s)\frac{\delta p}{p}\right)\sin(\alpha - \Delta\mu) = \frac{4\pi}{S}Q'\frac{\delta p}{p}\,. \qquad (3.17)$$

To superimpose the extraction separatrices, the momentum dependence has to be removed from (3.17). This is true if,

Hardt Condition: $$D_\mathrm{n}\cos(\alpha - \Delta\mu) + D_\mathrm{n}'\sin(\alpha - \Delta\mu) = -\frac{4\pi}{S}Q'\,. \qquad (3.18)$$

Neglecting, for the moment, the mathematically trivial case of zero dispersion and zero chromaticity that will be discussed in Section 3.13, the flexibility of this equation is somewhat limited.

- The dispersion function is a fundamental property of the lattice depending on the layout of the dipoles. If the lattice already exists, or is determined by other factors, this could be a very severe disadvantage.
- The choice of (α-$\Delta\mu$) is restrained by the geometry of the extraction. For optimised operation in the first or third quadrant (see Section 3.4), it can be shown that the phase (α-

$\Delta\mu$) is either 135° for particles with tune values below resonance ($\delta Q < 0$), or 315° for particles above resonance ($\delta Q > 0$).

- The sextupole strength cannot be used as a variable, since it determines the spiral step and spiral pitch and therefore the horizontal size of the extracted beam.
- For small, low-energy machines working below transition, the chromaticity should be negative to ensure the stability of the coasting beam [4]. However, the chromaticity can still be varied over a wide range and this is the main source of adjustment.

To understand better the Hardt Condition, the left hand side (LHS) of (3.18) can be regarded as a scalar product of two vectors,

$$\underbrace{\begin{pmatrix} D_\mathrm{n} & D'_\mathrm{n} \end{pmatrix}\begin{pmatrix} \cos(\alpha-\Delta\mu) \\ \sin(\alpha-\Delta\mu) \end{pmatrix}}_{\text{LHS}} = \underbrace{-\frac{4\pi}{S}Q'}_{\text{RHS}} \tag{3.19}$$

where the (cosine, sine) vector is the unit vector, perpendicular to the extraction separatrices as shown in Section 2.9. The LHS is zero, if these two vectors are perpendicular, which means that the normalised dispersion is parallel to the separatrices. In this case, it is clear that the separatrices cannot be superimposed for any finite value of chromaticity. (The case of a zero-chromaticity extraction is considered separately in Section 3.13). The absolute value of the LHS will be maximum for parallel vectors, which means the normalised dispersion is at right angles to the extraction separatrices and the shift of triangles with different momenta, relative to each other, is then the most effective. Thus, the underlying principle for an efficient lattice for the Hardt Condition is that the normalised dispersion vector should be at an angle close to 90° with respect to the extraction separatrix at the position of the electrostatic septum.

The RHS of (3.18) adjusts the scalar amplitude of the correction by varying the horizontal chromaticity. Providing the vectors forming the LHS are well positioned the correction will be possible without extreme values. Unfortunately the horizontal chromaticity is not entirely free, since its also affects the width of the stop band (the slope of the resonance line) and therefore the extracted momentum spread. Thus, there is a certain amount of trial-and-error adjustment needed to reach a satisfactory situation. In general, the extraction should be arranged such that the momentum spread of the extracted beam is small for two main reasons:

- The optics of the beam delivery system towards the patient is complicated and a small momentum spread reduces beam sizes in the bends. The momentum spread can be artificially increased at the end by a ridge filter to create a spread-out Bragg-peak.
- The transfer between electrostatic and magnetic septum is in general chromatic. These chromatic effects can be partially compensated by a higher voltage in the electrostatic septum, but are less important for a small momentum spreads.

The following Figures 3.13 and 3.14 illustrate how the Hardt Condition is applied at the electrostatic septum for first quadrant operation. The full lines represent the extraction separatrices corresponding to zero amplitude (zero-size triangle) and maximum amplitude (maximum-size triangle) particles. The dotted lines indicate the position of the maximum-size triangles if the dispersion function were zero. In Figure 3.13, the normalised dispersion vector is almost parallel to the separatrices, thus the absolute value of the LHS in (3.19) is small, whereas, in Figure 3.14, the normalised dispersion is perpendicular to the separatrices and the LHS is large. In Figures 3.13 (a) and 3.14 (a) the chromaticity is not adjusted to fulfil the Hardt condition (RHS of (3.19) is equal for both diagrams). In Figures 3.13 (b) and 3.14 (b), the chromaticity has been adjusted in order to superimpose the separatrices. In Figure 3.13 (b), the dispersion vector is not well suited; the final absolute chromaticity is small and the extracted momentum spread is large as can be seen from the Steinbach diagram. In Figure 3.14 (b) the normalised dispersion vector is perpendicular to the separatrices which are superimposed for a small momentum spread of the extracted beam and a large absolute chromaticity.

3.7 1ST AND 3RD QUADRANT OPERATION WITH THE HARDT CONDITION

Figures 3.15 and 3.16 summarise all the possible extraction layouts for optimised first and third quadrant operation. The arrows indicate the required direction of the normalised dispersion vector to fulfil the Hardt condition with the smallest possible momentum spread of the extracted beam. The dotted lines are the extraction separatrices for the zero-amplitude particles which are exactly on resonance ($\delta Q = 0$, $\delta p = 0$), the full lines correspond to maximum amplitude particles with a momentum deviation δp and a tune deviation $\delta Q = Q'\delta p/p$. The choice may appear large, but it quickly reduces in the following way:

- It was mentioned earlier that for small, low-energy machines that work below transition, the chromaticity should be negative in order to ensure the transverse stability of the beam. This constraint leaves only four possibilities for the extraction geometry at the electrostatic septum that are in Figures 3.15 and 3.16 in (b) and (d).

- In Figures 3.15 (b) and (d) the electrostatic septum is on the outside of the vacuum chamber, the dispersion required for the Hardt condition is $D_n > 0$ and $D'_n < 0$. In Figures 3.16 (b) and (d) the electrostatic septum is on the inside of the vacuum chamber, the dispersion required for the Hardt condition is $D_n < 0$ and $D'_n > 0$. In general, in small machines the dispersion is positive. This leaves only the two layouts, 3.15 (b) and (d), for the extraction.

- The position of the 'waiting' beam in the tune diagram determines the last choice between (b) and (d) in Figure 3.15. The PIMMS proposal is to place the 'waiting' beam above the diagonal ($Q_x = Q_z$) and above the resonance $Q_x = 5/3$. The last choice is indicated by the 'heavy' box in Figure 3.15 (d).

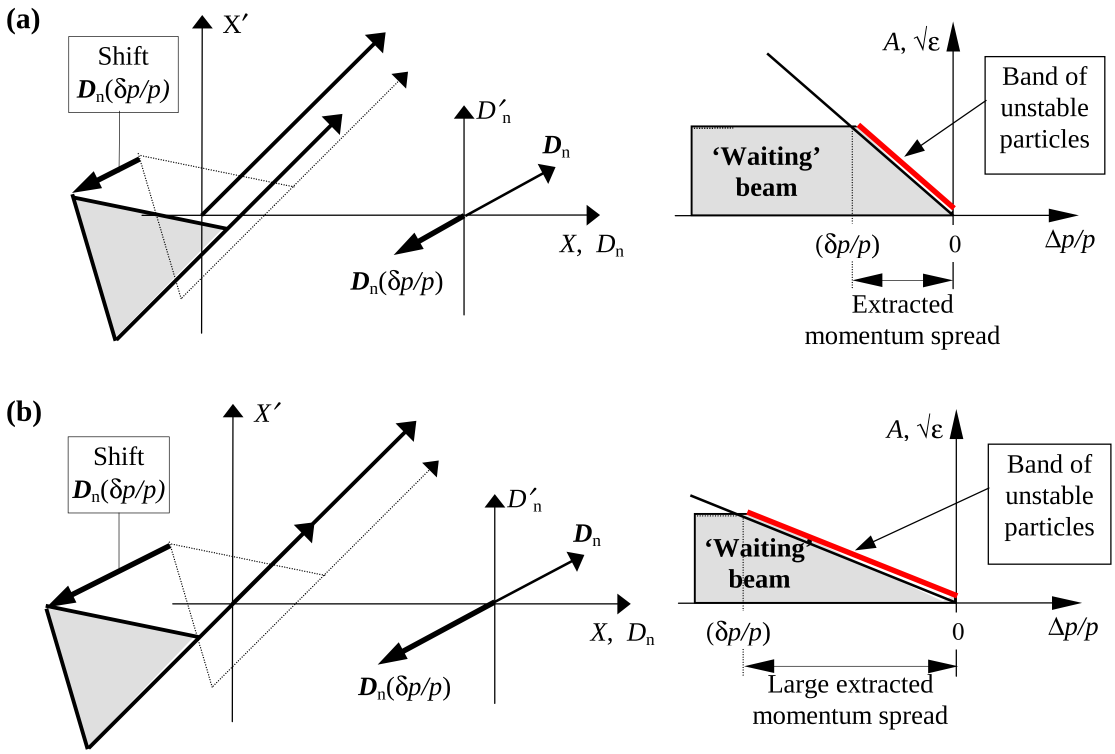

(a)
X′
Shift
Dn(δp/p)
D′n
Dn
X, Dn
Dn(δp/p)
A, √ε
Band of unstable particles
'Waiting' beam
(δp/p)
0
Δp/p
Extracted momentum spread
(b)
X′
Shift
Dn(δp/p)
D′n
Dn
X, Dn
Dn(δp/p)
A, √ε
Band of unstable particles
'Waiting' beam
(δp/p)
0
Δp/p
Large extracted momentum spread

Figure 3.13 Hardt condition for small chromaticity and large momentum spread of extracted beam

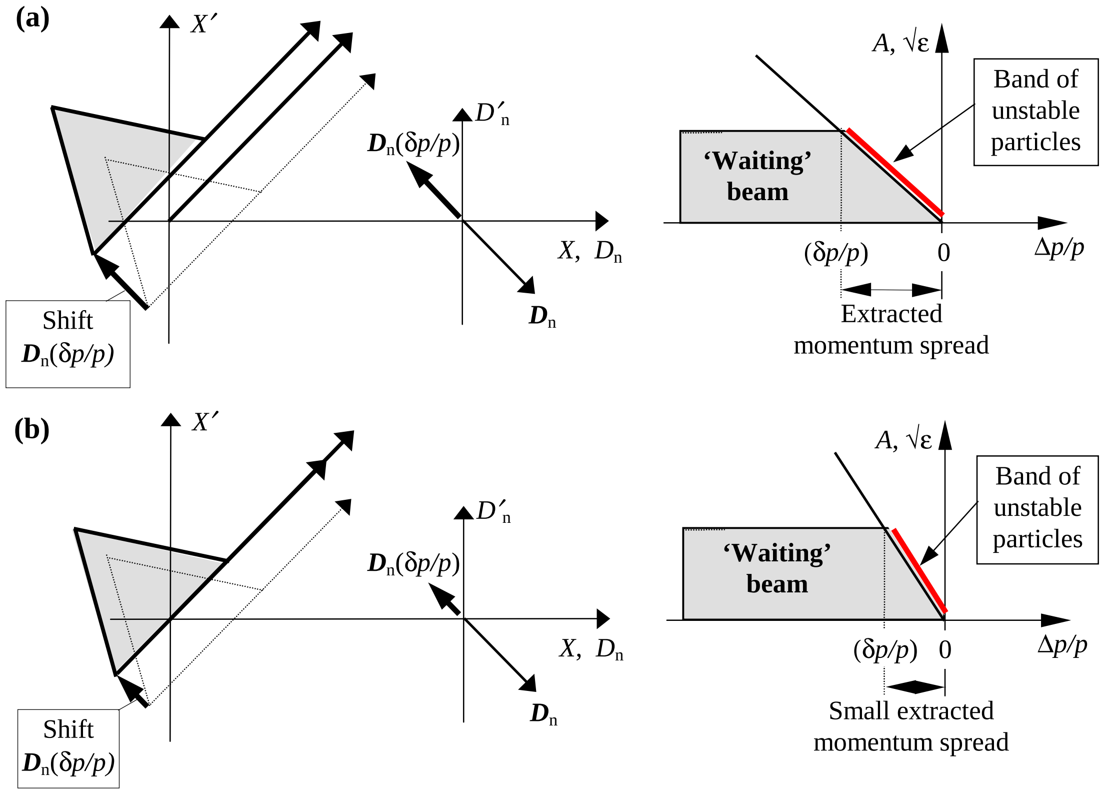

(a)
X′
D′n
Dn(δp/p)
X, Dn
Dn
Shift
Dn(δp/p)
A, √ε
Band of unstable particles
'Waiting' beam
(δp/p)
0
Δp/p
Extracted momentum spread
(b)
X′
D′n
Dn(δp/p)
X, Dn
Dn
Shift
Dn(δp/p)
A, √ε
Band of unstable particles
'Waiting' beam
(δp/p)
0
Δp/p
Small extracted momentum spread

Figure 3.14 Hardt condition for large chromaticity and small momentum spread of extracted beam

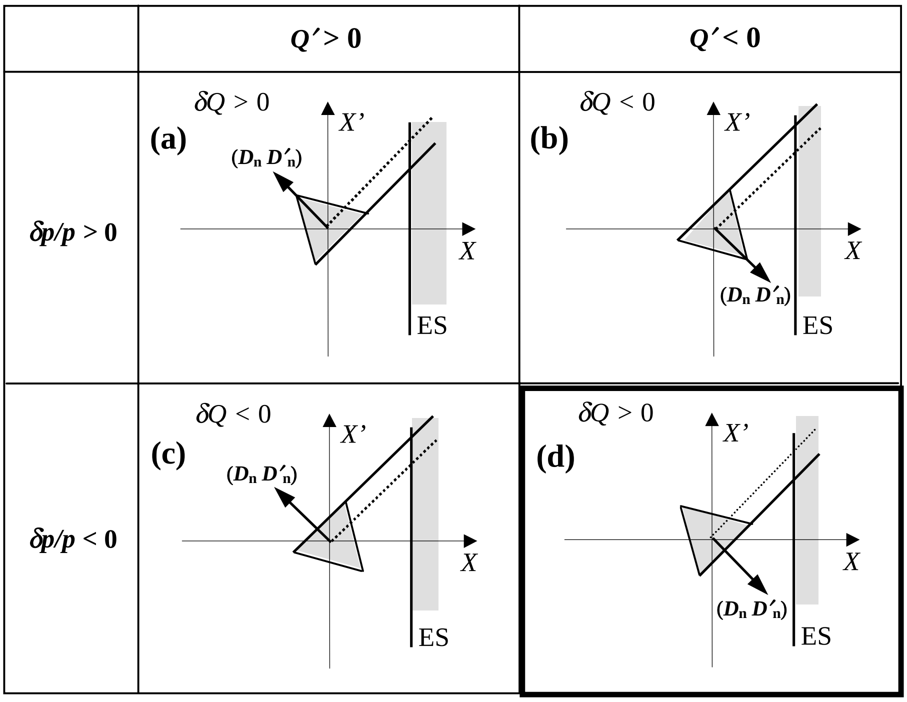

Q′ > 0
Q′ < 0
δp/p > 0
δp/p < 0
(a)
δQ > 0
(b)
δQ < 0
(c)
δQ < 0
(d)
δQ > 0
X’
X
(Dn D′n)
ES

Figure 3.15 Separatrix geometry for first quadrant operation with an ideal normalised dispersion vector

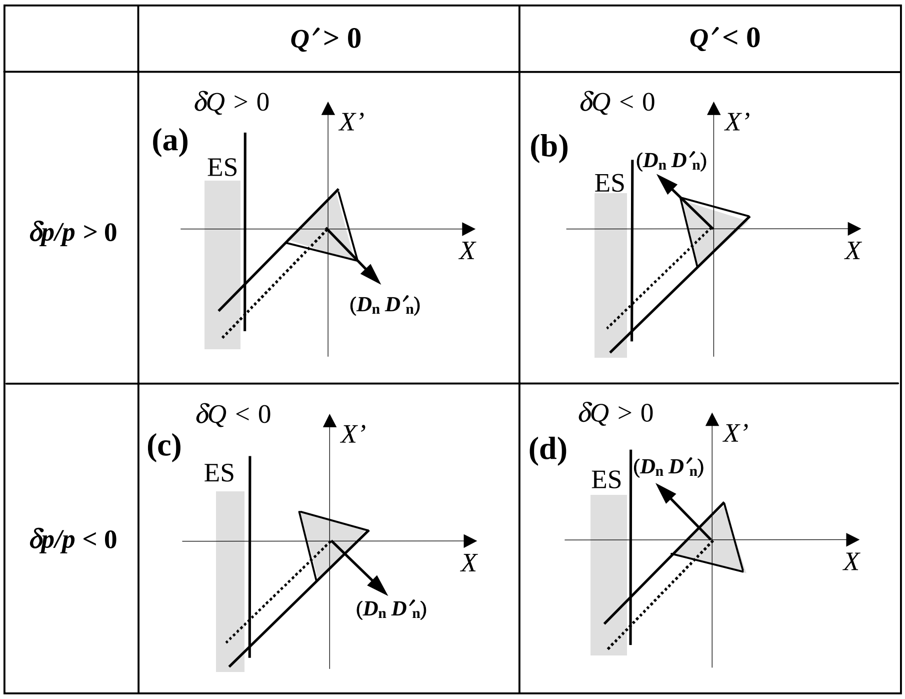

Q′ > 0
Q′ < 0
δp/p > 0
δp/p < 0
(a)
δQ > 0
(b)
δQ < 0
(c)
δQ < 0
(d)
δQ > 0
X’
X
(Dn D′n)
ES

Figure 3.16 Separatrix geometry for third quadrant operation with an ideal normalised dispersion vector

3.8 CHOOSING THE DISPERSION FUNCTION

Figure 3.17 shows the qualitative shape of the normalised dispersion curve in different types of lattices (i.e. a dispersion bump, an arc with regular distributed dipoles and a regular cell structure). In each case, the favoured position for the electrostatic septum is on the downward slope of the dispersion (shaded areas), where $D_n > 0$ and $D'_n < 0$.

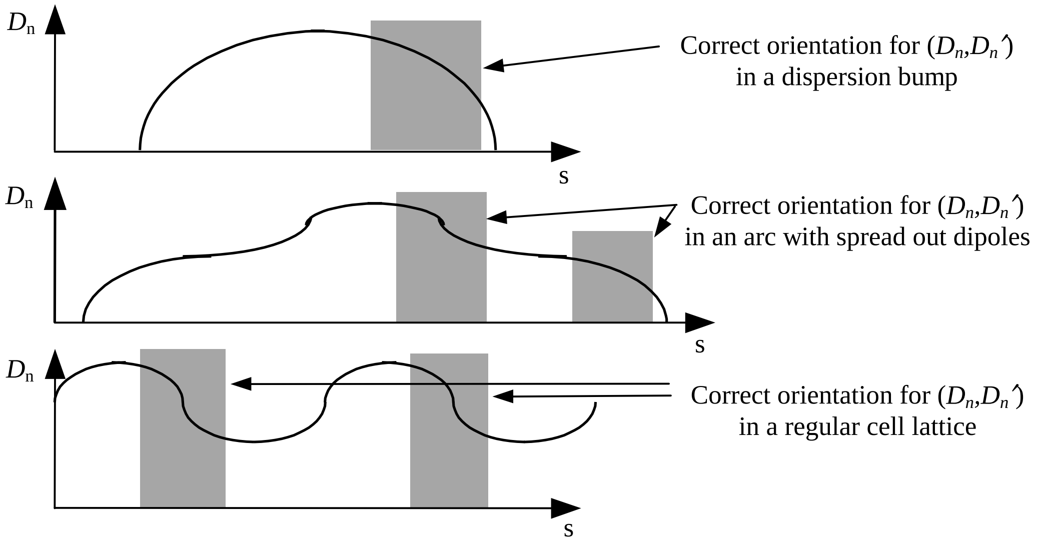

D_n
s
Correct orientation for (D_n,D_n') in a dispersion bump
D_n
s
Correct orientation for (D_n,D_n') in an arc with spread out dipoles
D_n
s
Correct orientation for (D_n,D_n') in a regular cell lattice

Figure 3.17 Qualitative shape of the normalised dispersion for typical lattice structures

3.9 POSITIONING THE ELECTROSTATIC AND MAGNETIC SEPTA

In Figures 3.15 (b) and (d) the electrostatic septum is in the outer half of the chamber. Earlier, in (3.12), it was shown that the ideal phase separation between electrostatic and magnetic septa is $\mu = 90° + n{\cdot}180°$. An even value for n (i.e. $\mu = 90°, 450°$ etc.) puts the two septa on the same side of the vacuum chamber and an odd number of n (i.e. $\mu = 270°, 630°$ etc.) puts the septa are on opposite sides of the vacuum chamber. Figure 3.18 shows these two situations schematically.

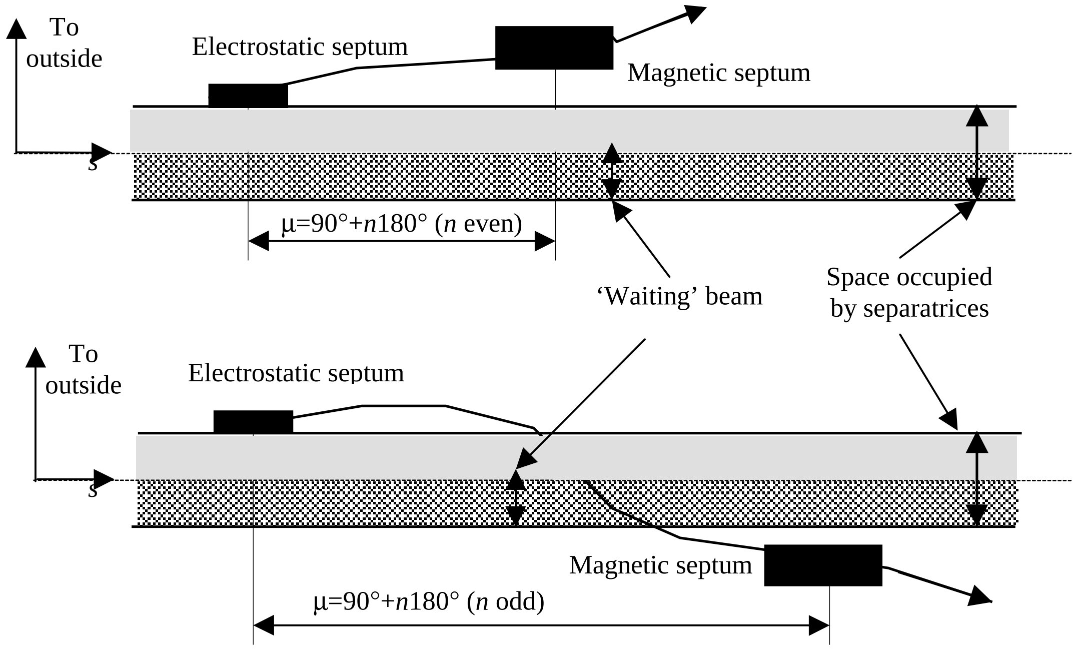

To outside
s
Electrostatic septum
Magnetic septum
μ=90°+n180° (n even)
'Waiting' beam
Space occupied by separatrices
To outside
s
Electrostatic septum
Magnetic septum
μ=90°+n180° (n odd)

Figure 3.18 Positioning the magnetic septum with respect to the electrostatic septum

The area occupied by the beam in Figure 3.18 is shared by the 'waiting' beam and the separatrices. The separatrices grow symmetrically from the resonance so the aperture is best used by placing the resonance at the centre. Intuitively, one feels that the aperture is likely to be restricted for the 'waiting' beam, if the septa are on opposite sides of the chamber. In addition, since the electrostatic septum will be in the outer half of the chamber, putting the magnetic septum in the inner half means that the extraction will be to the inside of the machine. This is undesirable because of the high magnetic rigidity of the beam. For these reasons, a phase separation of around 90° with both septa to the outside of the chamber has been chosen. A larger phase advance of $90°+ n\cdot 360°$ is less convenient as the extracted beam has to be transported for a longer distance inside the machine.

With these considerations, a schematic picture of the aperture starts to emerge (see Figure 3.19). During the extraction set-up, the beam has to be kept sufficiently far from the resonance not to be prematurely extracted. With both septa being on the outside of the vacuum chamber, it is natural to position the beam in the inner half of the chamber to avoid aperture limitations. In this case, the beam is below the resonance in momentum, but due to the negative chromaticity, it is above in tune. At extraction, the resonance is reached by accelerating the beam and reducing its tune. Figure 3.19 indicates two approximate dimensions. The electrostatic septum intercepts the growth of the separatrices at 30 to 40 mm from the axis and the spiral step is about 10 mm. Some justifications for these figures will be given later.

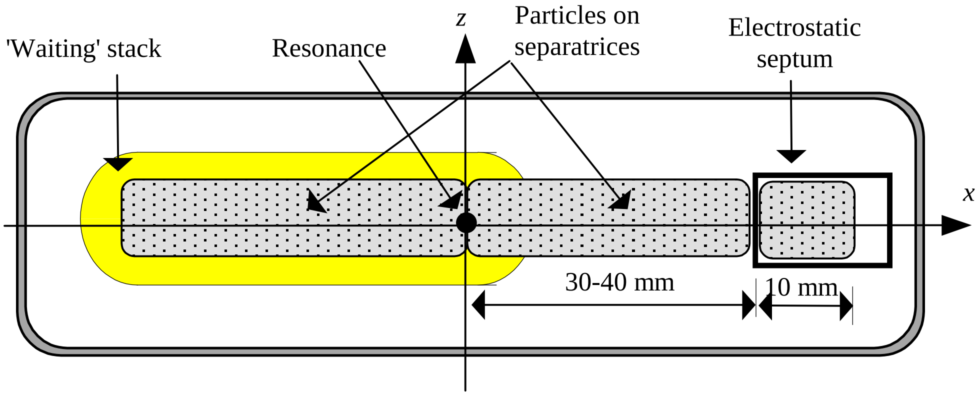


Figure 3.19 Schematic view of aperture

3.10 TRANSFER OF OFF-MOMENTUM PARTICLES BETWEEN SEPTA

If the Hardt Condition is fulfilled, all the extraction separatrices will be superimposed at the electrostatic septum and all particles will reach the septum on the same separatrix independent of their momentum. It would be ideal if this state of affairs could be maintained by an achromatic transfer from the electrostatic to the magnetic septum. The simplest way to achieve this is to avoid all magnetic elements and to place the two septa in the same drift space. Inevitably, the electrostatic septum then needs a bigger kick, because the phase advance between the septa will be smaller. This was already pointed out in Section 3.3.2 when the transfer of on-momentum particles was discussed. Owing to the necessity for a stronger kick in the electrostatic septum, this method is better suited to proton machines than light-ion machines.

The more difficult case where the beam is transported through the section of the machine between the septa will be considered here. If particles of different momenta arrive at the magnetic septum on different orbits, then there will in general be a reduction of the gap calculated in (3.12). This is the gap that has been opened between the main beam and the extracted segment of the separatrix in order to introduce the wall of the magnetic septum.

Consider the transfer of two off-momentum particles C and D that start from the radial position of the electrostatic septum, x_{ES} with the septum angle x'_{ES}, but only particle C receives the kick, φ, of the septum. In a linear lattice, the movement of particles with a momentum deviation can be described by a 3x3 transfer matrix formalism. The horizontal co-ordinates of the particles at the magnetic septum are derived as

Particle C:

$$\begin{aligned} x_{\mathrm{MS}} &= m_{11}\cdot x_{\mathrm{ES}} + m_{12}\cdot x'_{\mathrm{ES}} + m_{12}\cdot\varphi + m_{13}\cdot(\delta p/p) \\ x'_{\mathrm{MS}} &= m_{21}\cdot x_{\mathrm{ES}} + m_{22}\cdot x'_{\mathrm{ES}} + m_{22}\cdot\varphi + m_{23}\cdot(\delta p/p) \end{aligned} \tag{3.20}$$

and

Particle D:

$$\begin{aligned} x_{\mathrm{MS}} &= m_{11}\cdot x_{\mathrm{ES}} + m_{12}\cdot x'_{\mathrm{ES}} + m_{13}\cdot(\delta p/p) \\ x'_{\mathrm{MS}} &= m_{21}\cdot x_{\mathrm{ES}} + m_{22}\cdot x'_{\mathrm{ES}} + m_{23}\cdot(\delta p/p) \end{aligned}. \tag{3.21}$$

.

Comparison of (3.20) and (3.21) with (3.9) and (3.10) shows that the separation between circulating and extracted particles is the same as calculated for on-momentum particles, but the gap appears at a different position and angle. The shift in position reduces the effective gap width for the magnetic septum. Figure 3.20 shows the transfer between the septa for particles representing the full momentum spread of the extracted beam. The grey shaded areas represent particles with momenta in between these limits.

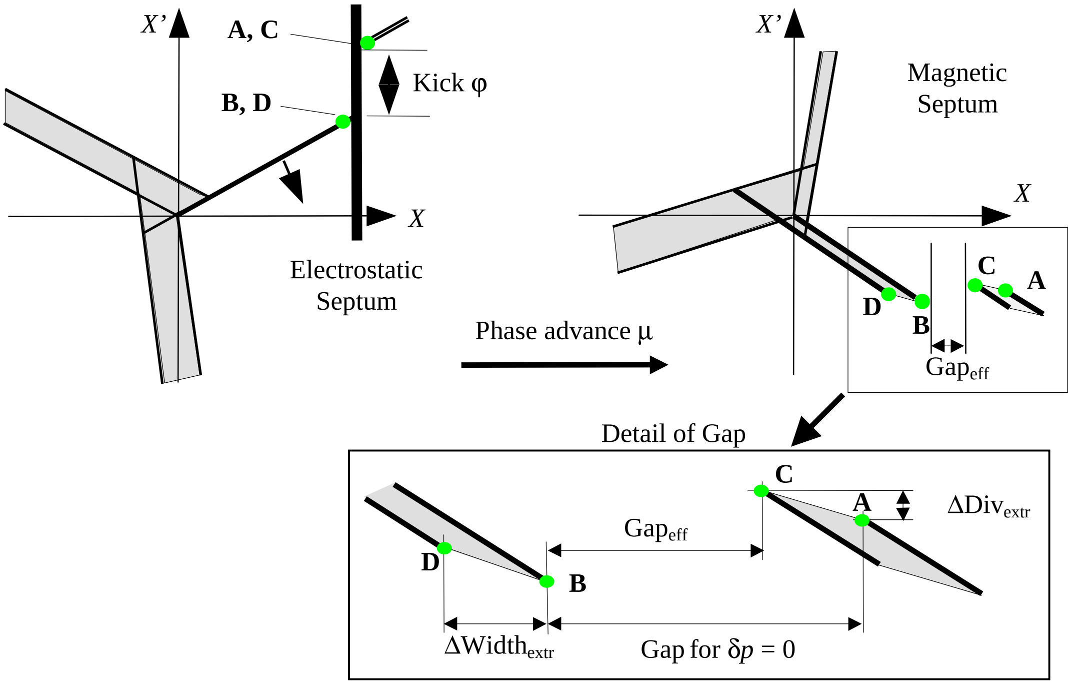


Figure 3.20 Transfer from electrostatic to magnetic septum

A non-zero m_{13} causes a loss of space for the magnetic septum and has to be compensated by a stronger kick from the electrostatic septum. The effective gap width is

$$\mathrm{Gap}_{\mathrm{eff}} = |m_{12} \cdot \varphi| - |m_{13} \cdot (\delta p / p)| . \tag{3.22}$$

The horizontal width of the extracted beam is also increased and, to avoid losses inside the magnetic septum, the horizontal aperture has to be enlarged.

$$\Delta\mathrm{Width}_{\mathrm{extraction}} = |m_{13} \cdot (\delta p / p)| . \tag{3.23}$$

A non-zero m_{23} is leading to a larger overall divergence of the extracted beam at the magnetic septum and also requires an enlarged horizontal aperture to avoid losses.

$$\Delta\mathrm{Divergence}_{\mathrm{extraction}} = |m_{23} \cdot (\delta p / p)| \tag{3.24}$$

At the electrostatic septum, any angle error of the extraction separatrices, however small, will directly increase particle losses, but at the magnetic septum, it is usual to foresee a small clearance of a few millimetres between beam and septum and therefore angular spreads up to 1mrad (approx.) can be tolerated without loss. For this reason, only the m_{13} term will be considered in the following discussion.

3.11 MINIMISATION OF CHROMATIC EFFECTS

By keeping the momentum spread of the extracted beam small, the above mentioned disadvantages can be minimised. However, the momentum spread is not a free parameter if the Hardt-condition is applied to avoid losses at the electrostatic septum. It is therefore important to optimise the lattice in such a way that the superposition of the extraction separatrices results in a small momentum spread of the extracted beam (as already discussed in Section 3.6). The explicit form of the element m_{13}, expressed in terms of normalised dispersion,

$$m_{13} = \sqrt{\beta_{\mathrm{MS}}} \cdot (D_{\mathrm{nMS}} - D_{\mathrm{nES}} \cdot \cos\mu - D'_{\mathrm{nES}} \cdot \sin\mu), \tag{3.25}$$

shows that the loss of space for the magnetic septum is proportional to $\sqrt{\beta_{\mathrm{MS}}}$. Although decreasing β_{MS} does reduce the influence of m_{13}, it also reduces the gap created by the electrostatic septum (3.12), since this is also proportional to $\sqrt{\beta_{\mathrm{MS}}}$. As a result, the effective gap (3.22) for the magnetic septum is in fact reduced by decreasing β_{MS}.

$$\begin{aligned}\mathrm{Gap}_{\mathrm{eff}} &= |m_{12}\varphi| - \left| m_{13} \frac{\delta p}{p} \right| \\ &= \sqrt{\beta_{\mathrm{MS}}} \left[\left|\sqrt{\beta_{\mathrm{ES}}} \sin\mu\right| - \left| (D_{n,\mathrm{MS}} - D_{n,\mathrm{ES}} \cos\mu - D'_{n,\mathrm{ES}} \sin\mu) \frac{\delta p}{p} \right| \right]\end{aligned} \tag{3.26}$$

The only effective way therefore to reduce chromatic effects is to minimise m_{13} directly.

3.12 ACHROMATIC TRANSFER BETWEEN EXTRACTION SEPTA

A momentum independent transfer between electrostatic and magnetic septa requires the chromatic term m_{13} in the general 3x3 transfer matrix to vanish. Inspection of (3.25) shows that the size of m_{13} is determined by the normalised dispersion functions at the septa. In order to reduce m_{13}, the lattice needs to provide suitable values of the dispersion at the two positions of the septa.

The shape of the dispersion function in a lattice is determined by the distribution of the dipole magnets. In the analysis below, the effect of the dipoles is approximated as point kicks. In this model, every passage of a dipole adds a kick $D'_{n,0}$ to the actual value of the normalised dispersion function, according to

$$\underbrace{\begin{pmatrix} D_n \\ D'_n \end{pmatrix}}_{\substack{\text{Entrance}\\ \text{to dipole}}} + \underbrace{\begin{pmatrix} 0 \\ D'_{n,0} \end{pmatrix}}_{\substack{\text{Point kick}\\ \text{of dipole}}} = \underbrace{\begin{pmatrix} D_n \\ D'_n + D'_{n,0} \end{pmatrix}}_{\substack{\text{Exit of}\\ \text{dipole}}}. \tag{3.27}$$

In a bending-free region, the dispersion function acts like a betatron oscillation of a particle and can be described with the Courant and Snyder matrix formalism. The transformation of the dispersion function between two lattice elements, denoted by the suffices 1 and 2, without crossing dipoles, is given by

$$\begin{pmatrix} D \\ D' \end{pmatrix}_2 = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \cdot \begin{pmatrix} D \\ D' \end{pmatrix}_1 . \tag{3.28}$$

In normalised phase space, the normalised dispersion vector simply rotates by the phase advance, μ_{12}, between the elements, according to

$$\begin{pmatrix} D_n \\ D'_n \end{pmatrix}_2 = \begin{pmatrix} \cos\mu_{12} & \sin\mu_{12} \\ -\sin\mu_{12} & \cos\mu_{12} \end{pmatrix} \cdot \begin{pmatrix} D_n \\ D_n' \end{pmatrix}_1 . \tag{3.29}$$

The combination of (3.27) and (3.29) allows an approximate analysis of the transfer between the extraction septa. Some particular cases are demonstrated below.

Both septa in a bending-free dispersion region

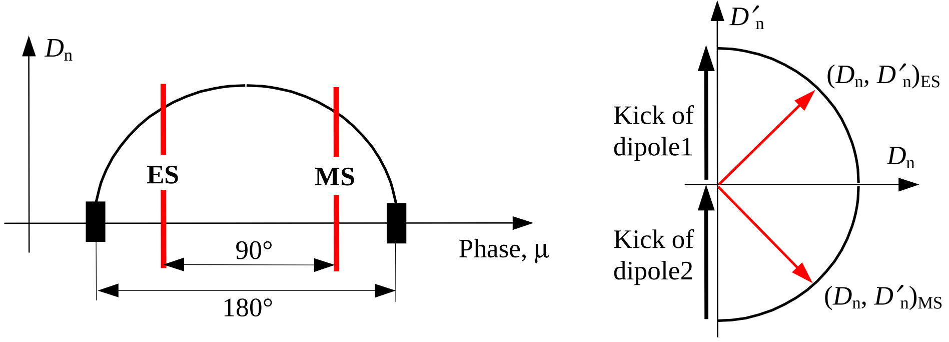


Figure 3.21 180°dispersion bump
(a) Normalised co-ordinates (μ, D_n), (b) normalised phase space (D_n, D'_n).

Figure 3.21 shows a dispersion bump as created by two dipole point kicks of identical strength, spaced by 180° in phase. The electrostatic septum is positioned in the first half of the bump and the magnetic septum 90° in the second half. To prove that in such a case the transfer is always achromatic, the closed orbit for particles with a momentum deviation δp is considered. The closed orbit at any position s in the machine is, to first order, defined by the product of dispersion function and momentum deviation,

$$x(s,\delta p) = D(s)\cdot(\delta p/p) \quad \text{and} \quad x'(s,\delta p) = D'(s)\cdot(\delta p/p). \tag{3.30}$$

A 3x3 transfer matrix links the orbit co-ordinates between two positions 1 and 2 in the machine according to

$$\begin{pmatrix} x_2(\delta p) \\ x_2'(\delta p) \\ \delta p/p \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1(\delta p) \\ x_1'(\delta p) \\ \delta p/p \end{pmatrix} \tag{3.31}$$

The insertion of (3.30) into (3.31), provides general expressions for m_{13} and m_{23}, in terms of the dispersion functions and the 2x2 matrix elements,

$$m_{13} = D_2 - m_{11}\cdot D_1 - m_{12}\cdot D'_1 \qquad \text{and} \qquad m_{23} = D'_2 - m_{21}\cdot D_1 - m_{22}\cdot D'_1\,. \tag{3.32}$$

In regions without bending, the dispersion function transforms according to (3.28) and, by inserting this into (3.32), it follows directly that m_{13} and m_{23} are zero and therefore:

- The transfer via a dispersion region without crossing bending magnets is always achromatic with respect to position and angle.

It should be noted that this result is exact and valid, not only inside a dispersion bump, but in any lattice section without bending.

A typical structure where this result can be applied is the so called ‘square’ lattice in which the dipole magnets are grouped in the corners, to create dispersion bumps on two opposite sides and dispersion-free straight sections on the remaining two sides. With the above principle that allows fully achromatic transfer between the septa in the straight sections with dispersion, it may appear that such a structure is ideally suited for slow extraction, but there are two major problems with a square lattice:

- The natural position for the electrostatic septum would be the first half of the dispersion bump, but this is the ‘worst’ place to put it when adjusting for the Hardt condition. It was shown in Section 3.8, that for ideal operation $D_n > 0$ and $D'_n < 0$, is required and this only occurs in the second half of the bump (see Figure, 3.17).

- The second problem is that in general there is too little useable phase advance inside the dispersion straight-section. An important fraction of the 180° is lost because of the finite length of the corner dipoles. To get the maximum phase separation between the septa, they have to be positioned close to the corners on either side of the straight section, but then the magnetic septum has to kick strongly to clear the dipole magnets of the downstream corner. If the magnetic septum is moved upstream then the electrostatic septum must kick harder.

Both septa in regions with dispersion and bending

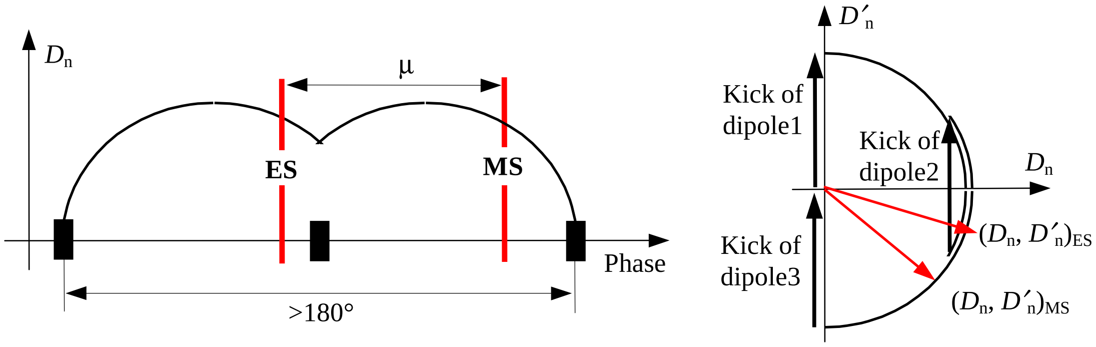


Figure 3.22 Extended dispersion bump
(a) normalised co-ordinates (μ, D_n), (b) normalised phase space (D_n,D'_n)

Figure 3.22 shows a so called *extended-dispersion-bump*. The free space within the dispersion bump is increased by adding a central dipole in order to overcome the problems of the 'square' lattice. The disadvantage of such a structure is that, in general, the transfer between the extraction septa is chromatic, as a bending magnet has to be crossed.

Some design guidelines can be found for an extraction layout with optimised phase advances of μ=90°+n·180° between the septa. Using expression (3.25) for the chromatic term,

$$m_{13} = \sqrt{\beta_{MS}} \cdot \left(D_{n,MS} - D_{n,ES} \cdot \cos\mu - D'_{n,ES} \cdot \sin\mu\right), \qquad (3.25)$$

and assuming a phase advance $\mu = 90° + n·360°$ between the electrostatic and magnetic septa gives

$$m_{13} = \sqrt{\beta_{MS}} \cdot \left(D_{n,MS} - D'_{n,ES}\right). \qquad (3.33)$$

To make the transfer achromatic (i.e. $m_{13} = 0$),

$$D_{n,MS} = D'_{n,ES} \qquad (3.34)$$

is required, but as shown earlier, one needs to work with a negative $D'_{n,ES}$ for adjusting the Hardt condition and therefore m_{13} can only be made zero by having negative dispersion at the magnetic septum.

For a phase advance of μ = 270° + n·360° (septa on opposite sides of the vacuum chamber) it follows that

$$m_{13} = \sqrt{\beta_{MS}} \cdot \left(D_{n,MS} + D'_{n,ES}\right) \qquad (3.35)$$

and therefore to make the transfer achromatic requires,

$$D_{\mathrm{n,MS}} = -D'_{\mathrm{n,ES}} \,. \qquad (3.36)$$

In this case, m_{13} can be made zero by having a positive $D_{\mathrm{n,MS}}$ and a negative $D'_{\mathrm{n,ES}}$ just as required for the Hardt Condition. A disadvantage of this solution might be that the particles which are extracted have to be transported for a longer distance inside the machine (e.g. crossing of sextupoles between the septa would be more difficult to avoid)[*].

Electrostatic septum in dispersion region and magnetic septum in zero-dispersion region

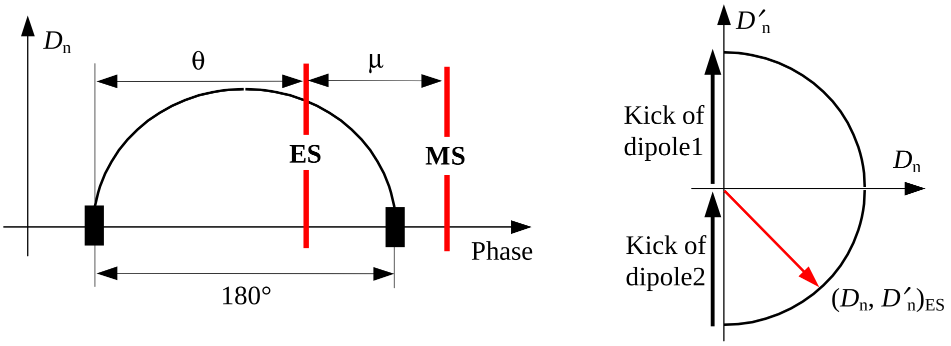


Figure 3.23 180°dispersion bump
(a) Normalised co-ordinates (μ, D_{n}), (b) normalised phase space (D_{n}, D'_{n}).

Figure 3.23 shows a 180° dispersion bump, but contrary to the extraction layout in Figure 3.21, the electrostatic septum is positioned in the second half of the bump where $D_{\mathrm{n}} > 0$ and $D'_{\mathrm{n}} < 0$, as required for the Hardt condition and the magnetic septum is positioned in the dispersion-free straight section after the dipoles that close the bump. As for the extended dispersion bump, the transfer between the septa is in general chromatic due to the passage through dipoles. Assuming that the dispersion bump is created by dipole point kicks, the normalised dispersion function can be described with (3.27) and (3.29) as

$$D_n(\vartheta) = D'_{n,0} \cdot \sin\vartheta \quad \text{and} \quad D'_n(\vartheta) = D'_{n,0} \cdot \cos\vartheta \,, \qquad (3.37)$$

where ϑ is the phase advance counted from the first dipole kick and $D'_{\mathrm{n},0}$ is the strength of the kick. With $D_{\mathrm{n,MS}} = 0$ in (3.25), the chromatic term m_{13} can be written as

$$m_{13} = -\sqrt{\beta_{\mathrm{MS}}} \cdot \left(D_{\mathrm{n,ES}} \cdot \cos\mu + D'_{\mathrm{n,ES}} \cdot \sin\mu\right), \qquad (3.38)$$

and by inserting (3.37) into (3.38) a simple expression for m_{13} is obtained,

---

[*] If a sextupole is crossed (either resonance or chromaticity) between the ES and the MS, then there is a variable optical element in the extraction channel. Any change in the Q' or resonance-strength alters the extraction geometry

$$m_{13} = -\sqrt{\beta_{\mathrm{MS}}} \cdot D_{\mathrm{n0,ES}} \cdot \sin(\vartheta + \mu). \qquad (3.39)$$

It follows immediately that the transfer between the septa will be achromatic, ($m_{13} = 0$), for

$$(\vartheta + \mu) = n \cdot 180^0. \qquad (3.40)$$

Obviously it is impossible to use exactly $n = 1$, since this gives the position of the dipole which is closing the dispersion bump. To keep m_{13} small, the magnetic septum has to be positioned as close to the dipole as possible. For larger n, there is again the problem of transporting the extracted part of the beam through a large distance in the machine.

Transfer for un-fulfilled Hardt Condition

Adjusting the Hardt Condition fixes the chromaticity and therefore the slope of the resonance line (3.14) and the momentum spread of the extracted beam (3.15). Alternatively, instead of superimposing the extraction separatrices at the electrostatic septum, the chromaticity can be used to superimpose the orbits of the different momenta at the magnetic septum. With a proper choice of starting conditions, the available space for the magnetic septum can be maximised even if the transfer between the septa is chromatic but, of course, the separatrices then arrive at the electrostatic septum with different angles. This method is used in the present CERN-PS slow extraction scheme [5].

Transfers between zero-dispersion regions

By inspection of the general expression (3.25) for m_{13}, it follows immediately that the transfer between zero-dispersion regions is always fully achromatic. However, with the electrostatic septum located in a zero-dispersion region, the separatrices can be superimposed only for zero chromaticity, which will be discussed in the next Section.

3.13 ZERO CHROMATICITY AND THE HARDT CONDITION

In Section 3.2, the extraction method of *Amplitude selection by amplitude growth* was described as having near-zero chromaticity, in Section 3.6 it was noted that a mathematically trivial solution for the Hardt Condition (3.17) was zero dispersion and zero chromaticity and finally, in Section 3.12, it was pointed out that an extraction made between two dispersion-free regions, or within a single dispersion-free region, would be achromatic for the transfer between the septa. The combination of these three points leads to a complete scenario for a zero-chromaticity extraction, with the Hardt Condition fulfilled and achromatic transfer between the septa.

References

[1] K. Hiramoto, M. Nishi, *Resonant beam extraction scheme with constant separatrix*, Nucl. Instr. and Meth. in Phys. Res.. A3320, (1992), p154.

[2] N. Araki, T. Kanai, M. Kanazawa, M. Kumada, S. Minohara ,K. Noda, H Ogawa, S. Sato, E. Takada, H. Tomura, M. Torikoshi,, S. Yamada, *Performance of a respiration-gated beam control system for patient treatment*, EPAC96, Barcelona, (Institute of Physics Publishing Bristol and Philadelphia), pp.2656-2658.

[3] G. Cesari, P. Lefèvre, D. Vandeplassche, *Feasibility study of a synchrotron for the european light ion medical accelerator*, CERN/PS/91-08 (DI), (April 1991).

[4] W. Schnell, B. Zotter, *A simplified criterion for transverse stability of a coasting beam and its application to the ISR*, CERN, CERN-ISR-GS-RF/76-26, (1976).

[5] C. Steinbach, *Beam optics at resonant extraction septa*, EPAC 1992, Berlin, (Editions Frontières), (1992), p857-9.

* * *

I-4 SPILL CHARACTERISTICS

A detailed knowledge of the extracted beam (or spill) is essential in order to obtain the quality that is required for high-precision active scanning. The momentum spread and emittance of the spill must be well understood to ensure a reproducible spot size and to correctly adapt the focusing and apertures in the transfer lines and gantries under all conditions. The transverse distribution of particles within the spill affects the dose calculations and the way in which the beam spot is scanned. The *transit time*, or *storage time*, in the resonance is the principal problem for feedback systems, but it also has the mitigating effect of smoothing the spill at high frequencies. The uniformity of the spill is probably the most important aspect of all and analysing the deleterious effect of tune ripple is a major aim of the present chapter. By first studying the behaviour of a single particle and then the *strip* of particles sitting along the edge of the stable triangle in phase space and finally the ‘band’ of particles that comprise all the strips of different momenta that become unstable at any one time, it has been possible to build up simulations of the micro-time profile of the spill under the influence of tune ripple [1,2]. These simulations have been made and compared for the different extraction techniques.

4.1 TRANSLATION OF THE HAMILTONIAN

In order to study the motion of a single particle as it leaves the machine under the influence of the resonance, it is convenient to shift the origin of the Kobayshi Hamiltonian (2.27) to the fixed point used for extraction.

Shifted Hamiltonian: $$\overline{\boldsymbol{H}} = \frac{S}{4}\left(4h^3 + 6h\overline{X}^{\,2} + 6\sqrt{3}h\overline{X}\,\overline{X}' + 3\overline{X}\,\overline{X}'^{\,2} - \overline{X}^{\,3}\right) \tag{4.1}$$

where $$h = \frac{2}{3}\frac{\varepsilon}{S} = \frac{4\pi}{S}\delta Q \tag{2.29}$$

and the ‘¯‘ indicates the variable is referred to new origin. However, to simplify the notation, the convention of adding the bar will be dropped and X' will be replaced by Y from now on. This situation is drawn in Figure 4.1.

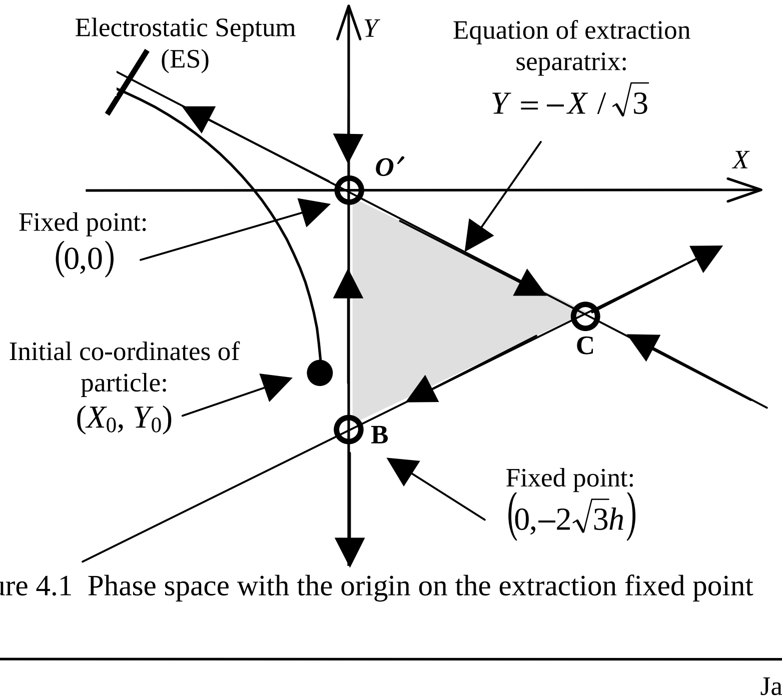


Figure 4.1 Phase space with the origin on the extraction fixed point

In the translated frame of reference, the equations of motion are:

$$\frac{dX}{dt} = \frac{\partial H}{\partial Y} = \frac{S}{4}\left(6\sqrt{3}hX + 6XY\right) \qquad (a)$$

$$\frac{dY}{dt} = -\frac{\partial H}{\partial X} = -\frac{S}{4}\left(12hX + 6\sqrt{3}hY + 3Y^2 - 3X^2\right) \qquad (b) \tag{4.2}$$

Time is dimensionless in (4.2) and is measured as a number of sets of three turns. This is a heritage from the derivation of the Kobayashi Hamiltonian that was derived by considering the displacement that a particle undergoes when viewed every third revolution. Thus, the elementary time d*t* corresponds to 3 T_{REV}, but it is dimensionless. Thus, if an extraction time is found to be 100, this means that the particle being considered needs 300 revolutions in the accelerator to reach the electrostatic septum and be extracted.

The time scale of the slow extraction process is such that the extracted particles will follow paths that are close to the separatrices and it is convenient to express the co-ordinates of these trajectories in terms of the apothem of the stable triangle, *h*, such that

$$X_0 = -\lambda_0 h \quad \text{and} \quad Y_0 = -\Lambda_0 Y_B = -2\sqrt{3}h\Lambda_0, \tag{4.3}$$

which effectively define the new quantities λ and Λ. Thus, the motion in *X* is measured in units of *h* and the motion in *Y* is measured in units of the length of the side of the stable triangle.

The strategy for finding the transit times is to first analyse the motion of particles travelling along the outgoing separatrix from O′ to the electrostatic septum under static conditions i.e. while the stable triangle remains constant in size and position. The next step is to find an expression for the travel time of particles moving close to the side of the stable triangle from an arbitrary point towards the stable fixed point O′, again under stable conditions. Finally, these expressions must be modified to take into account the dynamic conditions of a shrinking stable triangle. An important stage in this analysis is the definition of the 'hand-over' point between the trajectory associated with the side of the stable triangle and the trajectory associated with the outgoing separatrix. For convenience, these separate components of the travel times will be defined as follows:

$$T_{\text{static}} = T_{\text{corner, static}} + T_{\text{side, static}} \quad \text{and} \quad T_{\text{dynamic}} = T_{\text{corner, dynamic}} + T_{\text{side, dynamic}}$$

To help the understanding of some of the approximations made in the following, some tentative values of the main quantities are given:

Revolution time, $T_{\mathrm{REV}} = 0.5$ μs and Spill length = 0.5 s = $10^6 \times T_{\mathrm{REV}}$.
Tune separation from the resonance, $\delta Q = 3\times10^{-3}$ and $\varepsilon = 6\pi\delta Q = 5.65\times10^{-2}$.
$S = 36.7$ m$^{-1/2}$ and $h = (2/3)(\varepsilon/S) = 1.03\ 10^{-3}$ m$^{1/2}$.
$\lambda = 15\times10^{-6}$.
$\delta Q_{\mathrm{beam}} = 15\times10^{-3}$ and $dQ/dt = \delta Q_{\mathrm{beam}}/(10^6/3) = 45\times10^{-9}$

4.2 TRANSIT TIMES UNDER STATIC CONDITIONS

4.2.1 Transit time from the fixed point to the electrostatic septum

While δQ, S and $\Delta p/p$ are constant, the particle will follow a trajectory with constant, H. This trajectory will be given directly by (4.1):

$$\begin{aligned}&\frac{S}{4}(4h^3+6hX_0^2+6\sqrt{3}hX_0Y_0+3X_0Y_0^2-X_0^3)\\&=\frac{S}{4}(4h^3+6hX^2+6\sqrt{3}hXY+3XY^2-X^3)\end{aligned}. \tag{4.4}$$

In the vicinity of the fixed point O′, that is for $|X|$, $|Y| << h$, the third-order terms in X and Y can be neglected. Thus:

$$Y=\frac{X_0^2+\sqrt{3}X_0Y_0-X^2}{\sqrt{3}X}. \tag{4.5}$$

The substitution of (4.5) into the motion equation (4.2) (a), yields

$$\frac{\mathrm{d}X}{\mathrm{d}t}=\frac{S}{4}\left(6\sqrt{3}hX+\frac{6}{\sqrt{3}}X_0^2+6X_0Y_0-\frac{6}{\sqrt{3}}X^2\right). \tag{4.6}$$

Fortunately, this equation has a standard form* and can be integrated. Within the strict assumptions made above, the simplified trajectory equation (4.5) is only valid close to O′, but, since the particle approaches the separatrix asymptotically and $Y\Rightarrow-\frac{1}{\sqrt{3}}X$, the third-order terms in the Hamiltonian cancel out, so luckily they can also be neglected far from O′ along the outgoing separatrix and the integration can be extended right up to the electrostatic septum. Thus, the transit time $T_{\text{corner,static}}$ is

$$\begin{aligned}T_{\text{corner, static}}&=\int_{t_0}^{t_0+T_{\text{corner,static}}}\mathrm{d}t=\frac{2}{\sqrt{3}S}\int_{X_0}^{X_{\text{ES}}}\frac{1}{-X^2+3hX+X_0^2+\sqrt{3}X_0Y_0}\mathrm{d}X\\&\approx\frac{2}{3\sqrt{3}Sh}\left[\ln\left|\frac{-X-\dfrac{X_0^2+\sqrt{3}X_0Y_0}{3h}}{-X+3h+\dfrac{X_0^2+\sqrt{3}X_0Y_0}{3h}}\right|\right]_{X_0}^{X_{\text{ES}}}\end{aligned}. \tag{4.7}$$

with expansion of square roots to 1st order.

---

* $\int\frac{\mathrm{d}x}{ax^2+bx+c}=\frac{1}{\sqrt{b^2-4ac}}\ln\left|\frac{2ax+b-\sqrt{b^2-4ac}}{2ax+b+\sqrt{b^2-4ac}}\right|$

It is useful to express all the distances in terms of h. In correspondence with the definitions of λ and Λ in (4.3), the position of the electrostatic septum is written as $X_{ES} = -nh$, and the starting positions are written as $X_0 = -\lambda_0 h$ and $Y_0 = 2\sqrt{3}h\Lambda_0$. Both X_0 and Y_0 will be small and close to the stable fixed point. By neglecting the terms in λ_0 with respect to unity and $\lambda_0\Lambda_0$ with respect to λ_0 the final result obtained is,

Transit time from 'corner': (static conditions)

$$T_{\text{corner, static}} \approx \frac{1}{\varepsilon\sqrt{3}} \ln\left|\left(\frac{n}{n+3}\right)\frac{3}{\lambda_0}\right| \tag{4.8}$$

This formula works well for particles at positions up to $\Lambda_0 = 0.1$ and within this range the extraction time is independent of the initial Y_0 co-ordinate.

4.2.2 Extraction time from the side of the stable triangle

To evaluate the time needed to reach the electrostatic septum for a particle that has just become unstable, it is necessary to add the time spent moving along the side of the stable triangle to reach the fixed point O′ before moving along the outgoing separatrix. The motion in Y between a starting position close to the side of the triangle (so that X and X^2 can be neglected) and a point that is 'near' to O′ can be obtained from (4.2) (b).

$$\frac{\mathrm{d}Y}{\mathrm{d}t} \approx -\frac{S}{4}\left(6\sqrt{3}hY + 3Y^2\right). \tag{4.9}$$

Equation (4.9) has the same basic form as (4.7),

$$T_{\text{side, static}} = \int_{t_0}^{t_0+T_{\text{side, static}}} \mathrm{d}t = -\frac{4}{S}\int_{Y_0}^{Y_F} \frac{1}{\left(6\sqrt{3}hY + 3Y^2\right)}\mathrm{d}Y$$

which yields,

Transit time from 'side': (static conditions)

$$T_{\text{side, static}} \approx \frac{1}{\varepsilon\sqrt{3}} \ln\left|\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{1-\Lambda_F}{\Lambda_F}\right)\right| \tag{4.10}$$

where the integration has been made between the starting point $Y_0 = -2\sqrt{3}\Lambda_0 h$ and $Y_F = -2\sqrt{3}\Lambda_F h$[*] and the point defined by $-\Lambda_F$ has to be 'near enough to O' to allow the use of $T_{\text{corner,static}}$' (derived in the previous sub-section).

4.2.3 Matching the 'side' and 'corner' trajectories

In order to complete the present calculation, it is necessary to define a 'hand-over' point between the trajectory coming from the side of the stable triangle and the trajectory that extends out to the electrostatic septum. The general expression for the end-point co-ordinates of the first trajectory is obtained by equating the Hamiltonian at the start $(-\lambda_0 h, -2\sqrt{3}\Lambda_0 h)$ and the start and end of the trajectory.

---

[*] Note that in Ref. 1 the parameter $A_F = 2\sqrt{3}\Lambda_F$ is used to define the end point of the track.

$$\frac{S}{4}(4h^3 + 6hX_F^2 + 6\sqrt{3}hX_FY_F + 3X_FY_F{}^2 - X_F^3) =$$
$$= \frac{S}{4}\left[4h^3 + 6h(-\lambda_0 h)^2 + 6\sqrt{3}h(-\lambda_0 h)\left(-2\sqrt{3}\Lambda_0 h\right) + 3(-\lambda_0 h)\left(-2\sqrt{3}\Lambda_0 h\right)^2 - (-\lambda_0 h)^3\right]$$

By rewriting $X_F = \lambda_F h$ and $Y_F = -2\sqrt{3}\Lambda_F h$ and by neglecting the higher order terms in λ_0 and $\lambda_0\Lambda_0$, a relation is obtained between the starting co-ordinates and the finishing co-ordinates,

$$\Lambda_0(1-\Lambda_0)\lambda_0 \approx \lambda_F\Lambda_F\ . \tag{4.11}$$

At this stage, an expression for the X-position of the hand-over point, X_F, is noted for future use.

$$X_F = \frac{(1-\Lambda_0)}{\Lambda_F(1-\Lambda_F)}\lambda_0 h = \lambda_F h \tag{4.12}$$

The addition of (4.8) and (4.10) yields,

$$T_{\text{static}} = T_{\text{side, static}} + T_{\text{corner, static}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{1-\Lambda_F}{\Lambda_F}\right)\left(\frac{n}{n+3}\right)\frac{3}{\lambda_F}\right|$$
$$\approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{1}{\Lambda_F}\right)\left(\frac{n}{n+3}\right)\frac{3}{\lambda_F}\right|$$

And finally, with the introduction of (4.11)

Transit time : (static conditions)

$$T_{\text{static}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\underbrace{\frac{1}{(1-\Lambda_0)^2}}_{\substack{\text{Position on}\\ \text{side of}\\ \text{triangle}}}\ \underbrace{\left(\frac{n}{n+3}\right)}_{\substack{\text{Position of}\\ \text{electrostatic}\\ \text{septum}}}\ \underbrace{\frac{3}{\lambda_0}}_{\substack{\text{Distance}\\ \text{to stable}\\ \text{triangle}}}\right| \tag{4.13}$$

which no longer depends on Λ_F or λ_F.

4.3 TRANSIT TIMES UNDER DYNAMIC CONDITIONS

Resonant slow extraction can be 'activated' in many ways (see Section 3.2.3), but the easiest to understand is the variation of the tune Q of the machine, by changing the focusing quadrupoles, to bring the beam into resonance. An equivalent method, in terms of the present analysis, is to vary the momentum of the particles, as $\delta Q = Q' \Delta p/p$. Varying δQ, by either method, causes the stable triangle to shrink. The size of the triangle is directly proportional to the apothem h, which is proportional to the tune shift via $\varepsilon = 6\pi\delta Q$. Reducing the size of the triangle causes the particles on the largest orbits to pass from the stable region inside the triangle to the unstable region outside and thence to be extracted. A linear variation of the tune with time ($\dot{Q}$ = constant) will be considered. The main approximations used are the following:

- The relative variation in the size of the triangle during the extraction time will be small, that is $\Delta h \ll h$;

- Instead of considering the movement of the separatrices as the triangle shrinks, the relative movement of the separatrix with respect to the particle, will be considered as an additional contribution to the particle velocity and the triangle will be considered as fixed during the extraction time. This is illustrated in Figure 4.2, where the situation is sketched at two different times t_1 and t_2 with $t_2 > t_1$.

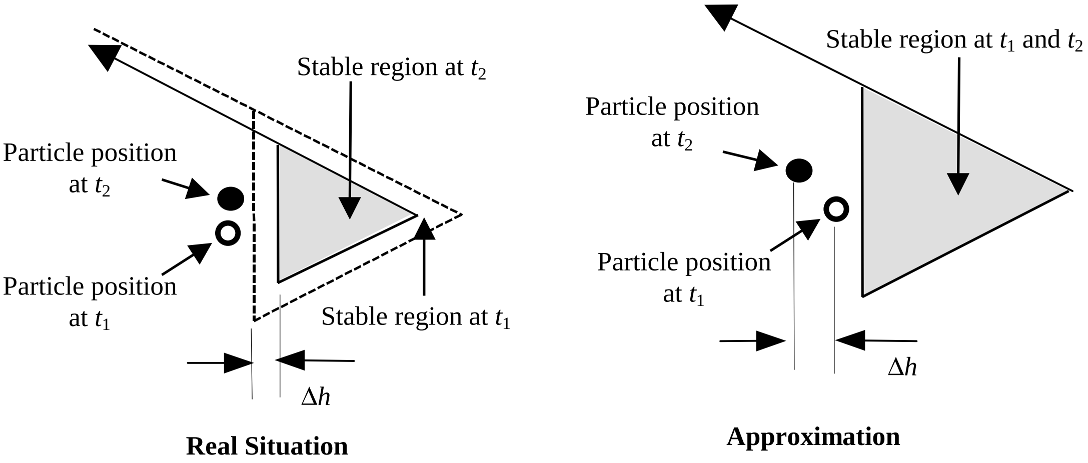

Stable region at t_2
Particle position at t_2
Particle position at t_1
Stable region at t_1
Δh
Real Situation
Stable region at t_1 and t_2
Particle position at t_2
Particle position at t_1
Δh
Approximation

Figure 4.2 The relative motion of an unstable particle and the stable triangle

4.3.1 Transit time from the fixed point to the electrostatic septum under dynamic conditions

It was shown in the static case in Section 4.2.1 that in this region the extraction time does not depend upon the initial value of Λ for values up to 0.1. Following this hint, Λ_0 is set to zero and the particle is assumed to move on the outgoing separatrix $Y = -\frac{1}{\sqrt{3}}X$. With this assumption equation 4.2(a) becomes:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\sqrt{3}S}{2}\left(3hX - X^2\right). \tag{4.14}$$

To this velocity, it is necessary to add the velocity of the separatrix as it recedes from the particle. It is easy to evaluate the velocity in the non-translated frame and to note that it has to be the same in both frames. In the non-translated frame, the separatrix equation is given by $X = -h$. Thus:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = -\frac{4\pi}{S}\frac{dQ}{dt} \tag{4.15}$$

so that:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\sqrt{3}S}{2}\left(3hX - X^2\right) + \frac{4\pi}{S}\dot{Q}\,. \tag{4.16}$$

This integral can be evaluated using the same standard form as in the static case, to give,

Transit time from 'corner':
(dynamic conditions)

$$T_{\text{corner, dynamic}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{n}{(n+3)}\frac{2\sqrt{3}}{\left(\frac{2}{\sqrt{3}}\lambda_0 - \frac{4\pi\dot{Q}}{\varepsilon^2}\right)}\right| = \frac{1}{\varepsilon\sqrt{3}}\ln\left|\underbrace{\frac{n}{(n+3)}}_{\text{Position of the electrostatic septum}}\ \underbrace{\frac{3}{\left(\lambda_0 - \frac{1}{\sqrt{3}\varepsilon}\frac{\dot{\varepsilon}}{\varepsilon}\right)}}_{\text{Effect of the separatrix motion}}\right| \tag{4.17}$$

which is obtained using the assumption $\Delta h << h$ which translates into

$$\left|\frac{4\pi\dot{Q}}{S}\frac{1}{\sqrt{3}\varepsilon}\right| << \left|\frac{2}{3}\frac{\varepsilon}{S}\right| \Rightarrow \left|\dot{Q}\right| << \frac{1}{2\sqrt{3}\pi}\varepsilon^2\,.$$

The term in λ_0 in the denominator has been kept in order to recover the static expression when $\dot{Q} = 0$.

4.3.2 Extraction time from the side of the stable triangle under dynamic conditions

As in the static case, it is necessary to evaluate the time spent moving along the side of the stable triangle. In equation (4.2)(b) for the motion in Y, the X and X^2 terms are neglected and the velocity of the separatrix (i.e. the approach velocity of the stable fixed point) is added to give:

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = -\frac{S}{4}\left(6\sqrt{3}hY + 3Y^2\right) - \frac{8\pi}{\sqrt{3}S}\dot{Q}\,. \tag{4.18}$$

When integrated between $Y_0 = -2\sqrt{3}\Lambda_0 h$ and $Y_{\text{F, dynamic}} = -2\sqrt{3}\Lambda_{\text{F}}h$, this gives:

Transit time from 'side':
(dynamic conditions)

$$T_{\text{side, dynamic}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{1-\Lambda_{\text{F}}}{\Lambda_{\text{F}}}\right)\right|, \tag{4.19}$$

which is exactly the same as in the case of the static situation in (4.10). This is reasonable, considering the approximation used in which the velocity in Y does not depend on X and in which the variation in the length of the triangle's side is negligible in the time considered.

4.3.3 Matching the ‘side’ and ‘corner’ trajectories under dynamic conditions

It now remains to evaluate the ‘hand-over’ point where the two trajectories coincide. This can be approximated by adding the X_F-position of the static case in (4.12) to the movement of the separatrix, which will be the product of the velocity of the separatrix and the time spent by the particle moving along the side of the stable triangle. The Y_F-position will be unchanged from that of the static case to this level of approximation.

$$X_{\mathrm{F,\,dyanmic}} = -\underbrace{\frac{\Lambda_0(1-\Lambda_0)}{\Lambda_{\mathrm{F}}(1-\Lambda_{\mathrm{F}})}\lambda_0 h}_{\text{Position in static case}} + \underbrace{\frac{1}{\varepsilon\sqrt{3}}\ln\left|\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\right)\right|\frac{4\pi}{S}\dot{Q}}_{\text{Movement of separatrix}} = -\lambda_{\mathrm{F,\,dynamic}} h \qquad (4.20)$$

Thus, the total time to reach the electrostatic septum for a particle starting far from O′, is the sum of (4.17) and (4.19) with the appropriate value of $\lambda_{\mathrm{F,\,dynamic}}$ from (4.20),

Transit time:
(dynamic conditions)

$$T_{\mathrm{dynamic}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\underbrace{\left(\frac{n}{n+3}\right)}_{\substack{\text{Position of}\\ \text{electrostatic}\\ \text{septum}}}\underbrace{\left(\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\right)}_{\substack{\text{Handover}\\ \text{point in } Y}}\underbrace{\left(\frac{\Lambda_0}{1-\Lambda_0}\right)}_{\substack{\text{Starting point}\\ \text{on stable}\\ \text{triangle}}}\underbrace{\left(\frac{3}{\lambda_{\mathrm{F,\,dynamic}} - \dfrac{2\pi\sqrt{3}}{\varepsilon^2}\dot{Q}}\right)}_{\substack{\text{Effect of separatrix}\\ \text{motion}}}\right|. \qquad (4.21)$$

Unfortunately in the dynamic case, Λ_F does not disappear as it did in the static case, but it turns out that particles with $0 < \Lambda_0 < 0.1$, come out approximately together. Thus, $T_{\mathrm{corner,\,dynamic}}$ can be considered to be correct up to $\Lambda_0 = 0.1$ and it seems natural to choose Λ_F to correspond to this value. The approximation of $\Lambda_F = 0.1$ is based on the simulations summarised in Table 4.1 that also compare the static and dynamic formulae to simulations over a wider parameter range.

Comparison of the various formulae for the transit time shows that they all have the form,

$$T = \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{\text{Other terms}}{\lambda_0 - \dfrac{1}{\varepsilon\sqrt{3}}\dfrac{\dot{\varepsilon}}{\varepsilon}}\right| \qquad (4.22)$$

and that the order of magnitude of the transit time is given by $1/\varepsilon$. The various transit time formulae have been checked against numerical simulation and have been found to be correct to within five percent, provided the relative variation of the parameters is small during the transit time. Table 4.1, shows a sub-set of the comparisons made with the static and dynamic formulae. The initial values for these calculations were $\lambda = 15 \times 10^{-6}$ $\varepsilon = 5.65 \times 10^{-2}$.

Comparison of numerical simulations and analytical results for the transit time			
Starting position, Λ_0 ($Y_0 = 2\sqrt{3}\Lambda_0 h$)	**Transit time (no. of revolutions)** Numerical simulation	$T_{\text{corner, static}}$	T_{static}
0	384	364	-
0.1	390	364	370
0.5	426		406
0.9	531		505
0.999	843		787
		$T_{\text{corner dynamic}}$	T_{dynamic}
0	306	290	-
0.1	309	290	290
0.5	327		322
0.9	381		375
0.999	531		501

Table 4.1 Numerical simulations and analytical results for the transit time

4.4 BEAM MODELS FOR THE SPILL

4.4.1 Time profile of a 'strip' [*]

In the previous sections, expressions have been derived for the time needed (transit time) for individual particles to leave the machine. The aim now is to use these basic results to predict the time profile of the spill from an elementary 'strip' of particles sitting along the side of the last stable triangle that are made unstable by the recession of the separatrix during three turns (see Figure 4.3). Ultimately, it will be possible to integrate over the elementary strips to form 'bands' that will include all the different momenta that become unstable at any one time. Once this is done, the full simulation of the time profile of a spill can be attempted.

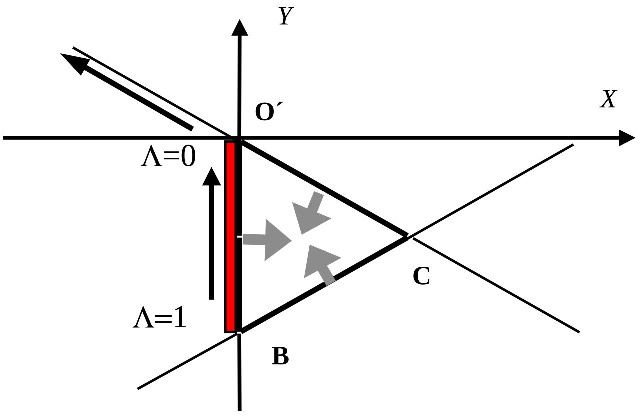


Figure 4.3 Shrinking the stable region leaves a narrow 'strip' of particles unstable
[In the figure, only one of the three sides of the triangle is considered and the motion of the particles and separatrices are shown by arrows.]

---

[*] The name 'strip' will be reserved for the narrow region of mono-energetic particles on the side of a stable triangle. The name 'band' will be introduced later for the series of strips in a beam with a momentum spread.

The first particle will reach the electrostatic septum at $t_0 = T(\Lambda{=}0)$ and the last one at $t_L = T(\Lambda{=}\Lambda_L)$. Note that the maximum time does not correspond to $\Lambda = 1$. This is due to the fact that the particles that start near to the stable point B will be overtaken by the inward movement of the separatrix BC. These particles will be extracted along the following separatrix. This is equivalent to considering them as starting near O´ (they are in fact replaced by particles from the separatrix OC that are overtaken by the movement of OB). The delay with which they appear near O´ will not be considered, for the moment. The value of Λ_L will be given later.

As noticed in Section 4.3.3, all the particles starting with $0<\Lambda<0.1$, reach the electrostatic septum at virtually the same moment, so the spill will start with a spike at $t_0 = T_{\text{corner, dynamic}}(\delta Q, \dot{Q})$. If $\rho(\Lambda)$ is the linear probability density of particles in the strip, and N_T is the total number of particles contained in the strip, the spike will contain N_{spike} particles, where:

$$N_{\text{spike}} = \underbrace{N_T \int_0^{0.1} \rho(\Lambda)\mathrm{d}\Lambda}_{\text{Particles in the vicinity of O'}} + \underbrace{N_T \int_{\Lambda_L}^{1} \rho(\Lambda)\mathrm{d}\Lambda}_{\text{Particles overtaken by the separatrix}} \tag{4.23}$$

where $\int_0^1 \rho(\Lambda)\mathrm{d}\Lambda = 1$. Thus the initial spike can be described by

Initial spike: $$N_T P_{\text{spike}}(t)\mathrm{d}t = N_{\text{spike}}\, \delta(t-t_0)\mathrm{d}t \tag{4.24}$$

After the initial spike, the spill shape can be evaluated by noticing that the particles coming out between $T(\Lambda)$ and $T(\Lambda)+\mathrm{d}t$ are the ones which started between Λ and $\Lambda+\mathrm{d}\Lambda$. This means

Spill tail: $$N_T P_{\text{tail}}(t)\mathrm{d}t = N_T P_{\text{tail}}(T(\Lambda))\mathrm{d}t = N_T \rho(\Lambda)\mathrm{d}\Lambda = N_T \rho(\Lambda)\frac{\mathrm{d}\Lambda}{\mathrm{d}t}\mathrm{d}t \tag{4.25}$$

for $t_0 < t < t_F$, which corresponds to $0.1 < \Lambda < \Lambda_L$. The time profile for the elementary strip is then given by the sum of P_{spike} and P_{tail}.

Strip profile: $$P_{\text{strip}}(t)\mathrm{d}t = P_{\text{spike}}(t)\mathrm{d}t + P_{\text{tail}}(t)\mathrm{d}t \tag{4.26}$$

Once Λ_L, $\rho(\Lambda)$ and $\Lambda(t)$ are known, the time profile of the elementary spill can be fully evaluated. The shape of this elementary spill is the key to calculating the spill profile for the whole beam and for including the influence of ripple.

Evaluation of the linear probability density $\rho(\Lambda)$

Assuming that the initial beam is smoothly distributed in phase space and that the resonance is applied adiabatically, then the density probability $\rho(\Lambda)$ in the strip will be proportional to $1/v(\Lambda)$, where $v(\Lambda) = \mathrm{d}Y/\mathrm{d}t$ and is the velocity in normalized phase space of the particles. With the Kobayashi Hamiltonian translated to the fixed point O´, $v(\Lambda)$ is given by (4.2)(b):

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = -\frac{S}{4}\left(12hX + 6\sqrt{3}hY + 3Y^2 - 3X^2\right). \tag{4.2)(b}$$

Restricting (4.2)(b) to trajectories close to the separatrix O´B and re-expressing the co-ordinates according to (4.3), but neglecting the terms in λ^2, yields,

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = v(\Lambda) = -3Sh^2\left(-\lambda - 3\Lambda + 3\Lambda^2\right), \tag{4.27}$$

so that

$$\rho(\Lambda) = k\frac{1}{v(\Lambda)} = k\frac{1}{-3Sh^2\left(-\lambda - 3\Lambda + 3\Lambda^2\right)}. \tag{4.28}$$

Since $\int_0^1 \rho(\Lambda)d\Lambda = 1$, the constant k is found to be , $k = \dfrac{-9Sh^2}{2\ln\frac{\lambda}{3}}$, so that

Line density:
$$\rho(\Lambda) = \frac{3}{2\left(3\Lambda^2 - 3\Lambda - \lambda\right)\ln\frac{\lambda}{3}}. \tag{4.29}$$

Inverting* T*(Λ)

The second element that is necessary for the evaluation of P_{tail}, is $\mathrm{d}\Lambda/\mathrm{d}t$. This will be derived from the time, T_{dynamic} needed for a particle to reach the electrostatic septum starting from $X_0 = -\lambda_0 h$ and $Y_0 = -2\sqrt{3}\Lambda_0 h$ close to the separatrix O´B. This transit time can be found in (4.21)

$$T_{\mathrm{dynamic}} \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\left(\frac{n}{n+3}\right)\left(\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\right)\left(\frac{\Lambda_0}{1-\Lambda_0}\right)\left(\frac{3}{\lambda_{\mathrm{F,\,dynamic}} - \frac{2\pi\sqrt{3}}{\varepsilon^2}\dot{Q}}\right)\right| \tag{4.21}$$

The re-arrangement of (4.21) to solve for Λ (the 0 subscript is now dropped as Λ becomes a variable) and the use of t instead of T_{dynamic} to stress that it is now the independent variable, yield

$$e^{\varepsilon\sqrt{3}t} = \frac{n}{n+3}\,\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\,\frac{\Lambda}{1-\Lambda}\,\frac{3}{\left(\lambda_{\mathrm{F,dynamic}} - \frac{1}{\varepsilon\sqrt{3}}\frac{\dot{\varepsilon}}{\varepsilon}\right)} \tag{4.30}$$

and

$$\frac{n+3}{3n}e^{\varepsilon\sqrt{3}t}=\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\frac{\Lambda}{1-\Lambda}\frac{1}{\frac{\Lambda(1-\Lambda)}{\Lambda_{\mathrm{F}}(1-\Lambda_{\mathrm{F}})}\lambda-\frac{1}{\varepsilon\sqrt{3}}\frac{\dot{\varepsilon}}{\varepsilon}\left(1+\ln\left|\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\frac{\Lambda}{1-\Lambda}\right|\right)} \tag{4.31}$$

As λ_0 is of the order of $-\frac{\dot{\varepsilon}}{\varepsilon}$, and $\varepsilon << 1$, the first term in the denominator of the last fraction in (4.31) can be neglected with respect to the second. Equation (4.31) can then be rewritten as

$$R=K+K\ln R \tag{4.32}$$

where

$$R=\frac{1-\Lambda_{\mathrm{F}}}{\Lambda_{\mathrm{F}}}\frac{\Lambda}{1-\Lambda};\quad K=-\frac{\dot{\varepsilon}}{\varepsilon^2\sqrt{3}}\frac{n+3}{3n}e^{\varepsilon\sqrt{3}t} \tag{4.33}$$

In order to invert $T(\Lambda)$, it is necessary to solve (4.32) with respect to R. This equation has real solutions only when $K \geq 1$, which corresponds to $t \geq t_0$. This is to be expected since no particle is extracted prior to t_0. Let $R = R(K)$ be the solution of equation(4.32), then:

$$\Lambda=\frac{g(K)}{1+g(K)} \tag{4.34}$$

where $g(K)=\frac{\Lambda_{\mathrm{F}}}{1-\Lambda_{\mathrm{F}}}R(K)$.

Referring back to equation (4.25) the aim is to evaluate $\rho(\Lambda)\frac{d\Lambda}{dt}$, which with (4.29) becomes,

$$\rho(\Lambda)\frac{\mathrm{d}\Lambda}{\mathrm{d}t}=\frac{3}{2\left(3\Lambda^2-3\Lambda-\lambda\right)\ln\frac{\lambda}{3}}\frac{\mathrm{d}\Lambda}{\mathrm{d}t}\approx\frac{3}{2\left(3\Lambda^2-3\Lambda\right)\ln\frac{\lambda}{3}}\frac{\mathrm{d}\Lambda}{\mathrm{d}t}. \tag{4.35}$$

The substitution of expression (4.34), yields:

$$\rho(\Lambda)\frac{\mathrm{d}\Lambda}{\mathrm{d}t}=-\frac{(1+g(K))^2}{2g(K)\ln\frac{\lambda}{3}}\frac{\dot{g}(K)}{(1+g(K))^2}=-\frac{1}{2\ln\frac{\lambda}{3}}\frac{\dot{g}(K)}{g(K)}=-\frac{1}{2\ln\frac{\lambda}{3}}\frac{\dot{R}(K)}{R(K)} \tag{4.36}$$

which no longer depends on Λ_{F}. Note that neglecting λ with respect to $(\Lambda^2 - \Lambda)$ in equation (4.35), means that $\Lambda >> \lambda$, which is always true for $\Lambda > 0.1$, and $(1 - \Lambda) >> \lambda$. This will be shown later to be also true for $\Lambda < \Lambda_{\mathrm{F}}$. The derivation of (4.32) with respect to time, gives

$$\dot{R}=\dot{K}+\dot{K}\ln(R)+K\frac{\dot{R}}{R}$$

$$\dot{R} = \frac{1+\ln(R)}{1-\dfrac{K}{R}}\dot{K}\,, \tag{4.37}$$

which, when substituted into (4.36), gives

$$\rho(\Lambda)\frac{\mathrm{d}\Lambda}{\mathrm{d}t} = -\frac{1}{2\ln\dfrac{\lambda}{3}}\,\frac{1+\ln(R)}{R-K}\,\dot{K} = \frac{1}{\ln\dfrac{\lambda}{3}}\,\frac{1+\ln(R)}{R-K}\,\frac{\pi\dot{Q}}{\varepsilon}\,\frac{n+3}{n}\,e^{\varepsilon\sqrt{3}t} \tag{4.38}$$

Note that neglecting the derivatives of n and ε was already included in the model when it was assumed that the stable region does not change during the extraction time. A plot of expression (4.38) is shown in Figure 4.4. When $K >> 1$, $R >> K$ and K can be neglected in the denominator of (4.38). From (4.32), $(1+\ln(R))/R = 1/K$ and for $K >> 1$, the asymptotic value of the spill density in the tail becomes,

Asymptotic value:

$$\rho(\Lambda)\frac{\mathrm{d}\Lambda}{\mathrm{d}t} \approx -\frac{1}{2\ln\dfrac{\lambda}{3}}\,\frac{\dot{K}}{K} = -\frac{\sqrt{3}\varepsilon}{2\ln\dfrac{\lambda}{3}} \tag{4.39}$$

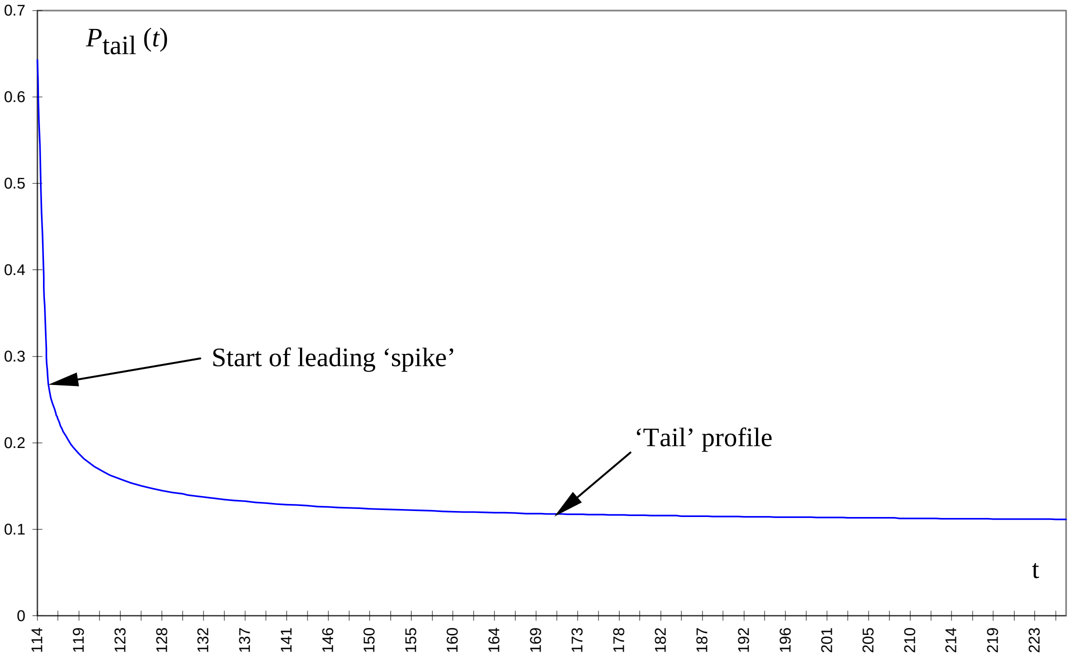


Figure 4.4 Shape of the tail profile $P_{\text{tail}}(t)$ for $\delta Q = 3\ 10^{-3}$ and $\dot{Q} = -9\ 10^{-9}$

Spill length

The end of the spill (and also the spill length) can be evaluated by noticing that the separatrix BC, in its movement, overtakes some particles with values of Λ sufficiently close to unity that their velocity is slower than the velocity of the separatrix itself. Those particles will then be extracted along the following separatrix. The effect is that some particle near B will disappear to reappear near O′. The spill therefore ends with the particle whose velocity is equal to the velocity of the separatrix. The Y co-ordinate of this particle is given by

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = \underbrace{-\frac{S}{4}\left(12hX + 6\sqrt{3}hY + 3Y^2 - 3X^2\right)}_{\text{Particle velocity}} = \underbrace{-\frac{8\pi}{\sqrt{3}S}\dot{Q}}_{\text{Separatrix velocity}} . \tag{4.40}$$

Neglecting the term in X and X^2 with respect to the right-hand side, and substituting $Y = -2\sqrt{3}\Lambda h$, yields

$$9Sh^2\Lambda^2 - 9Sh^2\Lambda - \frac{8\pi}{\sqrt{3}S}\dot{Q} = 0 \tag{4.41}$$

whose solutions are

$$\Lambda_{1,2} = \frac{1}{2}\left\{1 \mp \left(1 + \frac{16\pi\dot{Q}}{9\sqrt{3}S^2h^2}\right)\right\}. \tag{4.42}$$

The solution of interest is the one close to the stable point B, so that

$$\Lambda_\mathrm{L} = 1 + \frac{8\pi\dot{Q}}{9\sqrt{3}S^2h^2} = 1 + \frac{2\pi\dot{Q}}{\sqrt{3}\varepsilon^2} = 1 + \frac{\dot{\varepsilon}}{3\sqrt{3}\varepsilon^2} . \tag{4.43}$$

Note that 1-Λ_L >> λ. This justifies the omission of λ in equations (4.35) and (4.40). The spill will thus end at

$$T_\text{dynamic}(\Lambda_\mathrm{L}) \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{n}{n+3}\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\frac{\Lambda_\mathrm{L}}{1-\Lambda_\mathrm{L}}\frac{3}{\left(-\frac{2\pi\sqrt{3}}{\varepsilon^2}\dot{Q}\right)\left(1+\ln\left|\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\frac{\Lambda_\mathrm{L}}{1-\Lambda_\mathrm{L}}\right|\right)}\right| \approx$$

$$\approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{n}{n+3}\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\left(1+\ln\left|\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\right|\right)^{-1}\right|$$

where the term in λ in the last fraction has been neglected as in (4.31). Expressing this time in units of $t_0 = T_\text{corner, dynamic}(\delta Q, \dot{Q})$, yields

$$\frac{T_\text{dynamic}(\Lambda_\mathrm{F})}{T_\text{corner, dynamic}} \approx \frac{\frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{n}{n+3}\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\left(1+\ln\left|\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\ldots\right|\right)^{-1}\right|}{\frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{n}{n+3}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\right|} =$$

$$= \frac{\ln\left(\frac{n}{n+3}\right)+\ln\left(\frac{1-\Lambda_\mathrm{F}}{\Lambda_\mathrm{F}}\right)+2\ln\left|\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\right|-\ln\left|\left(1+\ln\left|\frac{1-\Lambda_F}{\Lambda_\mathrm{F}}\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\right|\right)\right|}{\ln\left(\frac{n}{n+3}\right)+\ln\left|\left(-\frac{3\sqrt{3}\varepsilon^2}{\dot{\varepsilon}}\right)\right|} \approx 2\,.$$

Where all the addenda have been considered negligible with respect to $\ln\left|\varepsilon^2/\dot{\varepsilon}\right|$. Hence this complicated derivation converges to a beautifully simple result,

Spill end: $$T_\mathrm{dynamic}(\Lambda_\mathrm{L}) = 2t_0\,. \tag{4.44}$$

Thus, the first particle reaches the electrostatic septum at t_0, the last particle arrives at $2t_0$, and the spill length is t_0.

Width of the initial spike.

It is now possible to evaluate the time needed for the last particle to be overtaken by the separatrix. Consider the velocity of the particle along the side of the stable triangle and add the velocity of the separatrix as in (4.18) and re-write using λ and Λ:

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = -3Sh^2\left(3\Lambda^2 - 3\Lambda - \lambda\right) + \frac{8\pi}{\sqrt{3}S}\dot{Q}$$

which yields,

$$\dot{\Lambda} = \frac{\varepsilon}{\sqrt{3}}\left(3\Lambda^2 - 3\Lambda - \lambda\right) - \frac{2\pi}{\varepsilon}\dot{Q}\,. \tag{4.45}$$

The integration of (4.45) gives the time for the separatrix to overtake the last particle,

$$T_\mathrm{i} = \int_{\Lambda_\mathrm{i}}^{1} \frac{1}{\varepsilon\sqrt{3}\Lambda^2 - \varepsilon\sqrt{3}\Lambda - \frac{\varepsilon}{\sqrt{3}}\lambda - \frac{2\pi}{\varepsilon}\dot{Q}}\,\mathrm{d}\Lambda \approx \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{\Lambda_\mathrm{i}\frac{2}{3}\frac{\dot{\varepsilon}}{\varepsilon}}{2\varepsilon\sqrt{3}(\Lambda_\mathrm{i}-1)+\frac{2}{3}\frac{\dot{\varepsilon}}{\varepsilon}}\right|. \tag{4.46}$$

As Λ_i approaches Λ_L, T goes to infinity as expected, since this is the time needed for the separatrix to reach a particle, which moves with the same velocity. It is therefore necessary to consider $\Lambda_\mathrm{i} = \Lambda_\mathrm{F} + 2\lambda/\sqrt{3}$, that is one step of the separatrix away from Λ_L. Then,

$$T = \frac{1}{\varepsilon\sqrt{3}}\ln\left|\frac{1}{6\varepsilon}\right|, \tag{4.47}$$

which may be a large fraction of t_0 for small ε. However, for the sake of simplicity, and considering that most of the particles are near the stable points and thus will be overcome in a few turns, this delay will be neglected and all the particles starting near B will be considered as if they were starting near O´.

Population of the spike and the tail of the elementary 'strip'

The evaluation of the fraction of the beam contained in P_{spike} and P_{tail} can be done in a straightforward way by integrating the density $\rho(\Lambda)$. It is easier to integrate P_{tail} between the hand-over point at $\Lambda = 0.1$ and the point at which the last particle escapes the advancing separatrix, $\Lambda = \Lambda_L$, as in this range λ can be neglected in the density. This results in,

Fraction in tail:

$$\int_{0.1}^{\Lambda_L} \rho(\Lambda)\,d\Lambda = \frac{\ln\left|\dfrac{\lambda}{27\varepsilon\sqrt{3}}\right|}{2\ln\dfrac{\lambda}{3}} \approx \frac{1}{2} \tag{4.48}$$

It is reasonable to neglect $\ln(27\sqrt{3}\,\varepsilon)$ and $\ln(3)$ with respect to $\ln(\lambda)$, which yields the very simple result that 1/2 of the spill is in P_{tail} and thus 1/2 in the initial spike.

This result is of consequence for the response to ripple and the efficiency of feedback systems. For perturbations up to frequencies corresponding to the width of the initial spike (of the order of 100 kHz for revolution periods of the order of 1 μs), half of the beam behaves coherently with a definite delay while the other half is extended over a period that will cause overlap with ripple frequencies as low as 1 kHz.

4.4.2 Time profile of a 'band' [*]

Up to now, only the microscopic behaviour of an elementary strip of particles that are marooned just outside the stable triangle has been considered. In general, the spill will contain a continuous range of momenta with different emittance triangles all contributing elementary strips into the spill. The continuous range of strips entering the resonance will be called a 'band'. A schematic view of this situation is shown in Figure 4.5 with an arbitrarily shaped beam.

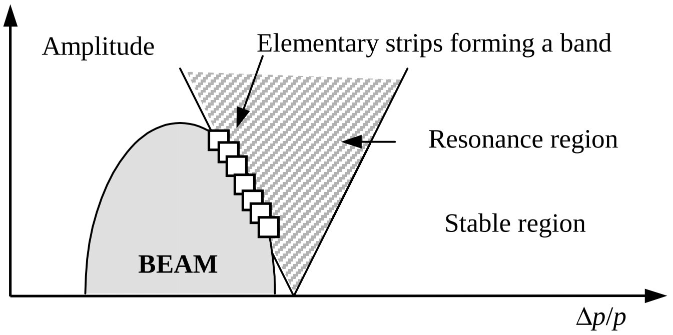


Figure 4.5 Elementary strips forming a 'band' in the Steinbach diagram

---

[*] Here the word 'band' is used to indicate the series of strips corresponding to different momentum and amplitude that become unstable simultaneously.

The total time profile of the spill arising from a band is obtained by summing the contributions of all the different elementary strips.

$$I(t) = \sum_i P_{\text{strip, i}}(t). \tag{4.49}$$

This is shown in a schematic way in Figure 4.6.

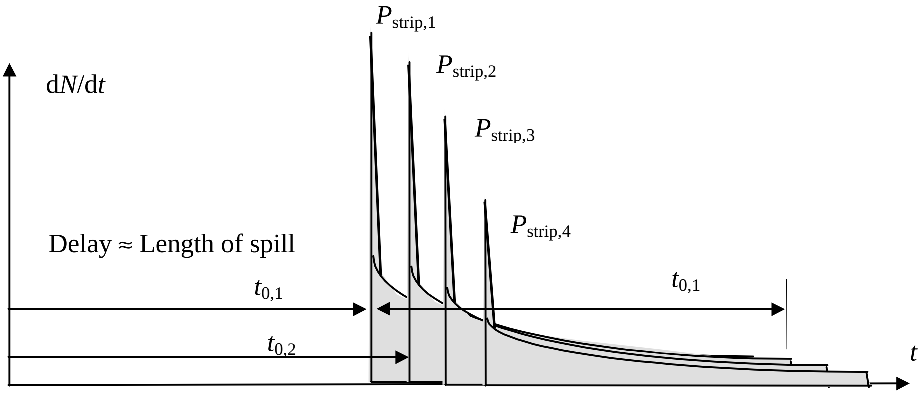


Figure 4.6 Summing strip spills to form a band spill

The evaluation of $I(t)$, can be greatly simplified when the total extraction time is much longer than the duration of a strip. In this case, all the terms $P_{\text{strip,i}}(t)$ are contributing to the total extracted current equally. The sum of all these terms is then given by the integral over one of them, which is simply the number of particles that become unstable as a function of time $N_T(t)$. This means that:

- If the extraction is performed in amplitude the time profile is the amplitude distribution,
- If the extraction is performed in momentum the time profile is the momentum distribution,
- Combinations of amplitude and momentum yield the combined distribution.

This is, of course, intuitively obvious for perfectly smooth spills. It will become clear later that the strip and band analyses are useful for understanding the response to ripple, but first the spill profile will be analysed for a band in the extraction method that moves the beam in momentum into a fixed resonance (i.e. 'Momentum selection by moving the beam' see Section 3.2.3).

Elementary 'band' for a wide momentum spread

Consider the momentum selection case, in which the beam is wide in tune so that a quasi stationary situation develops in which all betatron amplitudes become unstable simultaneously. In Figure 4.7, an amplitude-momentum selection extraction is shown and the elementary band of beam that becomes unstable in one step is highlighted.

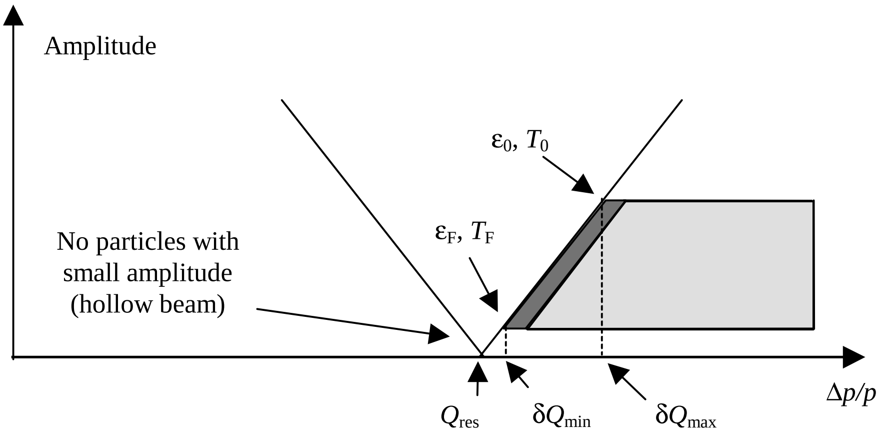

Amplitude
ε_0, T_0
ε_F, T_F
No particles with small amplitude (hollow beam)
Q_{res}
δQ_{min}
δQ_{max}
$\Delta p/p$

Figure 4.7 A band of particles that becomes unstable in one step
(Amplitude-momentum selection extraction)

A hollow beam is considered in order to set a minimum ε, which can be considered large with respect to its variation in the transit time. If the number of particles which would normally be in the hole is small, then a good representation for the beam with no hole will be found. This approximation is not so restricting since firstly the radial particle density across a beam rises from zero at the centre and secondly particles with a sufficiently small amplitude will cross the resonance without being extracted. Let the band that is extracted contain particles with tune shifts between δQ_{min} and δQ_{max}. The time profile of this band, is given by considering the time profile (4.26) for each ε value between $\varepsilon_F = 6\pi\delta Q_{min}$ and $\varepsilon_0 = 6\pi\delta Q_{max}$ and summing at each instant all these contributions, as shown in Figure 4.8

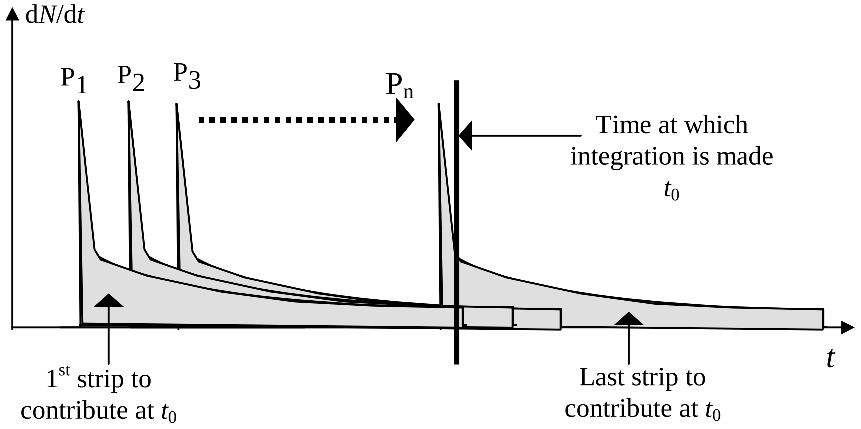

dN/dt
P_1
P_2
P_3
P_n
Time at which integration is made t_0
t
1st strip to contribute at t_0
Last strip to contribute at t_0

Figure 4.8 The total flux of particles is obtained summing the ‘strip profiles’ for all the amplitudes present

Band Profile:

$$N_B P_{band}(t) = \int_{\varepsilon_0}^{\varepsilon_F} N_T(\varepsilon) P_{strip}(t,\varepsilon)\mathrm{d}\varepsilon = \int_{\varepsilon_0}^{\varepsilon_F} N_T(\varepsilon)\big(P_{spike}(t,\varepsilon) + P_{tail}(t,\varepsilon)\big)\mathrm{d}\varepsilon \qquad (4.50)$$

where N_B is the total number of particles in the band and $N_T(\varepsilon)d\varepsilon$ is the number of particles marooned with detuning between ε and $\varepsilon + d\varepsilon$.

Simplified model

The exact integral (4.50) is too complicated to be evaluated analytically. So, to facilitate the task, each $P_{\mathrm{strip}}(t)$ is approximated by one delta function plus a rectangle representing P_{spike} and P_{tail} respectively. The width of the rectangles, as mentioned in a previous section, is equal to the time needed for the first particle to reach the electrostatic septum, that is $T_{\text{corner, dynamic}}$, and hence the height is $1/T_{\text{corner, dynamic}}$. Not all the elementary strips, $P_{\mathrm{strip}}(t, \varepsilon)$, are contributing at a given instant. Most of them are zero, either because they have not started, or because they have already finished. At time T, only the ε, for which $T/2 < T_{\text{corner, dynamic}}(\varepsilon) < T$ give a non-zero contribution. Thus the part of the integral for P_{tail}, becomes

$$N_{\mathrm{B}} P_{\text{band, tail}}(t) = \int_{\varepsilon(T)}^{\varepsilon(T/2)} \underbrace{\frac{N_{\mathrm{T}}(\varepsilon)}{2}}_{\substack{\text{Only half of the}\\ \text{particles are in the}\\ \text{rectangle (tail)}}} \frac{1}{T_{\text{corner, dynamic}}(\varepsilon)} \,\mathrm{d}\varepsilon\,. \tag{4.51}$$

To be correct the integration limits should be max(ε(T), ε_{L}), which corresponds to min(T, T_{L}), and min(ε(T/2), ε_0), which corresponds to max(T/2, T_0). After a change of the integration variable,

$$\int_{\max(\varepsilon(T),\varepsilon_{\mathrm{L}})}^{\min(\varepsilon(T/2),\varepsilon_0)} \frac{N_{\mathrm{T}}(\varepsilon)}{2} \frac{1}{T_{\text{corner, dynamic}}(\varepsilon)} \,\mathrm{d}\varepsilon = \int_{\min(T,T_{\mathrm{L}})}^{\max(T/2,T_0)} \frac{N_{\mathrm{T}}(\varepsilon(T))}{2} \frac{1}{T} \frac{\mathrm{d}\varepsilon}{\mathrm{d}T} \,\mathrm{d}T\,. \tag{4.52}$$

It is necessary to invert the expression for $T_{\text{corner, dynamic}}$ (4.17) to find ε(T),

$$T_{\text{corner, dynamic}} = \frac{1}{\sqrt{3}\varepsilon} \ln\left| \frac{n}{n+3} \frac{3}{\left(\lambda - \frac{1}{\sqrt{3}\varepsilon}\frac{\dot{\varepsilon}}{\varepsilon}\right)} \right|. \tag{4.17}$$

To simplify the calculation, it is assumed that the electrostatic septum is far away so that $n >> 1$ and $n/(n+3) \approx 1$. After neglecting λ and re-arranging into the form $V = \ln(U)/U$, the relation becomes,

$$\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}}\,T_{\text{cornar, dynamic}} = \frac{1}{\sqrt{-\frac{3\sqrt{3}}{\dot{\varepsilon}}\varepsilon}} \ln\left|\sqrt{-\frac{3\sqrt{3}}{\dot{\varepsilon}}}\varepsilon\right|. \tag{4.53}$$

A simple fit to the inversion, valid to within 10% in the range $4 < U < 30\,000$, is

$$\mathrm{U} = -1.42 \ln(V)/V. \tag{4.54}$$

Note that $U = 10$ corresponds to $\dot{\varepsilon}/\varepsilon^2 \approx 5\cdot 10^{-2}$ and values of U smaller than this should be avoided to stay within the validity limits of the transit time formulae

($\dot{\varepsilon}/\varepsilon^2$ measures the relative variation of the stable region during the transit time). Thus

$$\varepsilon = -1.42 \cdot \frac{2}{\sqrt{3}} \frac{\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)}{T} \tag{4.55}$$

from which

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}T} = -1.42 \cdot \frac{2}{\sqrt{3}} \frac{1}{T^2}\left(1 - \ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)\right). \tag{4.56}$$

It is now necessary to estimate $N_T(\varepsilon)d\varepsilon$. This is the number of particles that are in the border of thickness dh of a triangle of surface $3\sqrt{3}h^2 = 3\sqrt{3}\left(\frac{2}{3}\frac{\varepsilon}{S}\right)^2$. This border corresponds to the circular corona of width dR around a circle of the same surface πR^2 in the initial beam. Thus,

$$N_\mathrm{T}(\varepsilon)\mathrm{d}\varepsilon = N_\mathrm{B}\rho\left(R(\varepsilon)\right)\mathrm{d}\left(3\sqrt{3}H^2\right) = N_B\rho\left(\sqrt{\frac{3\sqrt{3}}{\pi}\frac{2}{3}\frac{\varepsilon}{S}}\right)\frac{8\sqrt{3}}{3S^2}\varepsilon\,\mathrm{d}\varepsilon, \tag{4.57}$$

where N_B is the total number of particles in the band. Two beam distributions (uniform and gaussian) will now be analysed using the simplified model described above.

Uniform distribution in phase space

In the first example, a uniform, phase-space distribution in the initial beam, of value ρ_0 = 1/Triangle Area = $\frac{3S^2}{4\sqrt{3}\varepsilon_0^2}$ is assumed and this yields for the integration of the tails from all the differential strips in the band,

$$P_\text{band, tail}(t) = \frac{1}{N_\mathrm{B}}\int_{\min(t,T_\mathrm{L})}^{\max(t/2,T_0)} \frac{N_\mathrm{T}(T(\varepsilon))}{2}\frac{1}{T}\frac{\mathrm{d}\varepsilon}{\mathrm{d}T}\mathrm{d}T =$$

$$= \rho_0 \frac{8\sqrt{3}}{3S^2}(-1.42)^2\frac{4}{3}\frac{1}{2}\int_{\min(t,T_\mathrm{L})}^{\max(t/2,T_0)} \frac{\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right) - \left[\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)\right]^2}{T^4}\mathrm{d}T$$

$$= \frac{4(-1.42)^2}{3\varepsilon_0^2}\int_{\min(t,T_\mathrm{L})}^{\max(t/2,T_0)} \frac{\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right) - \left[\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)\right]^2}{T^4}\mathrm{d}T$$

Fortunately, this integral can be evaluated using standard forms[*]

$$P_{\mathrm{band,tail}}(t)=\frac{4(-1.42)^2}{3\varepsilon_0^2}\times$$
$$\left[-\frac{\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)}{3T^3}-\frac{1}{9T^3}+\frac{\left[\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)\right]^2}{3T^3}+\frac{2\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}T}\right)}{9T^3}+\frac{2}{27T^3}\right]_{\min(t,T_\mathrm{L})}^{\max(t/2,T_0)}$$

and finally, for times $2T_0 < t < T_\mathrm{L}$, in the middle of the spill

$$P_{\mathrm{band,tai}l}(t)=\frac{4(-1.42)^2}{3\varepsilon_0^2}\times$$
$$\left[\frac{63\left[\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}t}\right)\right]^2-(144\ln(2)+21)\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}t}\right)+\left(72\ln(2)^2+24\ln(2)-7\right)}{27t^3}\right] \tag{4.58}$$

To correctly evaluate the initial and the final part of the band spill, the correct integration limits have to be considered. In fact, for times $t < T_0$, no particle has reached the septum and $P_{\mathrm{band,\ tai}l} = 0$. For times $T_0 < T < 2T_0$, the integral has to be performed between T and T_0 (there are no particles with $\varepsilon = \varepsilon(T/2)$). For the same reason, for times $T_\mathrm{L} < T < 2T_\mathrm{L}$, the integral has to be performed between T_L and $T/2$. Finally, for times greater than $2T_\mathrm{L}$, $P_{\mathrm{band,\ tail}} = 0$ again.

The contribution from the initial spikes of the strips in the band, gives,

$$P_{\mathrm{band,spike}}(t)=\int_{T_\mathrm{L}}^{T_0}\frac{N_\mathrm{T}(\varepsilon(T))}{2}\delta(t-T)\frac{\mathrm{d}\varepsilon}{\mathrm{d}T}\mathrm{d}T=$$
$$=-\frac{4(-1.42)^2}{3\varepsilon_0^2}\,\frac{\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}t}\right)-\left[\ln\left(\frac{1}{2}\sqrt{-\frac{\dot{\varepsilon}}{\sqrt{3}}t}\right)\right]^2}{t^3} \tag{4.59}$$

The graph of the spill shape for a particular case is drawn in Figure 4.9. This curve has a characteristic ‘shoulder’ that can be seen in measurements of spikes in slow extracted spills.

---

[*] $\int\frac{\ln(ax)}{x^\mathrm{p}}\mathrm{d}x=-\frac{\ln(ax)}{(p-1)x^{\mathrm{p}-1}}-\frac{1}{(p-1)^2x^{\mathrm{p}-1}}$ and

$$\int\frac{(\ln(x))^2}{x^\mathrm{p}}\mathrm{d}x=-\frac{(\ln(x))^2}{(p-1)x^{\mathrm{p}-1}}-\frac{2\ln(x)}{(p-1)^2x^{\mathrm{p}-1}}-\frac{2}{(p-1)^3x^{\mathrm{p}-1}}$$

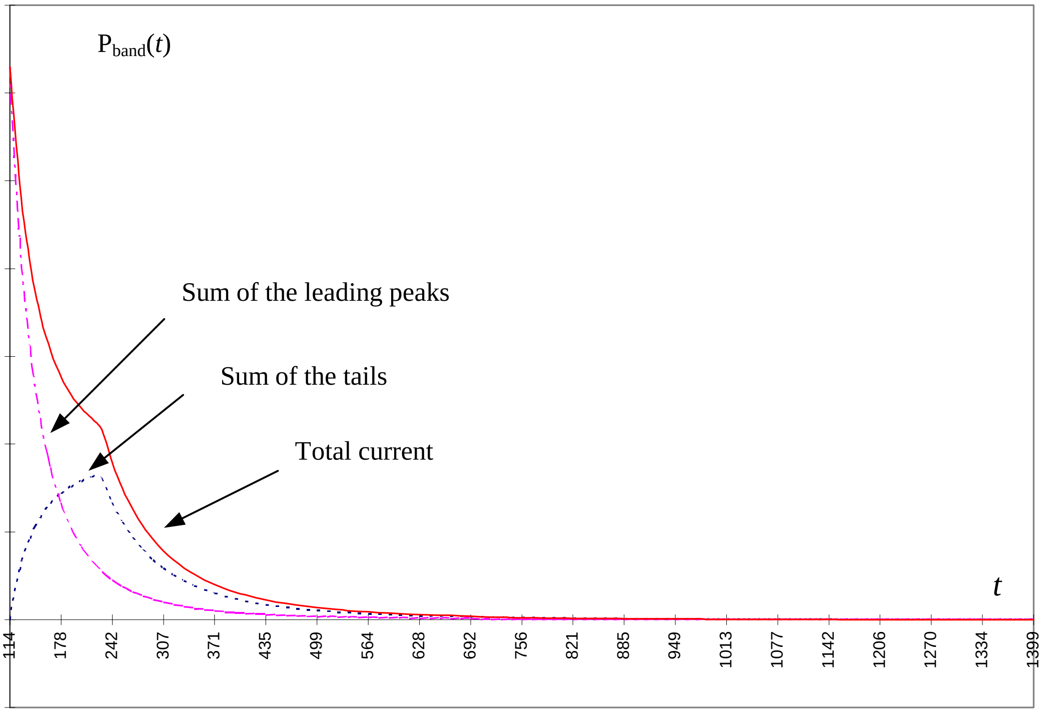


Figure 4.9 Spill profile for a 'band' from a uniformly distributed beam
[$\delta Q_{max} = 3\times10^{-3}$, $\delta Q_{min} = 3\times10^{-4}$, $\dot{Q} = -9\times10^{-9}$, $S = 15.33\ \mathrm{m}^{-1/2}$, $E_x = 10\ \pi$ mm mrad]

Gaussian distribution in phase space

In the second example, the frequently-used gaussian distribution in phase space is used, rather than the uniform distribution as assumed in the earlier example.

$$\rho(R) = \frac{1}{2E_x} e^{-\frac{\pi R^2}{2E_x}} \tag{4.60}$$

where E_x is the horizontal RMS emittance. Assuming that the area of the largest stable triangle (i.e. the acceptance imposed by the separatrices) corresponds to an emittance of the beam of n-σ, the one-σ emittance is given by,

$$E_x = \frac{Area}{n^2} = \frac{4\sqrt{3}\varepsilon_0^2}{3S^2n^2}$$

and (4.60) becomes,

$$\rho(\varepsilon) = \frac{3S^2n^2}{8\sqrt{3}\varepsilon_0^2} e^{-\frac{n^2\varepsilon^2}{2\varepsilon_0^2}}$$

$$N_\mathrm{T}(\varepsilon)\mathrm{d}\varepsilon = N_\mathrm{T} \frac{n^2}{\varepsilon_0^2} e^{-\frac{n^2\varepsilon^2}{2\varepsilon_0^2}} \varepsilon \mathrm{d}\varepsilon\,. \tag{4.61}$$

The resulting integral of (4.61)

$$P_{\text{band, tail}}(t) = \int_{\min(t,T_F)}^{\max(t/2,T_0)} \frac{N_{\text{T}}(T(\varepsilon))}{2} \frac{1}{T} \frac{\mathrm{d}\varepsilon}{\mathrm{d}T} \mathrm{d}T \qquad (4.62)$$

has to be integrated numerically (see Figure 4.10).

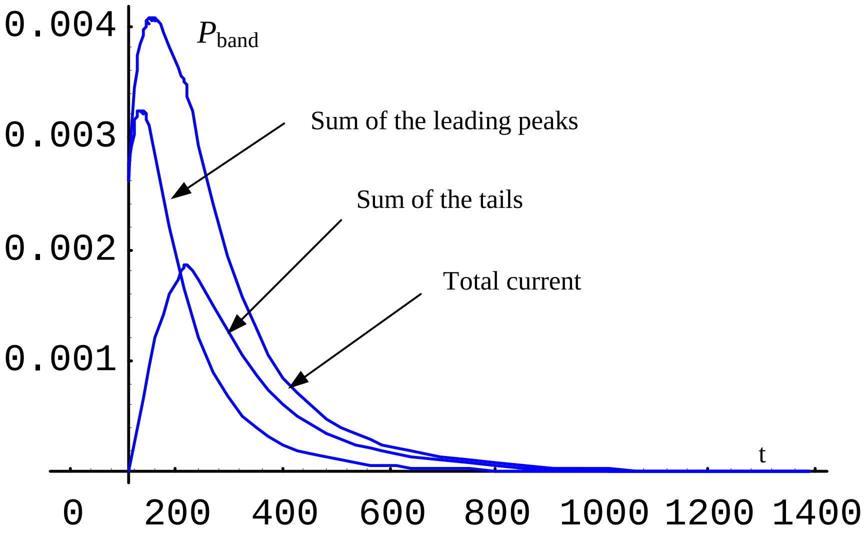


Figure 4.10 Spill profile for a 'band' from a gaussian beam
[$\delta Q_{\text{max}} = 3\times10^{-3}$, $\delta Q_{\text{min}} = 3\times10^{-4}$, $\dot{Q} = -9\ 10^{-9}$, $S = 15.33$ m$^{-1/2}$, $E_{\text{x}} = 10\ \pi$ mm mrad]

If the gaussian and the uniform cases are plotted on the same graph, the comparison in Figure 4.11 is obtained, which shows that the width is not very different.

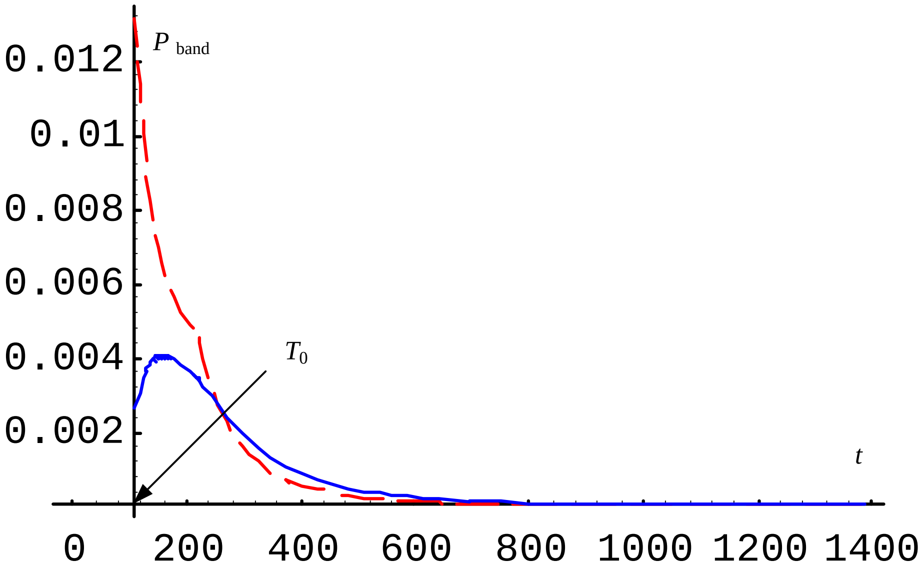


Figure 4.11 Band profiles for gaussian and uniform beams
[The 2σ emittance of the gaussian beam is set equal to that of the uniform beam]

From the plot above, it appears that the width of the 'band profile' is of the order of T_0. Thus, it can be expected that when the ripple period is of the order of T_0, the width of the profile partially fills the time interval during which no beam enters the resonance and smoothes the modulation of the extracted beam. This is an approximation that neglects the instantaneous variation of $\dot{Q}$ and uses an average value ($\dot{Q} = \dot{Q}_0 + \omega\,\delta Q_R \cos(\omega t) \approx \dot{Q}_0$). This is reasonable because $\dot{Q}$ appears in the expression for the transit time inside a logarithm. The relative insensitivity of the logarithm is illustrated in Figure 4.12 where the time profiles are compared for a variation in $\dot{Q}$ of 2 orders of magnitude.

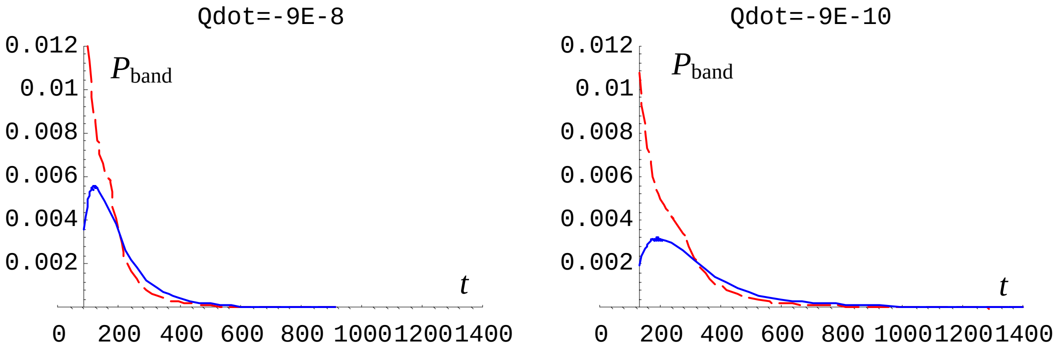


Figure 4.12 Time profiles for gaussian and uniform beams for different values of $\dot{Q}$

4.5 TUNE RIPPLE

The time profiles for the spills from elementary strips and bands provide the keys to understanding:

- How well feedback systems on the spill will work.
- The relative sensitivity of the different extraction methods to tune ripple.
- How to simulate the spill under specific conditions.

4.5.1 Comparison of the extraction methods

For simplicity and clarity, the three main extraction methods with diametrically opposed characteristics will be considered:

***Amplitude selection by moving the resonance*:** the beam is narrow in momentum and the betatron tune of the particles is changed by varying a quadrupole. The separatrix cuts through the beam from large to small amplitudes, see Figure 4.13(a).

RF knock-out: the chromaticity is zero (or near-zero), so that all particles have the same betatron tune. The particles are made unstable by blowing up the beam with transverse rf excitation. Thus, the resonance is reached in amplitude, see Figure 4.13(b).

Amplitude-momentum selection by moving the beam the beam is wide in momentum and the tune is changed via the chromaticity when accelerating the particles into the resonance (e.g. by a betatron core), see Figure 4.13(c).

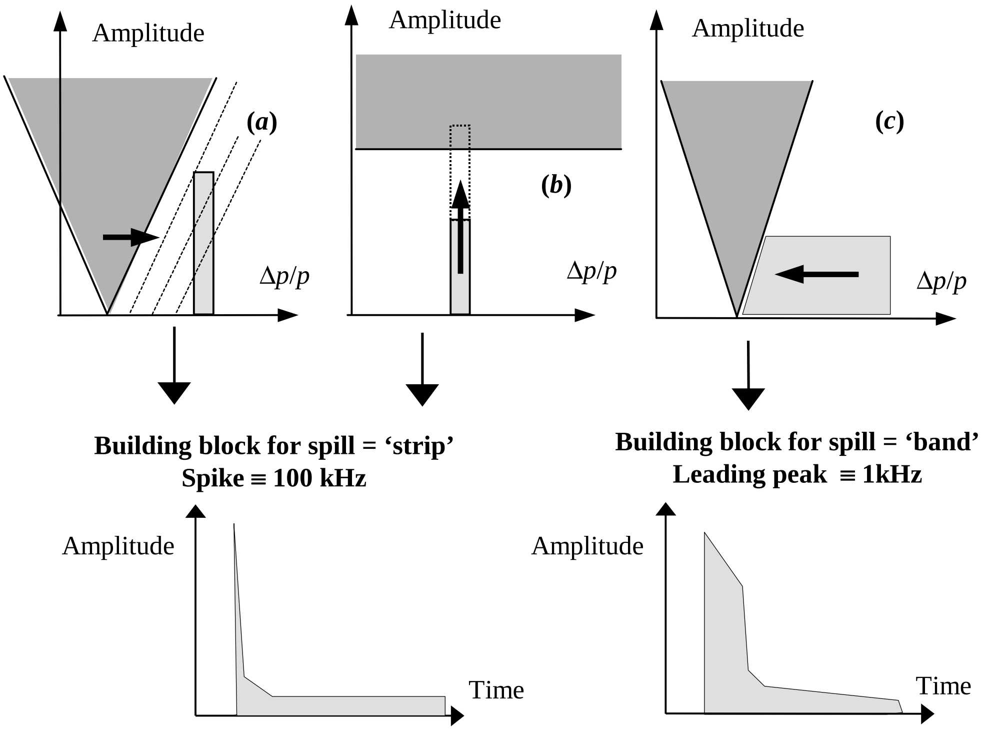


Figure 4.13 Comparison of the main extraction methods:
(*a*) Moving resonance, (*b*) Increasing particle amplitude, (c) Moving beam,

In the amplitude-selection case (a), the movement of the resonance maroons the particles with large betatron amplitudes first. Since the momentum spread is small, the basic element of spill marooned in the elementary extraction process (one movement of the resonance) is simply the 'strip' profile. So the 'building block' of this particular spill starts with a very narrow peak containing half of the particles involved. Moreover in the initial phase the transit time is very short (betatron amplitude large), the tails are equally short and sensitivity to the ripples is high. At the end of the spill, the transit times are longer (betatron amplitudes small), the particles in the tail are distributed over a longer period and ripple sensitivity reduces.

In the RF-knockout case (b), the blow-up velocity is fixed by the spill length. A ripple on the resonance excitation moves the resonance line up and down. As the particles enter the resonance at high amplitudes, and thus large tune distances, the transit time is short and the sensitivity to ripples is high. The situation is rather similar to the amplitude-selection case above, except that the sensitivity to ripple remains at a maximum throughout the spill because the extraction takes places constantly at maximum amplitude.

In the amplitude-momentum selection case (c), particles with all possible amplitudes become unstable at the same time. Thus the 'building block' of this particular spill is the 'band'. This leads to an enlarged leading peak, which tends to smooth ripple by filling the time intervals during which no beam enters the resonance in a much more efficient way.

4.5.2 Reducing the effect of ripple

This subject will be discussed later in more detail, but some comments can be made on the basis of the present knowledge:

- The amplitude-momentum selection case has the lowest intrinsic sensitivity to ripple
- Large betatron amplitudes means short transit times. This can be seen as a positive feature for a feedback system, but as a negative feature for the sensitivity to ripple. The present tendency is to use the beam feedback only for low frequencies (<1 kHz) and for the general spill shape. Other methods will be proposed to smooth ripple. Thus, small emittance beams appear preferable, which favours single-turn injection schemes over multi-turn injection schemes.
- A narrow resonance width (strength) means longer transit times and less sensitivity to ripple.

4.5.3 Spill simulations with ripple

To verify the analysis in the previous sections, numerical simulations have been made of the extraction processes with *amplitude selection* and with *amplitude-momentum selection*. The relevant numerical values are summarised in Table 4.2.

Values for numerical simulations	
Horizontal RMS emittance, E_x	$10\ \pi$ mm mrad
Normalised sextupole strength, S	$15.33\ \mathrm{m}^{-1/2}$
Revolution time	0.5×10^{-6} s
Max. detuning from resonance, δQ_{max}	$\frac{S}{4\pi}\sqrt{\frac{\pi}{3\sqrt{3}}}\sqrt{10^{-5}} = 3\cdot10^{-3}$
Min. detuning from resonance, δQ_{min}	3×10^{-4}
ε_{max}	$6\pi\delta Q_{max} = 5.65\times10^{-2}$
ε_{min}	$6\pi\delta Q_{min} = 5.65\times10^{-3}$
For transit time formulae to be valid $\frac{\dot{\varepsilon}}{\varepsilon^2} << 1$	$\frac{\dot{\varepsilon}}{\varepsilon_{min}^2} = 5\times10^{-2}$
Acceleration	$\dot{\varepsilon} = 1.6\cdot10^{-6}$ and $\dot{Q} = 8.47\cdot10^{-8}$
Tune step per turn	2.8×10^{-8}

Table 4.2 Values for numerical simulations

For the amplitude-momentum extraction, only a small slice of the beam of width $\Delta Q_{beam} = 3\times10^{-4}$, as illustrated in Figure 4.14, is considered. The number of revolutions needed to bring all the particles into the resonance is $3\frac{\Delta Q_{beam}}{\dot{Q}} = 10626$. A particle starting with the smallest amplitude that is considered has $\delta Q = 3\times10^{-4}$ and with the above set of parameters needs 1400 turns to reach the electrostatic septum. A maximum amplitude particle, starting with $\delta Q = 3\times10^{-3}$, needs only 260 turns. This large range of transit times is a feature that gives smoothing of the ripple.

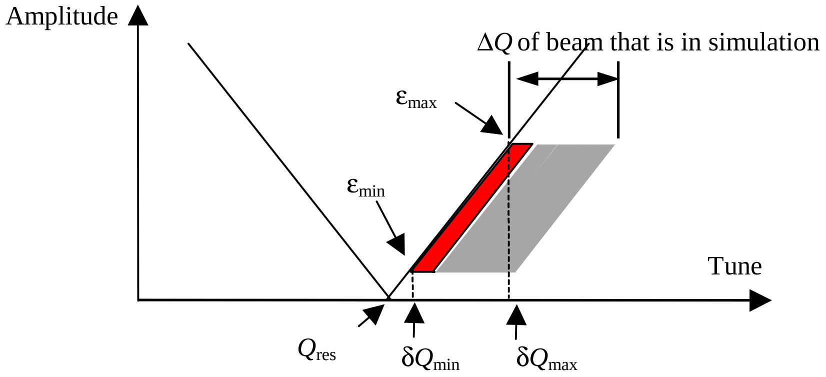


Figure 4.14 Initial beam distribution for amplitude-momentum selection simulations (not to scale)

A ripple frequency of 2 kHz has been considered. This is comparable to the longer transit times in the beam. The amplitude of a ripple has been adjusted to give 100% modulation.

Case 1 Amplitude-selection with a mono-energetic beam

In this case, only large amplitude particles, which have short transit times, are involved in the extraction simulation and only the first turns of the extraction are simulated. The modulation of the spill is 100%, as foreseen by the instantaneous transfer approximation. Figure 4.15 shows a beam without tune ripple and Figure 4.16 with ripple.

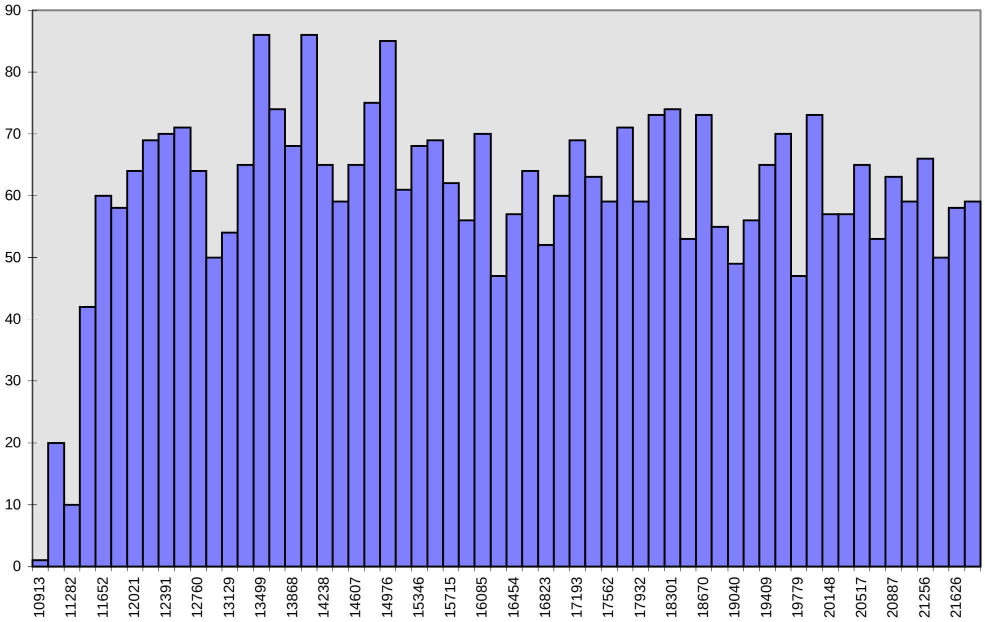


Figure 4.15 Spill for a mono-energetic beam without tune ripple

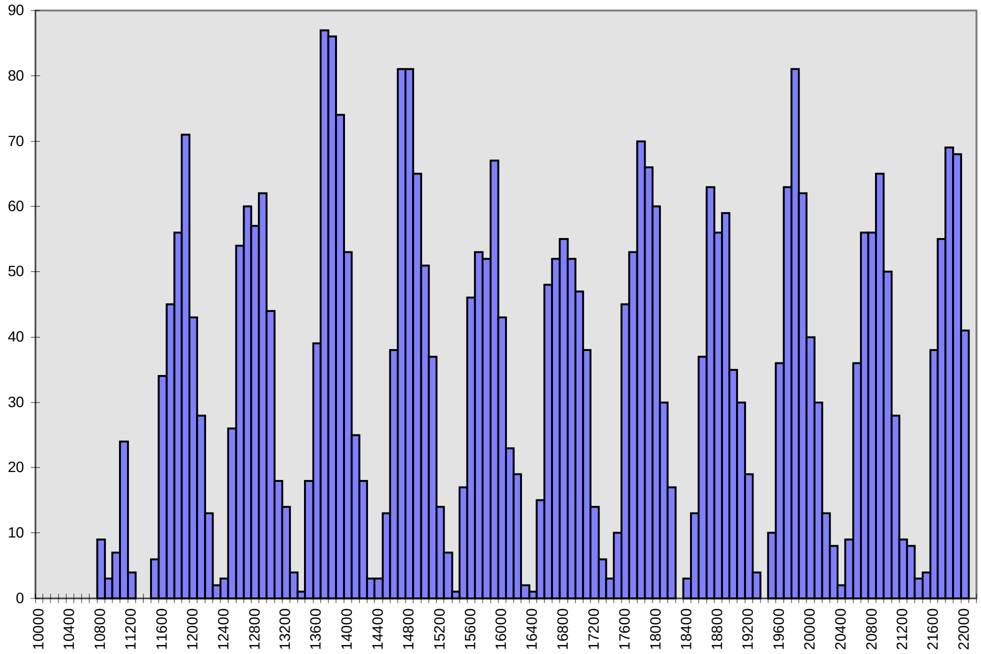


Figure 4.16 Spill for a mono-energetic beam with 2 kHz tune ripple

Case 2 Amplitude-momentum selection with a uniform beam distribution in phase space

In this simulation, the simultaneous extraction of all the amplitudes (band profile) smoothes the modulation in the extracted spill.

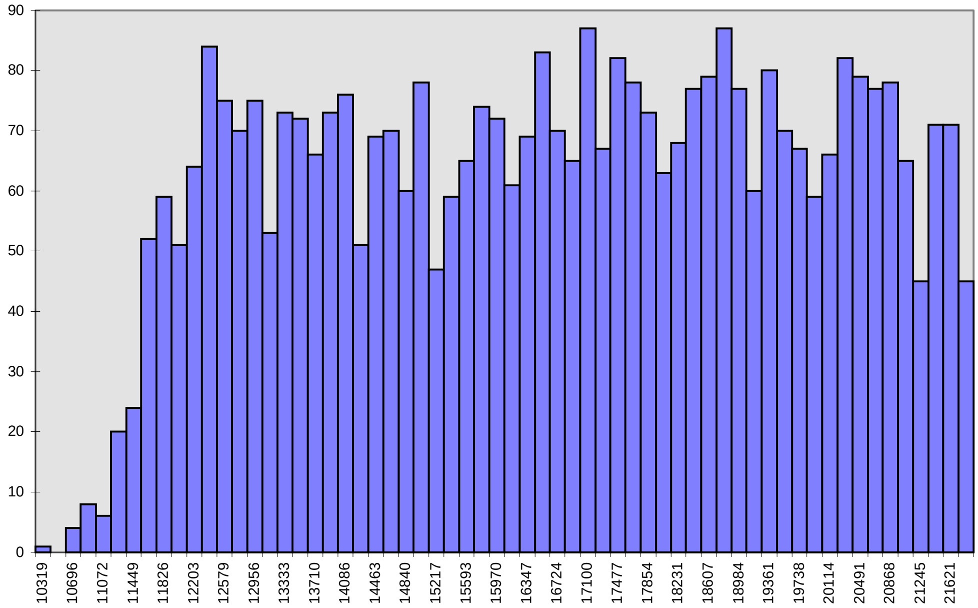


Figure 4.17 Spill for uniform distribution in phase space without ripple

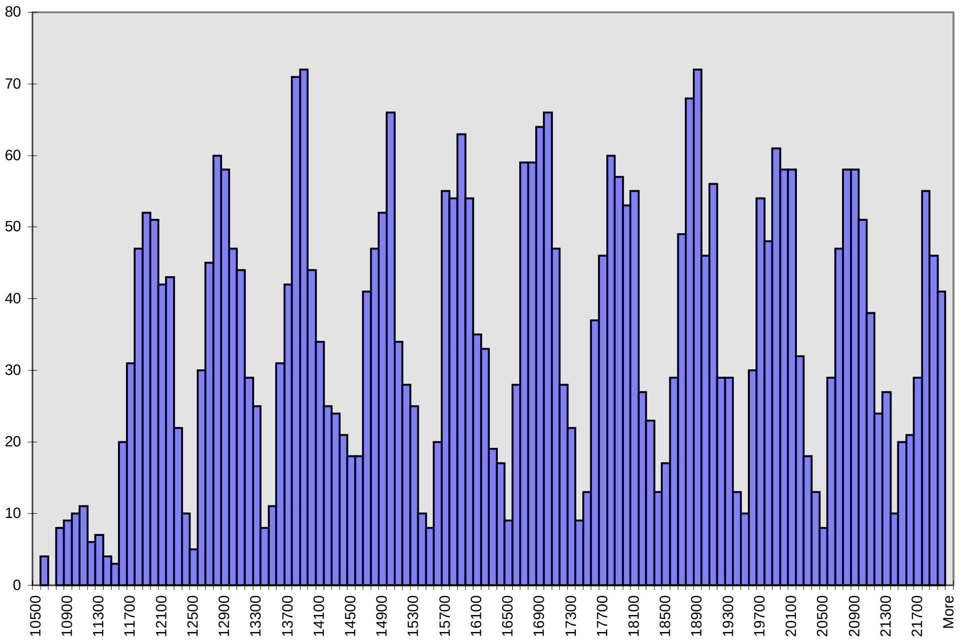


Figure 4.18 Spill for uniform distribution in phase space with 2 kHz ripple

Case 3 Amplitude-momentum selection with a gaussian beam distribution in phase space

The gaussian ‘band’ profile is slightly wider and less peaked than the uniform one. The result is that the modulation is less pronounced.

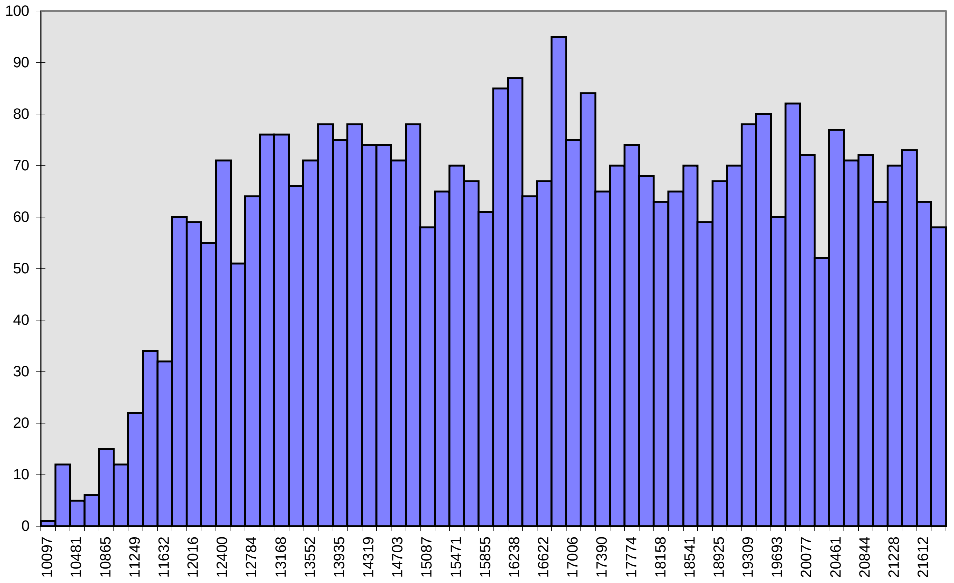


Figure 4.19 Spill for gaussian distribution in phase space without ripple

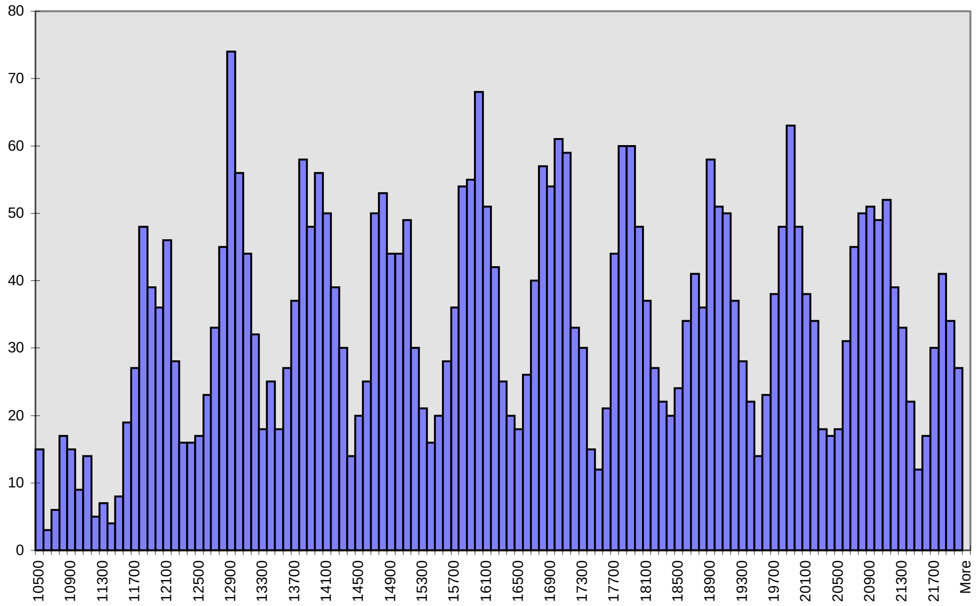


Figure 4.20 Spill for gaussian distribution in phase space with 2 kHz ripple

4.6 EMITTANCE OF THE EXTRACTED BEAM

4.6.1 A simple approach

The conservation of phase space (Liouville's Theorem) affords a simple and quick method for estimating and understanding the emittance of the extracted beam. In a linear machine, the (x, x'), (z, z') and ($\Delta p/p$,t) phase spaces are uncoupled and independently conserved, but the presence of a sextupole introduces two couplings:

(A) Firstly, between (x, x') and ($\Delta p/p$, t)
(B) Secondly, between (x, x') and (z, z').

(A) is an essential part of the mechanism of extraction and arises from the time variation of the positions of the separatrices, which 'cut' the phase space and 'peel' off the particles from the waiting beam. As mentioned in Section 3.1, the amplitude distribution (x, x') is converted into a momentum distribution ($\Delta p/p$).

(B) is an effect that couples the emittances of the two transverse planes. It can be minimised by making the vertical excursions of the particles small compared to the horizontal ones in the resonance and chromaticity sextupoles($\beta_z << \beta_x$). The effect is due to the high-order and cross terms in the magnetic fields (see equation (2.6).

To evaluate the emittance consider a coasting beam with a relative momentum spread of say $\Delta p/p = 0.005$ and let this beam be driven into the resonance over say 500 ms. Assume that the Hardt Condition is arranged so as to give a $\delta p/p$ for the extracted beam of 0.001. This determines the slope of the resonance line. Over the time of the extraction the separatrix acts like a knife shaving off the beam and the phase space. The transverse phase space and the longitudinal phase space ($\Delta p/p$, t) are jointly conserved as a phase-space volume. For the transverse plane, one can either neglect the coupling to the vertical plane, or consider the full volume (x, x', z, z'). By virtue of the phase space conservation, the effect of the extraction can be seen by equating the phase-space volumes before and after the extraction and then looking how the horizontal emittance in particular was affected. This illustrated schematically in Figure 4.21 and expressed quantitatively in (4.63).

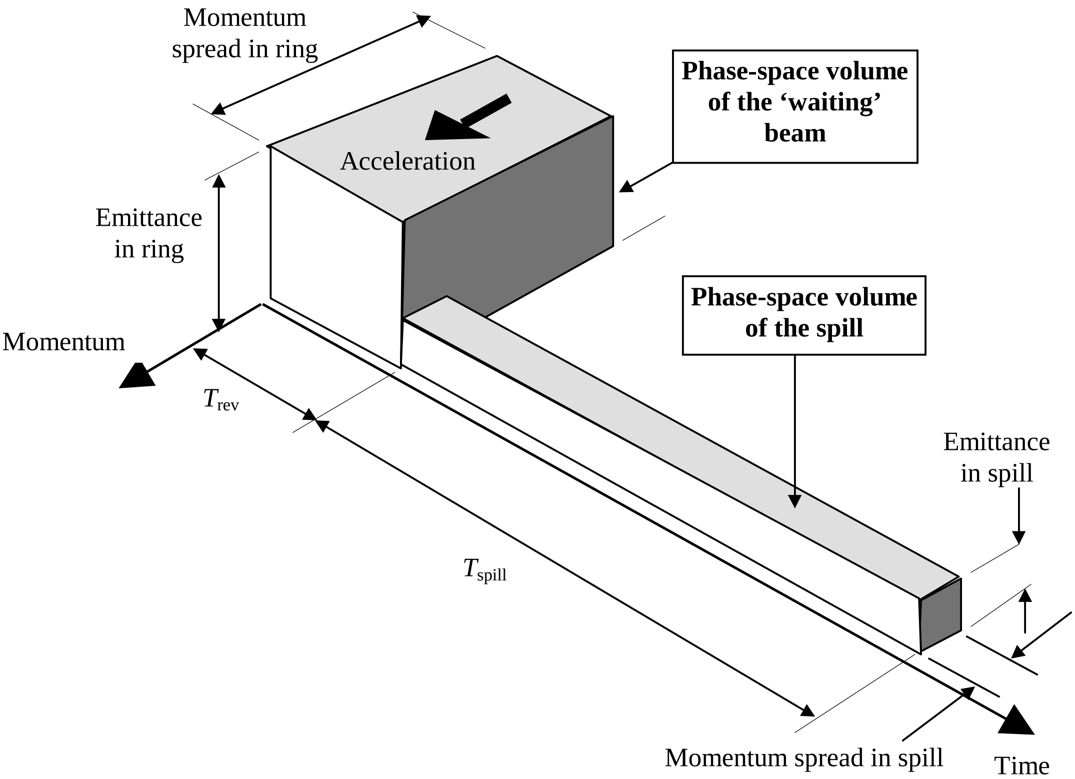

Momentum spread in ring
Phase-space volume of the 'waiting' beam
Acceleration
Emittance in ring
Momentum
Phase-space volume of the spill
T_{rev}
Emittance in spill
T_{spill}
Momentum spread in spill
Time

Figure 4.21 Schematic view of the phase-space volumes during extraction

$$E_{x,\text{stack}}\left(\frac{\Delta p}{p}\right)_{\text{stack}} T_{\text{rev}} = E_{x,\text{spill}}\left(\frac{\Delta p}{p}\right)_{\text{spill}} T_{\text{spill}}\ . \tag{4.63}$$

To give a quantitative impression, some orders of magnitude would be, $E_{x,\text{beam}} = 10\pi\times10^{-6}$ [m rad], $(\Delta p/p)_{\text{beam}} = 0.005$, $T_{\text{rev}} = 0.5\times10^{-6}$[s], $(\delta p/p)_{\text{spill}} = 0.001$, $T_{\text{spill}} = 500\times10^{-3}$s, so that $E_{x,\text{spill}} = 50\times10^{-12}\pi$ [m rad]. Thus, the long spill time compared to the short revolution time is balanced by the small transverse emittance of the spill compared to the much larger emittance of the 'waiting' beam. When dealing with small numbers of particles in irregularly shaped regions of phase space as occur in the spill, the statistical expression for the emittance, given below, is perhaps more meaningful, but the above can still be used as a guide to the extracted emittance.

$$E_x = \sqrt{\langle X^2\rangle\langle X'^2\rangle - \langle XX'\rangle^2} \tag{4.64}$$

Thus, the extracted emittance in the plane of the resonance will be extremely small. In fact, under ideal conditions (no noise and low coupling), it will be near zero.

4.6.2 The maximum emittance that can be extracted

As an illustrative exercise, it is interesting to calculate the maximum emittance that can be extracted. This is done by considering the rather impractical case of a mono-energetic beam that is exactly on the resonance tune and is suddenly subjected to the full sextupole field. The beam then finds itself instantaneously sitting on phase-space trajectories that leave the aperture as shown below. The stable triangle referred to earlier has collapsed infinitely quickly and left all the particles ‘marooned’ and unstable (see Figure 4.22).

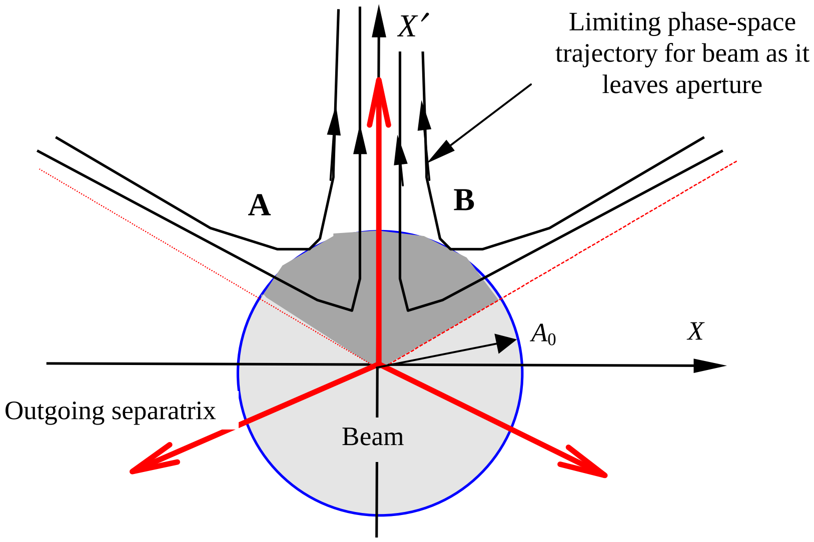

X′
Limiting phase-space trajectory for beam as it leaves aperture
A
B
A_0
X
Outgoing separatrix
Beam

Figure 4.22 Hypothetical beam sitting exactly on resonance

The phase space trajectories are given by the Kobayashi Hamiltonian (2.27),

$$\boldsymbol{H} = \frac{\varepsilon}{2}\left(X^2 + X'^2\right) + \frac{S}{4}\left(3XX'^2 - X^3\right). \tag{2.27}$$

Since it was assumed that the whole beam is exactly on resonance $\varepsilon = 0$. On the limiting phase-space trajectories, the Hamiltonian can be evaluated at the points A and B as the beam leaves the aperture.

$$\boldsymbol{H}_0 = \frac{S}{4}\left(3A_0 \cos\frac{\pi}{3} A^2{}_0 \sin^2\frac{\pi}{3} - A^3{}_0 \cos^3\frac{\pi}{3}\right) = \pm\frac{S}{4}A_0^3 \tag{4.65}$$

After the substitution of (4.65) into (2.27) the diagram in Figure 4.22 is rotated by 90°, for convenience and X becomes X' and X' becomes X. One third of the beam leaves along each separatrix (see Figure 4.23).

$$A_0^3 = \mp\left(3X'X^2 - X'^3\right) \tag{4.66}$$

The transverse emittance in the spill is found by integrating between X_0 and $X_0+\Delta R$ (the aperture of the electrostatic septum). Once the beam is as far out as the electrostatic septum $X' << X$ and the X'^3 term can be neglected. Thus,

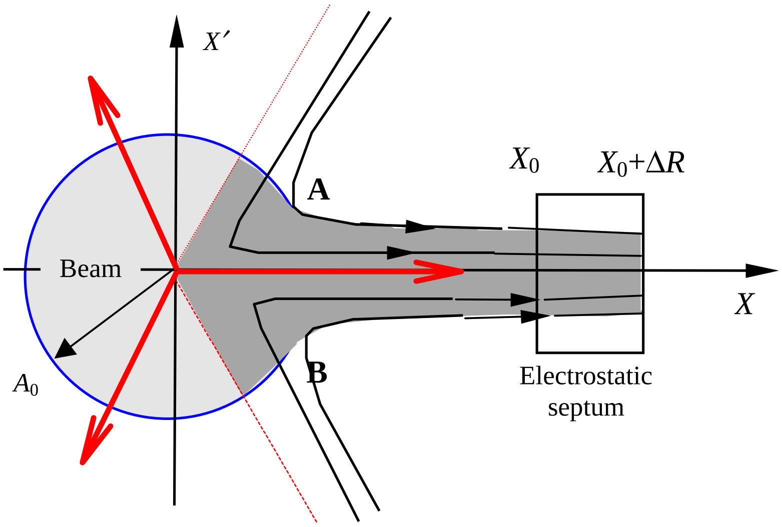


Figure 4.23 Hypothetical beam exactly on resonance being extracted exactly on resonance

$$\text{Area} = 2\int_{X_0}^{X_0+\Delta R}\left(\frac{A_0^3}{3X^2}\right)\mathrm{d}X = \frac{2}{3}A_0^3\left[\frac{1}{X_0} - \frac{1}{X_0+\Delta R}\right]. \qquad (4.67)$$

Now express this area as a fraction of the initial emittance, (πA_0^2) and make the approximation $(1+\Delta R/X_0)^{-1} = (1-\Delta R/X_0)$, to get

$$\text{Max. fraction of } E_x \text{ that can be extracted} \approx \frac{2A_0}{3\pi}\frac{\Delta R}{X_0^2}. \qquad (4.68)$$

Assuming the approximate dimensions for the spiral step (10 mm) and the position of the electrostatic septum (35 mm) given in Figure 3.19, then ΔR is $0.01/\sqrt{\beta}$ [$\mathrm{m}^{1/2}$] and X_o will be $0.035/\sqrt{\beta}$ [$\mathrm{m}^{1/2}$] or very close to these values. If A_0 is also chosen as $0.01/\sqrt{\beta}$ [$\mathrm{m}^{1/2}$] then the absolutely maximum emittance that can be extracted under these conditions is just 1.7% of the original beam. Thus, it appears that whatever is done there will be a strong asymmetry in the beam emittances.

References

[1] M. Pullia, *Transit time for third order resonance extraction*, CERN/PS 96-36 (DI).
[2] M. Pullia, *Time profile of the slowly extracted beam*, CERN/PS 97-50 (DI).

* * *

I-5 RIPPLE

There are at least two established conventions for quantitatively expressing the quality of a spill. The CERN convention uses a parameter of merit called the *Duty Factor*, which has the advantage of being analytic in form, but its name is an unfortunate accident of history and it should not be confused with the more widely used ‘Duty Factor’ and ‘Duty Cycle’ that describe the ratio of the ‘on’ to the ‘off’ time for linacs, etc. In addition, it does not give a very intuitive picture of the spill quality. The second convention is the GSI convention that is linked directly to the needs of an active scanning system and has the advantage that it gives a clearer picture of the quality. A third model is also used here that quotes equivalent sinusoidal modulation.

5.1 CONVENTIONS

5.1.2 Duty Factor, *F*

The irregularities of the spill are often quantified by a *Duty Factor, F,*

Definition:

$$F = \frac{\langle \phi \rangle^2}{\langle \phi^2 \rangle}, \text{ over time, } T_{\text{spill}} \tag{5.1}$$

where $\phi = \mathrm{d}N/\mathrm{d}t$, the particle flux in the spill (see Figure 5.1). In cases where it is more suitable to think of a continuous flux, ϕ will be used and (5.1) will be expressed in integral form. In cases where the granular nature of the beam is more evident, $\mathrm{d}N/\mathrm{d}t$ will be used and (5.1) will be expressed in differential form.

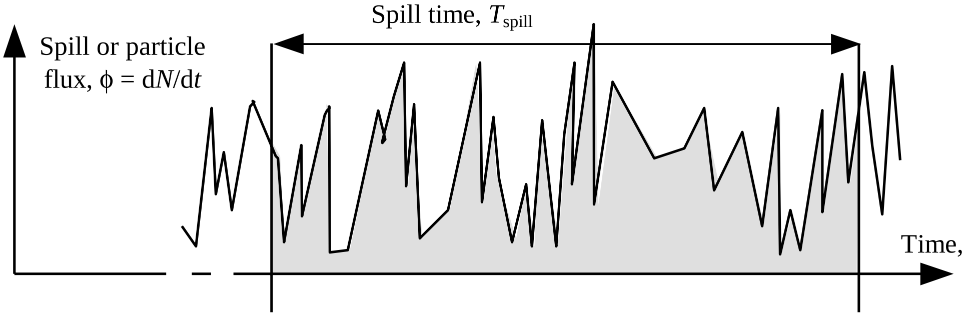

Spill time, T_{spill}
Spill or particle flux, $\phi = \mathrm{d}N/\mathrm{d}t$
Time, t

Figure 5.1 Particle spill

5.1.2 Spill Quality, *Q*

An alternative figure of merit used for a spill is the *spill quality, Q*, defined as

$$Q = \frac{\overset{\frown}{\phi}}{\overline{\phi}}, \text{ over time, } T_{\text{spill}} \tag{5.2}$$

where $\overset{\frown}{\phi}$ and $\overline{\phi}$ are peak and average values of the spill rate respectively.

5.1.3 Sinusoidal modulation of a spill

Unfortunately, the image conjured up by a certain duty factor of say 0.97 is not very clear, so to get some idea of what this means, consider a sinusoidal modulation of the spill,

$$\phi = \phi_0 + \phi_r \cos(\omega t). \tag{5.3}$$

where ϕ_0 is the unmodulated level of the spill and ϕ_r is the modulation amplitude. The duty factor in integral form is then,

$$F = \frac{\left\{\frac{1}{T_s}\int_{T_s}[\phi_0 + \phi_r \cos(\omega t)]\mathrm{d}t\right\}^2}{\frac{1}{T_s}\int_{T_s}[\phi_0 + \phi_r \cos(\omega t)]^2\,\mathrm{d}t}.$$

Duty Factor, $$F = \frac{\phi_0^2}{\phi_0^2 + \frac{1}{2}\phi_r^2} = \frac{1}{1 + \frac{1}{2}\frac{\phi_r^2}{\phi_0^2}}. \tag{5.4}$$

The evaluation of (5.4) for a range of sinusoidal modulations is tabulated in Table 5.1 to give some impression of the meaning of the duty factor. Although the duty factor does not easily invoke a mental picture of the spill quality it does have the advantage of being able to describe distributions in an analytical way.

Duty Factors for a sinusoidally modulated signal							
ϕ_r/ϕ_0	100%	80%	50%	20%	10%	5%	1%
F	0.67	0.76	0.89	0.98	0.995	0.999	0.99995

Table 5.1 Duty factors for a sinusoidal modulated signal

5.2 OVER-MODULATION OF A SPILL

If the modulation of the spill is greater than 100%, the spill will be intermittent, or 'chopped' and, in the extreme, it will become a series of spikes. Whatever the modulation level, the average number of particles leaving the resonance per cycle will remain the same, but the integral over the square of the spill intensity will change dramatically. Let the form of the spill be described by,

$$\phi = \phi_0 + \phi_r \cos(\omega t) \tag{5.3}$$

as before, except that now $\phi_r > \phi_0$. The ripple will clear a space between the beam and the resonance and the spill will only be active during the peaks of the oscillations (see Figure 5.2). The average flux rate in the spill will still be ϕ_0, but it will be concentrated in the spikes.

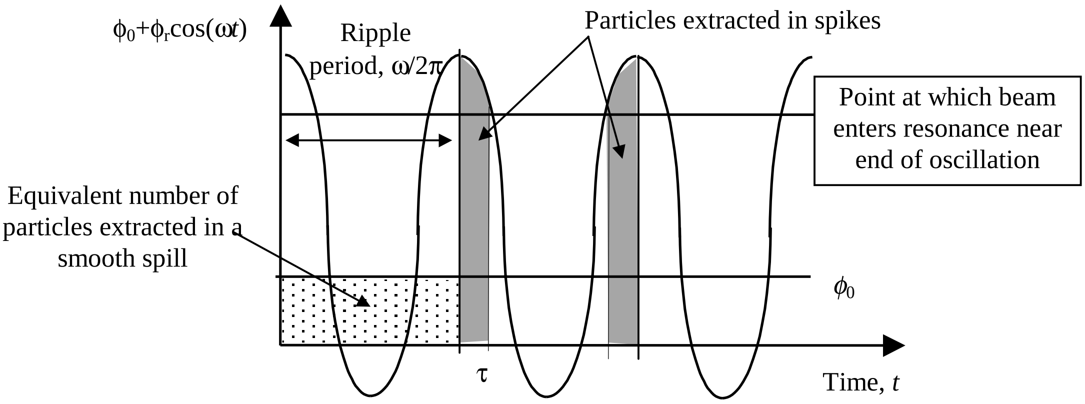

ϕ0+ϕrcos(ωt)
Ripple period, ω/2π
Particles extracted in spikes
Point at which beam enters resonance near end of oscillation
Equivalent number of particles extracted in a smooth spill
ϕ0
τ
Time, t

Figure 5.2 Formation of spikes in an over-modulated spill

The integrals needed for the duty factor are

$$\langle\phi\rangle^2 = \left\{\left(\frac{\omega}{2\pi}\right)2\int_0^\tau\left(\phi_0+\phi_r\cos(\omega t)\right)\mathrm{d}t\right\}^2 \qquad \text{(a)}$$

$$\langle\phi^2\rangle = \left(\frac{\omega}{2\pi}\right)2\int_0^\tau\left(\phi_0+\phi_r\cos(\omega t)\right)^2\mathrm{d}t \qquad \text{(b)} \qquad (5.6)$$

where the integration limit, τ, is found by equating the beam extracted in the peaks during one cycle (see Figure 5.2) to the beam extracted in a smooth spill. This leads to the relation,

$$\frac{2\pi}{\omega}\phi_0 = 2\int_0^\tau\left(\phi_0+\phi_r\cos(\omega t)\right)\mathrm{d}t$$

$$\pi-\omega\tau = \frac{\phi_r}{\phi_0}\sin(\omega\tau).$$

The integration of equations (5.6) (a) and (b) is straightforward. Equation (5.6)(a) reduces to ϕ_0^2 as expected and the duty factor becomes

$$F = \frac{\langle\phi\rangle^2}{\langle\phi^2\rangle} = \frac{\pi}{\omega\tau+2\dfrac{\phi_r}{\phi_0}\sin\omega\tau+\dfrac{1}{2}\left(\dfrac{\phi_r}{\phi_0}\right)^2\left(\omega\tau+\sin\omega\tau\cos\omega\tau\right)}\,. \qquad (5.7)$$

where $\omega\tau$ is defined above and has to be calculated numerically. This has been done for a range of over-modulation (ϕ_r/ϕ_0) from unity to ten and the results from (5.7) and the time for which the beam is switched off by the modulation are presented in Table 5.2. The last entry in Table 5.2 corresponds to a modulation amplitude ϕ_r that is ten times the constant level ϕ_0 and the beam is cut off for ~80% of the ripple cycle. If the spike were to be approximated by a triangular shape, then the average width would be 10% of the ripple period for a height of 10 times the dc level, which agrees well with the more detailed calculation. In practice, the situation of over-modulation is more likely to occur at kHz frequencies.

Duty Factors for a sinusoidally, over-modulated signal							
ϕ_r/ϕ_0	100%	200%	300%	400%	500%	700%	1000%
$\omega\tau$	π	0.396π	0.275π	0.212π	0.174π	0.128π	0.092π
F	0.667	0.386	0.272	0.212	0.173	0.128	0.092
$\Delta t/T_p$*	0.0	0.208	0.450	0.576	0.652	0.744	0.816

* where $\Delta t/T_p$ is the fraction of the ripple period for which the beam is switched off by the over-modulation.

Table 5.2 Duty factors for a sinusoidally, over-modulated signal

5.3 POISSON STATISTICS IN A SPILL

Even if all elements in the extraction system were perfect, there would still be natural fluctuations arising from the randomness of the particle distribution in the 'waiting beam'. Once all artificial effects have been removed, these fluctuations would obey Poisson statistics (see Appendix IA) and would determine the ultimate quality of a spill. A Poisson distribution has the very special characteristic that the mean of the distribution, $\bar{N}$, equals the square of the standard deviation of the distribution, σ,

$$\left\langle \frac{\mathrm{d}N}{\mathrm{d}t} \right\rangle = \left\langle \left(\frac{\mathrm{d}N}{\mathrm{d}t} \right)^2 \right\rangle - \left(\left\langle \frac{\mathrm{d}N}{\mathrm{d}t} \right\rangle \right)^2, \text{ or } \bar{N} = \sigma^2 . \tag{5.8}$$

The duty factor can now be re-expressed for the special situation of a Poisson spill as,

$$F = \frac{\left\langle \frac{\mathrm{d}N}{\mathrm{d}t} \right\rangle^2}{\left\langle \left(\frac{\mathrm{d}N}{\mathrm{d}t} \right)^2 \right\rangle} = \frac{\left\langle \frac{\mathrm{d}N}{\mathrm{d}t} \right\rangle}{\left\langle \frac{\mathrm{d}N}{\mathrm{d}t} \right\rangle + 1} = \frac{\bar{N}}{\bar{N}+1} . \tag{5.9}$$

The Poisson nature of a spill will only become apparent under two rather extreme conditions; firstly, for very low intensity spills and, secondly, when sampling a spill at very high frequencies. To give some quantitative idea of the above, the duty factor is evaluated using (5.9) in Table 5.3 for different expectations of the number of particles per measurement bin and, for comparison, the sinusoidal modulation amplitude that gives the same duty factor is also included.

Duty Factors for Poisson and sinusoidally modulated spills								
Particles per bin	1	2	3	4	10	20	100	250
Duty factor, F	0.5	0.67	0.75	0.80	0.91	0.95	0.99	0.996
Equiv. mod.ϕ_r/ϕ_0	-	100%	82%	35%	44%	32%	14%	9%

Table 5.3 Duty factors for Poisson and sinusoidally modulated spills

The spill shown in Figure 5.3 was recorded at GSI, Darmstadt. The spill is measured in 30 μs bins and the expected average spill rate was 15 particles per bin. From the above, the best possible spill quality would be $F = 0.9375$, which is equivalent to an amplitude modulation of only 36.5%. Clearly, the modulation is in excess of this estimate, so there is more than Poisson statistics at work in this spill. The spike structure indicates that there is strong over modulation in the kHz frequency range.

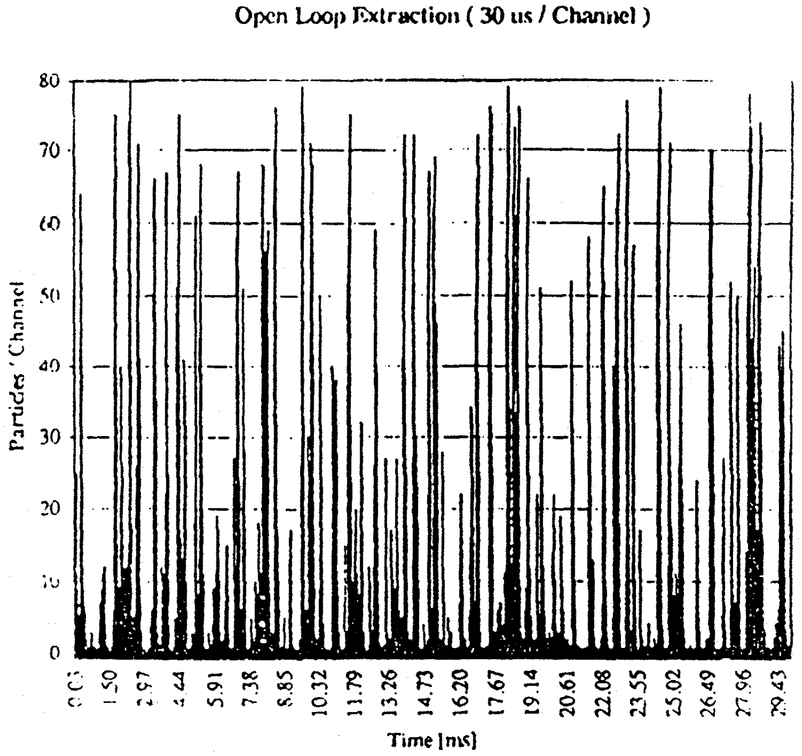


Figure 5.3 Slow extracted spill (Courtesy of GSI Darmstadt)

5.4 DUTY FACTOR AND FREQUENCY

Since all beam spills contain discrete particles, it is possible (in theory at least) to sample the beam at such a high frequency that the Poisson statistics become apparent. The duty factor will then deteriorate steadily as the frequency rises. This, however, is more of academic interest. In practice, frequencies in the spill above the sampling frequency (10 kHz) become increasingly unimportant. This statement is based on two approximate calculations:

- ***For voxel scanning*** the nominal time to fill a voxel is 5 ms, which would be 50 bins at a sampling frequency of 10 kHz. An error of one bin then corresponds to the nominal ±2% specification for precision in dose uniformity.

- ***For raster scanning***, the finite spot size and scanning speed of the spot combine such that any point will ‘see’ the beam for a nominal 5 ms. As before, this corresponds to 50 bins at a sampling frequency of 10 kHz.

Thus, it is more correct to define the duty factor for frequencies up to a certain maximum frequency, or for measurement bins down to a minimum time lapse. The specification and performance of the medical machine will be treated in this way.

5.5 EFFECT OF TUNE RIPPLE AT LOW FREQUENCIES

A simple analysis of the effects of ripple at low frequencies can be made by assuming that the particle flux that enters the resonance from the ‘waiting’ beam appears instantaneously in the observed spill. This is a fair approximation when the transit times in the resonance are small compared to the ripple frequency. Unfortunately the transit times depend upon the circumference of the machine, the strength of the resonance, the emittance and so on. For a small medical synchrotron, the transit times typically vary from 100 turns up to 4000 turns. There will always be a few particles outside this range, since particles can, in theory, take as little as a few tens of turns right up to an ‘infinite’ number of turns to leave the machine, but the range 100 to 4000 will include the majority of the beam. Now for ease of computation, assume that most particles leave the machine within 2000 turns. For a revolution time of 0.5 μs, this represents a generous delay of 1 ms. Thus, for 50 Hz, 100 Hz and 300 Hz ripples (common in power converters), it is reasonable to say that extraction is instantaneous. With this simplified picture in mind, consider the schematic model shown below in Figure 5.4, in which the ‘waiting’ beam is being consumed by a resonance. Whether the relative motion between the beam and the resonance is due to one partner, or both, is unimportant for the moment.

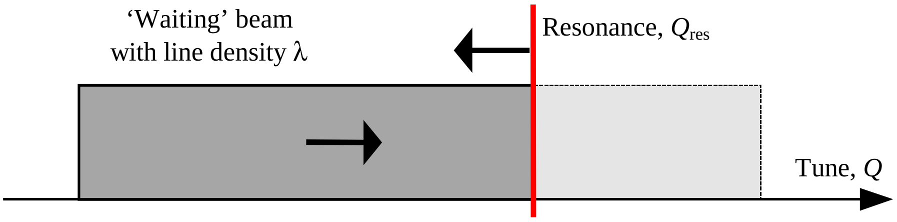


Figure 5.4 Simple model for ripple

The relative motion between the beam and resonance will comprise a constant velocity, $(\mathrm{d}Q/\mathrm{d}t)=\dot{Q}_0$ plus a ripple term that may come from either the position of the beam via the main dipole field or from the resonance via the tuning quadrupoles. Let the ripple have the form,

$$Q_r=\delta Q\sin(\omega t);\ \text{so that}\ \dot{Q}_r=\omega\delta Q\cos(\omega t). \tag{5.10}$$

The flux of particles entering the resonance will be,

$$\phi=\frac{\mathrm{d}N}{\mathrm{d}t}=\lambda\left(\dot{Q}_0+\dot{Q}_r\right)=\lambda\left(\dot{Q}_0+\omega\delta Q\cos(\omega t)\right) \tag{5.11}$$

where λ is the line density in the ‘waiting’ beam. The duty factor will be,

$$F=\frac{\left\{\dfrac{1}{T_s}\int\limits_{T_s}\lambda\left[\dot{Q}_0+\omega\delta Q\cos(\omega t)\right]\mathrm{d}t\right\}^2}{\dfrac{1}{T_s}\int\limits_{T_s}\lambda^2\left[\dot{Q}_0+\omega\delta Q\cos(\omega t)\right]^2\mathrm{d}t}\,.$$

Duty Factor for power converter ripple in the low-frequency régime (<1 kHz)

$$F = \frac{1}{1 + \frac{1}{2}\frac{(\omega\delta Q)^2}{\dot{Q}_0^2}} \tag{5.12}$$

The above assumes that $\dot{Q}_0 > \omega\delta Q$. In fact, the contrary can and very often, will exist. Once $\dot{Q}_0 \leq \omega\delta Q$, the resonance plunges periodically into the beam chopping the spill into a series of spikes. The critical ripple for this 'chopped' spill is given by,

$$(\omega\delta Q)_{\mathrm{critical}} = \dot{Q}_0 . \tag{5.13}$$

Equation (5.13) shows how delicately the resonance has to be handled. For example, the following parameters would be typical for a small machine,

$Q_{\mathrm{res}} = 1.666$ and $T_{\mathrm{spill}} = 1$ s.
Thus the tune shift to 'consume' the beam; $\Delta Q = 0.01$
And, the average tune speed, $\dot{Q}_0 = 0.01 s^{-1}$.
Relation of tune ripple to current ripple, $\Delta Q/Q = \Delta I/I$.

The relation (5.13) then shows that the spill will be 100% 'chopped' at 100 Hz for a tune ripple of approximately 10^{-5}. This translates directly into the current ripple in the quadrupoles[*], which represents a tight specification.

The critical values of the ripple at which the spill changes from being 100% modulated ($F = 0.67$) to being separate spikes are given in Table 5.4 for frequencies up to 1 kHz. The striking feature is the extremely low values of ripple that are required to stop the spill from being 'chopped'. The levels of ripple needed to ensure a reasonable duty factor (say better than 0.98 i.e. 20% modulation) would have to be 5 times lower still. This level of stability is not easily, or cheaply, achieved and clearly some form of protection other than the ripple filter of the converter itself is needed, if the spill is not to be chopped at frequencies above 100 Hz.

Conditions for 100% modulation of the spill at low frequencies

Frequency [Hz]	$(\delta Q/Q_{\mathrm{res}})_{\mathrm{critical}} = \Delta I/I$
50	2×10^{-5}
100	10^{-5}
300	3.2×10^{-6}
1000	9.6×10^{-7}

Table 5.4 Conditions for 100% modulation of the spill at low frequencies
[For $dQ_0/dt = 0.01$, $Q_{\mathrm{res}} = 1.666$]

---

[*] For the PIMMS design $\Delta Q/Q = 1.3\ \Delta I/I$.

5.6. EFFECT OF TUNE RIPPLE AT MEDIUM FREQUENCIES

As the stable triangle for a certain momentum slowly collapses, it maroons a narrow strip of particles in phase space outside the stable region after each turn. The line density of the particles in this strip will depend upon the position with respect to the three stable fixed points. In the region of the fixed points, the particle velocities are low and consequently the particle densities are high. The net result is;

- About 50% of the particles in the elementary strip leave from near a stable fixed point and will be the first to emerge in the spill in the form of a narrow spike.
- The remaining particles are spread, more or less evenly over the rest of the spill time. As a rule of thumb, the spill lasts as long again, as it took for the first particles to emerge.

In general, there will be a range of momenta with different oscillation amplitudes contributing to the spill, Thus at any given instant the spill will contain many elementary strip spills of the type described above. Large amplitude particles are quicker to emerge than small amplitude particles. The net result is that

- A spill will contain concentrations of ‘bands’ of elementary ‘strip’ spills.

The analysis of ‘strip’ and ‘band’ spills that form the basis of the simple model described above is given in Chapter 4.

Consider first an elementary strip spill as illustrated in Figure 5.5. For a small machine, the revolution time is in the range 0.5-1.0 μs and the number of turns before the particles appear (and also the length of the spill) is typically 1000-2000 turns, i.e 0.5-2 ms. Thus frequencies of typically 1 kHz and above will be affected by the transfer function of the resonance and the plateaux in the elementary spills will overlap, but because the leading peaks are narrow (~10 turns) they remain separate. For the leading peaks to overlap and smooth out, the frequency must be above 100 kHz.

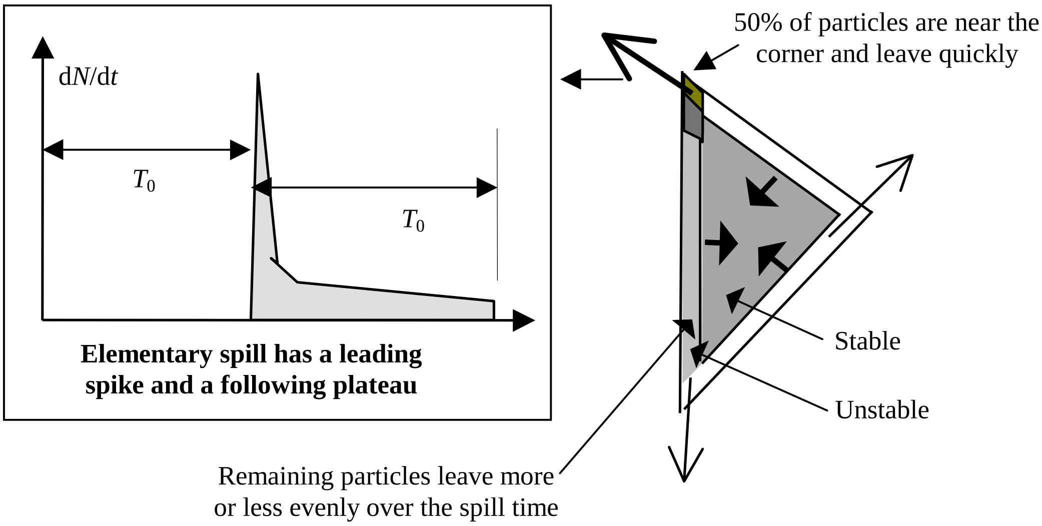

dN/dt
T_0
T_0
Elementary spill has a leading spike and a following plateau
50% of particles are near the corner and leave quickly
Stable
Unstable
Remaining particles leave more or less evenly over the spill time

Figure 5.5 An elementary ‘strip’ spill

The separation of an elementary spill into a narrow spike and a plateau can be incorporated into the simple ripple theory given in the last Section. The particles entering the resonance are, as before, equal to the line density multiplied by the relative velocity between the beam and the resonance.

Particles in; $$\left(\frac{\mathrm{d}N}{\mathrm{d}t}\right)=\lambda\left(\dot{Q}_0+\omega\delta Q\cos(\omega t)\right).$$ From (5.11)

$\downarrow$

Resonance delays all particles and then ejects
50% in a slow spill and 50% in a spike

$\downarrow$ $\downarrow$

$$\left(\frac{\mathrm{d}N}{\mathrm{d}t}\right)=\frac{\lambda}{2}\dot{Q}_0 \qquad \left(\frac{\mathrm{d}N}{\mathrm{d}t}\right)=\frac{\lambda}{2}\left(\dot{Q}_0+\omega\delta Q\cos(\omega t)\right)$$

The combination of these two terms can be used to evaluate the overall duty factor as,

$$F=\frac{\left\{\frac{1}{T_s}\int_{T_s}\left\{\frac{\lambda}{2}\dot{Q}_0+\frac{\lambda}{2}\left[\dot{Q}_0+\omega\delta Q\cos(\omega t)\right]\right\}\mathrm{d}t\right\}^2}{\frac{1}{T_s}\int_{T_s}\left\{\frac{\lambda}{2}\dot{Q}_0+\frac{\lambda}{2}\left[\dot{Q}_0+\omega\delta Q\cos(\omega t)\right]\right\}^2\mathrm{d}t}.$$

Duty Factor for power converter ripple in the medium-frequency régime (1-100 kHz)

$$F=\frac{1}{1+\frac{1}{4}\frac{(\omega\delta Q)^2}{\dot{Q}_0^2}} \tag{5.14}$$

The extra factor of two compared to the low-frequency formula (5.12) is very welcome, but of limited help. In fact, it only enters in the amplitude of the ripple as the square root of 2. Overall, the situation is worse than for the low-frequency régime for the same $\dot{Q}_0$, since the increase in ω is squared. The critical value of the ripple at which the spill changes from being modulated (F = 0.67) to being chopped is in Table 5.5 for frequencies between 1 and 100 kHz.

Conditions for 100% modulation of the spill at medium frequencies

Frequency [kHz]	$(\delta Q/Q_{\mathrm{res}})_{\mathrm{critical}}=\Delta I/I$
1	6.8×10^{-7}
10	6.8×10^{-8}
100	6.8×10^{-9}

Table 5.5 Conditions for 100% modulation of the spill at medium frequencies
[For $\mathrm{d}Q_0/\mathrm{d}t$ = 0.01, Q_{res} = 1.666]

5.7. EFFECT OF TUNE RIPPLE AT HIGH AND ULTRA-HIGH FREQUENCIES

Above 100 kHz the widths of the leading peaks of the elementary spills will also begin to overlap, which will finally smooth the ripple effects, but at 1 MHz and above the sampling rate becomes high enough to see the influence of Poisson statistics.

Duty Factor for power converter ripple in the high-frequency régime (>1MHz)

$$F = \frac{\bar{N}}{\bar{N}+1} \tag{5.15}$$

For therapy with carbon ions the lowest spill rate would be of the order of 5×10^6 particle/s. At a 1 MHz sampling rate the expected ‘granularity’ of the beam (5 particle/bin) would cause a minimum duty factor of 0.833, which would be equivalent to a sinusoidal spill modulation of 63%.

A reasonable sampling frequency for hadrontherapy is 10 kHz, which means that ‘high frequencies’, as defined here would be invisible to the measuring system. Even if the sampling frequency were to be increased, to alleviate problems due to the saturation of counters for example, some fluctuation might then be visible in the measurements but the relatively long integration times for the dose would make these fluctuations unimportant.

5.8 SUMMARY OF THE RIPPLE RÉGIMES

Low-frequency	**Medium frequency**	**High frequency**
0-1 kHz Modulation regime	1-100 kHz Over-modulation with partial smoothing	>100 kHz Poisson
$F = \frac{1}{1+\frac{1}{2}\frac{(\omega\delta Q)^2}{\dot{Q}_0^2}}$	$F = \frac{1}{1+\frac{1}{4}\frac{(\omega\delta Q)^2}{\dot{Q}_0^2}}$	$F = \frac{\bar{N}}{\bar{N}+1}$

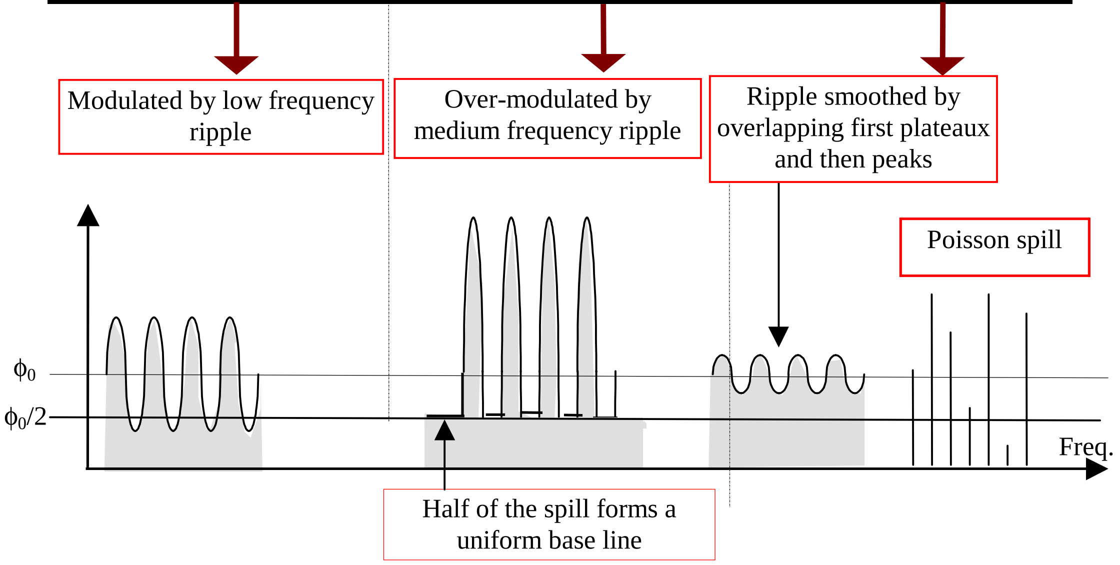

Modulated by low frequency ripple
Over-modulated by medium frequency ripple
Ripple smoothed by overlapping first plateaux and then peaks
Poisson spill
ϕ0
ϕ0/2
Freq.
Half of the spill forms a uniform base line

Figure 5.6 Summary of the spill régimes

Since many ripple frequencies will be present in the beam, there may also be enhancement of the lower frequencies by the higher ones. Since the average relative motion of the beam and resonance is unchanged by ripple, the average spill rate will be unchanged, so the combination of a high-frequency ripple with a lower one will chop the low-frequency modulated spill into bursts of taller spikes. This is illustrated qualitatively below.

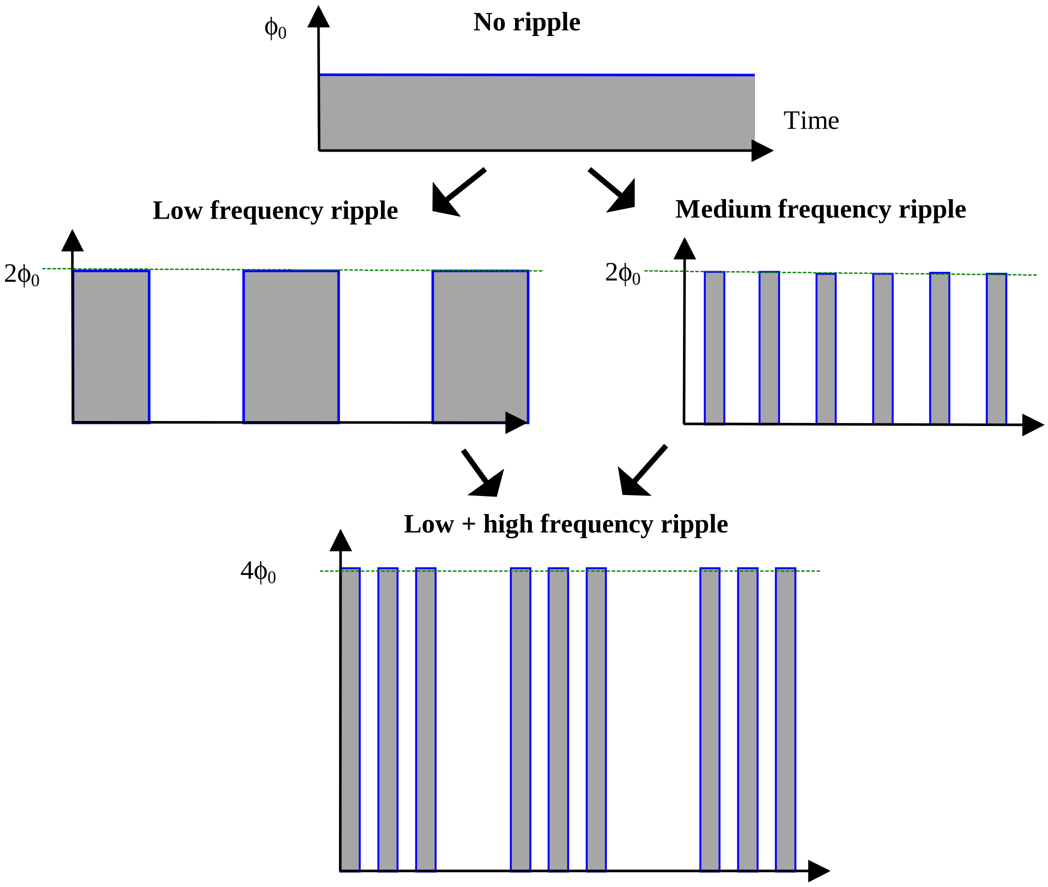


Figure 5.7 Interaction of the a low and higher frequency ripple

5.9 DEFENCE AGAINST TUNE RIPPLE

Extremely good quality power converters are needed

Whether it is better to choose conventional or switch-mode units must be decided case by case. If the switch-mode frequency can be placed above 20kHz, then the preference is for switch mode, but this is not always a practical proposition and depends on the power to be delivered and whether a bipolar supply is needed. It is also necessary to install DACs[*] with the highest number of bits possible (see Section 5.10).

Increase of $\dot{Q}_0$

The next step would be to increase the velocity of the particles as they enter the resonance. This can be done in a number of ways and will be described in Chapter

---

[*] DAC = Digital to Analogue Converter.

6. This is effective at low frequencies, but for higher frequencies the increase in ω will eventually outstrip the gain from the enhanced velocity.

The amplitude-momentum extraction technique

At medium frequencies, there is a distinct advantage in using the momentum-amplitude selection technique for entering the resonance as described in Chapter 4. This technique is intrinsically smoother in the most difficult frequency range of a few kHz.

On-line scanning control

The beam delivery system with the on-line dosimetry system is the last defence against a poor quality spill. If it is unable to cope with the fluctuations, then the spill intensity has to be reduced and the treatment time extended. Conversely, if the spill quality is good, the beam intensity can be raised and the treatment time reduced.

5.10 RAMPING POWER CONVERTERS

Conventional power converters are controlled via a DAC and the setting precision is determined by the number of bits in the DAC (see Table 5.6). A 12-bit DAC is commonplace, the 16-bit DAC was developed during the 1970s at the ISR [1] and is now available commercially. At the present time, the 18-bit DAC is still more of a development device than a commercial reality.

DAC precision			
DAC	12-bit	16-bit	18-bit
Precision	1 in 4’096	1 in 65’536	1 in 262’144

Table 5.6 DAC precision

To illustrate the potential problem posed by the DAC, consider that the resonance is being moved into the ‘waiting’ beam by a tune change driven by a series of tuning quadrupoles. Let the current range for the tuning quadrupoles be 50% of maximum and let the spill last 1 s for simplicity. The heights of the current steps and their frequencies are recorded in Table 5.7 for different DACs.

DAC frequencies and steps			
DAC	12-bit	16-bit	18-bit
No. of bits sent for spill (50%)	2’048	32’768	131’072
Bit frequency [kHz]	2	33	131
Height of step $\Delta I/I_{max}$	2.5×10^{-4}	1.5×10^{-5}	4×10^{-6}

Table 5.7 DAC frequencies and steps

The bit frequency should be compared to the sampling frequency of the on-line dose measurement system (typically 10 kHz). Frequencies below the sampling frequency are highly dangerous. Frequencies of 4-5 times this frequency are probably irrelevant, but the effect of those 2-3 times this frequency depends on their amplitude. To give some idea of the importance of the amplitude of current steps, Table 5.8 has been constructed from Tables 5.4 and 5.5 and the ripple formulae derived in the earlier sections of this Chapter. Table 5.8 lists the ripple amplitudes that just cause 'chopping' (100% modulation, i.e. $F = 0.67$) over a wide frequency range in an unprotected machine. This table is NOT a specification, it is only meant to give some feeling for the extreme sensitivity of the resonance to ripple.

Conditions for a 'chopped' beam in an unprotected machine
(F = 0.67 i.e. 100% modulation)

Frequency [Hz]	**$(\delta Q/Q_{\text{res}})_{\text{critical}}$ ***
50	2×10^{-5}
100	10^{-5}
300	3.2×10^{-6}
1000	9.6×10^{-7} (6.8×10^{-7})**
10'000	6.8×10^{-8}
100'000	6.8×10^{-9}

* For $dQ_0/dt = 0.01$, $Q_{\text{res}} = 1.666$. For PIMMS, $\delta Q/Q = \Delta I/I$ within a factor 2

** The first figure uses the low-frequency formula (5.12) and the second uses the medium-frequency formula (5.14).

Table 5.8 Conditions for a 'chopped' beam in an unprotected machine

The relative step height ($\Delta I/I_{\text{max}}$) of a single DAC bit in Table 5.7 should be compared to the critical values in Table 5.8. The relation between $\Delta Q/Q$ and $\Delta I/I$ will vary from one magnet chain to the next, but is typically close to unity. In all cases, the steps will cause 100% chopping of the beam. However, once the frequency is well above 10 kHz, whether the beam is chopped or not is unimportant. Thus, in the example chosen, a 16-bit DAC would be essential.

In the case of a dual species machine, the problem of the DAC is aggravated by the wide operational range needed to cover both protons and ions. For example, if in a proton-ion machine the ions at top energy require a current change of 100% then the lowest energy protons would only require 17%. The coarseness of the DAC then becomes very apparent and the DAC bit frequency may easily come down into the sub-kHz region.

The above considerations lead to the following recommendations:

- As few as possible power converters should be changing during the spill.
- Large inductive loads (e.g. a betatron core) are safer than light inductive loads (e.g. resonance quadrupoles).
- All stationary power converters should be 'locked' to prevent them from making single DAC steps (due to noise or drift in control circuit) during the spill.

The problem of DAC steps is one of the key reasons why the amplitude-momentum extraction technique using a betatron core to accelerate the beam has been chosen for PIMMS. This extraction technique makes it possible to maintain all ring power converters constant during the extraction, except for the betatron core. Since the betatron core is a single large inductance device it is possible to take extra care with this single power converter and with its DAC. By filtering the DAC output and applying a vector generator method that anticipates and smoothes out the DAC discontinuities, a factor of 100 can be achieved in the ripple [2].

5.11 DAMPING BY EDDY CURRENTS

The question arises as to how much of the current ripple actually appears as field ripple seen by the beam. This can be approached in the first instance in a very general way by postulating that the instantaneous derivative of a parameter is proportional to the separation from its equilibrium value and .

$$\frac{\mathrm{d}B}{\mathrm{d}t} = -\frac{1}{\tau}\left(B - B_0\right) \tag{5.16}$$

where $1/\tau$ is a constant and τ is better known as the time constant. If now the equilibrium value B_0 is replaced by a time varying equilibrium value $B_0\sin(\omega t)$, (5.16) becomes,

$$\frac{\mathrm{d}B}{\mathrm{d}t} + \frac{B}{\tau} = \frac{B_0}{\tau}\sin(\omega t).$$

This is a standard form and can be integrated to give,

$$B\mathrm{e}^{\frac{t}{\tau}} = \frac{B_0}{\tau}\int \mathrm{e}^{\frac{t}{\tau}}\sin(\omega t)\mathrm{d}t + const.$$

where the integral is again a standard form* giving,

$$\tau\frac{B}{B_0}e^{\frac{t}{\tau}} = e^{\frac{t}{\tau}}\left(\frac{\frac{1}{\tau}\sin\omega t - \omega\cos\omega t}{\frac{1}{\tau^2}+\omega^2}\right) + const.$$

For boundary conditions, set $B = 0$ at $t = 0$, then $const. = \left(\frac{\omega\tau}{1+\omega^2\tau^2}\right)$ and

$$B = \underbrace{B_0\left(\frac{\sin\omega t - \omega\tau\cos\omega t}{1+(\omega\tau)^2}\right)}_{\text{Steady}-\text{state term}} + \underbrace{B_0 e^{-t/\tau}\frac{\omega\tau}{1+(\omega\tau)^2}}_{\text{Transitory term}}. \tag{5.17}$$

---

* $I = \int \mathrm{e}^{at}\sin(bt)\mathrm{d}t = \mathrm{e}^{at}\left(\mathrm{a}^2 + b^2\right)^{-1}\left(a\sin(bt) - b\cos(bt)\right).$

The second term in (5.17) is transitory and will be put to zero. The first term in (5.17) can then be written as,

Field with damping:

$$B = B_0 \frac{1}{\sqrt{1+(\omega\tau)^2}} \sin(\omega t - \varphi) \tag{5.18}$$

where the phase constant φ equals $\tan^{-1}(\omega\tau)$. Thus, the field will lag behind the sinusoidal current with a phase angle of $\tan^{-1}(\omega\tau)$ and will be attenuated by the square root term in the denominator of (5.18), unless the time constants of the magnet and vacuum chamber are zero and then the attenuation and phase lag will also be zero.

The time constants of the magnets and vacuum chambers in the ring therefore have two conflicting requirements:

- Small time constants for low field distortion during ramping to the extraction energy.
- Large time constants to damp current ripple.

For the main ring magnets, the choice must lie with the field distortion during ramping, but for the betatron core the choice can be tuned for damping the ripple and, in particular, the damping of the frequencies introduced by the DAC steps.

Anticipating the results of the PIMMS design, a main dipole with its vacuum chamber will have a time constant of the order of 100 μs. Table 5.9 lists the attenuation factors to be expected with this time constant for ripple frequencies up to 10 kHz. It can be seen that at low frequencies (<1 kHz) the attenuation is negligible, but between 1 and 10 kHz the attenuation factor increases steadily to a useful value of 6.4 at the top frequency.

Attenuation factors for ripple in the PIMMS dipole

Time constant [μs]	100	100	100	100
Ripple frequency [Hz]	500	1000	5’000	10’000
Attenuation factor	1.05	1.2	3.3	6.4

Table 5.9 Attenuation factors for ripple in the PIMMS dipole

The main PIMMS quadrupole has a time constant closer to 50 μs and a very low field distortion during ramping. In this case, 3 mm thick laminations (rather than 1.5 mm) would bring its time constant up to that of the dipole and gain some smoothing. However, 3 mm laminations are difficult to stamp and the eddy current calculation is approximate. Thus, it might be more reasonable (both for the quadrupole and dipole) to add an ‘eddy-current shield’ to the vacuum chamber. This could then be adjusted experimentally to give a field quality that is just sufficient for ramping, while giving maximum damping for ripple. In the case of the betatron core, it is probably better to adapt the lamination thickness since the field quality is not an issue.

5.12 SPILL SPECIFICATION FOR VOXEL SCANNING

5.12.1 Basic strategy

In this section, a specimen specification for spill uniformity will be developed based on the *voxel* method of scanning [3]. In Chapter 11, similar specifications will be derived for the *mini-voxel* [4] and *'true' raster* scanning techniques.

In all three methods the particle flux will be monitored as it enters the patient and the dose will be controlled from this measurement. This can be done with an ionisation chamber working at 10 kHz or higher. It is assumed that measurements are possible down to a few 10^4 particle/bin.

The overall uniformity (precision) of the treatment plan is requested to be ±2.5%. This request will be interpreted in the strict sense as ±2.5% relative to the current dose rate and not ±2.5% of the maximum dose rate in the tumour. This will be implemented by:

- Tailoring the spill intensity and the dwell time for each voxel so that with nominal parameters the slightly better target precision of ±2% is obtained.
- A coarse variation of the spill intensity will be obtained by injecting more or less current in the machine. This will enable the average spill for a slice of the tumour to be matched to the average dose level for that slice.
- A finer adjustment will be obtained by varying the ramp rate of the betatron core.
- However, the ultimate adjustment must be made by the scanning system itself.

It will be clear from the following analysis that if the on-line measurement system demands a higher dose and the spill rate is left unchanged, then the relative precision improves, whereas, if the demand is for a lower dose the relative precision deteriorates. The aim will be to match the spill rate by the methods mentioned above to the prescribed dose closely enough that the margin of 0.5% between the specification and the target precision is not exceeded when the system calls for a lower dose. Within a given tumour the average dose levels required between distal and proximal slices may vary by a factor of 50 and, in practice, it may sometimes be difficult to reduce the spill intensity sufficiently to ensure the full relative precision for very small doses. However, in these cases, the absolute error will be so small that the relative error will only be of academic interest.

5.12.2. What duty factors are needed?

Nominal parameters assumed for a voxel system

In voxel scanning, the beam steering between voxels is carried out while the beam is switched off, so only the measurement 'quantisation' coming from the period of the sampling frequency[*] and the delay in cutting off the beam are of importance when estimating the dose uniformity. Consider a system with the following parameters:

---

[*] This is referred to as the 'clock' error in Chapter 11.

- A sampling frequency of 20 kHz.
- Adjustment of the spill intensity so that a nominal voxel requires 5 ms of beam time, i.e. 100 measurement bins of 50 µs.
- The maximum variation in the beam charge to be expected in one bin is 0 to 200% of the nominal value (i.e. 100% modulation).
- The order to switch-off the beam is given once the dose exceeds 98% of the desired dose.
- The time delay for switching off the beam is assumed to be 1 measurement bin of 50 µs.

The precision obtained is then explained by the following ‘worst-case’ scenarios:

- Let the previous measurement bin show an integrated dose of just less than 98%.
- The next bin can reach any value up to 100% (i.e. twice the expected value added).
- Now the instruction to switch off is given and one more time bin will pass.
- The final integrated dose can then rise to a maximum of 102%.
- Alternatively, let the previous measurement bin be fractionally over 98% and let several of the following bins be empty. In this case, the dose will be just 98%.
- Thus, for the assumptions given the ±2% precision can be achieved.

This scenario can be generalised by assuming that a maximum of one time bin is needed to detect a threshold value and that one more time bin is needed to switch off the beam. The precision of the voxel model is then contained in the relations,

$$\text{Precision}\,[\%] = \pm\left(\frac{100}{Nb_{\text{nom}}} + \frac{100A_{\text{r}}}{Nb_{\text{nom}}A_0}\right) \tag{5.20}$$

$$\text{Threshold}[\%] = \left(100 - \text{Precision}[\%]\right) \tag{5.21}$$

where Nb_{nom} is the nominal number of bins needed to fill the voxel and (A_{r}/A_0) is the ratio of the ripple amplitude to the dc amplitude of the spill. Equation (5.20) consists of two terms. The first term is independent of ripple and gives the ‘clock’ quantisation error and second adds the precision lost due to ripple. Thus, with the nominal 100 bins, the nominal spill rate and 100% modulation (i.e. 0-200% charge in a bin) the ±2% precision is obtained. If the nominal values were changed so that only 60 bins were needed to fill the voxel, then by (5.20), either the modulation would have to be reduced to 20% (i.e. 80-120%) to achieve the same precision of ±2%, or the precision would have to drop to ±3.33% and the threshold for cutting off the beam would be at 96.66% of the desired dose. If the nominal number of bins needed to fill the voxel is 50 and the precision is maintained at ±2%, then the ripple must be zero.

Equation (5.20) can also be used to evaluate the margin that the 0.5% (between the specification and the target precision mentioned earlier) gives in the spill rate. With 100% modulation, an error of 2.5% is obtained with 80 bins. Thus, the margin in the spill rate is +20%, i.e.,

Nominal spill rate should fill 1 voxel in 5 ms (-1 ms, +unlimited [*]). (5.22)

The above is rather general and a more detailed specification needs to be derived by combining sinusoidal modulations over the full frequency range.

---

[*] The system can always accept lower spill rates, or periods of no spill, because this only means waiting for more beam and it does not cause an error in the dose uniformity.

Low frequencies

A low frequency is defined as one where the half period is less than the voxel filling time of 5 ms. Thus, 100Hz would be the low frequency limit in this case. The effect of low frequency ripple is to cause a general increase (or decrease) of the spill intensity during the voxel filling time. This can be interpreted as a reduction (or increase) in the number of bins needed to fill the voxel and then equation (5.20) is applicable for calculating the precision. To keep within the ±2.5% tolerance, low frequency modulation must not exceed 20%. Note that the modulation contain several frequencies, but principally it covers the network frequency of 50 Hz. Since low frequencies are easier to control, this appears to be a reasonable limit.

High frequencies

High frequencies are above the sampling frequency. As the frequency rises above this threshold, the effect of the modulation is progressively dies away with periodic dips to zero when an integer number of ripple periods fits into one bin. The error peaks occur at integer numbers of half periods. The first and most important is at 1.5× the sampling frequency where the acceptable modulation is limited to 300%.

Medium frequencies

Medium frequencies fill the gap between the low and high regions. If single frequencies are considered, the modulation must not exceed 100%, which ensures that no bin receives more than twice the nominal charge (the original premise). However, the superposition of frequencies is important in this range. If the power converter frequencies (300, 600, 900, 1200 and 1500 Hz) and perhaps four DAC frequencies in the kHz range for changes during the extraction are considered, then nine frequencies could be mixed together. Assuming quadratic addition, this would lead to an upper limit on single-frequency modulation of 30%.

Specification

- Maximum of 20% modulation for all frequencies combined below 100 Hz. This is based on the general level of the spill and is independent of the higher frequency limits.
- Maximum single-frequency modulation of 30% in the range 100 Hz up to the sampling frequency. This assumes the random combination of up to 9 frequencies.
- Maximum of 300% modulation at one and a half times the sampling frequency and higher above.

Note that this is the specification for the spill and that this has to be converted to power converter current ripple specifications for a practical design.

References

[1] J. Pett, *Digital-Analogue Converter 16-bit*, ISR 6034.
[2] J. Bosser, *Rappel sur les systèmes linéaires échantillonnes*, SPS/ABM/Note/84-10.
[3] E. Pedroni, H. Blattmann, T. Böhringer, A. Coray, S. Lin, S. Scheib, U. Schneider, *Voxel scanning for proton therapy*, Proc. of the NIRS International Workshop on Heavy Charged Particle Therapy and Related Subjects, Chiba, Japan, (July 1991), 94-109.
[4] Th. Haberer, W. Becher, D. Schardt, G. Kraft, *Magnetic scanning system for heavy ion therapy*, Nucl. Instr. Methods in Physics Research, A330 (1993) p296-305,

* * *

I-6 'FEEDING' THE RESONANCE

How the beam is moved from the stable to the unstable region is of critical importance for the quality of the spill. The methods that can be used are described in Chapter 3 and the many of the advantages and disadvantages were discussed in Chapter 4. The focus will now be placed on the *amplitude-momentum method* of extraction and, more precisely, on the techniques that can be used to accelerate the beam into the unstable region. Accelerating the beam into the resonance has the considerable advantage that all the optical parameters of the machine can be kept constant during the spill. Another key feature of such acceleration techniques is the possibility of adding, what will be called, a 'front-end' acceleration to increase the speed in terms of the rate of change of tune with which the particles cross the separatrix. As was shown in Chapter 5, the higher the speed, the better the spill quality.

6.1 BETATRON CORE

Induction acceleration with a betatron core (as opposed to a betatron[*]) has been known for many years [1]. It has been used for extraction on a third-order resonance in Saclay [2] and, for purposes other than extraction, in Heidelberg [3]. More recently, it has been proposed to drive the slow-extraction process in a medical synchrotron [4].

A betatron core is a closed magnetic circuit in the form of a ferromagnetic ring through which the beam of a synchrotron passes. A coil wound on the ring controls the flux inside the circuit and variations in this flux induces an electric field on the axis that is 'felt' by the circulating beam and changes its kinetic energy (see Figure 6.1).

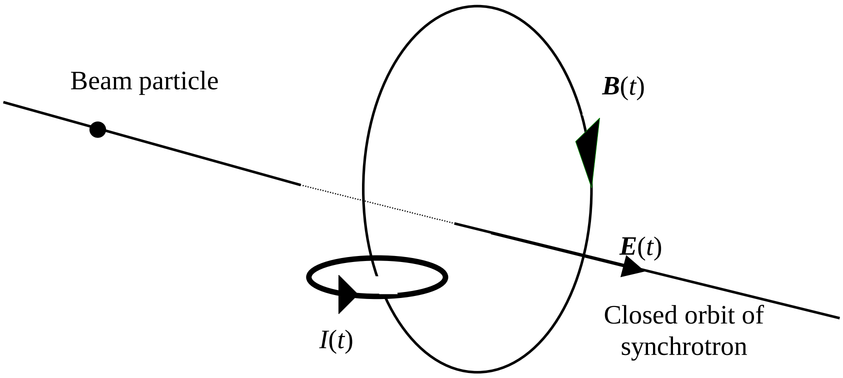


Figure 6.1 Working principle of the betatron core

The coil of the betatron core and the closed orbit of the particle act as the magnetically linked windings of a transformer. Starting from the Faraday-Neumann-Lenz law,

$$\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\frac{\partial}{\partial t} \iint \boldsymbol{B} \cdot d\boldsymbol{S} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}, \tag{6.1}$$

---

[*] The betatron core inverts the geometry of the betatron and so that the magnetic field forms a loop around the beam rather than the beam forming a loop about the magnetic field.

where Φ is the magnetic flux within the betatron core that is integrated over the cross-sectional area of the core $\boldsymbol{S}$. Now let $\overline{E}$ be the mean value of the electric field, $\boldsymbol{E}$, that is induced by the core, along the closed orbit of the beam, so that the average rate of change of momentum is given by,

$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = ze\overline{E}\ , \tag{6.2}$$

where *ze* is the net charge of the beam particle. The combination of (6.1), (6.2) and the standard relation for magnetic rigidity, $p = -zeB\rho$, yields

$$\Delta\Phi = CB\rho\frac{\Delta p}{p}\ , \tag{6.3}$$

where C is the machine circumference and $\Delta\Phi$ and Δp refer to the changes over one, or more, turns in order that the averages defined earlier have physical meanings. To obtain a uniform sweeping of the 'waiting' beam, which will be assumed to have a uniform density in $\Delta p/p$, a constant flux variation is needed during the extraction time, T_{ext},

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = CB\rho\frac{(\Delta p/p)_{\text{Waiting beam}}}{T_{\mathrm{ext}}}\ . \tag{6.4}$$

The theoretical expectation that a betatron core should give a smooth spill is supported by the experimental result shown in Figure 6.2 [5], which shows a slow extracted beam from the synchrotron SATURNE II of Saclay that uses a betatron core to extract protons at 2.4 GeV. The spill was measured with scintillators on the extraction line. The signal bandwidth, obtained by integrating the scintillator signals, is 2 kHz and the stability of the spill intensity is about ±20% (corresponding to a duty factor of 0.98).

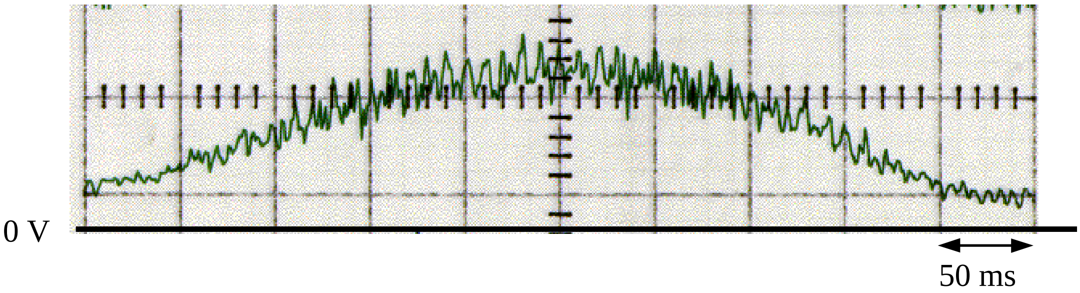


Figure 6.2 Spill extracted with the betatron core of Saturne II.
[Bandwidth of 2 kHz obtained by integrating a measurement bandwidth up to 20 kHz]

6.2 STOCHASTIC NOISE

6.2.1 The principle

Stochastic extraction was first proposed by S. van der Meer [6]. The potential of this technique as a low-ripple extraction was realised from the very beginning and is mentioned in the title of reference. The principle is to apply rf white noise over a frequency band Δf that overlaps the revolution frequencies of the 'waiting' beam, and/or harmonics of those frequencies. The noise can be applied by longitudinal kickers or rf cavities. Under the influence of the noise, the particles execute a random walk in the longitudinal phase space. This process is akin to diffusion and can be treated as such. There are three main applications,

- ***Beam shaping and homogenisation***. If the noise is rectangular in power density over the frequency band, the particles will diffuse towards a uniform distribution over that frequency band. This can be used to erase 'memories' from the beam and for making a rectangle distribution for a uniform spill (Figure 6.3).

- ***Stochastic resonant extraction***. The principle of this method consists of diffusing the particles towards and across the stability limit rather than driving them across as with conventional extraction. Particles are randomly accelerated and decelerated by noise, which causes a blow-up of the momentum distribution; the ones diffused as far as the resonance will be extracted. If the initial beam is gaussian, for example, and a particle sink (the resonance) is placed at one extremity of the frequency band, then, as the gaussian distribution spreads out, the particles will enter the resonance and be lost as illustrated in Figure 6.4. Another set-up for pure stochastic extraction is when the lower border of the noisy region is swept across the beam by slowly reducing the carrier frequency [7,8]. In this way, it is possible to vary the extracted spill length (up to 1 h, as in LEAR) and to compensate for a non-rectangular distribution of the 'waiting' beam.

- ***Combined resonant extraction***. The beam is driven towards the stability limit by another technique (for example, by using a betatron core, as described in Section 6.1) and rf noise is applied only in the vicinity of the resonance to hasten the transit of the particles into the unstable region. This so-called 'front-end' acceleration reduces the sensitivity to Q ripple (Figure 6.5).

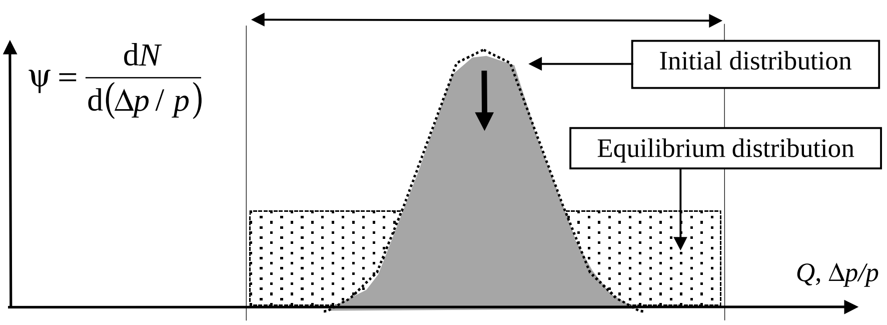

RF white noise with bandwidth Δf
$\psi = \frac{dN}{d(\Delta p/p)}$
Initial distribution
Equilibrium distribution
Q, Δp/p

Figure 6.3 Beam shaping with band-limited noise

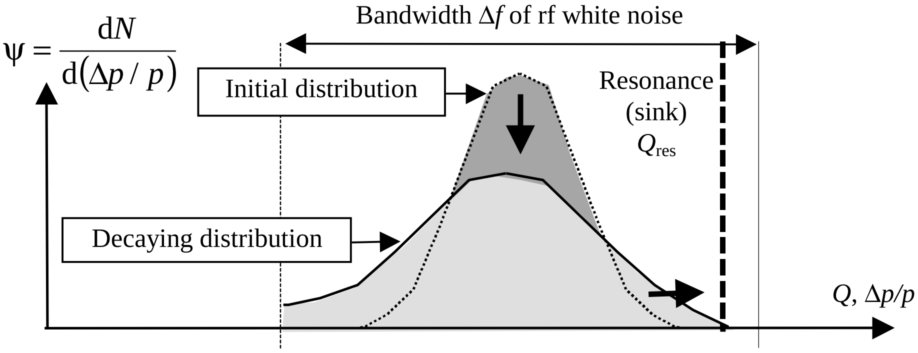


Figure 6.4 Stochastic resonant extraction with band-limited noise

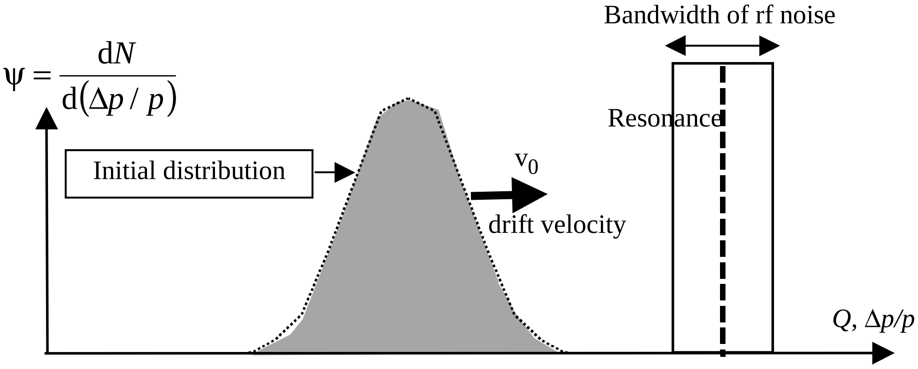


Figure 6.5 Combined resonant extraction with band-limited noise

6.2.2 Diffusion equation

The model of molecular diffusion can be used to describe stochastic diffusion under the influence of rf noise [9]. Consider the elementary cube shown in Figure 6.6 that is inside the diffusion volume. The net flow of particles will be from the high density regions to the low density regions.

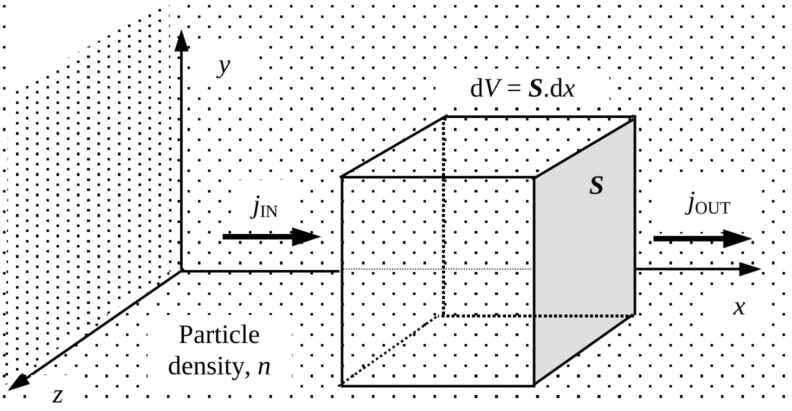


Figure 6.6 Description of molecular diffusion

Since diffusion by stochastic noise is a one-dimensional problem, this simplification will be exploited directly and it will be assumed that variations in density in Figure 6.6 only occur in the *x*-direction and that the particle densities are completely independent of the other transverse co-ordinates *y* and *z*. The accumulation rate of particles within the elementary volume, *V*, can then be related to the particle flows in the *x*-direction in and out of the volume and also to the density changes within the volume.

Accumulation in *V* by flow, $= \left(j_{\mathrm{in}} - j_{\mathrm{out}}\right) \cdot S = -\frac{\partial j}{\partial x} \cdot S \cdot \mathrm{d}x$ (6.5)

Accumulation in *V* by density, $= \frac{\partial n}{\partial t} \cdot \mathrm{d}V = \frac{\partial n}{\partial t} \cdot S \cdot \mathrm{d}x$ (6.6)

where *n* is the particle density per unit volume and *j* is the particle flow per unit area. The combination of (6.5) and (6.6) yields,

$$\frac{\partial n}{\partial t} = -\frac{\partial j}{\partial x}. \tag{6.7}$$

The flow *j* can be expressed as,

$$j = -D \cdot \frac{\partial n}{\partial x} \tag{6.8}$$

where *D* is known as the diffusion constant. The substitution of (6.8) into (6.7) gives,

Molecular diffusion equation, $\frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \cdot D \cdot \frac{\partial n}{\partial x}.$ (6.9)

The same analysis applies for stochastic noise with the substitutions and end result,

$$\boxed{\begin{aligned} &n \Rightarrow \psi; \quad x \Rightarrow \Delta p / p; \quad j \Rightarrow \phi = -D \frac{\partial \psi}{\partial(\Delta p / p)}; \\ &\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial(\Delta p / p)} \cdot D \cdot \frac{\partial \psi}{\partial(\Delta p / p)} \end{aligned} . \qquad (6.10)}$$

In the frequency domain, (6.10) has damped oscillatory solutions of the form [6],

$$\Psi(x,t) = \sum_{n=1.3.5}^{\infty} A_{\mathrm{n}} \cdot \cos\left(\frac{\pi n x}{2\Delta Q}\right) \exp\left[-\left(\frac{\pi n}{2\Delta Q}\right)^{2} Dt\right] \tag{6.11}$$

where the harmonic coefficients can be found by

$$A_{\mathrm{n}} = \frac{2}{\Delta Q} \int_{0}^{\Delta Q} \Psi(x,0) \cos\left(\frac{\pi n x}{2\Delta Q}\right) \mathrm{d}x. \tag{6.12}$$

Whereas in the time domain, (6.10) has solutions of the form [8]

$$\Psi(x,t)=\frac{1}{2\sqrt{\pi Dt}}\int\Psi(\rho,0)\exp\left(-\frac{(x-\rho)^2}{4Dt}\right)\mathrm{d}\rho\,. \tag{6.13}$$

If, for example, at $t = 0$ all N particles are concentrated at the origin ($\Psi(x,0)=\delta(x)$), the solution is then a Gaussian distribution

$$\Psi(x,t)=\frac{N}{2\sqrt{\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right). \tag{6.14}$$

If both drift and diffusion take place at the same time (for example, the combined resonant extraction), the total particle current is the sum of diffusion and drift current

$$j=-D\cdot\frac{\partial\Psi}{\partial x}+v_0\cdot\Psi \tag{6.15}$$

where v_0 is the drift velocity[*]. The Fokker-Planck equation, which describes the particle density in case of combined extraction, is obtained by substituting (6.15) in (6.7) and replacing n with Ψ as before.

$$\frac{\partial\Psi}{\partial t}=\frac{\partial}{\partial x}\cdot\left(D\cdot\frac{\partial\Psi}{\partial x}-v_0\cdot\Psi\right). \tag{6.16}$$

The stationary solution Ψ_0 within the noisy region, with the boundary condition $\Psi_0(0)=0$ at the resonance ($x=0$) is [10]

$$\Psi_0=\frac{\phi_0}{v_0}\left[1-\exp\left(\frac{v_0}{D}\right)x\right] \tag{6.17}$$

while outside the noisy region

$$\Psi_0=\frac{\phi}{v_0}\,. \tag{6.18}$$

If there is a ripple in the relative velocity between the beam and the resonance of angular frequency ω and amplitude r, the flux $\phi=\phi_0+\phi_\omega$ will be modulated from $\Psi=\Psi_0+\Psi_\omega$. An estimate for the modulation can be found by factorising Ψ_ω into a part periodic in time and an x-part

$$\Psi_\omega=\hat{\Psi}_\omega\exp(\mathrm{i}\omega t+\mu x). \tag{6.19}$$

---

[*] The drift velocity can be expressed in units of momentum or tune i.e. $x \Rightarrow \Delta p/p$ or $x \Rightarrow \delta Q$.

From (6.15):

$$\Psi_\omega = \hat{\Psi}_\omega \exp\left[\frac{x}{\lambda} + \mathrm{i}\left(\frac{x}{\lambda} + \omega t\right)\right] \tag{6.20}$$

where $\lambda = \sqrt{\frac{2D}{\omega}}$ is the damping length and $v_\omega = \sqrt{2D\omega}$ is the phase velocity for the diffusion waves. This results in a modulation of the spill rate $\phi = -D \cdot \left.\frac{\partial \Psi}{\partial x}\right|_{\Psi=0}$ with an amplitude to first order of

$$\hat{\phi}_\omega = \phi_0 \, \frac{\sqrt{2}\, r}{\lambda}\,. \tag{6.21}$$

6.2.3 Duty factor

The modulation of the spill rate may be described by the duty factor, defined as

$$F = \frac{\langle \phi \rangle^2}{\langle \phi^2 \rangle}\,. \tag{6.22}$$

For conventional extraction [6],

$$F = \frac{1}{1 + \frac{1}{2}\left(\frac{\omega\, r}{v_0}\right)^2}\,, \tag{6.23}$$

while for stochastic extraction from (6.21)

Duty factor:

$$F = \frac{1}{1 + \frac{1}{2}\,\frac{\omega\, r^2}{D}}\,. \tag{6.24}$$

The comparison of (6.23) and (6.24) shows that in the second case $v_\omega/\sqrt{2}$ replaces the former drift speed v_0. It is clear that by increasing the diffusion constant, D the phase velocity v_ω can be kept large without decreasing the spill temporal length. The maximum available rf power and the minimum tolerable bandwidth together with the time a particle takes to diffuse out of the resonance ($\tau_D \prec 1/D$) define a physical upper limit of the diffusion constant [9,11]. The choice of D is a compromise between insensitivity to ripple and extraction efficiency.

6.2.4 Expression for diffusion constant

An expression for the diffusion constant D can be found in analogy with Brownian motion [9]. Considering the probability of finding a particle at a certain time t at a distance between r and $r + \mathrm{d}r$, one gets:

$$D = \frac{1}{2} \cdot \frac{d\overline{r^2}}{dt} \tag{6.25}$$

with [D] in m^{2}s^{-1}. For stochastic extraction (diffusion in one dimension) and with the usual substitution $x \Rightarrow \Delta p / p$,

$$D = \frac{1}{2} \cdot \frac{d\overline{\left(\frac{\Delta p}{p}\right)^2}}{dt} \tag{6.26}$$

with [D] in s^{-1}. The expression for the one dimensional diffusion constant in $\Delta p/p$ space as a function of the rms noise voltage V_N is [6]

$$D_\mathrm{p} = \frac{1}{2} \cdot \frac{{V_\mathrm{N}}^2}{\Delta f} \cdot \frac{1}{(2\,\pi\,R\,B\,\rho)^2} \tag{6.27}$$

where Δf is the bandwidth covered by the noise spectrum, $B\rho$ the magnetic rigidity and R the machine radius. This expression is found considering the beam rms change in energy due to the uncorrelated kicks received by each particle on successive revolutions and using (6.26).

Analogue expressions can be found in tune space or frequency space using the following relations:

$$D_\mathrm{Q} = \left(\frac{\xi\,\mathrm{Q}}{p}\right)^2 D_\mathrm{p} \tag{6.28}$$

$$D_\mathrm{f} = \left(\frac{\eta\,\mathrm{f}_r}{p}\right)^2 D_\mathrm{p} \tag{6.29}$$

where ξ is the chromaticity, Q the tune, p the particle momentum, η the frequency dispersion and f the revolution frequency.

6.2.5 PIMMS rf noise parameters

The rf noise parameters (bandwidth and power) have been calculated for the PIMMS medical synchrotron at the maximum proton and carbon ion extraction energies using the above equations [12]. The situation considered is the one represented in Figure 6.7 with momentum spread of the beam $(\Delta p/p)_\mathrm{s} = 0.004$ and the momentum spread between the beam and resonance $(\Delta p/p)_\mathrm{s\text{-}r} = 0.001$ (for sake of simplicity the beam is drawn already shaped).

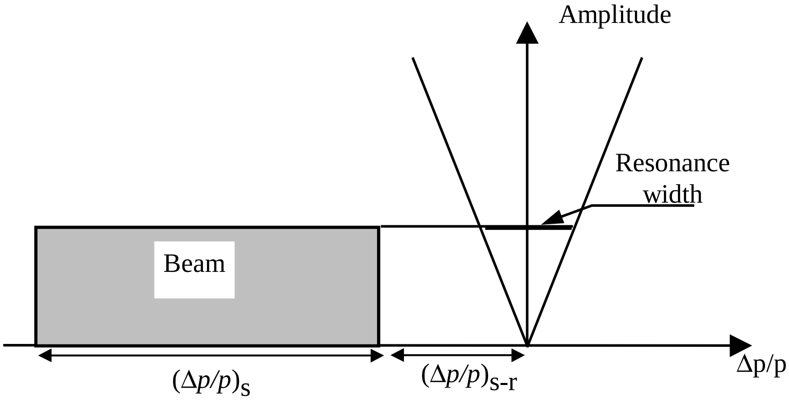


Figure 6.7 PIMM extraction configuration in $\Delta p/p$ space

The three cases described above have been considered. The results are listed in the Table 6.1.

PIMMS rf noise parameters		
	p @ 300 MeV	^{12}**C**$^{6+}$***@ 425 MeV/u***
Shaping		
Bandwidth Δf_S [kHz]	4.1	3.3
Voltage rms [V]	150	337
Power/Δf [W/Hz] (Z = 50 Ω)	0.12	0.7
Total power [W]	480	$2.3\cdot10^3$
Stochastic extraction		
Bandwidth Δf [kHz]	5	4.1
Voltage rms [V]	216	470
Power/Δf [W/Hz] (Z = 50 Ω)	0.18	1.09
Total power [W]	900	$4.4\cdot10^3$
Combined extraction (noise on the resonance)		
Bandwidth Δf_r [kHz]	1.2	1.1
Voltage rms [V]	52.4	120
Power/Δf [W/Hz] (Z = 50 Ω)	$45\cdot10^{-3}$	$270\cdot10^{-3}$
Total power [W]	55	288

Table 6.1 PIMMS rf noise parameters

In the case of shaping in Table 6.1, a shaping time of the order of 0.1 s (~10% of the flat top) was assumed, which gives $D = 0.8\times10^{-4}\ \text{s}^{-1}$. The values for stochastic extraction have been obtained in the hypothesis that the spill length is of the order of 1 s (which implies $D = 1.3\times10^{-4}\ \text{s}^{-1}$). For combined extraction, a modulation of the extracted spill of $\pm 20\,\%$ at 2 kHz (duty factor F = 0.98) with a ripple amplitude in $\Delta p/p$ units $r = 10^{-5}$ has been assumed [12]. In this case, $D = 0.8\times10^{-4}\ s^{-1}$. For all three cases, the values have been calculated with the hypothesis of working on the first harmonic to minimise the required total power. Nevertheless, the noise bandwidth should be big enough to avoid modulation of the extracted spill, because the noise signal behaves like a sine wave at the centre frequency, whose phase and amplitude remain coherent over times of the order $(\Delta f)^{-1}$. In order to reduce modulation from coherency, $(\Delta f)^{-1}$ should be small compared to the diffusion time. This limit can be a problem in the case of combined extraction with noise on the first harmonic, where the diffusion time is of the order of ms.

Two possible hardware systems to produce the noise have been investigated: a longitudinal ferrite kicker and an rf cavity. The first is the solution adopted in LEAR [13]; due to the shorter spill length and therefore the higher power needed for PIMMS medical synchrotron, high-power amplifiers are necessary to provide powers of 1 kW and more (Table 6.1). These are expensive and could cause distortions in the noise spectrum through inter-modulation.

The second possibility (rf cavity) is the cheaper solution, if it is possible to use the cavity foreseen for the acceleration of the beam. With the VITROVAC cavity it is possible to work up to the second harmonic, as the peak accelerating voltage is around 4 kV in the range 0.4 to 8 MHz [15]. For wider bandwidths, a second dedicated rf cavity working at higher harmonics should be foreseen.

The estimated cost for the kicker solution is 100 kCHF plus the price of the power amplifiers, which can be roughly assessed as 100 CHF/W per unit. For shaping, extraction and noise on the resonance at the first harmonic, this correspond to a total cost of the order of 700 kCHF. The cost for the third solution with a dedicated cavity is of the order of 800 kCHF.

When the cost estimates are considered together with the technological problems of:

- Working with such high power levels and
- Ensuring the stochasticity of the process during short shaping and extraction times,

the conclusion is that stochastic extraction is not the optimum solution for a medical spill of the order of 1 s.

6.3 PHASE-DISPLACEMENT ACCELERATION

6.3.1 Motion in longitudinal phase space

Phase displacement [16] is a technique based on the rf system. It is therefore useful to recall the theory of particle motion in the presence of an rf field. This motion can be described in terms of two first-order differential equations in the variables $\Delta E = E-E_s$, $\Delta\phi = \phi-\phi_s$ (see for example ref. [17]):

$$\frac{\mathrm{d}(\Delta\phi)}{\mathrm{d}t} = -\frac{h\omega_0\eta}{\beta^2 E}\cdot\Delta E \qquad (6.30)(a)$$

$$\frac{\mathrm{d}(\Delta E)}{\mathrm{d}t} = -\frac{\omega_0 zeV\cos\phi_s}{2\pi} \qquad (6.30)(b)$$

where ze is the charge, ω_0 is the revolution frequency, V is the rf voltage, $\beta = v/c$ is the particle normalised velocity, h is the harmonic number, $\eta = 1/\gamma^2 - 1/\gamma_t^2$ is the phase slip factor, γ is the relativistic mass factor, γ_t the γ at transition, E is the total energy, ϕ is the phase of the arbitrary particle and corresponds to the phase of the rf voltage, and the subscript ‘s’ refers to the synchronous particle. For small amplitude oscillations, from (6.30) (a) and (b):

$$\frac{\mathrm{d}^2(\Delta\phi)}{\mathrm{d}t^2} + \Omega_\mathrm{s}^2(\Delta\phi) = 0\,, \qquad \text{with:}$$

$$\Omega_\mathrm{s}^2 = \frac{qV\omega_0^2 h\eta\cos\phi_\mathrm{s}}{2\pi E\beta^2}\,.$$

The particles oscillate around the stable phase, describing ellipses in the phase plane. Ω_s is the frequency of the oscillations and is called the *synchrotron frequency*. For larger amplitudes:

$$\frac{\mathrm{d}^2(\Delta\phi)}{\mathrm{d}t^2} + \frac{\Omega_\mathrm{s}^2}{\cos\phi_\mathrm{s}}\cdot(\sin\phi - \sin\phi_\mathrm{s}) = 0\,.$$

Increasing the amplitude eventually makes the oscillations unstable. The set of trajectories representing stable oscillations is called the *rf bucket*. Its area is delimited by the last stable trajectory, called the *separatrix*. The area and the height of the *stationary bucket* (ϕ_s =0) are given by:

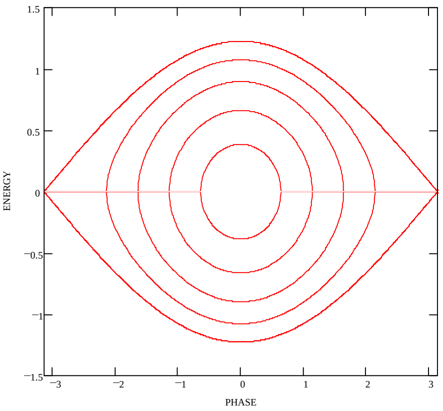
$$A_\mathrm{sb} = \frac{16\beta}{h}\sqrt{\frac{hzeVE}{2\pi|\eta|}} \quad [\mathrm{eV\,rad}], \quad H_\mathrm{sb} = \frac{2\sqrt{2}\beta}{h}\sqrt{\frac{hzeVE}{\pi|\eta|}} \quad [\mathrm{eV}]. \qquad (6.31)$$

The area and the height of the moving (accelerating or decelerating) bucket are related to the those of the stationary bucket by:

$$A_\mathrm{b} = \alpha(\Gamma)\cdot A_\mathrm{sb} \quad [\mathrm{eV\,rad}], \quad H_\mathrm{b} = \frac{\Upsilon(\Gamma)}{\sqrt{2}}\cdot H_\mathrm{sb} \quad [\mathrm{eV}], \qquad (6.32)$$

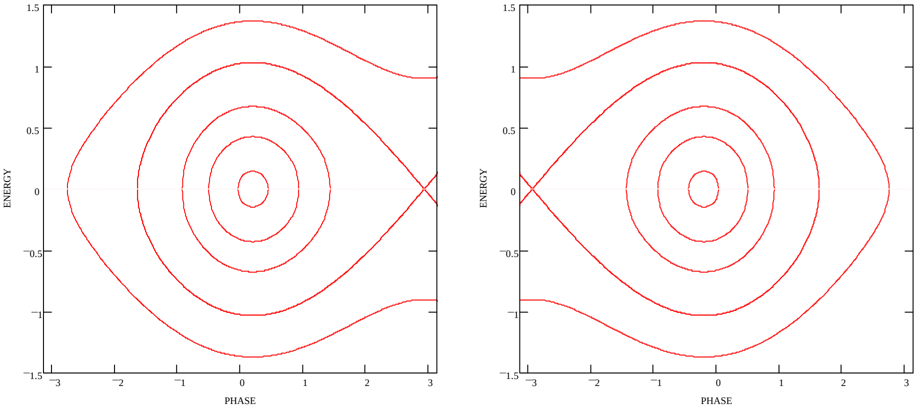
where $\Gamma = \sin\phi_\mathrm{s}$, and the functions $\alpha(\Gamma)$ and $\Upsilon(\Gamma)$ can be found from tables [18].

Figure 6.8 shows some trajectories in longitudinal phase space for stationary ($\phi_\mathrm{s} = 0$) and moving ($\phi_\mathrm{s} \neq 0$) buckets.



(a) Stationary bucket $\phi_s = 0$

[If below transition the motion is anti-clockwise, if above it is clockwise]



Moving bucket $\phi_s > 0$

[If below transition the bucket is accelerating and the particle motion is anti-clockwise, if above transition the bucket is decelerating and the motion is clockwise]

Moving bucket $\phi_s < 0$

[If below transition the bucket is decelerating and the particle motion is anti-clockwise, if above transition the bucket is accelerating and the motion is motion clockwise]

Figure 6.8 Stationary and moving rf buckets

6.3.2 The principle

Phase displacement acceleration was extensively used at the CERN ISR in the early 1960s [19]. It allows the acceleration (or deceleration) of a stacked coasting beam using a relatively simple rf system. Empty rf buckets are created outside the beam, and then the bucket energy is decreased (or increased) so that it traverses the beam. During the traversal, the particles are forced to change their phase and energy in order to turn around the bucket, without (ideally) entering it. By Liouville’s theorem, phase space density must be conserved, so that the beam will be displaced upwards (or downwards), such that the area of the displacement is equal to the bucket area. (see Figure 6.9). Multiple traversals can be made so as to bring the beam to the desired final energy. If the beam is close to the resonance energy, this method can be used to ‘feed’ the resonance.

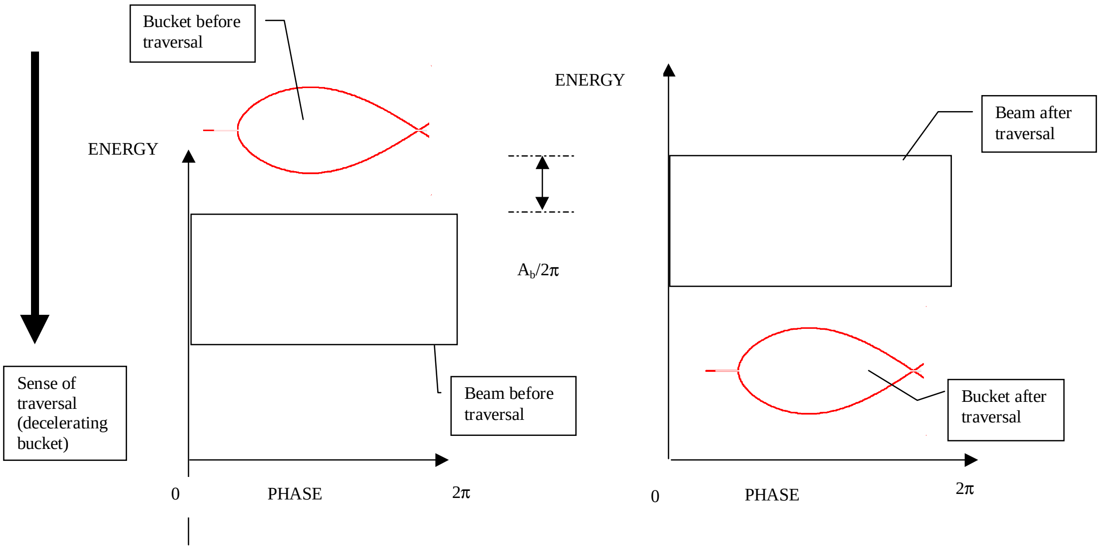

Bucket before traversal
ENERGY
Sense of traversal (decelerating bucket)
Beam before traversal
$A_b/2\pi$
0
PHASE
2π
ENERGY
Beam after traversal
Bucket after traversal
0
PHASE
2π

Figure 6.9 Principle of phase displacement

The variation of the energy of the beam due to one traversal is given by:

$$\Delta E_{\text{sweep}} = \frac{A_{\text{b}}}{2\pi}\,. \tag{6.33}$$

To traverse the beam, the frequency variation Δf_{rf} (i.e. the sweep) that must be provided by the rf system is given by:

$$\Delta f_{\text{rf}} = \frac{\eta}{\beta^2} f_{\text{rf}} \cdot \frac{\Delta E}{E} \tag{6.34}$$

where $\Delta E/E$ is the energy spread of the beam and from (6.30)(b) the speed of variation of the sweep is obtained as:

$$\frac{\mathrm{d}f_{\text{rf}}}{\mathrm{d}t} = \frac{\eta \cdot f_{\text{rf}}^2}{\beta^2 Eh} \cdot zeV\Gamma\,. \tag{6.35}$$

In the approximation that the sweeps are narrow in energy, all the parameters are referred to the 'nominal' resonance energy. Typical spill lengths for medical machines are of the order of 1 s and the spill should be continuous and as smooth as possible. At typical extraction energies and for typical values of V and Γ, a single sweep would last a few milliseconds. This means that it is not possible to accelerate the beam by a single sweep. Several sweeps with small buckets of a few volts rf voltage would be needed, if the extraction time should be about 1 s. The number n of sweeps needed is given by:

$$n = \frac{\Delta E}{\Delta E_{\text{sweep}}} = \frac{\Delta E}{E}\frac{2\pi}{A_{\text{b}}} E\,, \tag{6.36}$$

where ΔE is the total energy spread of the beam. Each sweep increases the energy of the beam bringing the edge of the beam into the resonance. For a continuous spill, it is not possible to wait until the first sweep has ended, but the sweeps must continue one after the other. According to CERN ISR experience[19], the minimum distance in energy between buckets, for the effect of the preceding bucket to be over, is twice the bucket height. This means that, to have a continuous spill, the resonance should have a minimum width, in order to always contain a moving bucket (Fig 6.10). Moreover, this sets the minimum distance at which the voltage of the buckets has to be switched on and off.

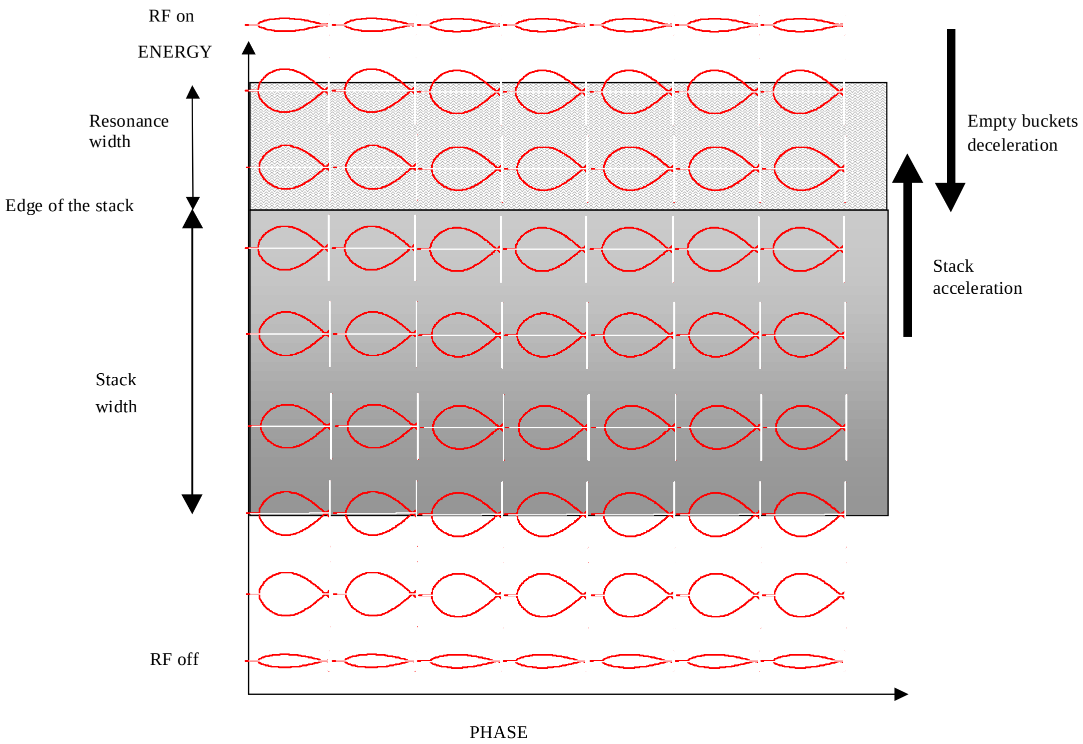


Figure 6.10 ‘Feeding’ the resonance by phase displacement

If the $\Delta E/E$ of the beam is fixed, then it is necessary to choose the rf voltage and the speed of the sweep. This would immediately give the energy increase per sweep $\Delta E_{\mathrm{sweep}}$ and Γ. The choice should be made with consideration for the following points:

‘Scattering’

During the sweep the particles vary their trajectories in phase space, since they are accelerated by different amounts. This results in a ‘scattering’ of the particles; the degree of which is a function of Γ. This will increase the machine aperture occupation (i.e. the $\Delta E/E$). The increase in the root-mean square energy of the coasted beam is given by [20]:

$$\Delta E_{\mathrm{rms}} = \Gamma \cdot \frac{A_{\mathrm{bs}}}{2\pi} \tag{6.37}$$

and after n sweeps the total increase in energy spread $\Delta E_{\mathrm{T,rms}}$ can be calculated by successive iterations with the formula:

$$\left(\Delta E_{\mathrm{T,rms}}\right)_{\mathrm{n}}^{2} = \left(\Delta E_{\mathrm{T,rms}}\right)_{\mathrm{n-1}}^{2} + \left(\Delta E_{\mathrm{rms}}\right)_{\mathrm{n}}^{2} .$$

The recommendation is to keep low Γ (i.e. quasi-stationary buckets) and a minimum number of sweeps.

Variation of bucket parameters

Even if the rf voltage is kept constant, the bucket area is not constant during the sweep. If the area increases, some particles may enter the bucket and be lost. If it decreases, the stack width increases. These effects are small, but become evident after a large number of sweeps. In the hypothesis that the machine is operating far from transition, the variation of the area of the stationary bucket is given by:

$$\frac{\Delta A_{\mathrm{bs}}}{A_{\mathrm{bs}}} = \frac{1}{2}\frac{\Delta E}{E} - \frac{1}{2}\frac{\Delta\eta}{\eta} \cong \frac{1}{2}\frac{\Delta E}{E} \pm \frac{1}{2}\frac{2\Delta E}{E} ,$$

where the plus sign is for the case below transition and the minus is for above. This means that the area of the stationary bucket increases with increasing energy below transition, whereas it decreases with increasing energy above transition. The variation of $\alpha(\Gamma)$ is related to the that of Γ by:

$$\frac{\Delta\alpha(\Gamma)}{\alpha(\Gamma)} = G(\Gamma)\cdot\frac{\Delta\Gamma}{\Gamma} ,$$

where $G(\Gamma)$ is a function defined in the above equation and $\Delta\Gamma/\Gamma$ is given by (recalling equation (6.35), and again the plus sign is for the case below transition, the minus is for above transition):

$$\frac{\Delta\Gamma}{\Gamma} = \frac{\Delta E}{E} - 2\frac{\Delta f_{\mathrm{rf}}}{f_{\mathrm{rf}}} - \frac{\Delta\eta}{\eta} \cong \frac{\Delta E}{E} \pm \frac{2\Delta E}{E} ,$$

which means that below transition Γ increases with increasing energy (i.e. A_{b} decreases) and vice-versa above transition. (Δf_{rf} is small compared to ΔE). The corresponding variation in $\Delta\alpha(\Gamma)/\alpha(\Gamma)$ is given by:

$$\frac{\Delta\alpha(\Gamma)}{\alpha(\Gamma)} = G(\Gamma)\cdot\left(1\pm 2\right)\frac{\Delta E}{E} .$$

The variation of the area of the moving bucket can be set to zero, i.e.

$$\frac{\Delta A_{\mathrm{b}}}{A_{\mathrm{b}}} = \frac{\Delta A_{\mathrm{bs}}}{A_{\mathrm{bs}}} + \frac{\Delta\alpha(\Gamma)}{\alpha(\Gamma)} = 0 ,$$

which gives, $G(\Gamma) = -\frac{1}{2}.$

Using tabulated values of $\alpha(\Gamma)$ [18], the required value of Γ is found to be $\Gamma = 0.25$.

Unwanted excitation of betatron resonances

Coupling between harmonics of the rf frequency and the betatron frequencies of the 'waiting' beam has to be avoided during the sweep, because this would cause strong beam losses. Resonance is excited when:

$$m \cdot f_{\mathrm{rf}}(E_{\mathrm{sweep}}) = (mh \pm Q) \cdot f_0(E_{\mathrm{beam}})$$

where Q is the tune, $f_0(E_{\mathrm{beam}})$ is the revolution frequency for a particle inside the beam at energy E_{beam}, $f_{\mathrm{rf}}(E_{\mathrm{sweep}})$ is the sweeping frequency corresponding at the energy E_{sweep} and m is the integer at which the equation is satisfied. This can be rewritten as:

$$\left|Q \cdot f_0(E_{\mathrm{beam}})\right| = m \cdot \left(hf_0(E_{\mathrm{beam}}) - f_{\mathrm{rf}}(E_{\mathrm{sweep}})\right).$$

A careful choice of the rf parameters is needed to avoid, or at least to reduce, this effect.

Another consideration is needed for rf phase noise. Phase modulation causes diffusion of the particles across the bucket separatrix. This results in an increase of the beam width, a variation of its average energy, and a loss of particles that leave beam inside the bucket. To keep phase noise down to acceptable levels, a phase-lock servo loop is needed, which is, in fact, standard practice.

It becomes clear that a compromise has to be taken, in the choice of the parameters, in order to minimise the 'scattering', minimise the variation of bucket parameters and avoid excitation by unwanted resonances. The choice of the parameters is then tested with computer simulations. Tracking programs such as ESME [21] are commonly used for this purpose.

6.3.3 Advantages and disadvantages of phase displacement for 'feeding' the resonance

Advantages:

- No machine parameters are varied, only the rf voltage and the rf frequency change.
- It is easy and quick to start and stop.

Disadvantages:

- To have a quasi-continuous spill, the rf system is quite complicated and expensive since several frequencies have to be excited at the same time.
- In the case of voxel scanning, the beam has to be switched on and off during the spill [13]. If it were sufficient to slow down the repetition frequency of the rf sweeps, then the hardware would be less complicated.
- A strong structure in the spill at the repetition frequency of the rf sweep as well as at the rf frequency cannot be avoided.

- One of the most important characteristics of the spill for a medical machine is uniformity (low ripple). This can be obtained with a fast crossing from the stable to the unstable region. With the phase displacement system, there is no means of improving the average speed of the crossing, i.e. the d*p*/d*t* at the resonance, which can be very low for some particles.

6.4 UNSTACKING

6.4.1 The principle

Consider a uniform coasting beam (the stack) from which a small slice of the energy spread is to be taken and brought to a different energy. Small rf buckets can be created (rf voltage of a few tens of volts) with a high harmonic rf system at the edge of the stack. A small fraction of the stack is then trapped and accelerated inside the small buckets to a different energy. This method is called *unstacking*. If the stack is close to the resonance, this method can be used to transport small numbers of particles to the resonance energy, see Figure 6.11.

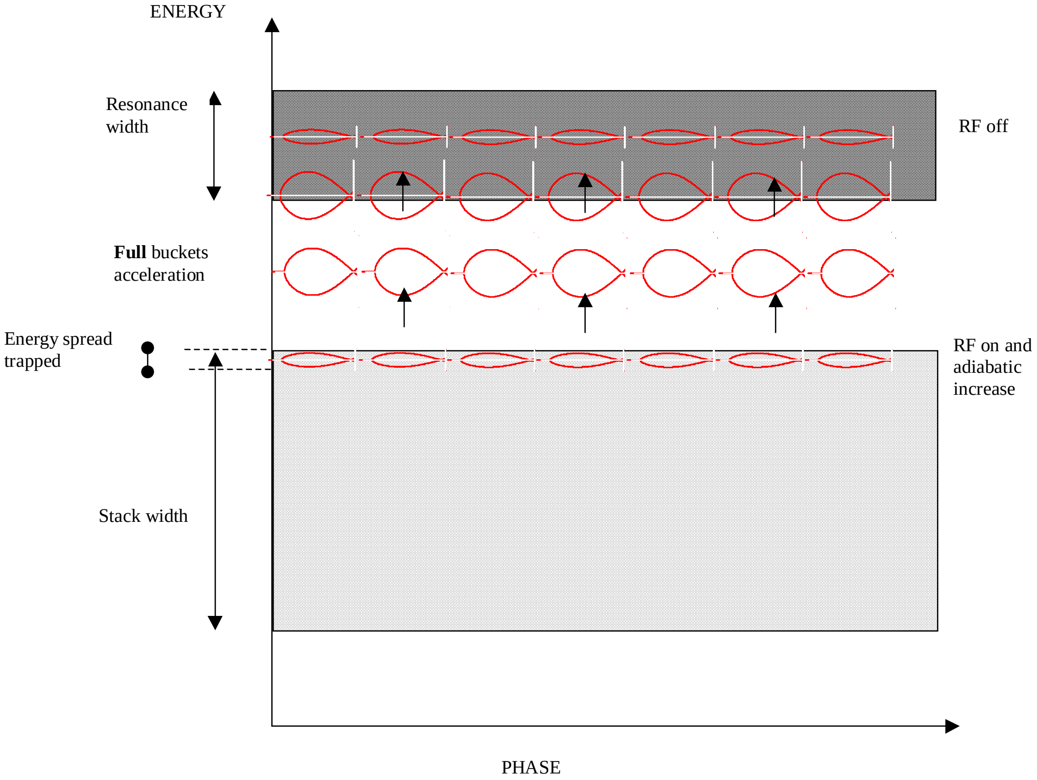


Figure 6.11 ‘Feeding’ the resonance by unstacking

In order to have a theoretical 100% capture and to minimise the longitudinal emittance dilution during the creation of the small stationary buckets, the rf voltage increase should follow the so-called iso-adiabatic law. In practice, this means that the increase should be slow (ideally infinitely slow) with respect to the synchrotron frequency Ω_s. The law that has to be followed is [22]:

$$V_{\mathrm{rf}}(t) = \frac{V_{\mathrm{final}}}{\left[\left(\dfrac{V_{\mathrm{final}}}{V_{\mathrm{initial}}}\right)^{1/2} - \dfrac{\alpha}{T_{\mathrm{s}}}t\right]^{2}} \tag{6.38}$$

where and V_{initial} and V_{final} are the initial and final rf voltages to be applied, $T_{\mathrm{s}} = 1/\Omega_s$, and α is the adiabatic factor, defined in terms of A_{b} the bucket area as

$$\alpha = \left| \frac{\dfrac{\mathrm{d}A_{\mathrm{b}}}{\mathrm{d}t} T_{\mathrm{s}}}{A_{\mathrm{b}}} \right| .$$

It can be seen from the definition of α that it should be kept small (ideally zero) in order to have an adiabatic process. The parameters in equation (6.38) are chosen by trying to optimise the capture efficiency and to minimise the longitudinal emittance dilution, which would enlarge the energy spread. The choice has to be tested by computer simulations.

Once the desired energy spread is trapped, the acceleration starts and transports the particles, inside the buckets, to the extraction energy. The stack has to be positioned as close as possible to the resonance area, in order to avoid large frequency swings that could cause particle losses (see paragraph 6.3.2). In this case, dilution due to 'scattering' is of no concern because the stack is not traversed*. If large frequency swings cannot be avoided, when the stack is no longer influenced by the rf bucket, (minimum distance 2 bucket heights), the rf voltage can be increased (always following an iso-adiabatic law) before continuing the sweep up to the resonance energy. Verification by computer simulations is always needed.

6.4.2 Advantages and disadvantages of unstacking for 'feeding' the resonance

Advantages:

- No machine parameters are varied; only the rf system is working.
- It is easy and quick to start and stop.
- Several small bucket trains can be performed at the same time in order to obtain a continuous spill.
- The particles can be transported deeply into the resonance region, thus avoiding the possibility of them being liberated and then 're-eaten' by the movements of the resonance separatrix. This cannot be avoided with the phase-displacement method.

Disadvantages:

---

* There are two possibilities: *unstacking from the top* and *unstacking from the bottom* In one case, the stack will be traversed and in the other not. When the stack is not traversed, then the starting energy (frequency) has to be stepped on each cycle, but if the stack is traversed then the phase displacement moves the stack down (up) to fill the space left by the accelerated (decelerated) beam.

- Unfortunately this method suffers from the same limitations encountered with phase-displacement (see Section 6.3.3): a complicated rf system and a strong modulation of the spill at the repetition rate of the sweeps.
- Unfortunately, there is no mean of improving the speed at the crossing of the resonance separatrix, which is important for the uniformity of the spill. This will depend on the position of the particle in the rf bucket, and for some particles the d*p*/d*t* can be very low.
- A resonance width less than the minimum distance between trains (2 bucket heights) is not acceptable, since the spill would become chopped.

6.5 ‘FRONT-END’ ACCELERATION BY EMPTY RF BUCKET CHANNELLING

Phase displacement is not a good candidate for accelerating the ‘waiting’ beam into the resonance, as explained in Section 6.3, due to the poor uniformity of the resulting spill, but paradoxically, it can be very useful for making the beam particles cross more quickly from the stable to the unstable region. The technique is called *empty rf bucket ‘channelling’*. This turns out to be a good candidate for attenuating the effects of low-frequency tune ripple on spills of the order of one second in length. Experimental evidence of the improvement has been reported at CERN PS [23] and at IHEP Protvino, Russia [24]. First consider the beam spill *S(t)* that can be written as:

$$S(t) = \frac{\mathrm{d}N}{\mathrm{d}t} = \frac{\mathrm{d}N}{\mathrm{d}Q} \cdot \frac{\mathrm{d}Q}{\mathrm{d}t} \tag{6.39}$$

where N is the number of particles and Q the horizontal tune. The tune change can be expressed as the sum of two components, $\dot{Q}_0$ the constant component and $\dot{Q}_\mathrm{r}$ the component due to unwanted ripple, so that:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \dot{Q}_0 + \dot{Q}_\mathrm{r} = \dot{Q}_0 \cdot \left(1 + \frac{\dot{Q}_\mathrm{r}}{\dot{Q}_0}\right). \tag{6.40}$$

The betatron core (see Section 6.1) is used to provide the constant component of the acceleration $\dot{Q}_0$. This high inductance device is well suited to delivering a smooth spill and, since the energy stored is high, it has the characteristic of responding slowly to transients that could give unintentional beam spikes to the patient.

For a uniform spill *S*(*t*), the product of d*N*/d*Q* and d*Q*/d*t* must be kept constant. The form of the stack determines d*N*/d*Q*, but this is a slow variation and can be controlled by feedback on the acceleration rate from the betatron core. Unfortunately, d*Q*/d*t* is affected strongly at all frequencies by $\dot{Q}_\mathrm{r}$. The purpose of the rf bucket channelling is to reduce the effect of this contribution.

Inspection of (6.40) shows that the contribution of $\dot{Q}_\mathrm{r}$ can be reduced, if $\dot{Q}_0$ is increased. However, $\dot{Q}_0$ cannot be changed for the whole stack, since it is fixed by the spill time, but it can be increased in particular phase-space regions, if the density of the particles d*N*/d*Q* in those regions is decreased accordingly, so as to keep a constant *S(t)* (see equation (6.39)). The scheme shown in Figure 6.12 has a region of high speed and low density created close to the resonance.

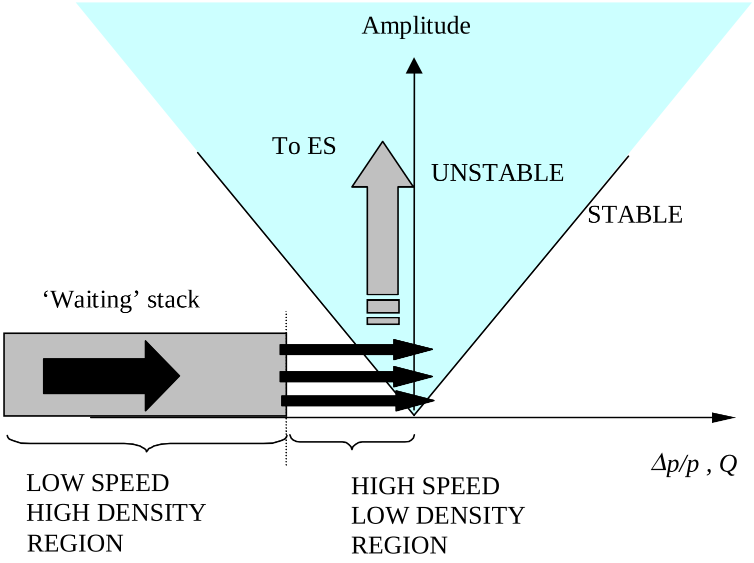


Figure 6.12 Extraction with a high-speed, low-density region close to resonance

The parameter that indicates the quality of the extracted spill is the *duty factor*, see Section 5.1, and is given by,

$$F = \frac{\left(\int_{T_{\text{spill}}} \phi(t)dt \right)^2}{T_{\text{spill}} \int_{T_{\text{spill}}} \phi^2(t)dt} .$$

Note that $\phi = \mathrm{d}N/\mathrm{d}t$ is used here in order to correspond to Chapter 5 and should not be confused with the rf phase. The substitution of (6.40), which is explained more fully in Section 5.5, gives

$$F = \frac{1}{1 + \frac{1}{2}\left(\frac{\dot{Q}_{\text{r}}}{\dot{Q}_0} \right)^2} \quad \text{for} \quad \frac{\dot{Q}_r}{\dot{Q}_0} \leq 1.$$

Thus, the duty factor can be improved by an increase in $\dot{Q}_0$. The increase in $\dot{Q}_0$ need only be an increase in the 'local' acceleration (in the vicinity of the resonance) and will be represented by a factor *K*. Finally, the ratio $\frac{\dot{Q}_r}{\dot{Q}_0}$ can be re-placed by $\frac{\dot{p}_r}{\dot{p}_0}$ to give

$$F = \frac{1}{1 + \frac{1}{2}\left(\frac{\dot{p}_{\text{r}}}{\dot{p}_0} \right)^2} \qquad (6.41)$$

where $\dot{p}_0$ and $\dot{p}_r$ are the rates of change in momentum corresponding to $\dot{Q}_0$ and $\dot{Q}_r$.

6.5.1 General theory and description

An empty bucket is created at the resonance frequency corresponding to the resonance energy E_{res}. The hardware should be set in order to keep the bucket frequency 'fixed' at the resonance frequency for the whole spill time. At the starting time, the beam is out of the resonance. The flux of the extracted particles is set by the acceleration imposed by the betatron core.

In the longitudinal phase plane, the particles turn around the rf bucket, in the same way as happens when accelerating a stack by rf phase displacement, except that on this occasion the stack moves and the bucket remains fixed. At the energy of the resonance, the particles are swept between the buckets in the small phase interval $\Delta\phi$ limited by the bucket separatrices. In this channel, dN/dQ is reduced and dQ/dt correspondingly increased. In other words, the empty bucket creates a 'bottle neck' in the phase space, through which the particles are swept with increased velocity $K\dot{Q}_0$ see Figure 6.15.

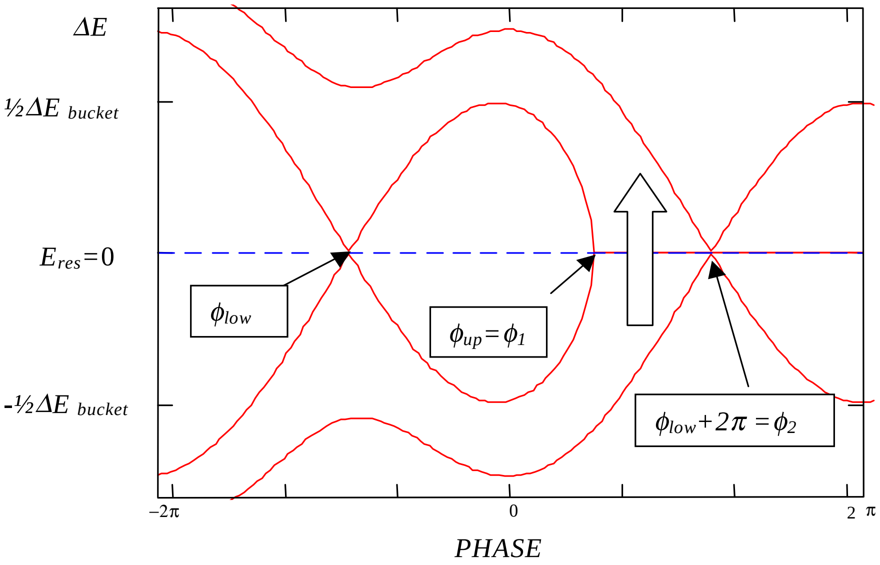


Figure 6.13 Particle channelling between buckets (arrow), case of a decelerating bucket below transition, or an accelerating bucket above transition, and $\phi_s < 0$
[E is the total energy, $E_{\mathrm{res}} = 0$ is the energy at which the empty bucket is positioned and ½ $\Delta E_{\mathrm{bucket}}$ is the bucket half height. In the case $E = E_{\mathrm{res}}$ (i.e. $\Delta E = 0$), the bucket limits in phase are ϕ_{low} and ϕ_{up} and the channel width is $\Delta\phi = \phi_2 - \phi_1$]

6.5.2 Average multiplying factor *K*

Since dQ/dt is proportional to dp/dt, and dp/dt is proportional to dE/dt, it is sufficient to calculate the improvement of dE/dt when particles cross the resonance energy (separatrix). Using the Hamiltonian formalism (see for example [17]), the motion in the longitudinal phase plane can be described in terms of two first order differential equations in the conjugate variables ($\Delta E/h\omega_0$, $\Delta\phi = \phi - \phi_s$):

$$\frac{\mathrm{d}\left(\dfrac{\Delta E}{h\omega_0}\right)}{\mathrm{d}t} = \frac{zeV}{2\pi h}\left(\sin\phi - \sin\phi_s\right) \tag{6.42}$$

$$\frac{\mathrm{d}(\phi-\phi_s)}{\mathrm{d}t} = -\omega_0^2 h^2 \frac{\eta}{\beta^2 E}\left(\frac{\Delta E}{h\omega_0}\right). \tag{6.43}$$

These equations are equivalent to (6.30) (a) and (b), except that they use conjugate variables. The corresponding Hamiltonian ***H*** is:

$$\boldsymbol{H} = \frac{1}{2}\frac{\eta}{\beta^2 E}(\Delta E)^2 + \frac{zeV}{2\pi h}\left[\sin\phi_s(\phi_s - \phi) + \cos\phi_s - \cos\phi\right]. \tag{6.44}$$

With the hypothesis that the acceleration is smooth and continuous over one turn,

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{zeV}{2\pi}\omega_0\left(\sin\phi - \sin\phi_s\right)$$

where the bucket is empty, but the particles outside are still affected. A particle that crosses the resonance when $\phi = 0$ is not affected by the rf voltage. Its energy will vary by:

$$\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_0 = -\frac{zeV}{2\pi}\omega_0 \sin\phi_s\,,$$

which is exactly equal to the d*E*/d*t* given by the betatron core. Other particles will be affected in different ways depending on their arbitrary phase and energy. On average:

$$\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{AV}} = \frac{1}{\phi_2 - \phi_1}\int_{\phi_1}^{\phi_2}\frac{zeV}{2\pi}\omega_0\left(\sin\phi - \sin\phi_s\right)\mathrm{d}\phi \tag{6.45}$$

where ϕ_1 and ϕ_2 are the minimum and maximum phases of the channel (see Figure 6.14). The average multiplying factor is given by:

$$K = \frac{\left(\dfrac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{AV}}}{\left(\dfrac{\mathrm{d}E}{\mathrm{d}t}\right)_0} = \frac{-1}{(\phi_2 - \phi_1)\sin\phi_s}\int_{\phi_1}^{\phi_2}\left(\sin\phi - \sin\phi_s\right)\mathrm{d}\phi \tag{6.46}$$

which gives:

$$K = \frac{\Gamma\Delta\phi + \cos\phi_2 - \cos\phi_1}{\Gamma\Delta\phi} = 1 + \frac{\cos\phi_2 - \cos\phi_1}{\Gamma\Delta\phi} \tag{6.47}$$

where $\Gamma = \sin \phi_s$ and $\Delta\phi = \phi_2 - \phi_1$.

Intuitively, it can be seen that the more the bucket obstructs the available phase space, the faster the particles must move. Thus, the closer the bucket becomes to a stationary bucket the greater the particle velocity enhancement. For this reason, quasi-stationary buckets will be studied more closely.

In the general case, it is necessary to solve equation (6.47), with values of ϕ_1 and ϕ_2 depending on the value of ΔE (see Figure 6.14). Only $0 < \Delta E < \Delta E_{\text{bucket}}/2$, will be considered because this is the useful range (see Section 6.4.3). It is useful to calculate the Hamiltonians H_1 and H_2 corresponding to ϕ_1 and ϕ_2, the phase limits of the channel. The Hamiltonian H_1 corresponds to the 'internal' separatrix, and the Hamiltonian H_2 to the 'external' separatrix. Making use of (6.44) and of the bucket relation linking ϕ_1 and ϕ_2 in the case $\Delta E = 0$ [25]:

$$\phi_1 \sin\phi_\text{s} + \cos\phi_1 = \phi_\text{u} \sin\phi_\text{s} + \cos\phi_\text{u} \tag{6.48}$$

H_1 and H_2 for the accelerating bucket are found to be:

$$H_1 = \frac{qV}{2\pi h}\left[2\cos\phi_\text{s} + (2\phi_\text{s} - \pi)\sin\phi_\text{s}\right],\quad H_2 = \frac{qV}{2\pi h}\left[2\cos\phi_\text{s} + (2\phi_\text{s} + \pi)\sin\phi_\text{s}\right]$$

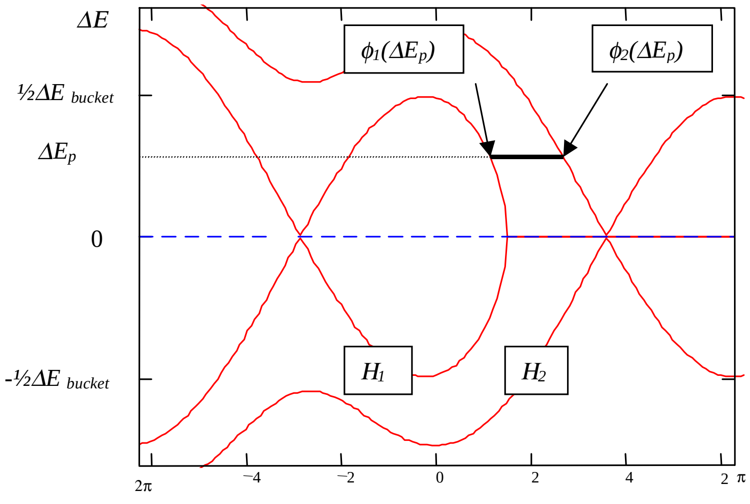


Figure 6.14 Position and parameters of the channel for a general $\Delta E = \Delta E_\text{p} \neq 0$
[H_1 is the Hamiltonian corresponding to the 'internal' separatrix,
H_2 is the Hamiltonian corresponding to the 'external' separatrix]

For the decelerating bucket H_1 becomes H_2 and vice versa.

ϕ_1 (ϕ_2) is the root of the equation:

$$H_{1,2} - \frac{1}{2}\frac{\eta}{\beta^2 E}\Delta E^2 - \frac{zeV}{2\pi h}\left[\sin\phi_\text{s}(\phi_\text{s} - \phi) + \cos\phi_\text{s} - \cos\phi\right] = 0 \tag{6.49}$$

for $$0 < \Delta E < \frac{\Delta E_{\text{bucket}}}{2}.$$

The solution of equation(s) (6.49) can be found numerically and used to find the value of the general multiplying factor K as a function of ΔE, using

$$K(\Delta E) = 1 + \frac{\cos[\phi_2(\Delta E)] - \cos[\phi_1(\Delta E)]}{\Gamma[\phi_2(\Delta E) - \phi_1(\Delta E)]}. \qquad (6.50)$$

In particular, for the case $\Delta E = 0$, Table 6.2 gives the values of the parameters concerned for the four different acceleration conditions:

	Below transition		**Above transition**	
ϕ_s ~0 (Γ=sin ϕ_s)	Accelerating bucket ϕ_s >0	Decelerating bucket ϕ_s <0	Accelerating bucket ϕ_s <0	Decelerating bucket ϕ_s >0
$\Delta\phi = \phi_2-\phi_1$	$\phi_1-(\phi_u-2\pi)$	$(\phi_1+2\pi)-\phi_u$	$\phi_1-(\phi_u-2\pi)$	$(\phi_1+2\pi)-\phi_u($
cos ϕ_s	$1-\Gamma^2/2$	$1-\Gamma^2/2$	$1-\Gamma^2/2$	$1-\Gamma^2/2$
ϕ_1	~ $-\pi$	$-\pi-\phi_s$	$-\pi-\phi_s$	~ $-\pi$
ϕ_u	$\pi-\phi_s$	~π	~π	$\pi-\phi_s$
cos ϕ_1	$-1+\frac{(\phi_1+\pi)^2}{2}$	$-\cos\phi_s$	$-\cos\phi_s$	$-1+\frac{(\phi_1+\pi)^2}{2}$
cos ϕ_u	$-\cos\phi_s$	$-1+\frac{(\phi_u-\pi)^2}{2}$	$-1+\frac{(\phi_u-\pi)^2}{2}$	$-\cos\phi_s$

Table 6.2 Some useful bucket parameters
[cos ϕ_s, cos ϕ_1, and cos ϕ_u are given by the Taylor series expansions]

From Table 6.2 and equation (6.45)

$$\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{AV}} = -\frac{zeV}{2\pi}\,\omega_0\,\frac{2\pi\sin\phi_s}{\Delta\phi}$$

and for the average multiplying factor, K

$$K = \frac{\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{AV}}}{\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_0} = \frac{2\pi}{\Delta\phi} = \frac{2\pi}{2\pi - \text{bucket width}} \qquad (\Delta E = 0).$$

The last formula has an easy geometrical interpretation;: in the case $\Delta E = 0$, K is given by the ratio between the whole phase segment 0-2π and the sub-segment containing the allowed phase-space trajectories. In the hypothesis (usually true) that $\phi_s \ll 2\pi$, using (6.48) and Table 6.2 the channel width becomes,

$$\Delta\phi = \phi_1 - \phi_u + 2\pi \approx 2\sqrt{\pi\Gamma}$$

and the multiplying factor,

$$K = \frac{2\pi}{\Delta\phi} = \sqrt{\frac{\pi}{\Gamma}} \quad (\Delta E = 0). \tag{6.51}$$

This agrees with the particular case given in reference [23] for the CERN/PS synchrotron, where an empty bucket is created above transition.

6.5.3 Positioning the rf bucket

It can be seen directly from Figure 6.12 that the resonance energy of the particles depends on their betatron amplitude, the higher the betatron amplitude the lower the resonance energy, and vice versa. Qualitatively, it can be stated that in order to obtain a high multiplying factor for all the particles, two conditions have to be fulfilled. One on the bucket height and the other on the bucket position.

First condition

states that the beam energy spread engaged in the resonance has to be smaller than the bucket half height, i.e.

$$\Delta E_{\text{spread of resonance}} < \frac{1}{2} \cdot H_{\text{b}} \tag{6.52}$$

Since the height of the rf bucket is given by:

$$\frac{1}{2} \cdot H_{\text{b}} = \frac{\beta}{h} \cdot Y(\phi_{\text{s}}) \cdot \sqrt{\frac{hzeVE}{\pi|\eta|}}$$

where $Y(\phi_{\text{s}}) = \sqrt{\sin\phi_{\text{s}}(2\phi_{\text{s}} \pm \pi) + 2\cos\phi_{\text{s}}}$, where the ‘+’ sign is valid when $\phi_{\text{s}} > 0$. This condition imposes a constraint on the rf voltage:

$$V > \frac{1}{\beta^2} \frac{1}{[Y(\phi_{\text{s}})]^2} \frac{h\pi|\eta|}{zeE} (\Delta E_{\text{spread of resonance}})^2 . \tag{6.53}$$

Second condition

relates the position of the bucket to the resonance energy. In order to have a positive improvement for all the betatron amplitudes, one should position the bucket as shown in Figure 6.15. This is the case of a stack starting from energies lower than the resonance energy and below transition.

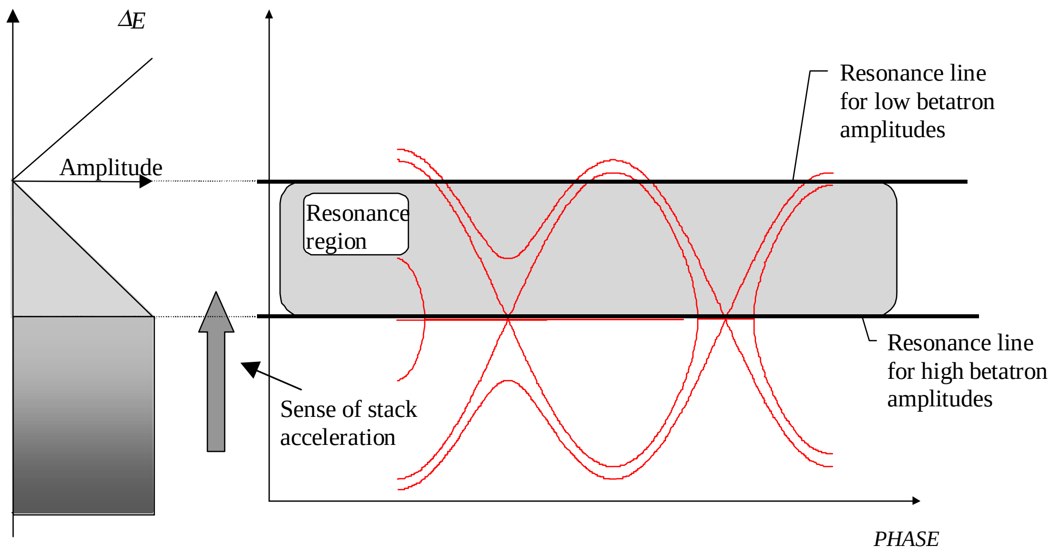

ΔE
Amplitude
Resonance region
Resonance line for low betatron amplitudes
Resonance line for high betatron amplitudes
Sense of stack acceleration
PHASE

Figure 6.15 Position of the bucket with respect to the resonance region (in the case of PIMMS)

6.5.4 Adjusting the rf voltage

The parameters of the rf bucket are determined by considering a fictitious particle trapped in the bucket at the synchronous phase. The rf frequency will be the revolution frequency, or the harmonic, needed to keep the particle energy constant at the synchronous phase (in this case d*B*/d*t* = *0*). The energy losses or gains in the bucket will be those needed to compensate the changes that take place in the machine (in this case the energy gain in the betatron) i.e. (see [26]):

$$\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{cavity}} = -\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{betatron}\quad\text{core}} ,$$

which can be written as (see [17]):

$$V \cdot |\Gamma| = 2\pi R \frac{\mathrm{d}(B\rho)}{\mathrm{d}t} \qquad (6.54)$$

where R is the mean radius of the orbit and $B\rho$ is the beam rigidity. The link with the beam momentum p is given by [27]:

$$p = \frac{ze}{A} \cdot \frac{c}{10^9} \cdot B\rho \qquad (6.55)$$

where A is the atomic mass number, p is the beam momentum and c is the velocity of the light. The combination of (6.54) with (6.55) gives:

$$|\Gamma| = \frac{A}{ze} \cdot \frac{2\pi \cdot 10^9}{c} \cdot \frac{R}{V} \cdot \frac{\mathrm{d}p}{\mathrm{d}t} . \qquad (6.56)$$

Thus the rf voltage is proportional to the energy change in the core and inversely proportional to $\Gamma = \sin \phi_s$. In order to have a large multiplying factor, *K*, the rf voltage should be kept high in order to have a low Γ, since this shrinks the width of the channel. The limit will be fixed by the maximum rf voltage available.

6.5.5 Dependence of *K* on the ripple and improvement in duty factor

The amplitude of the ripple and its frequency dependence degrade the improvement in duty factor, as explained in [26]. With the hypothesis of a smooth and continuous momentum increase during extraction,

$$\dot{p}_0=\frac{1}{T_{\mathrm{spill}}}\frac{\Delta p}{p}\,p_0 \tag{6.57}$$

where p_0 is the momentum of the particle (without ripple), $\Delta p/p$ is the momentum spread at extraction energy, T_{spill} is the spill time in seconds. The ripple contribution to the momentum at the frequency ω can be written as:

$$p_{\mathrm{r}}(\omega)=p_{\mathrm{r0}}+p_{\mathrm{r}}\ \sin\omega t \tag{6.58}$$

where p_{r0} is the constant component, and $p_{\mathrm{r}\omega}$ is the modulated component of the amplitude of the ripple at the particular frequency considered. By defining:

$$\alpha_{\mathrm{r}}=\frac{p_{\mathrm{r}\omega}}{p_0+p_{\mathrm{r0}}}\approx\frac{p_{\mathrm{r}\omega}}{p_0}\qquad(\text{for}\quad p_{\mathrm{r0}}<<p_0)$$

$\dot{p}_{\mathrm{r}}$ becomes:

$$\dot{p}_r(\omega)=\omega p_{\mathrm{r}\omega}\cos\omega t=\omega\alpha_{\mathrm{r}}p_0\cos\omega t\,.$$

The total momentum derivative during extraction is therefore:

$$\dot{p}=\dot{p}_0+\dot{p}_{\mathrm{r}}(\omega)=\frac{1}{T_{\mathrm{spill}}}\frac{\Delta p}{p}\,p_0+\omega\alpha_{\mathrm{r}}p_0\cos\omega t\,. \tag{6.59}$$

In the general case, ($0<\Delta E<\Delta E_{\mathrm{bucket}}/2$) the average multiplying factor is given by (6.46). Inserting in (6.46) the results of (6.55) and (6.58) gives the new multiplying factor (for simplicity $\dot{p}_r(\omega)=\dot{p}_r$):

$$K(\Delta E,\phi)=1+\frac{\cos\phi_2-\cos\phi_1}{\Delta\phi\cdot\left|\dfrac{A}{ze}\cdot\dfrac{2\pi\cdot10^9}{c}\cdot\dfrac{R}{V}\cdot p_0\cdot(v_0+\omega\alpha_{\mathrm{r}}\cos\omega t)\right|}$$

where: $v_0=\dfrac{1}{T_{\mathrm{spill}}}\dfrac{\Delta p}{p}$.

The duty factor is given by equation (6.41), inserting the new value for *K*:

Duty Factor, $$F(\Delta E,\omega)=\frac{1}{1+\dfrac{1}{2\cdot K^{2}(\Delta E,\omega)}\cdot\left(\dfrac{\dot{p}_{\mathrm{r}}}{\dot{p}_{0}}\right)^{2}}\,. \tag{6.60}$$

In the case of $\Delta E = 0$, a new expression for K is found. Let K_0 be the average improving factor coming from (6.51). After the insertion in (6.51) of the results of (6.56) and (6.59)

$$K(0,\omega)=\sqrt{\frac{ze}{A}\cdot\frac{c}{2\cdot 10^{9}}\cdot\frac{V}{R}\cdot\frac{1}{\dot{p}_{0}}}\cdot\sqrt{\frac{\dot{p}_{0}}{\dot{p}_{\mathrm{r}}+\dot{p}_{0}}}=K_{0}\cdot\sqrt{1-\frac{\dot{p}_{\mathrm{r}}}{\dot{p}_{0}+\dot{p}_{\mathrm{r}}}} \tag{6.61}$$

and the duty factor can be written as:

$$F(0,\omega)=\frac{1}{1+\dfrac{1}{2}\cdot\left(1+\dfrac{\dot{p}_{\mathrm{r}}}{\dot{p}_{0}}\right)\cdot\left(\dfrac{\dot{p}_{\mathrm{r}}}{K_{0}\dot{p}_{0}}\right)^{2}}\,.$$

It is also interesting to know the minimum K value and hence the poorest value to be expected for F. The minimum K and F values are calculated for $\Delta E = 0$ and for maximum $\dot{p}_r$ and are given by:

$$K_{\min}(0,\omega)=K_{0}\sqrt{1-\frac{\alpha_{\mathrm{r}}\omega}{\dfrac{1}{T_{\mathrm{spil}l}}\dfrac{\Delta p}{p}+\alpha_{\mathrm{r}}\omega}}\,,\qquad F_{\min}(0,\omega)=\frac{1}{1+\dfrac{1}{2}\cdot\left(1+\dfrac{\omega\alpha_{\mathrm{r}}}{v_{0}}\right)\cdot\left(\dfrac{\omega\alpha_{\mathrm{r}}}{K_{0}v_{0}}\right)^{2}}\,.$$

It can be seen that the multiplying factor decreases if the frequency of the ripple increases, or if the spill time increases. As an example, the multiplying factor for a 100% modulation (i.e. $\dot{p}_r = \dot{p}_0$) from (6.60) leads to:

$$K_{\min}(0,\omega)=\frac{K_{0}}{\sqrt{2}}$$

6.5.6 Recommendations

In conclusion, some practical recommendations should be kept in mind:

- The bucket half height should be higher than the energy spread engaged in the resonance, which sets a minimum value needed for the rf voltage.
- The bucket should be properly positioned with respect to the resonance region.
- The rf voltage value will depend on the machine, on the available hardware, and on the beam parameters.
- The empty bucket must always give the same magnitude but opposite acceleration rate as that given by the betatron core.

Empty rf bucket channelling improves the spill quality during slow extraction by increasing the d*p/*d*t* at the resonance crossing. The improvement is not the same for all particles, but depends on their betatron amplitude. Furthermore, it depends on the amplitude and the frequency of the ripple. The method becomes less effective as the ripple frequency increases.

Another characteristic of this technique is that the particles are extracted in a small interval of the longitudinal phase interval 0-2π. This results in a modulation of the spill at harmonics of the rf frequency (few MHz). In the case of medical machines, the degradation of the spill quality at frequencies above 10 kHz is of no concern, since the modulation coming from such high frequencies is averaged by the slow extraction process, the physical spot size and the integration time in the on-line dosimetry system [28].

References

[1] K. Johnsen, *Betatron core for the electron storage ring*, PS/Int. AR/60-29.

[2] J.C. Ciret, *Extraction du faisceau de Saturne II par accélération bétatronique le Gephyrotron, GERMA* 76.02/1E-117 (1976).

[3] Ch. Ellert, D. Habs, E.Jaeschke, T. Kambara, M. Music, D. Schwalm, P.Sigray, A. Wolf, *An induction accelerator for the Heidelberg Test Storage Ring TSR*, Nucl. Instr. And Meth. A314 (1992) 399-408.

[4] Private communication, Ch. Steinbach, *Betatron acceleration during extraction*, Minutes of the Meeting on Slow Extraction from Synchrotrons for Cancer Therapy, CERN 13-14 February 1996.

[5] L. Badano, S. Rossi, *Characteristics of a betatron core for extraction in a proton-ion medical synchrotron*, CERN/PS 97-19(DI).

[6] S. van der Meer, *Stochastic extraction, a low ripple version of resonant extraction*, CERN/PS/AA Note 78-6.

[7] R. Cappi et al., *Ultraslow extraction*, CERN/PS/DL/LEA 82-3.

[8] G. Molinari, H. Mulder, *The improved ultra slow extraction noise system at LEAR*, Proc. of the 4th European Particle Accelerator Conference, London, 1994, pp. 2376-2378.

[9] D. Boussard et al., *Slow extraction at 400 GeV/c with stochastic RF noise*, SPS Improvement Report No. 179.

[10] W. Hardt, *Moulding the noise spectrum for much better ultra slow extraction*, CERN/PS/DL/LEAR Note 84-2.

[11] W. Hardt, *Remarks on stochastic extraction*, CERN/PS/DL Note 78-5.

[12] Private communication, L. Badano, *Feeding the resonance with noise*, Minutes of the 2nd Meeting on Slow Extraction from Synchrotrons for Cancer Therapy, CERN 2-3 September 1996.

[13] S. Rossi, *Elements for a comparison between active scanning systems: raster and voxel scanning*, Minutes of the PAC Meeting, CERN, December 1995.

[14] G. Molinari, *L'extraction stochastique de LEAR*, CERN/PS/AR Note 92-05.

[15] M. Crescenti, G. Primadei, and A. Susini, *A New Compact Large Frequency-Swing RF System for Hadron Acceleration: Test Results*, CERN/PS 97-60 (DI).

[16] K.R. Symon, A.M. Sessler, *Methods of radio frequency acceleration in fixed field accelerators with applications to high current and intersecting beams of particles, Proc. CERN Symposium on High Energy Accelerators and Pion Physics*, Geneva, 1956 (CERN, 1956), Vol. 1, 44-58.

[17] W. Pirkl, *Longitudinal Beam Dynamics*, Proceedings of CAS Rhodes, CERN 95-06 Vol. 1 pp. 233-257.

[18] C. Bovet, R. Gouiran, I. Gumowski, K.H. Reich, *A Selection of Formulae and Data Useful for the Design of A.G. Synchrotrons*, CERN/MPS-SI/Int. DL/70/4.

[19] E. Ciapala, *Stacking and Phase Displacement Acceleration*, Proceedings of CAS Paris, CERN 85-19 Vol.1 pp. 195-225.

[20] E.W. Messerschmid, *Dispersion of stacked protons in synchrotron phase space by a modulated radio-frequency voltage*, CERN-ISR-RF/72-28

[21] J. MacLachlan, *User's Guide to ESME v.8.2 (1996)*. S. Hancock, private communication.

[22] M. Crescenti and S. Rossi, *PIMMS RF Programme*, Minutes of the Meeting on Proton Ion Medical Machine Study, CERN February 1997.

[23] R. Cappi and C. Steinbach, *Low Frequency Duty Factor Improvement for the CERN PS Slow Extraction Using RF Phase Displacement Techniques*, 1981 Particle Accelerator Conference, Washington, March 11-13, 1981.

[24] A. Maksimov, IHEP Protvino, Russia, private communication.

[25] J. Le Duff, *Longitudinal Beam Dynamics in Circular Accelerators*, Proceedings of CAS Paris, CERN 85-19, Vol.1 p.137.

[26] M. Crescenti, *RF Empty Bucket Channelling Combined With a Betatron Core to Improve Slow Extraction in Medical Synchrotrons,* CERN/PS 97-68 (DI).

[27] D. Carey, *The Optics of Charged Particle Beams,* Harwood Academic Publishers 1987, p.3.

[28] U. Amaldi and M. Silari (ed.), *The TERA Project and The Centre for Oncological Hadrontherapy,* Frascati INFN Publisher, Second Edition Vol. 2 Chap. 7, (1995).

* * *

I-7 LATTICES

It is useful to review briefly the types of ring lattice that can be used for the slow extraction schemes discussed in the earlier chapters, as well as some of the basic design choices. Note that the three examples of ring lattices are all drawn to same scale.

7.1 REGULAR-CELL LATTICE

The simplest structures are built by concatenating several identical cells. The individual cells may be *FODO, doublet*, *triplet* or a variant of these. Figure 7.1 is an example of such a lattice taken from Reference 1. This example is based on a doublet cell. The natural choice for this lattice (and that chosen in Ref. 1) is to use a tuning quadrupole to move the resonance into the beam as illustrated in Figure 3.2 (IV).

The main advantages are:

- Simplicity.
- Small numbers of components.

The main disadvantages are:

- Equipment for injection, extraction, rf etc. has to be adapted to fit the rigid layout of straight sections.
- No dispersion-free regions in which the resonance sextupole and the rf cavity can be installed.
- With the quadrupole extraction, the lattice functions are changing during the spill.

The movement of the separatrix during the extraction makes the Hardt Condition inapplicable, achromatic transfer between the electrostatic and magnetic septa would require an undesirably large phase advance between septa (see Section 3.12) and ‘front-end’ acceleration is not easily applied. This type of machine would be adequate for a beam delivery using passive spreading, but would be less suitable for active scanning.

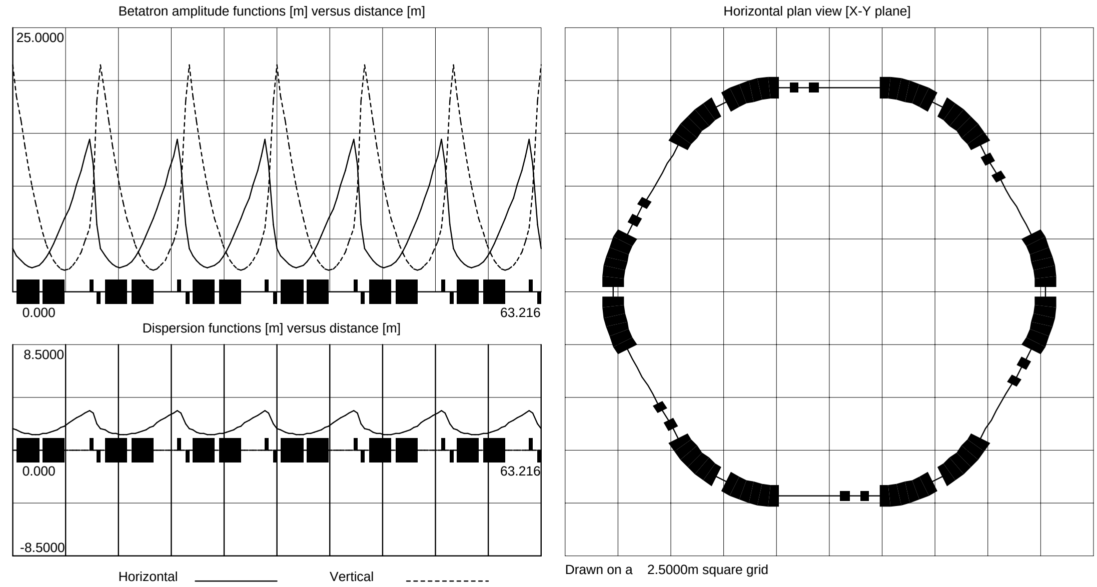


Figure 7.1 A regular cell lattice for a medical machine [1]

7.2 ‘SQUARE’ RING LATTICE

The primary problem with the regular-cell lattice is the inflexibility of its layout. Ideally, one would like to adapt the drift spaces to the tasks they are required to perform. A first step in this direction is to create the ‘square’ lattice, in which the dipole magnets are collected into four ‘corners’ separated by four straight sections with focusing. Figure 7.2 shows the example of the EULIMA ring taken from Reference 2. The sides of the ‘square’ are one pair of long straight sections (on opposite sides) and one pair of short straight sections. The lattice is symmetric about the centres of the long and short straight sections. Injection and extraction are placed in the long straight sections with the two remaining short straight sections being used for rf, diagnostics, etc. There is no zero dispersion region in this example. The underlying cell structure is a *split FODO*.

The main advantages are:

- The straight sections can be designed with longer or shorter lengths (in opposing pairs) with relative ease.
- The dipoles are usually combined into either four 90° units, or perhaps eight 45° units, which makes a compact structure.

The main disadvantages are:

- The magnet yokes and the coils are more difficult to manufacture than say the twelve smaller units of the regular-cell lattice in Figure 7.1.
- Extracting the beam from within a single straight section may require rather extreme fields for the extraction elements (since the phase advance is low), which implies that this type of structure is better suited to protons than to light ions.

The Hardt condition can be arranged and the achromatic transfer between the electrostatic and magnetic septa is automatically satisfied while the extraction is confined to a single straight section. All extraction schemes could be applied in such a lattice.

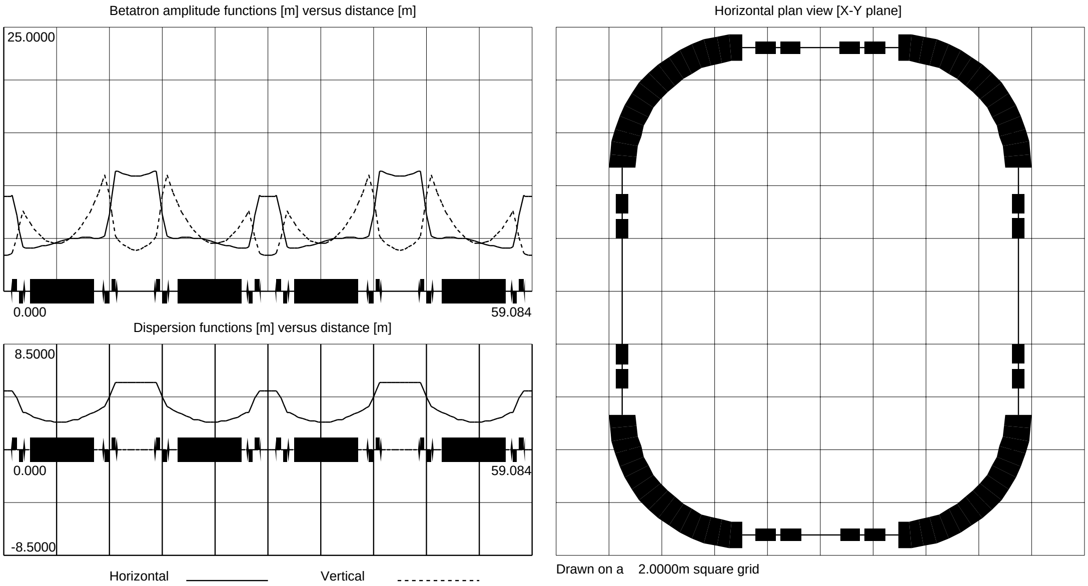


Figure 7.2 EULIMA - A ‘square’ ring lattice [2]

7.3 CUSTOMISED RING LATTICE

As the optics becomes more sophisticated, the need to customise the lattice becomes stronger. Inevitably, the ‘wish’ list for the lattice becomes very long and, in the case of a medical machine, the main points would be:

- The resonance sextupole should be in a dispersion-free region with the horizontal betatron amplitude function large and the vertical betatron amplitude function small.
- The phase advances from the sextupole to the electrostatic septum and the electrostatic septum to the magnetic septum should be chosen according to the extraction scheme.
- The lattice at the septa should have small vertical betatron amplitude functions for aperture reasons and large horizontal betatron amplitude functions for efficiency reasons.
- The dispersion function needs to be shaped for the Hardt Condition (most probably $D>0$ and $D'<0$) at the electrostatic septum.
- If possible, the transit between the two extraction septa should be achromatic.
- To ensure that the longitudinal and transverse oscillations remain uncoupled, rf cavities should be sited in dispersion-free regions.
- Chromaticities are better made negative below transition and positive above transition for stability.
- The co-existence of the separatrices and the ‘waiting’ beam in the same machine imposes more than the usual constraints on the positions of obstacles in the aperture such as collimators, dumps and septa, but primarily, it is necessary to arrange for the separatrices to grow in a balanced way in the aperture.

Whenever a conflict occurs in the design, the overriding need is that of a stable reproducible spill and this will dominate the design choices at all levels. For example, the overall closed-orbit distortion is not a critical factor, but the local distortions at the sextupoles and septa are critical.

Operational considerations favour separate-function lenses for focusing and the provision of separate lenses for correction systems, although this may be counter to economic and/or space considerations. The magnets will be ramped and will therefore have to be laminated to ensure a sufficiently good field quality, but, for reasons of sensitivity to ripple, it is counter-productive to make the time constants too short (see Section 5.11). In general, precision and reproducibility are improved by avoiding backleg windings and trimming supplies on main lattice magnets. Individual correctors are preferable for operation.

Arranging small vertical beam sizes in the main lattice dipole saves ampere-turns and hence power. Equally, small vertical beam sizes in septa and kickers can help the designs considerably.

With the above points in mind, a customised lattice for PIMMS was designed and is shown in Figure 7.3. The underlying cell structure is a partially split FODO where the F is split but the D is not. This could also be viewed as a triplet structure.

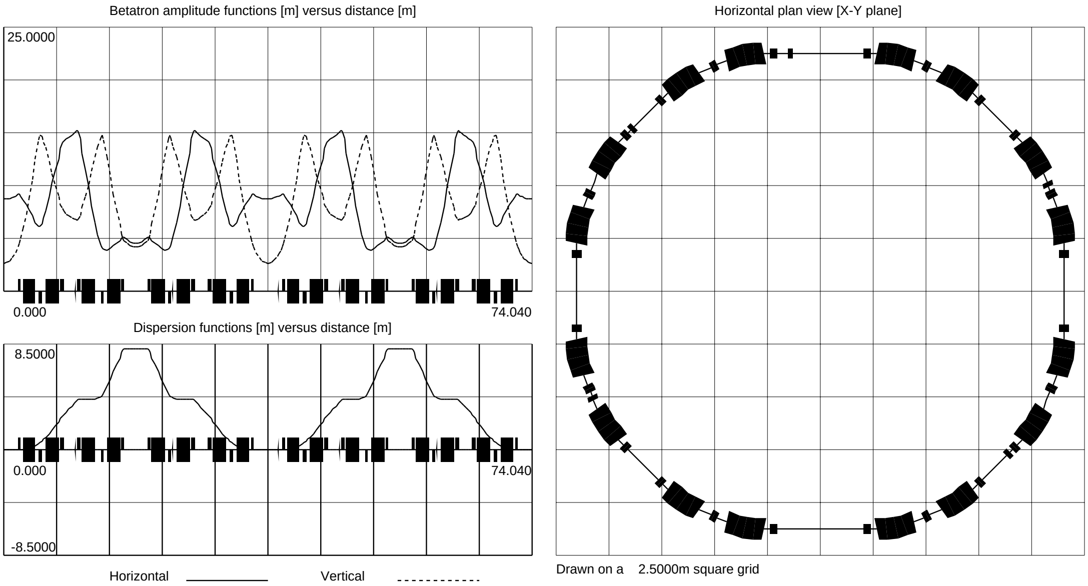
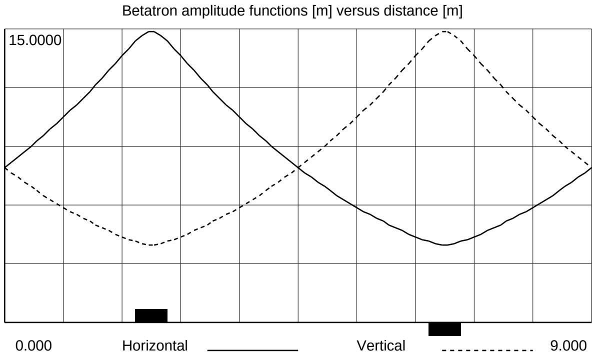


Figure 7.3 The PIMMS customised lattice

7.4 BASIC CELL TYPES

Figure 7.4 summarises four types of lattice cells upon which ring designs are usually based. It is difficult to compare these cells in an absolute way, but, to give some idea of the general properties, all the cells have been scaled to give 72° phase advance in both planes over a distance of 9 m. Beside each diagram, the data for the integrated, normalised quadrupole gradients (MAD sign convention) and the maximum and minimum betatron amplitudes functions are noted. In terms of the maximum betatron amplitude functions, the cells vary over a range of 1.30, but this is not so much when it is realised that the beam sizes are proportional to the square roots of these values and vary only by a factor of 1.14. The integrated gradient strengths vary much more and change by a factor 2.3 between the economical FODO cell (I) and the triplet (III). However, the tuning ranges of these cells are so large that it is unwise to pay too much attention to these generalisations. In most cases, it is better to choose the cell type upon which to a base a lattice according to the requirements of the user. The PIMMS lattice, for example, resembles the split FODO inasmuch as the F is split to produce quasi-constant lattice functions inside septa. The structure also resembles the triplet and this aspect provides some of the smallest vertical beam sizes in the machine inside the extraction and injection equipment where large gaps would be expensive in terms of power and equipment. It is true that one would expect a lower overall power consumption in the lattice, if the structure could be modified to a FODO, but then the septa would not be so well treated and the gradients at injection, that are already uncomfortably low, would become even lower and more prone to fluctuations from remanent fields. In this case, it is felt that the power gain would be marginal and the disadvantages too important to neglect, but of course in a different machine, with a different purpose, the final choice could be very different.

(I) FODO

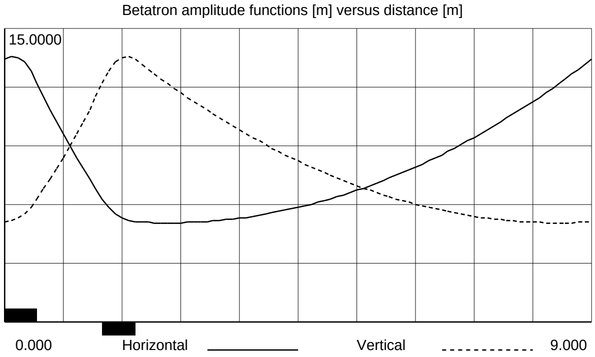


$\Delta\mu_x = 72°$
$\Delta\mu_z = 72°$

$L.k_{F,D} = \pm 0.27027\ \mathrm{m}^{-1}$

$\beta_{x,z,max} = 14.85\ \mathrm{m}$
$\beta_{x,z,min} = 3.98\ \mathrm{m}$

(II) Doublet

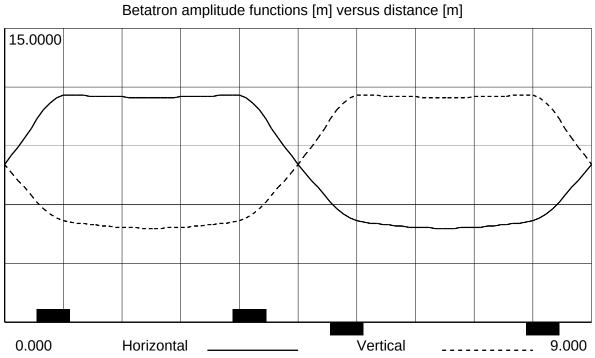


$\Delta\mu_x = 72°$
$\Delta\mu_z = 72°$

$L.k_{F,D} = \pm 0.3748\ \mathrm{m}^{-1}$

$\beta_{x,z,max} = 13.56\ \mathrm{m}$
$\beta_{x,z,min} = 5.08\ \mathrm{m}$

(III) Split FODO, or FOFDOD

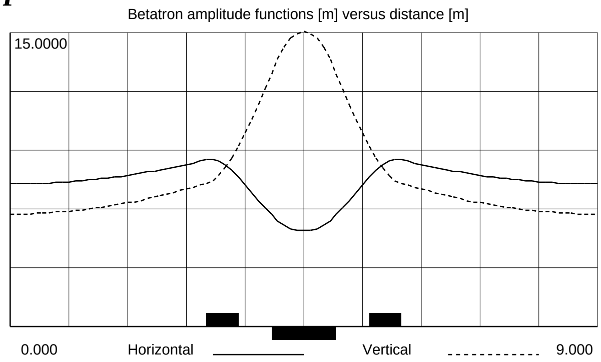


$\Delta\mu_x = 72°$
$\Delta\mu_z = 72°$

$L.k_{F,D} = \pm 0.4863$

$\beta_{x,z,max} = 11.61\ \mathrm{m}$
$\beta_{x,z,min} = 4.80\ \mathrm{m}$

(IV) Triplet



$\Delta\mu_x = 72°$
$\Delta\mu_z = 72°$

$L.k_F = +0.6402$
$L.k_D = -0.6046$

$\beta_{x,z,max} = 8.54/15.05\ \mathrm{m}$
$\beta_{x,z,min} = 4.88/5.74\ \mathrm{m}$

Figure 7.4 Lattice cell types

- ***The FODO cell (I)*** is perhaps the best known and provides the most efficient focusing system. The optimum phase advance in a FODO cell for beam size is ~76°, which is close to the case shown.

- ***The doublet structure (II)*** conveniently provides one long and one short drift space, but since the quadrupoles are closer together they must be stronger and the power consumption increases.

- ***The split FODO, or FOFDOD, structure (III)*** provides a flexible lattice in which the FOF and the DOD spacing can be varied over a wide range without instability. Of the examples shown, it has the smallest beam sizes and the quasi-constant regions of the betatron amplitude functions, in which one plane is higher than the other, can be used to create a semi-independence between the planes even over long pieces of equipment. However, the power consumption is again higher with respect to the FODO.

- ***The triplet structure (IV)*** provides, a low-β-shape for the lattice functions in both planes in the long drifts. This is especially useful for equipment that needs small apertures in both planes. The peak that occurs at the centre of the triplet is usually arranged to be in the vertical plane so that the horizontal betatron amplitude function is kept small at all positions. In this way, space is created for the horizontal dispersion function and the two planes end up on aggregate as being well balanced from the point of view of aperture.

7.5 COMMENTS

- ♦ The general consensus of opinion is that proton machines for passive spreading should be made as small as possible with as few dipoles as possible, which favours the ‘square’ ring with either a quadrupole or an rf-driven extraction. The lower magnetic rigidity of the protons reduces the overall size of the dipoles to a point where the extra complexity of the coil and yoke design for the large bending angles is not so important.

- ♦ When active scanning is considered, there is an advantage in using light ions because these particles scatter far less in the body and the spot size can therefore be controlled more effectively than for protons. A synchrotron is then the preferred machine and a customised lattice adapted to the Hardt Condition with the amplitude-momentum extraction scheme driven by a betatron core and some form of front-end acceleration will give the best results as regards a smooth spill.

- ♦ This leaves the case of protons with active scanning. The recommendation would be to use a synchrotron with a customised lattice that can apply the Hardt Condition, a betatron core and some form of front-end acceleration. If the equipment already exists, or space is a critical problem, then the betatron core should take priority. If the betatron core cannot be installed, then a quadrupole extraction is the next choice, but the quadrupole should have a time constant of at least 100 μs in order to gain some smoothing for the multi-kHz ripple.

- The lattice should NOT be chosen without first choosing the extraction method and the type of treatment to be performed. The best results are obtained by making an integrated choice.

The example ring lattices given in this chapter were all machines designed for use with light ions. For proton operation only, the lattices would have been smaller. The progression from the regular-cell lattice to the customised lattice is accompanied by an increase in the dispersion function. This is to be expected. In the customised lattice, the dispersion function is forced to zero, which means that there is an oscillation about the average value of 3 to 4 m with troughs going to zero and peaks going to about twice the average value. If the lattice were much larger, it would be possible to use a dispersion suppressor and to avoid this oscillation.

References

[1] Chen Wuzhong, *Design of a light ion medical synchrotron*, GSI-92-24 (October 1992).
[2] G. Cesari, P. Lefèvre, D. Vandeplassche, *Feasibility study of a synchrotron for the European light ion medical accelerator*, CERN/PS/91-08(DI).

* * *

I-8 MATCHING TO GANTRIES

In the extraction line, a rather special situation is met in which a fixed transfer line must be matched to a section of line, called the gantry, that has to be able to rotate through a full 360° without affecting the beam spot at the patient [1] (see Figure 8.1). Emittance inversion insertions are fairly commonplace in accelerator laboratories, but not lines with full rotational optics as required in this case. In cyclotron-based facilities, the problem is solved by making the beam distribution rotationally symmetric at an interface between the fixed and rotating lines. A rotationally symmetric beam requires equal emittances, equal lattice parameters and zero dispersion. Unfortunately, the slow-extracted beam from a synchrotron, as described in Chapter 4, is far from fulfilling these requirements. The problem is soluble in an elegant way, but it requires the extraction line(s) and the gantry(ies) to be designed according to an overall plan. Finally, the high magnetic rigidity of the carbon beam makes it desirable to investigate forms of gantries other than the iso-centric gantry shown in Figure 8.1.

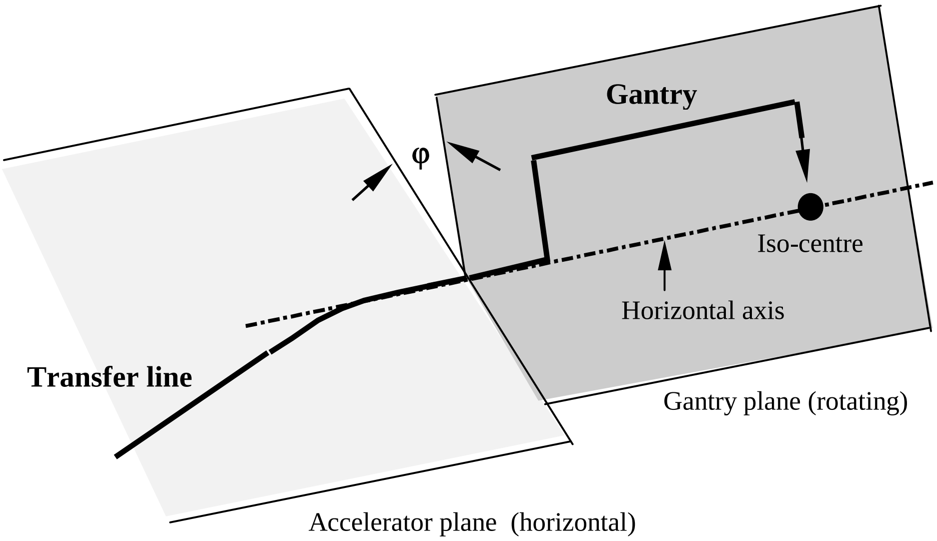


Figure 8.1 Schematic view of an iso-centric gantry

8.1 MATCHING METHODS

A gantry makes it possible to deliver the radiation dose to the tumour from different angles and hence to spread out the entry dose and to avoid critical organs. The value of the gantry is clear, but it raises the problem of how to make the optics of the transfer line and gantry completely independent of the gantry rotation, so that there is:

- No change in spot size and shape at the treatment volume.
- No correlation between momentum and position.
- No change of the beam optics inside the gantry.

There appear to be only two methods of matching, that satisfy all the above requirements, and a third method that partially fulfills them. A brief description of these methods is given below before the tools are developed for analysing them.

Symmetric-Beam Method:

This method requires a fully symmetric beam at an interface point between the fixed beam line and the gantry. A fully symmetric beam would have gaussian or KV[*] distributions with equal emittances and lattice functions and zero dispersion functions. These requirements automatically mean that the beam will be rotationally symmetric in real space.

Round-Beam Method:

In this method the phase advances in the gantry are made multiples of π in both transverse planes. This matrix family has the top right element zero and acts as a telescope giving constant magnifications of the betatron amplitude functions. In the general case, the alpha functions are dependent on the incoming betatron amplitude functions, which may be acceptable in a gantry design, but is not aesthetic. In most cases, the entry and exit alpha functions would be made equal and most probably zero, so that the transfer matrices for each plane will have all zero off-diagonal terms. The one-to-one, or identity matrix is then a special member of this family. This type of matrix will map the beam directly to the patient at the exit with a constant magnification. If the input beam is round then the output beam will also be round, but scaled by the magnification, and the rotation of the gantry will not be evident. In this case, a round beam means $E_x\beta_x = E_z\beta_z$ with gaussian or KV distributions, which is a slightly less stringent condition than that required for the symmetric beam method. The dispersion function must also be zero as for the symmetric beam method. The only problem that occurs is that the optics inside the gantry changes with rotation, which could affect the beam steering/scanning in gantries that incorporate scanning in their optics. Although each case has to be studied separately, it is often possible to freeze the optical parameters in the last section of the gantry where the steering occurs.

Rotator Method:

The complete solution to this problem is the ‘rotator’ which is a section of quadrupole lattice with phase advances of 2π and π in the transverse planes that is placed just upstream of the gantry [2]. This module has to be physically rotated by half of the gantry angle. Mathematically, the mapping from the fixed line to the rotating gantry is exactly one-to-one and one-to-minus-one and there are no restrictions on the lattice functions or the dispersion vectors. As will be explained later, the practical design does require some care and beam sizes may impose practical limits.

8.2 TOOLS FOR DESCRIBING THE MATCHING METHODS

8.2.1 σ - Matrix Formalism [3]

Statistical average of a distribution

Let $\boldsymbol{y}$ be a vector containing the transverse phase-space co-ordinates of a particle. The statistical averages that describe a distribution of particles in phase space are then contained in the co-variance σ-matrix defined as,

---

[*] K-V stands for Kapchinskij-Vladimirskij and refers to distributions that are correlated such that E_x+E_z = constant for each particle. This form of correlation erases any indication of rotation asymmetry without skewing the 4-D ellipsoid i.e. $<xx'>$ and similar cross terms remain zero.

Definition of co-variance matrix, $\sigma = \left\langle \boldsymbol{y}\boldsymbol{y}^{\mathrm{T}} \right\rangle$ (8.1)

The elements of the σ-matrix are therefore given by,

$$\sigma = \left\langle \begin{pmatrix} x \\ x' \\ z \\ z' \end{pmatrix} \begin{pmatrix} x & x' & z & z' \end{pmatrix} \right\rangle = \begin{pmatrix} \langle x^2 \rangle & \langle xx' \rangle & \langle xz \rangle & \langle xz' \rangle \\ \langle x'x \rangle & \langle x'^2 \rangle & \langle x'z \rangle & \langle x'z' \rangle \\ \langle zx \rangle & \langle zx' \rangle & \langle z^2 \rangle & \langle zz' \rangle \\ \langle z'x \rangle & \langle z'x' \rangle & \langle z'z \rangle & \langle z'^2 \rangle \end{pmatrix} \tag{8.2}$$

where the < >brackets indicate the estimators for the expectation values for moments of the *N* particles of,

$$\langle x \rangle = \frac{1}{N}\sum_{i=1}^{n} x_{\mathrm{i}}\,; \quad \langle x^2 \rangle = \frac{1}{N-1}\sum_{\mathrm{i}=1}^{n}\left(x_{\mathrm{i}} - \langle x \rangle\right)^2; \quad \langle xz \rangle = \frac{1}{N-1}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left(x_{\mathrm{i}} - \langle x \rangle\right)\left(z_{\mathrm{i}} - \langle z \rangle\right); \quad \text{etc.}$$

Transformation properties of the σ-matrix

Let ***M*** represent a linear transformation, so that,

$$\boldsymbol{y}_2 = \boldsymbol{M}\boldsymbol{y}_1\,. \tag{8.3}$$

The use of this linear transformation with the definition of the σ-matrix gives[*] ,

$$\sigma_2 = \left\langle \boldsymbol{y}_2 \boldsymbol{y}_2^{\mathrm{T}} \right\rangle = \left\langle \boldsymbol{M}\boldsymbol{y}_1 (\boldsymbol{M}\boldsymbol{y}_1)^{\mathrm{T}} \right\rangle = \boldsymbol{M}\underbrace{\left\langle \boldsymbol{y}_1 \boldsymbol{y}_1^{\mathrm{T}} \right\rangle}_{\sigma_1}\boldsymbol{M}^{\mathrm{T}}\,.$$

Hence,

Transformation $\sigma_2 = \boldsymbol{M}\sigma_1\boldsymbol{M}^{\mathrm{T}}\,.$ (8.4)

The matrix ***M*** can either be a transfer matrix for any linear lattice, or a rotation matrix, describing the rotation of a section of beam line. The knowledge of the σ-matrix at one point in a linear lattice is therefore sufficient to calculate the σ-matrix anywhere else provided the transfer matrices are known. This applies equally to lattices where sections have rotated about their axes.

Invariance

Consider, $\boldsymbol{W} = \boldsymbol{y}^{\mathrm{T}}\sigma^{-1}\boldsymbol{y}$, providing σ^{-1} (known as the error matrix) exists. The evaluation of ***W*** can be made at two positions related by $\boldsymbol{y}_2 = \boldsymbol{M}\boldsymbol{y}_1$ so that,

$$\boldsymbol{W}_2 = \boldsymbol{y}_2^{\mathrm{T}}\sigma_2^{-1}\boldsymbol{y}_2 = (\boldsymbol{M}\boldsymbol{y}_1)^{\mathrm{T}}\sigma_2^{-1}(\boldsymbol{M}\boldsymbol{y}_1) = \boldsymbol{y}_1^{\mathrm{T}}\boldsymbol{M}^{\mathrm{T}}\sigma_2^{-1}\boldsymbol{M}\boldsymbol{y}_1\,. \tag{8.5}$$

With (8.4) and the standard relationship $(a\,b\,c)^{-1} = c^{-1}b^{-1}a^{-1}$,

---

[*] $(a.b)^{\mathrm{T}} = b^{\mathrm{T}}\cdot a^{\mathrm{T}}$

$$\sigma_2^{-1} = \left(\boldsymbol{M}\sigma_1\boldsymbol{M}^{\mathrm{T}}\right)^{-1} = \left(M^{\mathrm{T}}\right)^{-1}\sigma_1^{-1}\boldsymbol{M}^{-1}.$$

The substitution of this result in (8.5) gives,

$$\begin{aligned}\boldsymbol{W}_2 &= \boldsymbol{y}_2^{\mathrm{T}}\sigma_2^{-1}\boldsymbol{y}_2 = \boldsymbol{y}_1^{\mathrm{T}}\underbrace{\boldsymbol{M}^{\mathrm{T}}\left(\boldsymbol{M}^{\mathrm{T}}\right)^{-1}}_{\boldsymbol{I}}\sigma_1^{-1}\underbrace{\boldsymbol{M}^{-1}\boldsymbol{M}}_{\boldsymbol{I}}\boldsymbol{y}_1 \\ &= \boldsymbol{y}_1^{\mathrm{T}}\sigma_1^{-1}\boldsymbol{y}_1 = \boldsymbol{W}_1\end{aligned}$$

Thus, ***W*** is an invariant of the beam,

Invariant, $$\boldsymbol{y}_2^{\mathrm{T}}\sigma_2^{-1}\boldsymbol{y}_2 = \boldsymbol{y}_1^{\mathrm{T}}\sigma_1^{-1}\boldsymbol{y}_1 = \boldsymbol{W} \tag{8.6}$$

An uncoupled beam

An uncoupled beam is one that has no correlation between the two transverse phase spaces so that all elements of the matrix that couple the horizontal and the vertical phase spaces vanish to give the form shown below.

$$\sigma_{\mathrm{uc}} = \begin{pmatrix} \langle x^2\rangle & \langle xx'\rangle & \langle xz\rangle & \langle xz'\rangle \\ \langle x'x\rangle & \langle x'^2\rangle & \langle x'z\rangle & \langle x'z'\rangle \\ \langle zx\rangle & \langle zx'\rangle & \langle z^2\rangle & \langle zz'\rangle \\ \langle z'x\rangle & \langle z'x'\rangle & \langle z'z\rangle & \langle z'^2\rangle \end{pmatrix} = \begin{pmatrix} \langle x^2\rangle & \langle xx'\rangle & 0 & 0 \\ \langle x'x\rangle & \langle x'^2\rangle & 0 & 0 \\ 0 & 0 & \langle z^2\rangle & \langle zz'\rangle \\ 0 & 0 & \langle z'z\rangle & \langle z'^2\rangle \end{pmatrix}. \tag{8.7}$$

The invariant ***W*** for an uncoupled beam is formed as before,

$$\boldsymbol{W} = \boldsymbol{y}^{\mathrm{T}}\sigma^{-1}\boldsymbol{y},$$

$$\boldsymbol{W} = \begin{pmatrix} x & x' & z & z' \end{pmatrix}\begin{pmatrix} \langle x'^2\rangle & \langle -xx'\rangle & 0 & 0 \\ \langle -x'x\rangle & \langle x^2\rangle & 0 & 0 \\ 0 & 0 & \langle z'^2\rangle & \langle -zz'\rangle \\ 0 & 0 & \langle -z'z\rangle & \langle z^2\rangle \end{pmatrix}\begin{pmatrix} x \\ x' \\ z \\ z' \end{pmatrix}$$

$$\boldsymbol{W} = \begin{pmatrix} x & x' & z & z' \end{pmatrix}\begin{pmatrix} x\langle x'^2\rangle - x'\langle xx'\rangle \\ -x\langle xx'\rangle + x'\langle x^2\rangle \\ z\langle z'^2\rangle - z'\langle zz'\rangle \\ -z\langle zz'\rangle + z'\langle z^2\rangle \end{pmatrix}.$$

Thus, for an uncoupled beam the general invariant, ***W***, separates into two independent invariants, $\boldsymbol{W}_{\mathrm{x}}$ and $\boldsymbol{W}_{\mathrm{z}}$

$$\boldsymbol{W} = \underbrace{x^2\left\langle x'^2\right\rangle - 2xx'\left\langle xx'\right\rangle + x'^2\left\langle x^2\right\rangle}_{\boldsymbol{W}_\mathrm{x}} + \underbrace{z^2\left\langle z'^2\right\rangle - 2zz'\left\langle zz'\right\rangle + z'^2\left\langle z^2\right\rangle}_{\boldsymbol{W}_\mathrm{z}} \tag{8.8}$$

This can also be written as,

$$\boldsymbol{W}_\mathrm{uncoupled} = \boldsymbol{y}^\mathrm{T}\sigma^{-1}\boldsymbol{y} = \underbrace{\boldsymbol{x}^\mathrm{T}\sigma_\mathrm{x}{}^{-1}\boldsymbol{x}}_{\boldsymbol{W}_\mathrm{x}} + \underbrace{\boldsymbol{z}^\mathrm{T}\sigma_\mathrm{z}{}^{-1}\boldsymbol{z}}_{\boldsymbol{W}_\mathrm{z}} \tag{8.9}$$

8.2.2 Link between the σ - matrix and the Courant and Snyder formulation

If the beam is uncoupled, it is sufficient to consider just one transverse plane. The derived invariant for the *x*-plane from (8.8)

$$\boldsymbol{W}_\mathrm{x} = \left\langle x'^2\right\rangle x^2 - 2\left\langle xx'\right\rangle xx' + \left\langle x^2\right\rangle x'^2 \tag*{(8.8)$'$}$$

is strongly reminiscent of the Courant-Snyder Invariant

$$E_\mathrm{x} = \pi\left(\gamma_\mathrm{x}x^2 + 2\alpha_\mathrm{x}xx' + \beta_\mathrm{x}x'^2\right). \tag{8.9}$$

With the help of the definition of the statistical emittance,

$$E_\mathrm{x} = \pi\sqrt{\left\langle x^2\right\rangle\left\langle x'^2\right\rangle - \left\langle xx'\right\rangle^2}\ , \tag{8.10}$$

the invariant $\boldsymbol{W}_\mathrm{x}$ can then be rewritten as,

$$\boldsymbol{W}_\mathrm{x} = E_\mathrm{x}\left\{x^2\,\frac{\left\langle x'^2\right\rangle}{E_\mathrm{x}} - 2xx'\,\frac{\left\langle xx'\right\rangle}{E_\mathrm{x}} + x'^2\,\frac{\left\langle x^2\right\rangle}{E_\mathrm{x}}\right\}. \tag{8.11}$$

By comparison with (8.9), the Courant and Snyder parameters can now be defined as,

$$\gamma_\mathrm{x} = \pi\frac{\left\langle x'^2\right\rangle}{E_\mathrm{x}},\ \alpha_\mathrm{x} = -\pi\frac{\left\langle xx'\right\rangle}{E_\mathrm{x}},\ \beta_\mathrm{x} = \pi\frac{\left\langle x^2\right\rangle}{E_\mathrm{x}}. \tag{8.12}$$

It is quickly verified that these definitions satisfy the Courant and Snyder relationship, $\gamma = \dfrac{\left(1+\alpha^2\right)}{\beta}$.

Finally, the bridge between the σ-matrix and the Courant and Snyder parameters is completed by writing the σ-matrix for an uncoupled beam in terms of the Courant and Snyder parameters,

$$\sigma_{\mathrm{uc}} = \frac{1}{\pi}\begin{pmatrix} E_{\mathrm{x}}\beta_{\mathrm{x}} & -E_{\mathrm{x}}\alpha_{\mathrm{x}} & 0 & 0 \\ -E_{\mathrm{x}}\alpha_{\mathrm{x}} & E_{\mathrm{x}}\gamma_{\mathrm{x}} & 0 & 0 \\ 0 & 0 & E_{\mathrm{z}}\beta_{\mathrm{z}} & -E_{\mathrm{z}}\alpha_{\mathrm{z}} \\ 0 & 0 & -E_{\mathrm{z}}\alpha_{\mathrm{z}} & E_{\mathrm{z}}\gamma_{\mathrm{z}} \end{pmatrix}. \tag{8.13}$$

Hence it is possible to evaluate the σ-matrix for the uncoupled beam, extracted from the accelerator, in terms of the usual lattice parameters and then to find the σ-matrices at any position downstream (e.g. in the rotator, gantry, etc.) by the use of the appropriate transfer and/or rotation matrices. Once the beam is coupled, either after a rotation or after a coupling element such as a skew quadrupole, the Courant and Synder formalism no longer applies, but the sigma matrix maintains a full description of the beam and the 1-σ beam sizes in the planes of the normal modes can be found very easily from the square roots of the matrix elements σ_{11} and σ_{33}.

8.3 SYMMETRIC-BEAM METHOD

In the symmetric beam method, the gantry is matched directly to the fixed beam line coming from the accelerator. This situation is shown schematically in Figure 8.2.

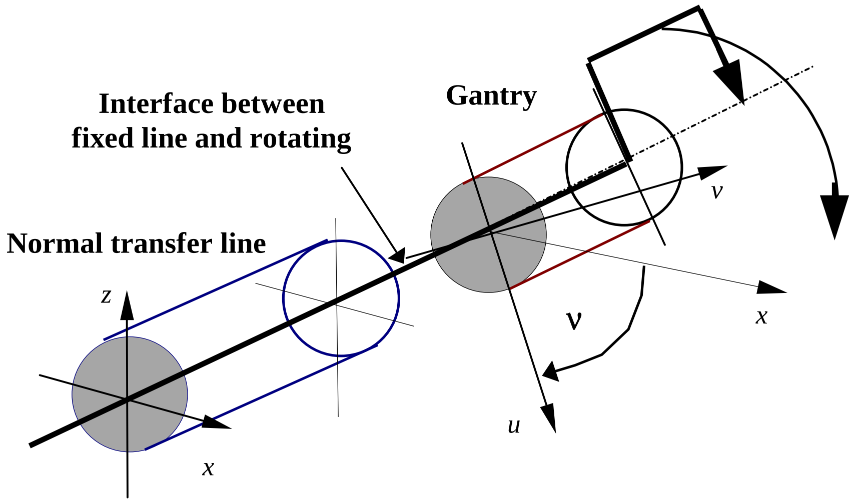


Figure 8.2 Equipment layout for the symmetric beam method

Assume that the beam coming from the accelerator is uncoupled. The σ-matrix at the end of the fixed beam line, just in front of the gantry is then given by,

$$\sigma_{\mathrm{uc}} = \frac{1}{\pi}\begin{pmatrix} E_{\mathrm{x}}\beta_{\mathrm{x}} & -E_{\mathrm{x}}\alpha_{\mathrm{x}} & 0 & 0 \\ -E_{\mathrm{x}}\alpha_{\mathrm{x}} & E_{\mathrm{x}}\gamma_{\mathrm{x}} & 0 & 0 \\ 0 & 0 & E_{\mathrm{z}}\beta_{\mathrm{z}} & -E_{\mathrm{z}}\alpha_{\mathrm{z}} \\ 0 & 0 & -E_{\mathrm{z}}\alpha_{\mathrm{z}} & E_{\mathrm{z}}\gamma_{\mathrm{z}} \end{pmatrix}_{\mathrm{Interfacee}} . \tag{8.14}$$

If the gantry is rotated by an angle ν, the positions and angles of all particles just at the junction can be described in the rotated co-ordinate system (*u*,*v*) by multiplication with the rotation matrix $\boldsymbol{R}_{\nu}$,

$$\boldsymbol{R}_{\upsilon} = \begin{pmatrix} \cos\upsilon & 0 & \sin\upsilon & 0 \\ 0 & \cos\upsilon & 0 & \sin\upsilon \\ -\sin\upsilon & 0 & \cos\upsilon & 0 \\ 0 & -\sin\upsilon & 0 & \cos\upsilon \end{pmatrix}. \qquad (8.15)$$

The σ-matrix after a rotation can be derived from the original matrix by using the transformation according to (8.4) to give,

$$\sigma_2 = \boldsymbol{R}_{\upsilon}\sigma_1\boldsymbol{R}_{\upsilon}{}^{\mathrm{T}} = \frac{1}{\pi}\times$$

$$\begin{pmatrix} c^2E_{\mathrm{x}}\beta_{\mathrm{x}} + s^2E_{\mathrm{z}}\beta_{\mathrm{z}} & -c^2E_{\mathrm{x}}\alpha_{\mathrm{x}} - s^2E_{\mathrm{z}}\alpha_{\mathrm{z}} & sc(-E_{\mathrm{x}}\beta_{\mathrm{x}} + E_{\mathrm{z}}\beta_{\mathrm{z}}) & sc(E_{\mathrm{x}}\alpha_{\mathrm{x}} - E_{\mathrm{z}}\alpha_{\mathrm{z}}) \\ -c^2E_{\mathrm{x}}\alpha_{\mathrm{x}} - s^2E_{\mathrm{z}}\alpha_{\mathrm{z}} & c^2E_{\mathrm{x}}\gamma_{\mathrm{x}} + s^2E_{\mathrm{z}}\gamma_{\mathrm{z}} & sc(E_{\mathrm{x}}\alpha_{\mathrm{x}} - E_{\mathrm{z}}\alpha_{\mathrm{z}}) & sc(-E_{\mathrm{x}}\gamma_{\mathrm{x}} + E_{\mathrm{z}}\gamma_{\mathrm{z}}) \\ sc(-E_{\mathrm{x}}\beta_{\mathrm{x}} + E_{\mathrm{z}}\beta_{\mathrm{z}}) & sc(-E_{\mathrm{x}}\alpha_{\mathrm{x}} + E_{\mathrm{z}}\alpha_{\mathrm{z}}) & s^2E_{\mathrm{x}}\beta_{\mathrm{x}} + c^2E_{\mathrm{z}}\beta_{\mathrm{z}} & -s^2E_{\mathrm{x}}\alpha_{\mathrm{x}} - c^2E_{\mathrm{z}}\alpha_{\mathrm{z}} \\ sc(-E_{\mathrm{x}}\alpha_{\mathrm{x}} + E_{\mathrm{z}}\alpha_{\mathrm{z}}) & sc(-E_{\mathrm{x}}\gamma_{\mathrm{x}} + E_{\mathrm{z}}\gamma_{\mathrm{z}}) & -s^2E_{\mathrm{x}}\alpha_{\mathrm{x}} - c^2E_{\mathrm{z}}\alpha_{\mathrm{z}} & s^2E_{\mathrm{x}}\gamma_{\mathrm{x}} + c^2E_{\mathrm{z}}\gamma_{\mathrm{z}} \end{pmatrix} (8.16)$$

where $s = \sin\nu$ and $c = \cos\nu$ and ν = the angle of rotation.

The matrix σ_2 gives the various averages in the rotated gantry co-ordinate system (u,v) in terms of the beam parameters of the fixed beam line co-ordinate system (x,z).

If the gantry is not rotated ($\nu = 0$ and $\sin\nu = 0$), all terms in the off-axis quadrants are zero and the main terms regain the form of the uncoupled matrix. Similarly, for a rotation $\nu = 90°$, the matrix σ_2 is of the same form as the uncoupled matrix but with x and z interchanged.

For the beam to be symmetric with respect to rotations, the matrix σ_2 must be independent of the rotation angle ν and the following constraints have to be fulfilled at the entry to the gantry:

$$E_{\mathrm{x}}\beta_{\mathrm{x}} = E_{\mathrm{z}}\beta_{\mathrm{z}}; \quad E_{\mathrm{x}}\alpha_{\mathrm{x}} = E_{\mathrm{z}}\alpha_{\mathrm{z}}; \quad E_{\mathrm{x}}\gamma_{\mathrm{x}} = E_{\mathrm{z}}\gamma_{\mathrm{z}},$$

which is equivalent to,

Symmetric beam conditions,

$$E_{\mathrm{x}} = E_{\mathrm{z}}$$
$$\beta_x = \beta_z;\ \alpha_x = \alpha_z;\ \gamma_x = \gamma_z \qquad (8.17)$$
$$D_{\mathrm{x}} = D_{\mathrm{z}} = 0; \quad D'_{\mathrm{x}} = D'_{\mathrm{z}} = 0$$

To these conditions, it is necessary to add that the beam distributions in x and z must be either gaussian or K-V in order that the rotation of the beam spot is not made visible by a structure in the beam that does not affect the statistical expectations[*] .

---

[*] For example, uncorrelated, truncated gaussians would form a square image in real space that would make rotation visible, although the emittances etc. could still be equal.

To have the same lattice parameters in both planes is not very limiting in practice, but the constraint of equal emittances is a more severe problem, especially for resonant extraction from a synchrotron. Entering and leaving with zero dispersion functions is also a fairly strict optical condition for the gantry.

The above conditions demand that the beam be symmetric in real space and in the two transverse phase spaces. It is not sufficient to have only a physically 'round' beam at the interface point. This is quickly seen by considering a simple case of a beam with $\alpha_x = \alpha_z = 0$ and $E_x = n \cdot E_z$. To make this beam round in real space, $E_x\beta_x = E_z\beta_z$ is required and from this it follows that $\beta_x = \frac{1}{n}\beta_z$ and because of $\alpha_{x,z}$ being zero $\gamma_x = n \cdot \gamma_z$. From the above conditions $E_x\gamma_x = E_z\gamma_z$ must also be fulfilled and it follows that:

$$E_x\gamma_x = nE_z n\gamma_z = n^2 E_z\gamma_z = E_z\gamma_z \Leftrightarrow n = 1 \Rightarrow E_x = E_z\,,$$

which forces the choice of *n* to be unity for equal emittances.

8.4 ROUND-BEAM METHOD

In this method, the equipment layout is the same as for the symmetric-beam method (see Figure 8.2), but additional constraints are placed on the gantry optics in order to relieve partially the constraints on the beam at the interface point.

The general transfer matrix for a section of line can be written as,

$$\begin{pmatrix} \sqrt{\frac{\beta_2}{\beta_1}}(\cos\Delta\mu + \alpha_1 \sin\Delta\mu) & \sqrt{\beta_1\beta_2}\sin\Delta\mu \\ -\frac{1}{\sqrt{\beta_1\beta_2}}\left[(1+\alpha_1\alpha_2)\sin\Delta\mu + (\alpha_2 - \alpha_1)\cos\Delta\mu\right] & \sqrt{\frac{\beta_1}{\beta_2}}(\cos\Delta\mu - \alpha_2\sin\Delta\mu) \end{pmatrix} \tag{8.18}$$

where the subscripts 1 and 2 refer to the entry and exit parameters respectively and $\Delta\mu$ is the phase advance through that section. If the phase advance is adjusted to an integral number of π, then the matrix simplifies to,

$$\begin{pmatrix} \pm\sqrt{\frac{\beta_2}{\beta_1}} & 0 \\ \mp\frac{1}{\sqrt{\beta_1\beta_2}}(\alpha_2 - \alpha_1) & \pm\sqrt{\frac{\beta_1}{\beta_2}} \end{pmatrix}. \tag{8.19}$$

Although this form could be accepted as a gantry matrix, it is far more likely that the alpha functions would be made equal at the entry and exit and very probably zero. Equation (8.19) would then be rewritten as,

$$\begin{pmatrix} \pm m & 0 \\ 0 & \pm m^{-1} \end{pmatrix} \tag{8.20}$$

where m is the magnification factor equal to the square root of the ratio of the betatron amplitude functions at entry and exit. This type of matrix is known as a ‘telescope’ and will be discussed further in Section 9.2 The identity, or one-to-one matrix also belongs to this family.

The rotation of the gantry is described with a rotation matrix $\boldsymbol{R}_\nu$ that was given in (8.15). Thus the overall transfer matrix from the end of the fixed beam line to the treatment volume is then given by the product of (8.15) and (8.20) assuming equal magnifications in the two planes for simplicity:

$$\boldsymbol{M}_0 = \boldsymbol{M}_{\text{gantry}} \boldsymbol{R}_\upsilon = \begin{pmatrix} m\cos\upsilon & 0 & m\sin\upsilon & 0 \\ 0 & m^{-1}\cos\upsilon & 0 & m^{-1}\sin\upsilon \\ -m\sin\upsilon & 0 & m\cos\upsilon & 0 \\ 0 & -m^{-1}\sin\upsilon & 0 & m^{-1}\cos\upsilon \end{pmatrix} \tag{8.21}$$

where an even number of π have been chosen for the phase advances for simplicity.

At the treatment volume, the spot size and shape have to be independent of the rotation angle. To derive the required beam parameters at the matching point, it is sufficient to consider only the 2x2 transfer matrix T_0 for x and z, (the matrix for x', z' being the similar):

$$\boldsymbol{T}_0 = m \begin{pmatrix} \cos\upsilon & \sin\upsilon \\ -\sin\upsilon & \cos\upsilon \end{pmatrix}. \tag{8.22}$$

If now the incoming beam had an asymmetry e.g. $|x| < x_0$ and $z = 0$, then at the treatment volume the distribution can be described in the gantry co-ordinate system (u,v) with the matrix $\boldsymbol{T}_0$ as,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \boldsymbol{T}_0 \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} = m \begin{pmatrix} x \cdot \cos\upsilon \\ -x \cdot \sin\upsilon \end{pmatrix} \tag{8.23}$$

This means that the beam, seen from the gantry co-ordinate system, is rotated at the treatment volume by the negative gantry angle ν (see Figure 8.3).

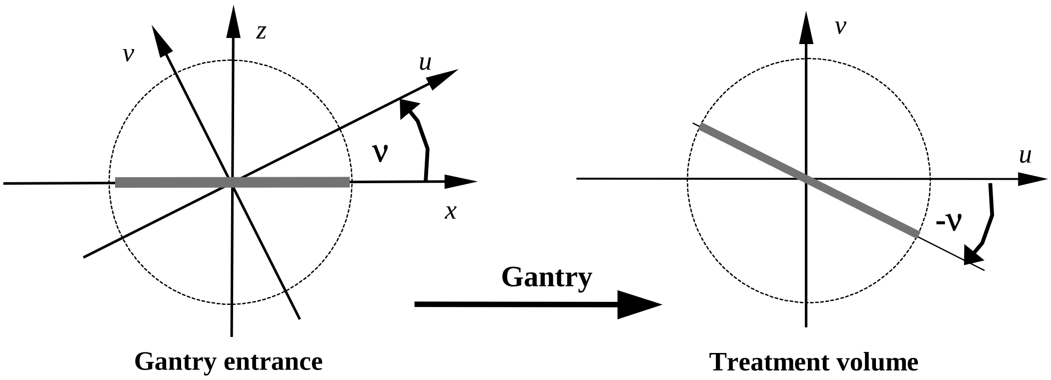


Figure 8.3 Rotation of the beam in the round-beam method

The rotation of the beam at the patient cannot be avoided, but it can be made of no consequence by entering the gantry with a ‘round’ beam, so that the patient always ‘sees’ the same particle distribution independent of the gantry angle. In this case, a ‘round’ beam in real space requires,

Round beam conditions,

$$E_x\beta_x = E_z\beta_z \qquad (8.24)$$

$$D_x = D_z = 0; \quad D'_x = D'_z = 0$$

at the interface point with gaussian or K-V distributions.

This constraint is slightly less strict than that required for the symmetric beam in Section 8.3, but the optics (beam sizes etc.) are not constant inside the gantry with rotation. For gantries using passive spreading after the last bend, the variable optics are of no consequence, but for gantries which incorporate active scanning, this could be a problem.

8.5 ROTATOR METHOD

The only truly rotational solution that also includes the dispersion vectors is the rotator method. Consider a section of bending-free transfer line with a betatron phase advance of 360^0 in the x-plane and 180^0 in the z-plane and $\nu_{\mathrm{entry}} = \nu_{\mathrm{exit}}$. The transfer matrix for this line will be;

$$M_{\mathrm{rotator}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \qquad (8.25)$$

Now let this line be rotated by an angle ν with respect to the normal fixed transfer line and let ν be just half of the rotation angle of the gantry, 2ν. The layout of the equipment is schematically shown in Figure 8.4.

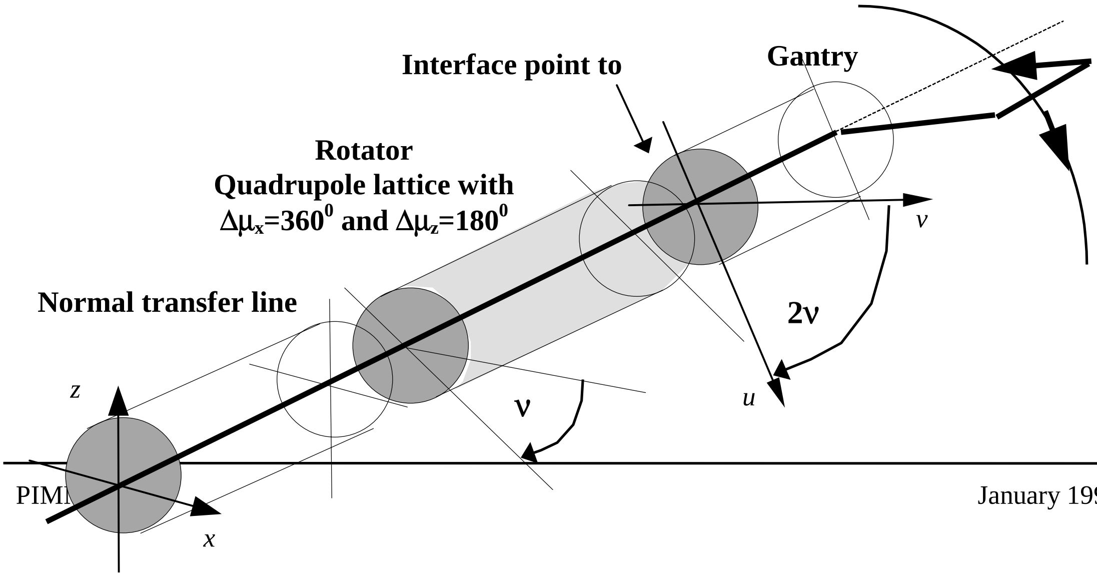


Figure 8.4 Equipment layout for the rotator method

The transfer matrix from the exit of the fixed beam line to the interface point at the entry to the gantry is found by multiplying the rotation matrix for the angle ν, the rotator and a second rotation matrix also for the angle ν.

$$\begin{pmatrix} \cos\upsilon & 0 & \sin\upsilon & 0 \\ 0 & \cos\alpha & 0 & \sin\upsilon \\ -\sin\upsilon & 0 & \cos\upsilon & 0 \\ 0 & -\sin\alpha & 0 & \cos\upsilon \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\upsilon & 0 & \sin\upsilon & 0 \\ 0 & \cos\upsilon & 0 & \sin\upsilon \\ -\sin\upsilon & 0 & \cos\alpha & 0 \\ 0 & -\sin\upsilon & 0 & \cos\upsilon \end{pmatrix}$$

$$\begin{pmatrix} \cos^2\upsilon+\sin^2\upsilon & 0 & \sin\upsilon\cos\upsilon-\cos\upsilon\sin\upsilon & 0 \\ 0 & \cos^2\upsilon+\sin^2\upsilon & 0 & \sin\upsilon\cos\upsilon-\cos\upsilon\sin\upsilon \\ -\sin\upsilon\cos\upsilon+\cos\upsilon\sin\upsilon & 0 & -\cos^2\upsilon-\sin^2\upsilon & 0 \\ 0 & -\sin\upsilon\cos\upsilon+\cos\upsilon\sin\upsilon & 0 & -\cos^2\upsilon-\sin^2\upsilon \end{pmatrix}$$

Thus, the final overall transfer matrix maps the incoming normal modes directly to those of the gantry without any cross-coupling and independently of the rotation angle.

Rotator

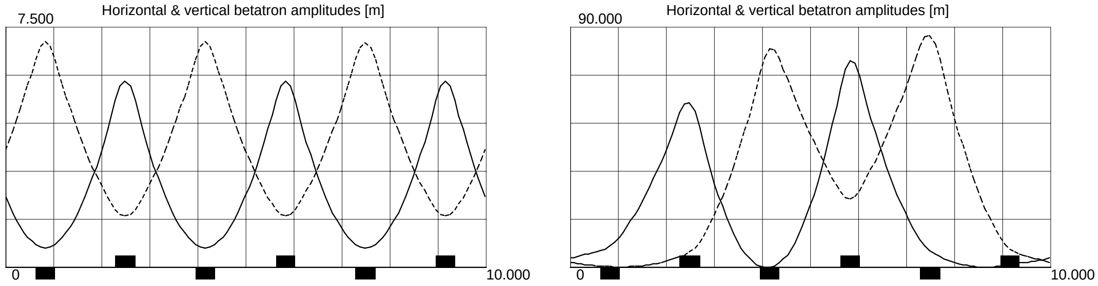
$$\begin{pmatrix} u \\ u' \\ v \\ v' \end{pmatrix}_{\text{Gantry}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ z \\ z' \end{pmatrix}_{\text{Input}} \qquad (8.26)$$

Since a dispersion vector in a bending-free region behaves as a betatron oscillation to first order, the rotator will automatically match the dispersion vector (*D,D'*) into the gantry at the same time as it matches the normal modes. Thus there are no fundamental limitations on the beam symmetry, or on the dispersion functions or any changes to the optics inside the gantry with rotation angle. The fact that the rotator can match a dispersion vector into the gantry opens the possibility of a simplified gantry design with fewer quadrupole magnets and with the dispersion bump closed in the fixed part of the beam line. Note, however, that the beam is coupled inside the rotator.

The beam distribution is also of less consequence with a rotator. For example, although two truncated, uncorrelated gaussians would form a square boundary to the beam spot, the spot would always have the same orientation with respect to the gantry system. This point is relevant to the present study, since the slow-extracted beam will have sharp parallel edges in one plane and gaussian edges in the other. The spot will therefore not be round and the invariance of the orientation inside the gantry system will be of paramount importance for the scanning system.

8.6 ROTATOR DESIGN EXAMPLES

So far, the rotator has been represented by its 4x4 transfer matrix and the only constraint for its design was that the phase advances in the transverse planes must be 2π and π. The overall transfer matrix, however, cannot give any information about beam sizes, chromatic effects etc. inside the structures. In the following Figures 8.5 to 8.7, the lattice functions and structures are shown for three different rotators at the 0° and 90° points.



(a) Not rotated (b) Rotated by 90°

Figure 8.5 Three-cell regular FODO rotator (a) not rotated, (b) rotated by $\pi/2$
[Three cells of phase advance per cell $\mu_p=120^0$, $\mu_q=60^0$, entry values of $\beta_p=2.23\text{m}$, $\beta_q=3.69\text{m}$, $\alpha_p=1.77$, $\alpha_q=-1.85$, two quadrupole families]

A rotation by 90° is equivalent to changing all the focusing lenses of the lattice into defocusing lenses. When viewed in this way, it is not surprising that the regular FODO rotator leads to very large fluctuations of the lattice functions. In the unrotated FODO $\frac{\widehat{\beta}_p}{\breve{\beta}_p}=10$ and $\frac{\widehat{\beta}_q}{\breve{\beta}_q}=4.5$, whereas in the rotated case these ratios increase alarmingly to $\frac{\widehat{\beta}_p}{\breve{\beta}_p}=2000$ and $\frac{\widehat{\beta}_q}{\breve{\beta}_q}=220$. This effect is strong when the entry values for the alpha functions are large and have opposite signs. This gives the first hint that rotators are best made with zero alpha functions at the entry and exit. This is further supported by the triplet rotator of Figure 8.6, in which the ratios of the betatron amplitudes functions at 0° and 90° remain exceedingly well behaved at $\frac{\widehat{\beta}_p}{\breve{\beta}_p}=4$ and $\frac{\widehat{\beta}_q}{\breve{\beta}_q}=4$. The *p*-plane and *q*-planes swap at 90°.

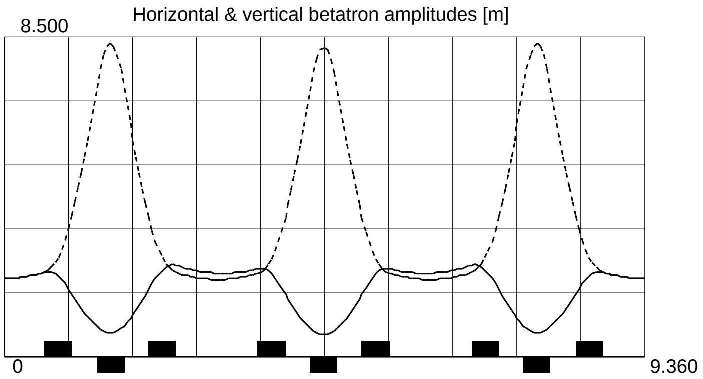


Figure 8.6 Three-cell triplet rotator
[Three cells of phase advance per cell $\mu_p=120^0$, $\mu_q=60^0$, entry values of $\beta_p=\beta_q=4$m, $\alpha_p=\alpha_q=0$, two quadrupole families]

In Figure 8.7 a doublet matches symmetric lattice parameters into a FODO channel that gives the required phase advance. The lattice functions at 0° and 90° are again well behaved with $\frac{\widehat{\beta}_p}{\breve{\beta}_p}=9$ and $\frac{\widehat{\beta}_q}{\breve{\beta}_q}=9.7$. The exotic behaviour of the FODO rotator is thus avoided by entering with equal alpha-values set to zero.

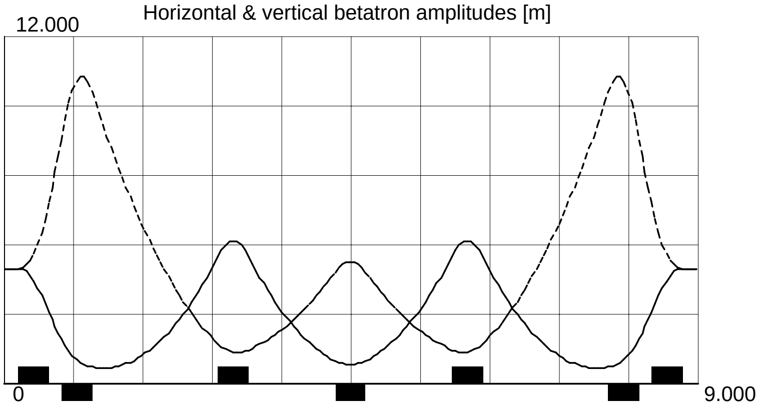


Figure 8.7 Doublet-FODO rotator
[Phase advances of $\mu_p=360^0$, $\mu_q=180^0$, entry values of $\beta_p=\beta_q=4$m, $\alpha_p=\alpha_q=0$, four quadrupole families]

Thus rotators should be designed to have equal alpha functions equal to zero at entry and exit. The complete evaluation of beam sizes at all rotation angles will be dealt with in Section 8.8.

8.7 LENGTH SCALING

So far gantry structures and rotators have been mentioned as examples of lattice modules with phases advances that are an integral number of π. In the next Chapter, similar modules will be proposed for closed-dispersion bends and extension modules. Since it is very often that a lattice has to fit an existing building, it is interesting to know how these structures can be scaled while maintaining their special phase advance properties. Starting from the equation of motion for the horizontal plane:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}s^2} + K(s)x = 0, \tag{8.27}$$

where the general focusing constant, $K(s) = h^2(s) - k(s)$ and $h(s)$ is the inverse of the local radius of curvature and $k(s)$ is the local normalised gradient. The substitution of the solution $x(s) = \sqrt{\beta(s)}\cos\mu(s)$ leads to the well-known equation of the betatron amplitude function,

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\sqrt{\beta} + K(s)\sqrt{\beta} = \beta^{-3/2}. \tag{8.28}$$

This equation determines the β-function for the whole structure and therefore also the phase advance $\mu(s) = \int \beta^{-1}\mathrm{d}s$. To scale a structure, equation (8.28) is rewritten with the inclusion of scaling factors (κ, for the length; λ for the focusing and τ for the β function), so that

$$\frac{\mathrm{d}^2}{\kappa^2\mathrm{d}s^2}\sqrt{\tau\beta} + \lambda K(s)\sqrt{\tau\beta} = (\tau\beta)^{-3/2}.$$

After some re-arrangement, this equation yields the relations between the scaling factors

$$\left(\frac{\tau^2}{\kappa^2}\right)\frac{\mathrm{d}^2}{\mathrm{d}s^2}\sqrt{\beta} + \lambda\tau^2 K(s)\sqrt{\beta} = \beta^{-3/2} \quad \text{so that } \tau = \kappa,\ \lambda = \kappa^{-2}. \tag{8.29}$$

Scaling with constant phase advance

$$s \Rightarrow \kappa s; \quad K \Rightarrow \kappa^{-2}K; \quad \beta \Rightarrow \kappa\beta.$$

From $\alpha = -\frac{1}{2}\frac{\mathrm{d}\beta}{\mathrm{d}s}$, it follows that α remains unscaled and that γ scales inversely with β.

The quadrupole apertures are determined by the beam size. Since the emittance is constant the apertures scale like: $A = \sqrt{E\beta} \Rightarrow \sqrt{\kappa}\sqrt{E\beta}$ the gradients scale like $k \Rightarrow \frac{1}{\kappa^2}k$ .and, finally, it follows for the pole tip field: $B_{\text{pole}} \Rightarrow \kappa^{-3/2}B_{\text{pole}}$.

8.8 BEAM SIZES IN ROTATORS

8.8.1 With zero dispersion

The σ-matrix formalism can be used to study the beam sizes inside the rotator. It is assumed that the initial beam coming from the accelerator is uncoupled. The σ-matrix then comes directly from (8.7),

$$\sigma = \begin{pmatrix} \langle x^2 \rangle & \langle -xx' \rangle & 0 & 0 \\ \langle -x'x \rangle & \langle x'^2 \rangle & 0 & 0 \\ 0 & 0 & \langle z^2 \rangle & \langle -zz' \rangle \\ 0 & 0 & \langle -z'z \rangle & \langle z'^2 \rangle \end{pmatrix} \tag{8.7}$$

The transformation of the σ-matrix through the lattice was given in (8.4) as $\sigma_2 = \boldsymbol{M}\sigma_1\boldsymbol{M}^{\mathrm{T}}$, where ***M*** can be written as,

$$\boldsymbol{M} = \begin{pmatrix} C_{\mathrm{x}} & S_{\mathrm{x}} & 0 & 0 \\ C_{\mathrm{x}}' & S_{\mathrm{x}}' & 0 & 0 \\ 0 & 0 & C_{\mathrm{z}} & \mathrm{S}_{\mathrm{z}} \\ 0 & 0 & C_{\mathrm{z}}' & \mathrm{S}_{\mathrm{z}}' \end{pmatrix} \tag{8.30}$$

where *C* and *S* are sometimes known as the principal trajectories. The exact forms are given below and can be found by comparison with the well-known general transfer matrix given in (8.18). The subscripts 1 and 2 denote the initial (rotator entrance) and final (a point downstream in the rotator) values respectively, Δμ is the phase advance through the lattice and all four equations exist in both planes with the appropriate lattice parameters.

$$C = \sqrt{\frac{\beta_2}{\beta_1}}\left(\cos\Delta\mu + \alpha_1 \sin\Delta\mu\right) \tag{8.31}$$

$$C' = -(\beta_1\beta_2)^{-1/2}\left[(1+\alpha_1\alpha_2)\sin\Delta\mu + (\alpha_2 - \alpha_1)\cos\Delta\mu\right] \tag{8.32}$$

$$S = (\beta_1\beta_2)^{1/2}\sin\Delta\mu \tag{8.33}$$

$$S' = \sqrt{\frac{\beta_1}{\beta_2}}\left(\cos\Delta\mu - \alpha_2 \sin\Delta\mu\right). \tag{8.34}$$

Let the initial σ-matrix in the normal transfer line before the rotator be uncoupled. The effect of the rotation can be expressed as,

$$\sigma_2 = \boldsymbol{R}\sigma_1\boldsymbol{R}^{\mathrm{T}}.$$

The matrix σ_2 can now be transferred along the rotator using the transfer matrix ***M***:

$$\sigma_3 = \boldsymbol{M}\underbrace{\boldsymbol{R}\sigma_1\boldsymbol{R}^{\mathrm{T}}}_{\sigma_2}\boldsymbol{M}^{\mathrm{T}}$$

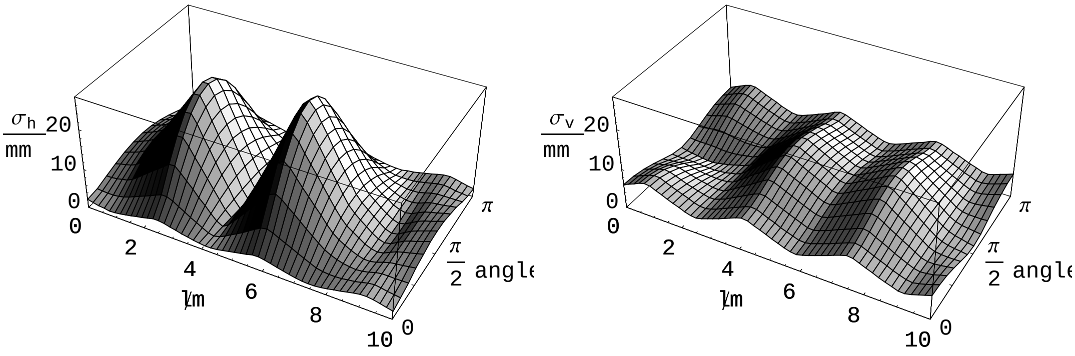
where ***M*** is evaluated using the principal trajectories of (8.30).

The full expansion of the matrix σ_3 can be avoided if only the beam size is wanted. The $\sigma_{1,1}$ and $\sigma_{3,3}$ terms give the rms beam sizes $<p^2>$ and $<q^2>$ in the co-ordinate system of the rotator, but expressed in terms of the original beam parameters. The two terms in the final σ-matrix are given below.

$$\begin{aligned}\sigma_{1,1} &= \left\langle p^2 \right\rangle \\ &= \frac{1}{\pi}\left\{ C_\mathrm{p}{}^2\left(c^2 E_\mathrm{x}\beta_\mathrm{x} + s^2 E_\mathrm{z}\beta_\mathrm{z}\right) + S_\mathrm{p}{}^2\left(c^2 E_\mathrm{x}\gamma_\mathrm{x} + s^2 E_\mathrm{z}\gamma_\mathrm{z}\right) - 2C_\mathrm{p}S_\mathrm{p}\left(c^2 E_\mathrm{x}\alpha_\mathrm{x} + s^2 E_\mathrm{z}\alpha_\mathrm{z}\right)\right\}\end{aligned} \tag{8.35}$$

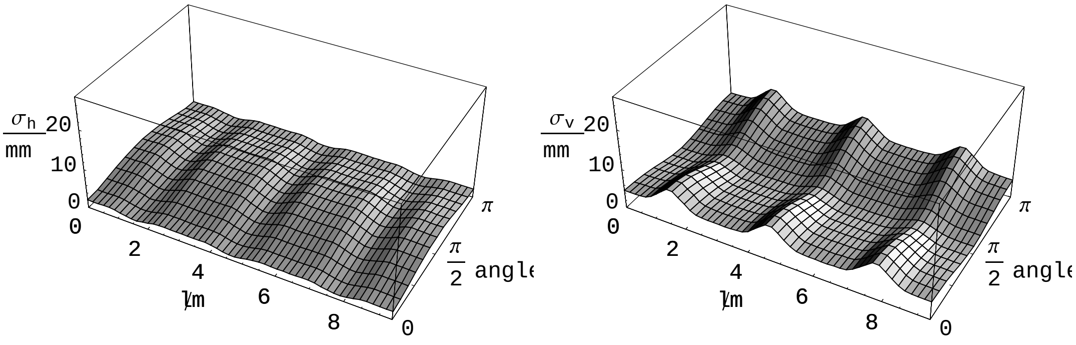
$$\begin{aligned}\sigma_{3,3} &= \left\langle q^2 \right\rangle \\ &= \frac{1}{\pi}\left\{ C_\mathrm{q}{}^2\left(s^2 E_\mathrm{x}\beta_\mathrm{x} + c^2 E_\mathrm{z}\beta_\mathrm{z}\right) + S_\mathrm{q}{}^2\left(s^2 E_\mathrm{x}\gamma_\mathrm{x} + c^2 E_\mathrm{z}\gamma_\mathrm{z}\right) - 2C_\mathrm{q}S_\mathrm{q}\left(s^2 E_\mathrm{x}\gamma_\mathrm{x} + c^2 E_\mathrm{z}\gamma_\mathrm{z}\right)\right\}\end{aligned} \tag{8.36}$$

where, s and c are sine and cosine of the rotation angle and $C_{\mathrm{p,q}}$ and $S_{\mathrm{p,q}}$ are the principal trajectories of the rotator. The following Figures 8.8-8.10 show the beam sizes inside the three rotators as seen in the co-ordinate system of the rotator as a function of the rotation angle. Beam sizes in these examples are calculated for $E_\mathrm{x} = 2$ [π mm mrad] and $E_\mathrm{z} = 10$ [π mm mrad].



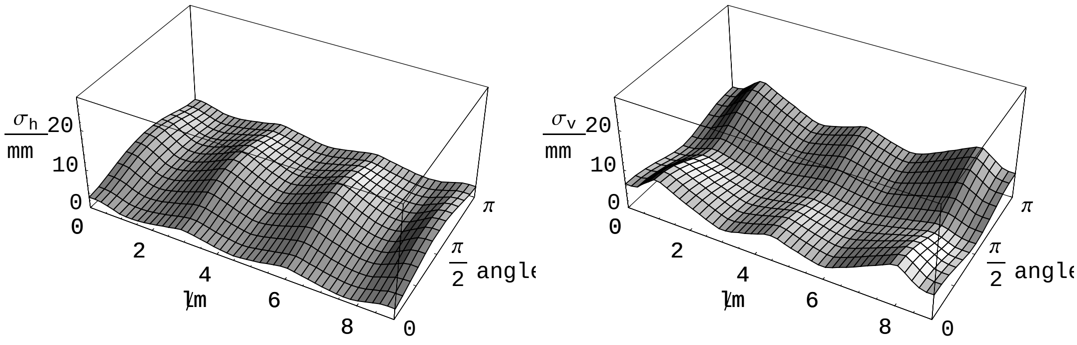
(a) (b)

Figure 8.8 Horizontal (a) and vertical (b) beam sizes inside a FODO rotator when rotating from 0 to π (the form repeats for π to 2π)



(a) (b)

Figure 8.9 Horizontal (a) and vertical (b) beam sizes inside a triplet rotator when rotating from 0 to π (the form repeats for π to 2π)



(a) (b)

Figure 8.10 Horizontal (a) and vertical (b) beam sizes inside a Doublet/FODO rotator when rotating from 0 to π (the form repeats for π to 2π)

8.8.2 With finite dispersion

When calculating the beam size from the σ-matrix using equations (8.35) and (8.36), the average position of the beam is ignored. Fortunately, this omission is easily corrected by tracking the single particle that is exactly on the central or closed orbit of the beam. For simple orbit distortion, the normal 2×2 matrices with the appropriate rotation matrices can be used. For off-momentum beams, the normal 3×3 transfer matrices must be used that include the momentum effects.

8.9 SPREADING AND SCANNING SYSTEMS

The last action of the optical system is to ‘spread’ or ‘scan’ the beam over the tumour. Beam spreading is the more conventional approach, in which the tumour is treated either as a volume or in thick slices. Beam scanning divides the tumour into thin slices with many pixels and adapts the beam to treat each pixel separately. These two main approaches are summarised in Figure 8.11. Note that Figure 8.11 is far from being comprehensive and that many variants and combinations exist. In particular, no mention is made of systems that move the patient as part of the scanning procedure or displace steering magnets to economise on magnet apertures.

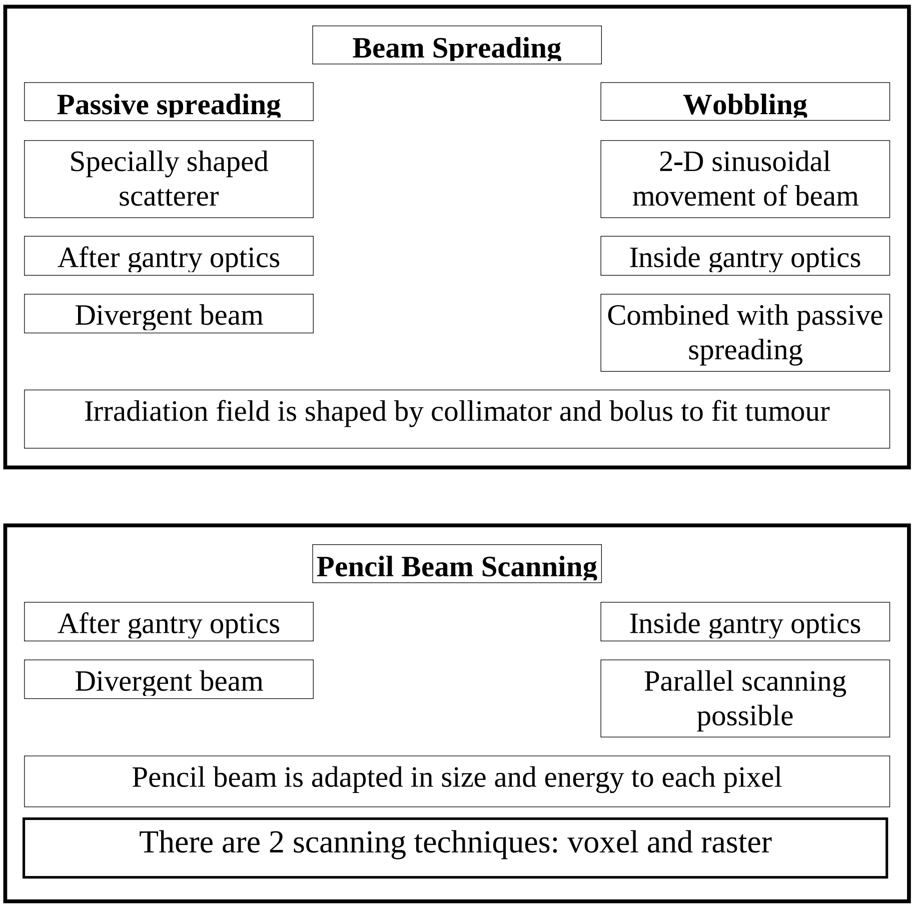

Beam Spreading
Passive spreading
Wobbling
Specially shaped scatterer
2-D sinusoidal movement of beam
After gantry optics
Inside gantry optics
Divergent beam
Combined with passive spreading
Irradiation field is shaped by collimator and bolus to fit tumour
Pencil Beam Scanning
After gantry optics
Inside gantry optics
Divergent beam
Parallel scanning possible
Pencil beam is adapted in size and energy to each pixel
There are 2 scanning techniques: voxel and raster

Figure 8.11 Beam spreading and scanning

Source to surface distance (SSD)

Passive spreading and divergent beam scanning both suffer from a finite source-to-surface distance (SSD) that leads to an enhancement of the surface dose (see Figure 8.12). The maximum angle of divergence is a matter of discussion. In reference 4, the maximum divergence is quoted as ±1° and this criterion is used to relax on the strict requirement of parallel scanning. However, for passively spread beams in Ref. 5 a minimum SSD of 2m is quoted. If the field of irradiation is ±10 cm then a SSD of 2 m corresponds to ±2.9°.

Positioning of the scanning magnets

After the last gantry magnet. In this case, the scanning is forcibly divergent, but this layout has the great advantage that the gantry magnets do not need enlarged apertures. Thus, the weight and cost of the gantry are reduced.

Incorporated in the gantry optics. Once the scanning magnets are moved further upstream into the gantry the focusing of the last dipole and possibly some quadrupoles can be used to create a parallel scanning system. If spaces are available at the correct phase advances, then parallel scanning is possible with a single magnet in each plane, otherwise two magnets are needed per plane. However, a usual compromise is one magnet per plane with approximately the correct phase advance. The loss of parallel scanning is only partial and is offset by the advantage in size and cost gained by reducing the gap in the last dipole magnet of the gantry.

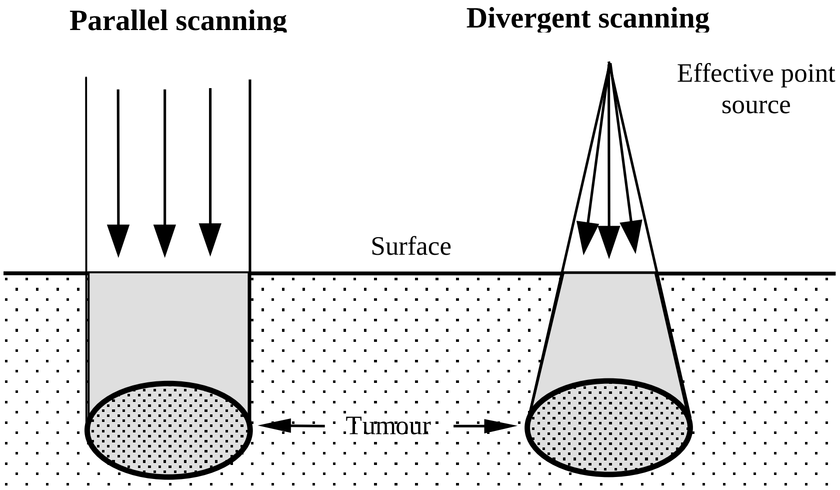


Figure 8.12 Enhancement of surface dose with divergent beams

References

[1] M. Benedikt, C. Carli, *Matching to gantries for medical synchrotrons*, Part. Accel. Conf. 1997, Vancouver, (1997) and M. Benedikt, C. Carli, *Optical design of a beam delivery system using a rotator*, CERN/PS 96-041 (1996).

[2] L.C. Teng, Private communication, Laboratory notebook (Jan. 1970) and Int. Rep. LL-134 (Oct. 1986).

[3] J. Buon, *A statistical description of a particle beam subjected to a linear coupled betatron motion*, LAL/RT 96-03, (April 1996).

[4] M. Pavlovic, *Beam-optics study of the gantry beam delivery system for light-ion cancer therapy*, GSI-Preprint-95-83, (Dec.1995).

[5] U. Amaldi, M. Silari (editors), *The TERA project and the centre for oncological hadrontherapy*, The TERA Foundation, (INFN, Frascati, 1995, II Ed.), Vol II, p276.

* * *

I-9 EXTRACTION LINES

The particular ‘footprint’ of the slow-extracted beam segment in phase space and the need to control the beam sizes in a lattice with rotating optical elements create a rather special problem for the extraction transfer line. The design concept presented here regards the whole line from the electrostatic septum inside the ring to the patient as an integrated system built mainly from *‘telescope’* modules with integer π phase advances. The beam size in the plane of the extraction is controlled at the patient by altering the phase advance in the line in order to rotate the extracted beam segment in phase space at the patient. The vertical beam size is controlled by stepping the vertical betatron amplitude function over a range of values and passing the changed beam size from ‘hand-to-hand’ through the ‘telescope’ modules to the patient. The matching to the gantry is assured by a module called a *rotator*. An example of this optics is given with a preliminary design for the optics of the so-called *Riesenrad gantry*.

9.1 DESIGN CONCEPT

The principal ingredients of the design of the transfer line and delivery system can be summarised as:

- Matching the unequal emittances and non-zero dispersion functions to a rotating gantry by the use of a rotator.
- Using ‘telescope’ modules that have *‘one-to-one’*, *‘one-to-minus one’*, or a fixed magnification with integer π phase advances. The rotator already fits this category. A module from the same family with a 2π phase advance is ideal for embedding a bend with a closed dispersion bump and this type of structure can also be adapted to the Riesenrad gantry [1].
- Exploiting the ‘bar’ of charge to create an independent control of the horizontal beam size by rotating the bar in an unfilled phase-space ellipse using a phase shifter at the entry to the line.
- Controlling the vertical beam size using a *‘stepper’* that steps the betatron amplitude function over a range of values, whilst keeping all the parameters in the horizontal plane constant, and then hands the chosen beam size through the telescope modules all the way to the patient.
- Placing the phase shifter and stepper at the exit to the accelerator so that they can act for all gantries in the complex.
- Avoiding unnecessarily large beam sizes in the rotator, for example, by using the telescope modules to de-magnify and then re-magnify the betatron amplitude functions in the vertical plane.

This general strategy has been adapted to an example lattice for a cancer therapy facility. The layout also includes such practical features as a beam ‘chopper’, space for diagnostics in a dispersion-free region, a long drift space for a vehicle track around the ring and a modular layout of the treatment rooms, see Figure 9.1. The medical specifications (see Table 1-1) require spot sizes 4-10 mm full width at half height with either protons between 60 and 250 MeV, or carbon ions between 120 and 400 MeV/u. The beam at the patient should also be achromatic.

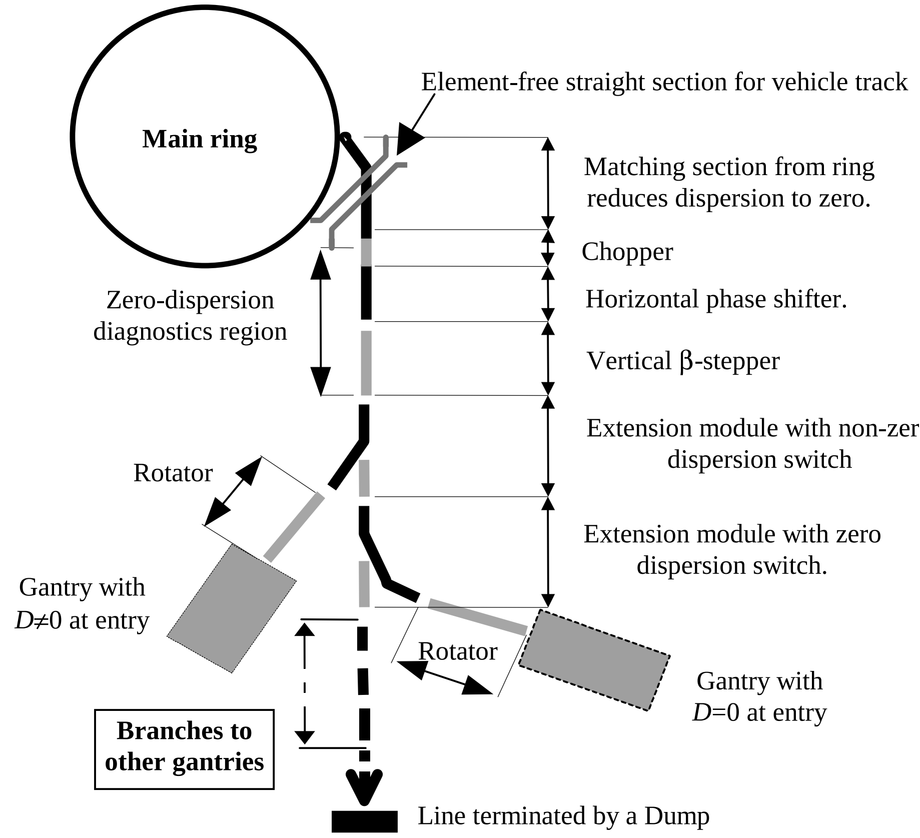

Main ring
Element-free straight section for vehicle track
Matching section from ring reduces dispersion to zero.
Chopper
Horizontal phase shifter.
Vertical β-stepper
Zero-dispersion diagnostics region
Extension module with non-zero dispersion switch
Rotator
Extension module with zero dispersion switch.
Gantry with D≠0 at entry
Rotator
Gantry with D=0 at entry
Branches to other gantries
Line terminated by a Dump

Figure 9.1 Schematic view of an example extraction layout

9.2 ‘TELESCOPES’

Before describing the individual modules, it is useful to review the properties of ‘telescopes’ with integer π phase advances. The general transfer matrix from (8.18) can be written as,

$$\begin{pmatrix} C & S \\ C' & S' \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{\beta}{\beta_0}}\left(\cos(\Delta\mu)+\alpha_0\sin(\Delta\mu)\right) & \sqrt{\beta\beta_0}\sin(\Delta\mu) \\ -\frac{1}{\sqrt{\beta\beta_0}}\left[(\alpha-\alpha_0)\cos(\Delta\mu)+(1+\alpha\alpha_0)\sin(\Delta\mu)\right] & \sqrt{\frac{\beta_0}{\beta}}\left(\cos(\Delta\mu)-\alpha\sin(\Delta\mu)\right) \end{pmatrix} \quad (9.1)$$

where *C* and *S* are known as the principal trajectories and the other symbols have their usual meaning. The family of telescope modules of interest is characterised by having:

Definition of a ‘telescope’ module (I): $S = 0$ and $C' = 0$. (9.2)

From inspection of (9.1) and since β can never be zero, Δμ must be *n*π to satisfy $S = 0$ and α must equal α_0 to satisfy $C' = 0$. Note that both conditions are independent of the initial values of α_0 and β_0. Thus in the Courant and Snyder formalism, Definition of a ,'telescope', module (II) $\Delta\mu = n\pi$ and $\alpha_0 = \alpha$. (9.3)

The lattice functions can be transmitted through a structure by a standard matrix expression,

$$\begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} C^2 & -2CS & S^2 \\ CC' & CS'+SC' & -SS' \\ C'^2 & -2C'S' & S'^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \alpha_0 \\ \gamma_0 \end{pmatrix}. \tag{9.4}$$

This quickly simplifies if $S = 0$ (9.2),

$$\begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} C^2 & 0 & 0 \\ CC' & CS' & 0 \\ C'^2 & -2C'S' & S'^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \alpha_0 \\ \gamma_0 \end{pmatrix}. \tag{9.5}$$

When the lattice is matched for $C' = 0$ (9.2) and $\alpha = \alpha_0$ (9.3), CS' will be unity and (9.5) simplifies further to:

$$\begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} C^2 & 0 & 0 \\ 0 & CS'=1 & 0 \\ 0 & 0 & S'^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \alpha_0 \\ \gamma_0 \end{pmatrix}. \tag{9.6}$$

Thus for any incoming set of lattice parameters, it follows that

'Telescope' magnification: $\frac{\beta}{\beta_0} = C^2$ and $\alpha = \alpha_0$. (9.7)

9.3 EXAMPLE OPTICS WITH A RIESENRAD GANTRY

Figure 9.2 shows the ring and extraction line of the example optics. The line is modularised with very specific functions for each module. The base condition for handing the beam over from one module to the next is $\beta_x = \beta_z = 3$ m, $D_x = D_z = 0$ and $\alpha_x = \alpha_z = 0$. However, the 'telescope' modules will locally modify the betatron amplitude functions by some magnification factor. As shown in Figure 9.1 the idea is to match out of the accelerator into a long straight line from which the gantry lines are derived. The gantry lines can turn to the left or the right and extra units can be inserted at will. It is stressed that the layout in Figure 9.2 is an example and that in principle there is considerable flexibility within the limits of the modular design.

9.3.1 Initial conditions

To evaluate the beam characteristics at the entrance to the line, simulations of the extraction process have been made for the PIMMS synchrotron. By recording the position of the particles at the entrance to the electrostatic septum, the expected beam distribution can be studied and used to define the initial conditions at the entry to the extraction line. Figure 9.3, is the result of such a simulation in phase space of the plane of extraction. The stable triangle can be seen situated at the vacuum chamber with one separatrix reaching out towards the electrostatic septum. The segment of separatrix that is cut off by the septum is indicated. The procedure followed here to situate the segment in an unfilled ellipse rather than fit an ellipse directly to the segment is treated in the Section 9.3.2.

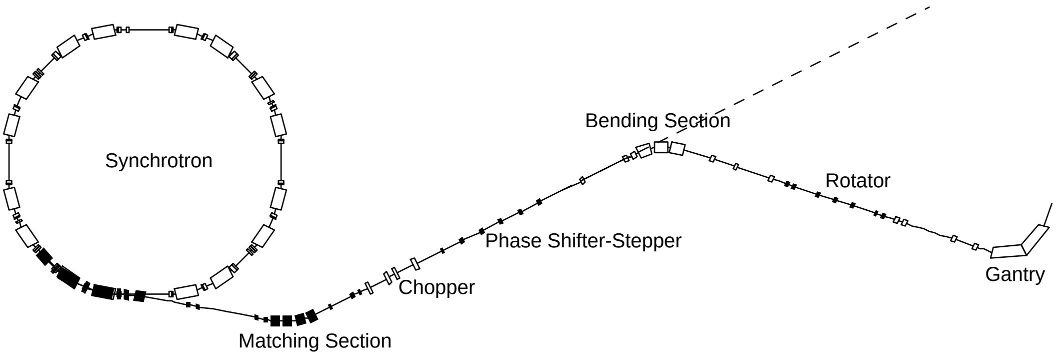


Figure 9.2 Layout of example optics for ring and extraction line
[Horizontal plane drawn on a 9 m square grid]

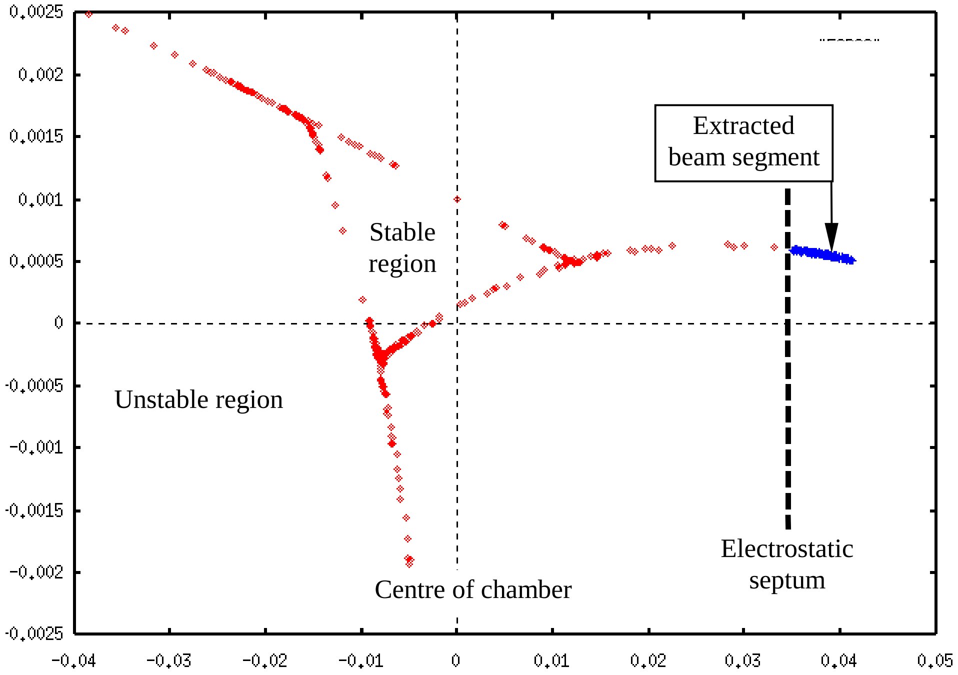


Figure 9.3 Simulation of the extraction phase space at the electrostatic septum

The extracted segment in Figure 9.3 appears as a simple line, but in practice this line is composed of many lines coming from the stable triangles of different momenta (emittances). The thickness of these lines depends upon the emittance in the orthogonal plane and the coupling from the sextupoles (mainly the resonance sextupole). Figure 9.4 shows in more detail the extracted segments for maximum amplitude (emittance) in *x* (short lines) and zero amplitude in *x* (long lines) and for maximum amplitude in *z* (thick lines) and zero amplitude in *z* (narrow lines). The ellipses shown in the same figure are chosen to contain the segments and to have reasonable lattice functions. The segments can then be considered as rotating in the unfilled ellipses as the beam goes along the extraction line.

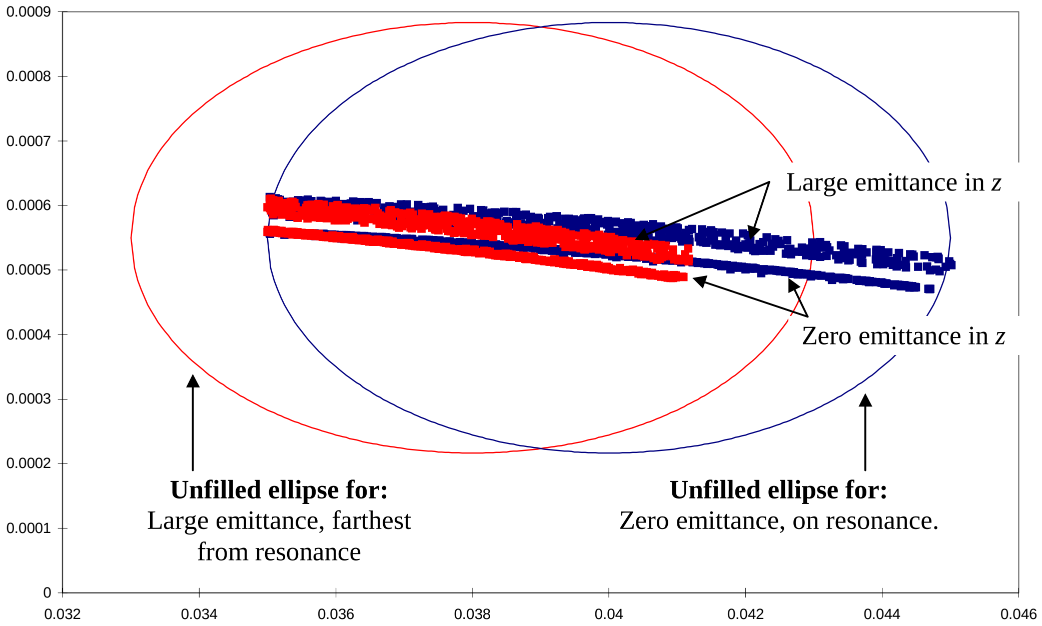


Figure 9.4 Simulation of the extracted beam and ‘unfilled ellipse’ at the electrostatic septum

The difference in the spiral step for particles with different momentum deviation (different emittance) causes a shift in the centre of the segment for a different momenta. This can be taken into account with an initial dispersion and derivative of the dispersion defined by the distance between the centres divided by the momentum difference. The unfilled ellipses are chosen such that the dimension perpendicular to the segment is large, so that the width of the segment is unimportant.

Finally the set of initial conditions chosen for the PIMMS extraction line in this simulation is the following:

$\boldsymbol{\beta_x = 15}$ **m**	$\boldsymbol{\beta_z = 6.77}$ **m**
$\boldsymbol{\alpha_x = 0}$	$\boldsymbol{\alpha_z = -0.0814}$
$\boldsymbol{D_x = 2.09}$ **m**	$\boldsymbol{D_z = 0}$ **m**
$\boldsymbol{D'_x = -0.0174}$	$\boldsymbol{D'_z = 0}$
$\boldsymbol{\varepsilon_{x,total} = 1.65 \times 10^{-6} \pi}$ **m rad**	$\boldsymbol{\varepsilon_z}$ **= varies with energy**

Note that in the horizontal plane the segments, and therefore also the unfilled ellipses, are determined geometrically by the resonance and are invariant with beam energy. In the vertical plane, the emittance comes directly from the circulating beam (with a small perturbation from coupling) and is beam energy dependent. This introduces the rather strange situation where the extracted segment is reasonably normal in the vertical plane, but in the horizontal plane, it is invariant with beam energy in one dimension and partially adiabatically damped via coupling from the vertical plane in the other dimension.

9.3.2 Direct and indirect fitting of an ellipse

The simulations presented in the previous section show that the phase-space 'footprint' of the slow-extracted beam from a synchrotron differs markedly from the fast-extracted beam, or a beam from a cyclotron:

- In the plane of extraction, the phase-space shape is close to being a narrow rectangle (known as the 'bar' of charge) that corresponds to the segment of the outward-spiralling separatrix that is deflected out of the machine by the electrostatic septum.

- In the orthogonal plane, the phase-space shape is the same as that of the circulating beam to first-order.

Figure 9.5 shows this situation schematically for a mono-energetic beam at the entrance to the extraction channel.

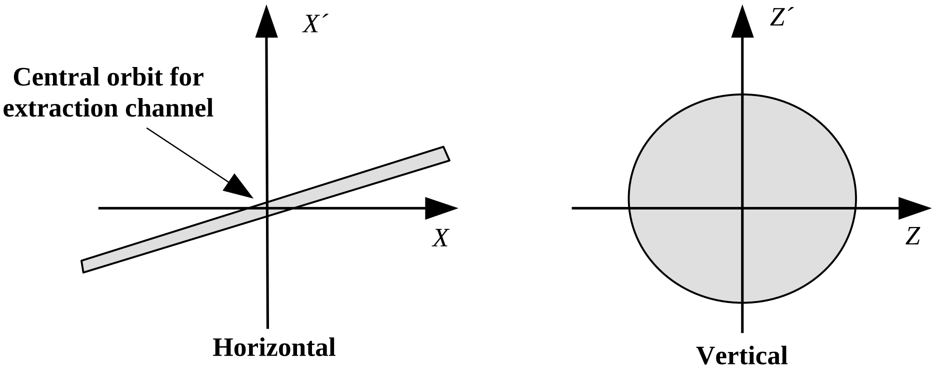


Figure 9.5 Schematic view of the phase-space 'footprint' of the beam in the extraction channel

If the severed segment of the separatrix were viewed at the entry to the electrostatic septum and if it were rotated until it was horizontal, it would be:

- 5-10 mm long depending on the momentum.
- ~0.00005 rad wide.
- The particle density for a given momentum would be quasi-constant over the 'rectangle'.

Courant and Snyder representation by 'direct fitting'

The conventional approach to representing beams is to find a closely fitting ellipse. In the case of a rectangle, the fitted ellipse would be somewhere in the range between the inscribed ellipse and the circumscribed ellipse. This is shown schematically in Figure 9.6 and the emittances and fitted lattice functions are summarised in Table 9.1.

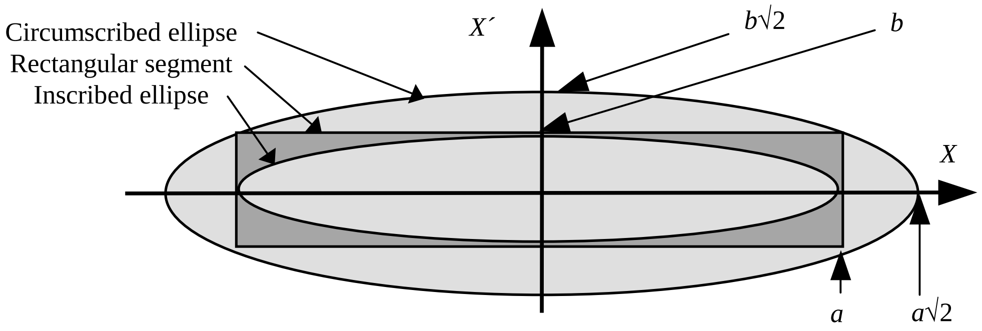


Figure 9.6 Direct fitting ellipse
[‘Beam-rectangle’ (full width 10 mm, full divergence 0.05 mrad) = 0.16 π mm mrad]

Direct fitting of inscribed and circumscribed ellipses

Rectangular beam segment

Full width 10 mm, full divergence 0.05 mrad E_x = 0.16 π mm mrad]

	Circumscribed ellipse,	**Inscribed ellipse,**
Lowest extraction energy	E_x = ~0.25 [π mm mrad] β_x = ~200 [m]	E_x = ~0.125 [π mm mrad] β_x = ~200 [m]
Highest extraction energy	E_x = ~0.13 [π mm mrad] β_x = ~400 [m]	E_x = ~0.063 [π mm mrad] β_x = ~400 [m]

Table 9.1 Direct fitting of inscribed and circumscribed ellipse

The range in emittance between the inscribed and circumscribed ellipses is already a factor of two which indicates that the fitting is not a natural action in this case. Furthermore, adiabatic damping in the vertical plane reduces the coupling to the horizontal plane (effect is proportional to the vertical emittance and the resonance sextupole strength) as the energy increases. Thus, at the top extraction energy, the horizontal emittance is reduced to approximately 50% of its value at the lowest extraction energy, which is reflected in a parallel change in the betatron amplitude function from 200 m to 400 m (see Table 9.1). This is an extremely strong energy dependence to take into account in the optics.

The length of the extracted beam segment is well known, but the divergence is more difficult to calculate and very difficult to measure. Firstly, it is very small and, secondly, it is very sensitive to vertical closed-orbit distortions at the resonance sextupole. A distortion of 4 mm can increase the beam divergence by a factor 4, which in turn reduces the betatron amplitude of the fitted ellipse by a factor 4. Thus, the fitted ellipse is strongly dependent on machine conditions such as the closed orbit.

The central orbit of the extraction channel is defined as the central momentum of the extracted beam. In Table 9.2, the results of a first-draft-lattice design show the approximate situation at the lowest extraction energy, where the spiral step variation with momentum alters the length of the extracted segment. The fitted ellipse must follow these changes and this results in an extremely strong chromatic effect in the betatron amplitude function.

Variation of the extracted beam segment with momentum deviation

	Full length [mm]	**Full divergence [mrad]**	**β_x [m]**
$\Delta p/p$ = -0.0006	6	0.05	120
$\Delta p/p$ = 0.0	8	0.05	160
$\Delta p/p$ = 0.0006	10	0.05	200

Table 9.2 Variation of the extracted beam segment with momentum deviation

Thus, the chromaticity in the betatron amplitude function is,

$$\Delta\beta/(\Delta p/p) = 6.7 \times 10^4 \tag{9.8}$$

The direct approach of closely fitting an ellipse to the extracted beam therefore leads to a number of undesirable effects:

- A β_x that changes with energy (factor 2), because the horizontal size of the beam is constant while the beam divergence undergoes a quasi-adiabatic damping.
- The absolute value of β_x is not well known. In practice, the coupling from the vertical plane is very sensitive to vertical closed-orbit distortions at the resonance sextupole and a distortion of 4 mm can give a factor of four increase in the emittance and a factor of 4 reduction in the horizontal betatron function.
- The variation in spiral step with momentum deviation causes a strong chromatic effect in β_x.

Courant and Snyder representation by 'indirect fitting'

The extracted beam can be treated as part of an ‘unfilled’ ellipse in the horizontal phase space. Since the transfer line is short, the motion of the beam will be coherent and its phase-space shape will be preserved. The orientation in the chosen ellipse and hence the beam width will be given by the phase advance down the line (see Figure 9.7).

The ‘unfilled’ ellipse is chosen from the infinite family of ellipses that have the ‘bar of charge' as a diameter. The choice is best made so that,

- The divergence of the ellipse is far greater than that of the beam. The match between the ‘bar’ of charge and the diameter of the ellipse is then insensitive to the changes in the 'bar’s' divergence. The larger the divergence of the ellipse the narrower the bar can be made when turned upright in the ellipse.
- The β_x-value of the ‘unfilled’ ellipse should be adjusted so that maximum beam sizes in the extraction line at the gantry exit are convenient values.

The different momenta in the beam can be transported in similar ellipses, but of different emittances to reflect the different lengths of the ‘bar’. In this scheme, it is essential to be able to control the horizontal phase advance in order to know and to be able to adjust the horizontal beam size at the patient.

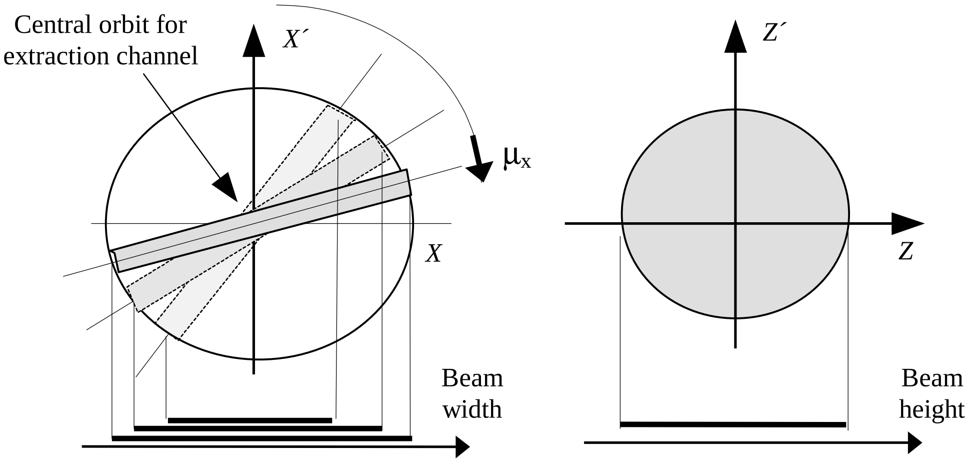


Figure 9.7 Rotation of the 'bar of charge' with phase advance

9.3.3 Matching section

The first part of the extraction line, from the electrostatic septum to the magnetic septa, is inside the main ring, but off-axis. This is known as a ‘distorted’ orbit and has been calculated and introduced as a special section of lattice in which the quadrupoles, for example, are re-expressed as combined function magnets with edge angles calculated from the particle trajectory. After passing the second magnetic septum, there is a doublet and a long drift space before a final doublet and the bending magnets needed to close the dispersion bump from the ring. The long drift has two functions:

- The small β_x provides the phase advance to bring the dispersion function to zero.
- The vacuum pipe can be removed to allow the passage of a vehicle round the ring.

After matching the dispersion to zero, three quadupoles are inserted to match the lattice functions to the input of the next section while leaving a long drift at the end to provide space for the chopper. When the dispersion is zero, the different bars of charge corresponding to different momenta are overlapped at the centre. The geometry and the optical functions in this section of lattice are shown in Figures 9.8 and 9.9.

9.3.4 ‘Chopper’

The ‘chopper’ (see Figure 9.10) is a system comprising four dipoles powered in series that allows the switching of the beam on and off. When the dipoles are not powered, the beam is absorbed in a dump. On the contrary, when the dipoles are powered, the beam is translated parallel to itself by the first two dipoles and is brought back to the axis by the second two. This circumvents the dump and allows transmission of the beam. As the dipoles are powered in series the beam stays on axis downstream of the chopper at all times during the flat top and the rise and fall. Thus the beam position at the patient can be guaranteed as soon as transmission starts and is independent of fluctuations and the ‘rounding’ in and out of the power converter on the ramp. This device can be used for scheduled beam interruptions as well as emergency interruptions.

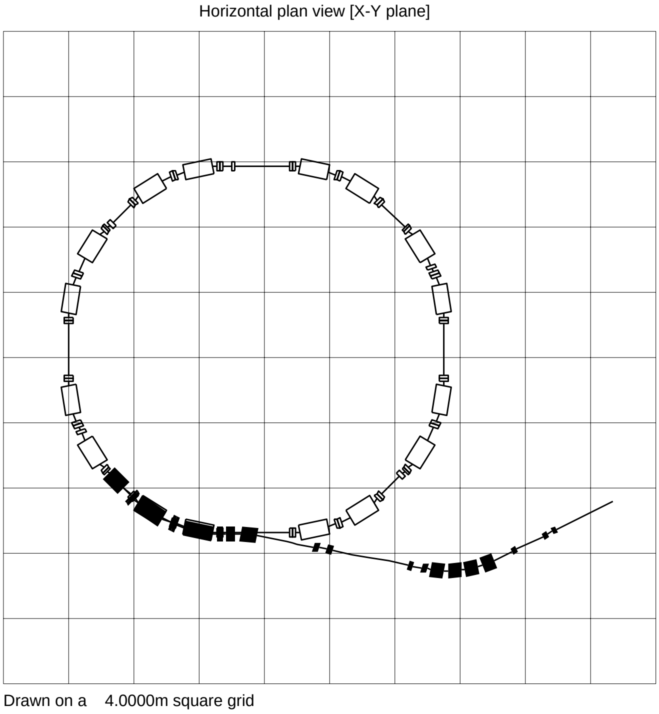

Horizontal plan view [X-Y plane]
Drawn on a 4.0000m square grid

Figure 9.8 Plan view of the matching section and the ring

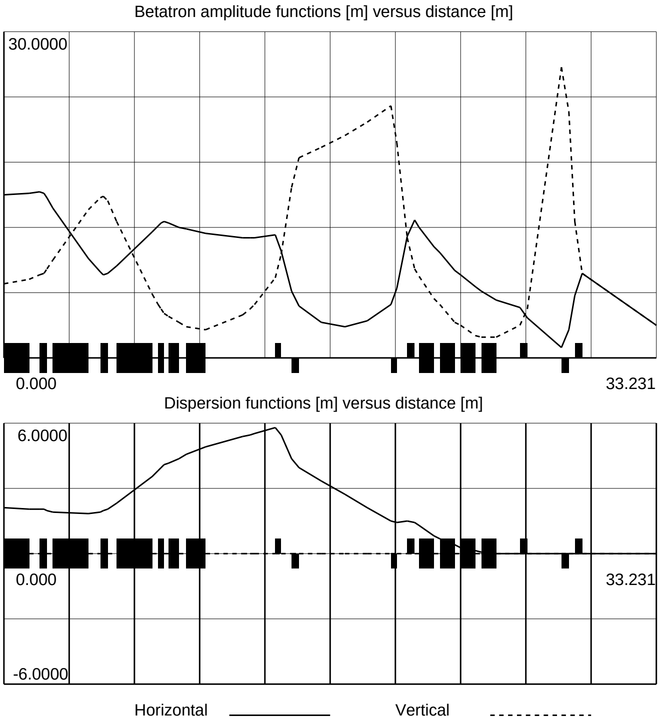


Figure 9.9 Optical functions in the matching section

The chopper is not a module in the sense of the optics modules, such as the rotator. It requires a few metres of free space with a small betatron amplitude function at the position of the dump block in the direction of the deflection. In the present scheme, it is placed between the matching section from the ring and the phase shifter. Figure 9.11 shows the phase space distributions in the two planes at that position.

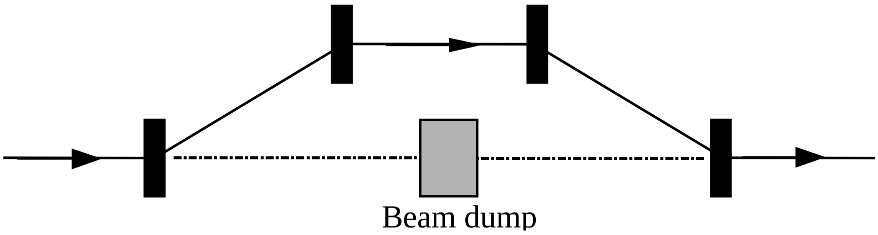


Figure 9.10 Schematic drawing of the chopper

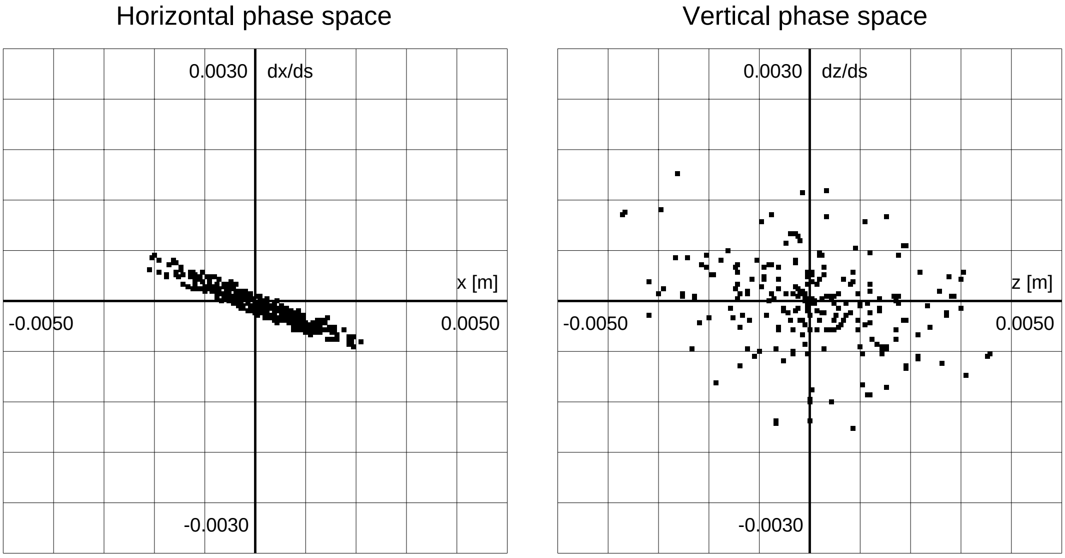


Figure 9.11 Phase space plot of the beam at the centre of the chopper

9.3.5 Phase Shifter

As mentioned above and discussed in Section 9.3.3, the horizontal beam envelope is not adequately described by the betatron amplitude function. This implies that the standard way of changing the beam size through focusing is not suitable in this case. Fixing the values of β_x, α_x, and ε_x fixes the unfilled ellipse in which the ‘bar’ of charge is contained, but the ‘bar’s’ orientation depends on the phase advance. This situation can be used to provide an independent method for changing the beam size at the patient by varying the phase advance while keeping the lattice functions β_x and α_x constant in the line. A dedicated insertion that changes the phase advance without modifying the downstream lattice parameters has therefore been designed and will be referred to as the ‘phase shifter’. To fully exploit the variation in beam size, at least 90 degrees of variation in the phase advance is needed (if the footprint of the beam in the horizontal phase space is parallel to one of the axes at the minimum or maximum phase advance). Since the ellipse in the vertical phase space is filled, the vertical phase advance is unimportant. In Figure 9.12, the optical functions inside the phase shifter are shown for the phase advance $\Delta\mu_x$ between 2.2 and 3.9 radians (97.5°), which allows the choice of any orientation of the bar of charge at the end of the line. The beta functions outside the module do not change, whatever phase advance is selected. In order to obtain a smooth change in the lattice functions, the vertical phase advance has been kept constant in Figure 9.12, but this parameter could be varied to help the matching if needed.

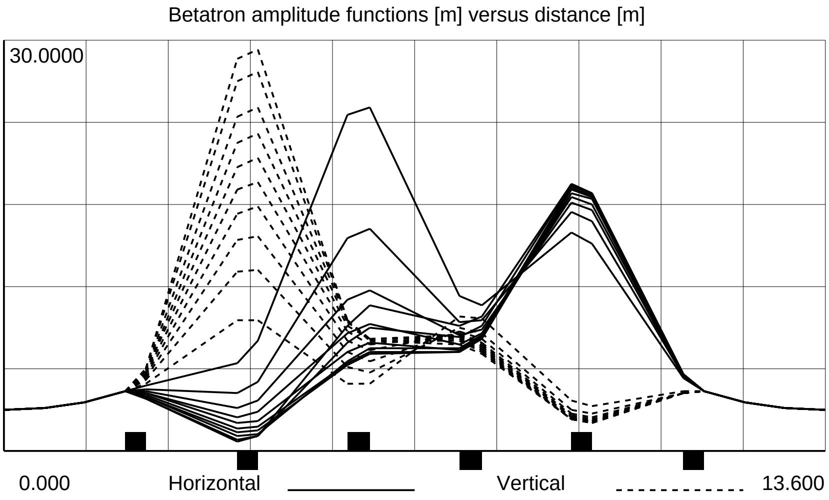


Figure 9.12 Optical functions in the phase shifter
[$\Delta\mu_x$ = 3.9, 3.8 then in steps of 0.2 down to 2,2 radian,
$\Delta\mu_z$ = 2.9 radian, $\beta_x = \beta_z = 3$ and $\alpha_x = \alpha_z = 0$ at entry and exit]

9.3.6 'Stepper'

To change the beam size in the vertical plane, a dedicated module, called the 'stepper', has been designed to vary the vertical betatron amplitude function over a wide range while keeping all the optical parameters in the horizontal plane, including horizontal phase advance, constant. Assuming that at the minimum extraction energy the vertical, rms, geometrical emittance of the circulating beam is the same for protons and ions and is equal to 1.65 π mm mrad, then the range needed at the patient for β_z is 2-23 m. Since the gantry and the deflection section out of the main extraction line will be telescopes in the *z*-plane in order to contain the betatron amplitude functions in the rotator, it is possible, to a certain extent, to choose the range of the stepper in β_z. The ratio β_{max}/β_{min} does not change, but the absolute value depends on the magnification factor chosen. In the gantry, a factor 3 will be applied, while in the deflection section a factor 0.5 is used. This implies that the stepper has to cover the range 2/3 β_{min} to 2/3 β_{max}, that is 1.333-15.333 m. In Figure 9.13, the β_z covers the range 1 to 17 and at the end of the module with an entry value of 3 m, $\beta_x = 3$ m at entry and exit, $\alpha_x = \alpha_z = 0$ at entry and exit and $\Delta\mu_x = \pi$. During the matching, the vertical phase advance was held constant at 2.9 radian for β_z = 1-11 m, but then to facilitate the minimisation the phase advance was allowed to vary, which interrupts the continuity between the traces in Figure 9.13.

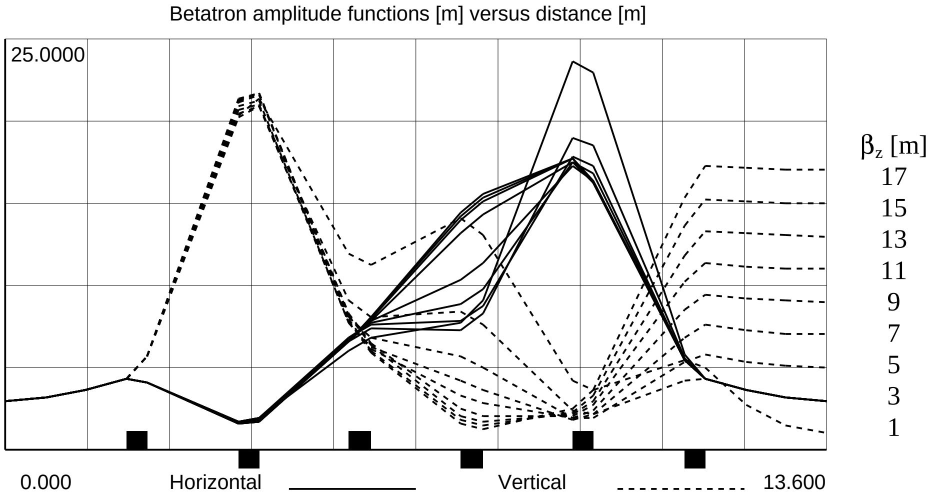


Figure 9.13 Optical functions in the 'stepper'
[β_z = 3 m at the entry and steps from 1 to 17 m in steps of 2 m at the exit. $\Delta\mu_x = \pi$ radian, $\Delta\mu_z$ varies from 2.9 to 3.23 radian, β_x = 3 and $\alpha_x = \alpha_z = 0$ at entry and exit]

9.3.7 Phase shifter - 'Stepper'

The modules shown in Figures 9.12 and 9.13 are in fact identical and it is possible to combine their functions into a single unit. This is inconvenient inasmuch as a single module has to span over a two dimensional parameter space, which makes the operation more complicated and may reduce the global ranges, but it represents a considerable saving in space. Figure 9.14, shows the beta functions in the 'phase shifter-stepper' for four extreme cases in the parameter range.

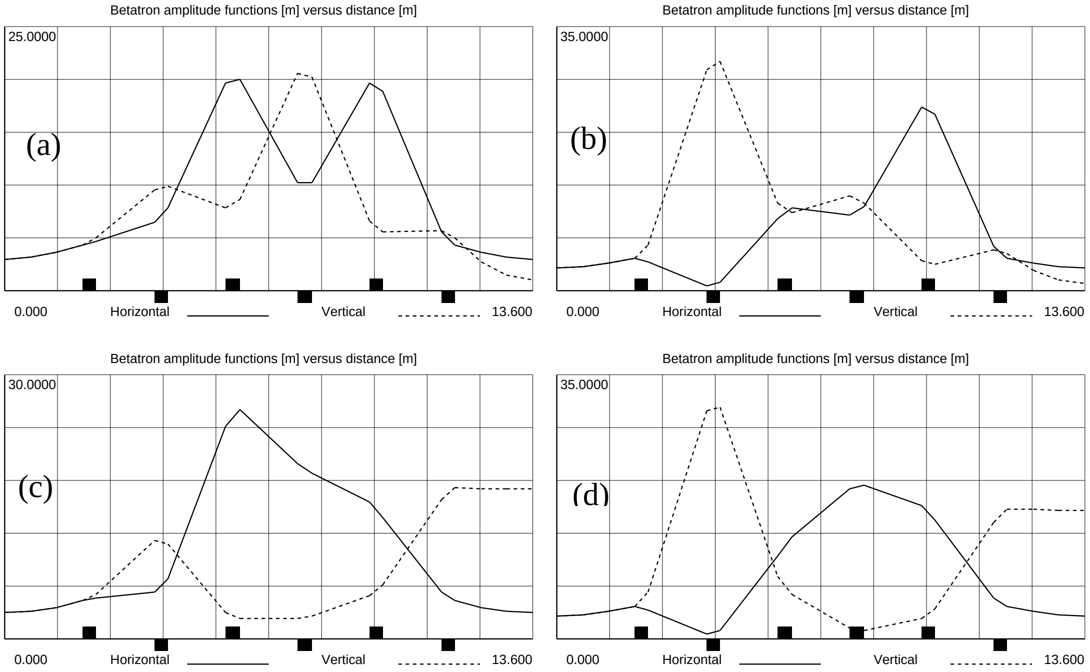
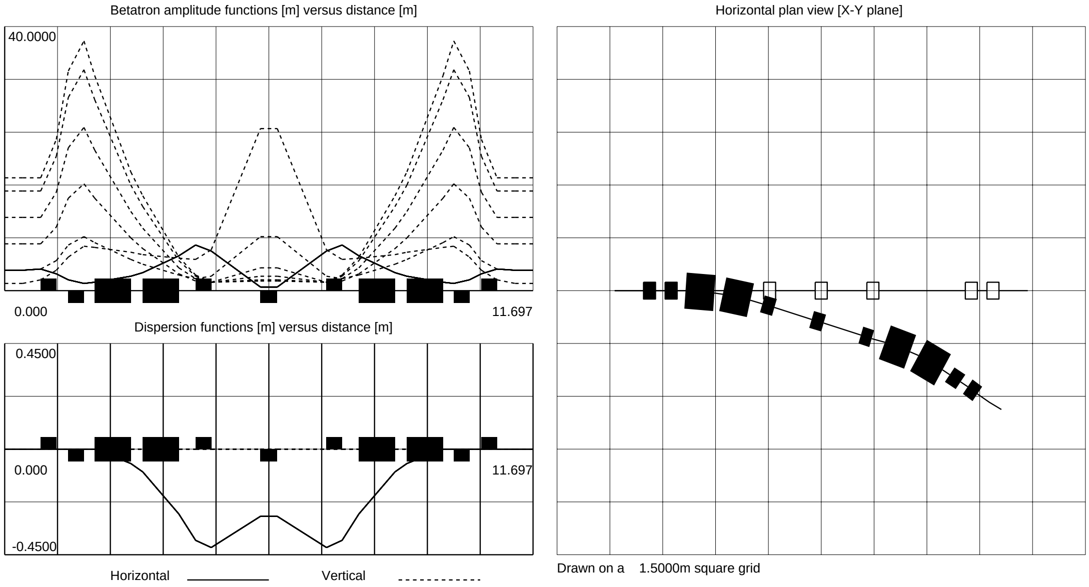


Figure 9.14 Extreme optical functions in the combined 'phase shifter-stepper'
[(a), (b) β_z = 1 m and $\Delta\mu_x$ = 2.2 and 3.9 radian; (c), (d) β_z = 17 m and $\Delta\mu_x$ = 2.2 and 3.9 radian]

9.3.8 Closed-dispersion bend

A one-to-one structure with a 2π phase advance is a very convenient structure in which to embed a bend with a closed-dispersion bump. Figure 9.15(a) shows the lattice functions of such a lattice module and 9.15(b) shows the geometry. The lattice functions are shown with the range of 1 to 17 m in the vertical plane. Since the bend is designed as a one-to-one module the maximum values are over 35 m. This module would be used with a gantry that requires zero dispersion at its entry and has a magnification of 1.35. In the next section, an example has been chosen where the bend is designed to de-magnify the vertical betatron amplitude function in order to limit the large vertical beam sizes.

In Figure 9.15(b), the outline is shown of the extension module that would continue to the next gantry. The lattice functions in this module would be very similar to the those in 9.15(a).



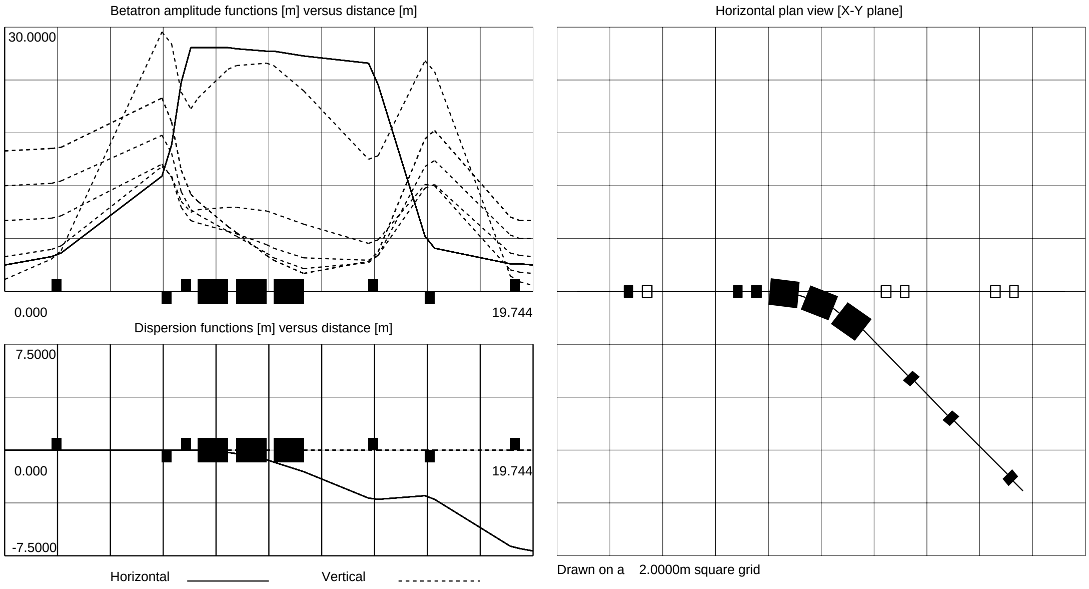
(a) (b)

Figure 9.15 Deflection section in the form of a one-to-one module

9.3.9 Open-dispersion bend

An open-dispersion bend has to be designed for a specific task, which in the present example is the matching to a Riesenrad gantry. As anticipated in the stepper section, a magnification factor of 0.5 has been chosen for the deflection module in the vertical plane. In the horizontal plane, the lattice is one-to one for $\beta_x = 3$ m $\alpha_x = 0$ to $\beta_x = 3$ m $\alpha_x = 0$. The dispersion has to be matched such that the dispersion bump closes in the gantry. Figure 9.16(a) shows the lattice functions in the bend for the full range of the vertical amplitude function and Figure 9.16(b) shows the geometry.

In Figure 9.16(b), the outline is shown of the extension module that would continue to the next gantry. The lattice functions in this module would be very similar to those in 9.16(a) although an extra quadrupole has been added before the bend.



(a) (b)

Figure 9.16 Open-dispersion bend
[Initial optical functions, β_z = 1.333, 4, 8, 12 and 16 m]

9.3.10 Riesenrad gantry

The Riesenrad gantry inverts the conventional iso-centric gantry geometry by placing the heavy accelerator equipment on the axis and positioning the patient on the outside as shown schematically in Figure 9.17. This configuration is felt to be a possible solution for an ion gantry and will be discussed in more detail in Part II.

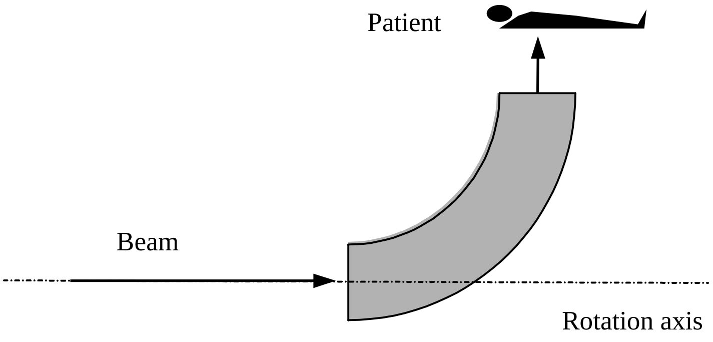
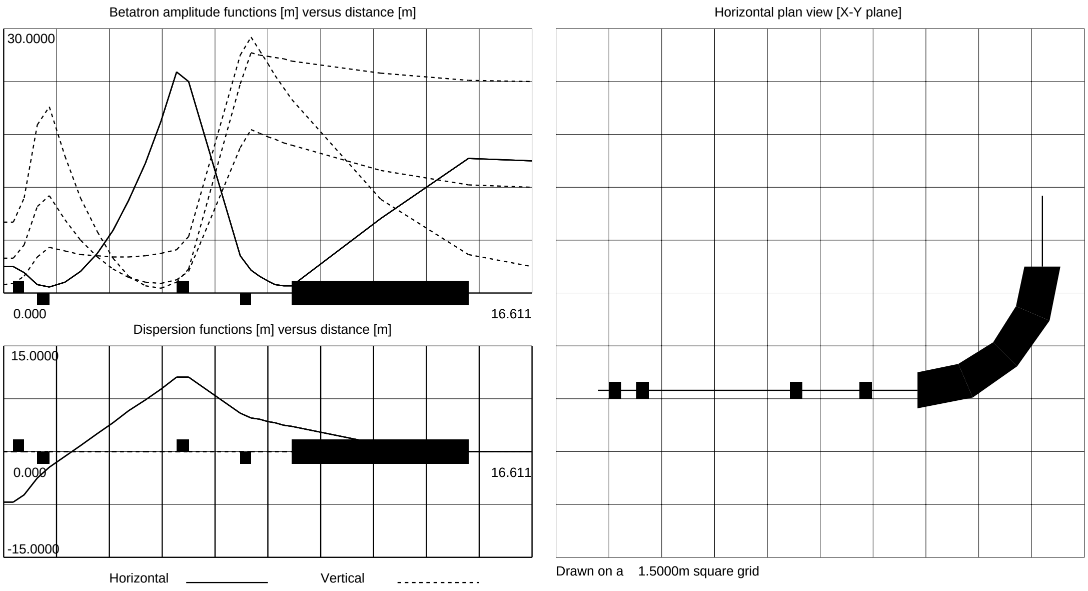


Figure 9.17 Schematic view of a Riesenrad gantry

The optical structure is essentially a single bending magnet preceded by a some quadrupoles to match the optical constraints. The Riesenrad gantry cannot have a closed-dispersion bump because there is only one dipole, but, thanks to the rotator, the dispersion bump can be closed in the deflection section. Thus, the remaining constraints are to:

- Maintain reasonable values of β_x and β_z inside the structure.
- Obtain the desired values for β_x and α_x.
- Obtain an $n\pi$ phase advance in the vertical plane.

As anticipated, a magnification factor of 3 has been chosen and an example lattice is shown in Figures 9.18(a) and (b).



(a) (b)

Figure 9.18 Plan view (a) and optical functions (b) in the Riesenrad gantry

9.3.11 Verification by tracking

Finally, the whole line has been tested by tracking the extracted particle distribution with different settings for the phase shifter-stepper. The phase-space footprints at the patient are shown in Figures 9.19 and 9.20 for $\beta_z = 1.5$ m and the two extreme horizontal phase advances that corresponding to vertical and horizontal positions of the 'bar' of charge. These two figures demonstrate the efficiency of the chosen system for varying the horizontal dimension of the beam. The vertical beam size is adjusted by varying β_z in the stepper that is situated upstream near the extraction from the synchrotron. In this way, a single module can control the both beam sizes for all the gantries in a complex.

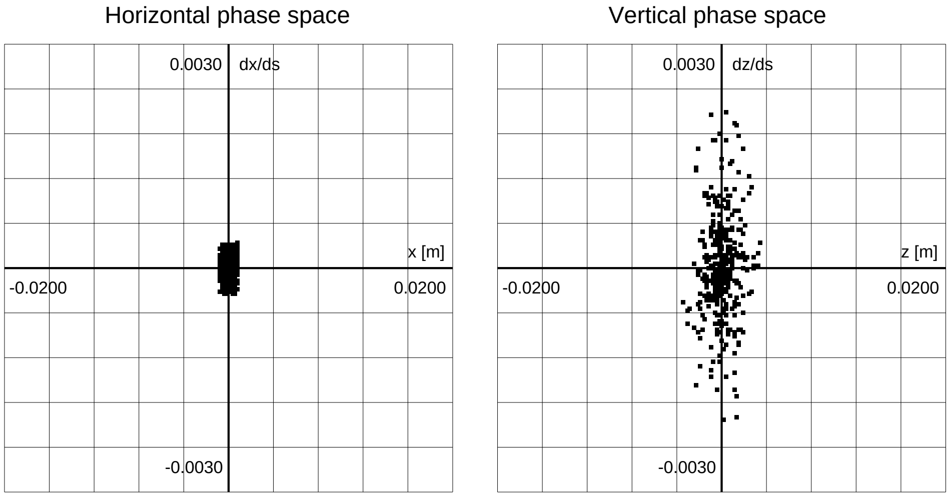


Figure 9.19 Phase space footprint for $\Delta\mu_x = 2.39$ rad in the phase shifter

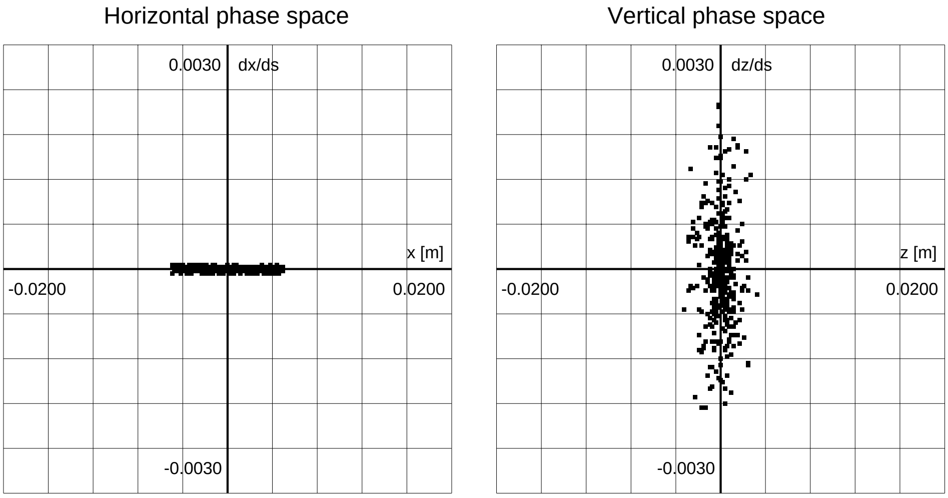


Figure 9.20 Phase space footprint for $\Delta\mu_x = 3.96$ rad in the phase shifter

References

[1] U. Amaldi, Private communication, The gantry design referred to here has been named the ‘Riesenrad’ and was first suggested during the preparation of the TERA Blue Book (*The TERA Project and Centre for Oncological Hadrontherapy*, ed. U. Amaldi and M. Silari, INFN, Frascati, 1995, II Ed.) and was later mentioned by E. Pedroni in the Grey Book (*Hadrontherapy in Oncology*, Como Proceedings, Excerpta Medica, International Congress Series, 1077, Elsevier, Amsterdam (1994), p434).

* * *

I-10 MULTIPLE SCATTERING

Multiple coulomb scattering is important in medical machines in several contexts. First, it plays a role in the stripping foil before injection into light-ion machines. In general, the emittances coming from the linac are smaller than those required to generate the spot sizes at the patient with reasonable values of the betatron amplitude function and this opens the possibility of using the stripping foil to dilute the emittance before injection. In the main ring, scattering and electron capture are concerns for the quality of the vacuum. In the beam delivery system, scattering is used to prepare the beam for the patient. It is usual to use a ridge filter to increase the momentum spread to widen the otherwise narrow Bragg peak and, in passive beam delivery systems, the irradiation field is made uniform by a specially designed double scatterer. Finally, the scattering in the patient's body adversely affects the spot size, which is more noticeable for protons than light ions.

In this chapter, the basic theory for multiple scattering will be reviewed with the emphasis on the practical aspects of applying the theory. The aim is to incorporate multiple scattering in the traditional Courant and Snyder theory for beam optics, as well as the usual Monte Carlo single-particle tracking.

10.1 CHARACTERISTIC MULTIPLE SCATTERING ANGLE

10.1.1 Highland's formula

Charged, high-energy particles traversing an absorber are liable to small deflections due to attractive and repulsive electrical forces of the orbital electrons and the nucleus of the absorber. This single scattering may occur many times during the traversal of the particle through the absorber and can add up to an appreciable net deviation from its original path. In order to calculate how much a particle may be scattered, it is necessary to know a characteristic scattering angle and its distribution.

There is common agreement in the literature that multiple scattering is best described by the theory of Molière [1] with the corrections made by Bethe [2], Fano [3] and Scott [4]. For an extensive review of the different approaches to multiple scattering see Ref. [5]. The distribution in Molière's theory is approximately gaussian for small angles, but for larger angles it behaves more like Rutherford scattering with larger tails than those of a gaussian distribution.

Unfortunately, Molière's theory is technically complicated and not easy to apply, but it is often sufficient to make a gaussian approximation for cases where the tails of the distribution are not of particular interest. This is true for most accelerators and medical applications where the effects of small scattering angles are dominant. In most cases, particles with large scattering angles will, in any case, be lost on the vacuum chamber walls.

A good approximation for the *characteristic scattering angle* θ_0 in a gaussian distribution is the approximation developed by Highland [6]. The Highland formula gives the RMS value of the distribution of scattering angles when projected onto a plane. The formula appears in the literature in a number of forms that fit different situations with varying degrees of precision, usually in the range of a few per cent. Highland's original formula appeared as,

Original:
$$\theta_0 \quad \frac{14.1[\mathrm{MeV}]}{p\quad c[\mathrm{MeV}]}\, z_{\mathrm{inc}} \sqrt{\frac{L}{L_{\mathrm{r}}}} \quad 1 \quad \frac{1}{9}\log_{10} \quad \frac{L}{L_{\mathrm{r}}} \quad [\mathrm{radian}] \tag{10.1}$$

where L is the scatterer thickness and L_{r} is the radiation length. The length units must be consistent and are usually either [g/cm^2] or [cm]. z_{inc} is the charge number of the incident particle, p is its momentum and its relativistic beta value. This formula is a fit that is meant to take into account the full thickness of the scatterer and the energy loss.

10.1.2 Improved Highland equation

Another popular description which can be found in [7] was given by Lynch and Dahl [8], sometimes called the *improved Highland equation*. The gain in improvement is achieved by applying a gaussian fit to a central region of a Molière distribution. They found that the best results can be obtained for the central 98% of such a distribution. This results in:

Improved:
$$\theta_0 \quad \frac{13.6\,[\mathrm{MeV}]}{p\quad c\,[\mathrm{MeV}]}\, z_{\mathrm{inc}} \sqrt{\frac{L}{L_{\mathrm{r}}}} \quad 1 \quad 0.038\ln \quad \frac{L}{L_{\mathrm{r}}} \quad [\mathrm{radians}]. \tag{10.1a}$$

According to [5], there is no real advantage in this approach and since the results from equations (10.1) and (10.1a) only differ for large scatterers no use will be made of (10.1a) here. A drawback of using the Highland formula is that it was obtained using data for 1 GeV protons, yet the literature claims that the accuracy is better than 5% in the range $10^{-3} \leq L/L_{\mathrm{r}} \leq 10$, except for very light elements and very low velocities. However, in the energy range of protons used for hadron therapy the error is much larger even for relatively small absorbers (< $2L_{\mathrm{r}}$). Figure (10.1) shows experimental data taken from [5] for 160 MeV protons scattered in lead. The dashed line shows the prediction by the original Highland formula.

The situation can be vastly improved by modifying equation (10.1) by taking into account the rapid change of energy of the incident proton, especially for thick absorbers. This yields in a generalised Highland equation [5],

Generalised:
$$\theta_0 \quad 14.1[\mathrm{MeV}]\, z_{\mathrm{inc}} \quad 1 \quad \frac{1}{9}\log_{10}\frac{L}{L_{\mathrm{r}}} \sqrt{\int_0^L \quad \frac{1}{p\quad c[\mathrm{MeV}]} \quad ^2 \frac{\mathrm{d}L'}{L_{\mathrm{r}}}}\,. \tag{10.2}$$

The integral under the square root in Equation (10.2) has to be evaluated numerically, yet it is sufficient to perform the evaluation using Simpson's rule,

$$\int_a^b f(x)\mathrm{d}x \quad \left[f(a) \quad 4f\left(\tfrac{a+b}{2}\right) \quad f(b)\right]\frac{b \quad a}{6} \tag{10.3}$$

for a number of slices. The number of slices depends on thickness of the scatterer. Absorbers where the average energy loss is smaller than 50% of the original particle energy can be calculated in a single slice. For scatterers larger than this limit extra slices should be introduced, but since $\left(\frac{1}{p\,c}\right)^2$ changes rapidly for a particle towards the end of its path, the slices should not be of equal size, but should be chosen in such a way that the integral steps are roughly equal. A good choice seems to be to make the following slice 1.6 times smaller than the previous slice. In total, no more than 8 slices are needed for any material for absorbers up to 97% of the path-length of an incident particle.

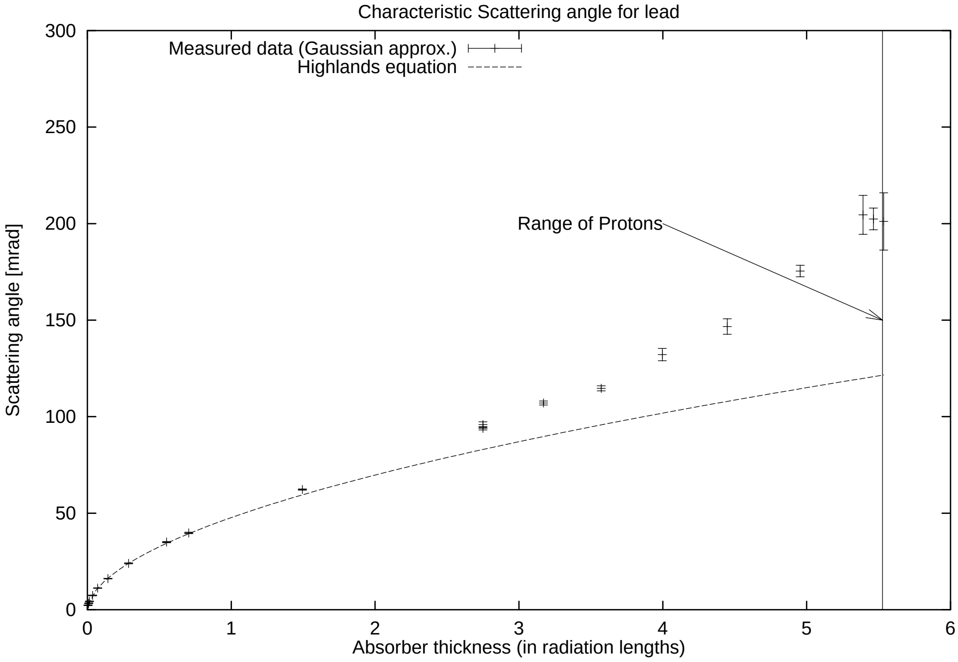


Figure 10.1 Comparison between measured data and the original Highland formula for 160 MeV protons

The term $\left(\frac{1}{p\,c}\right)^2$ is dependent on the energy of the particle, which is governed by the Bethe-Bloch Equation (10.18), which will be discussed later. Integrating (10.18) is complicated, however, and it is a lot easier and more accurate to use range tables to calculate the energy loss, which is also discussed later on. This makes the approach of using the Simpson rule for (10.3) for integrating (10.2) even simpler, since only the initial, final and midpoint energies have to be known.

The improvement of the generalised Highland equation (10.2) can be seen in Figure 10.2 which shows the same data as Figure 10.1, but with the prediction of Equation (10.2).

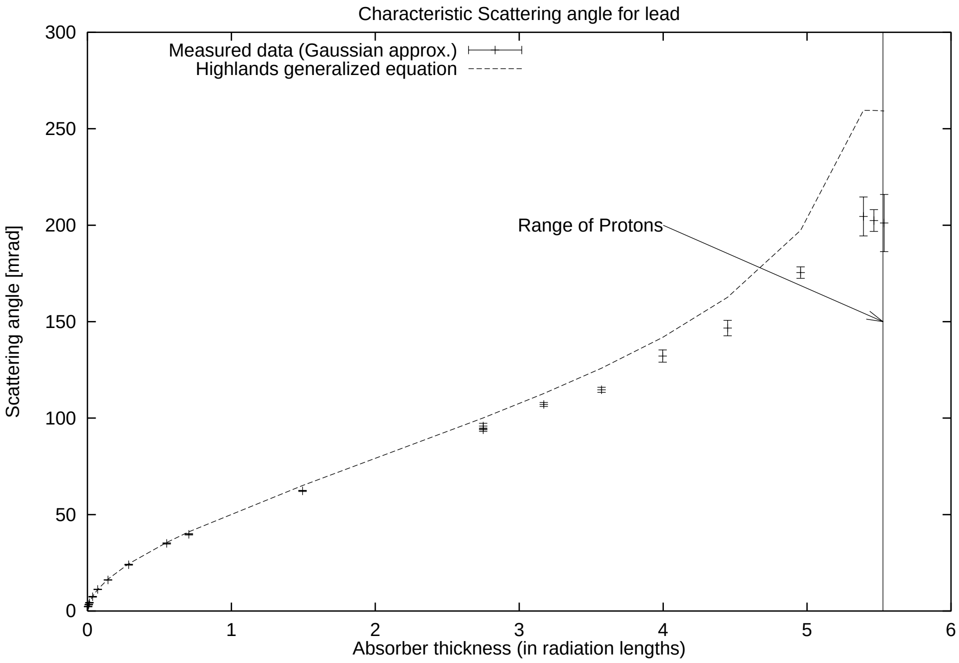


Figure 10.2 Comparison between measured data and the generalised Highland equation for 160 MeV protons

10.1.3 Sub-dividing scatterers

In general, it is a bad idea to subdivide a scatterer for numerical calculations. It neither improves the accuracy, nor does it save any time.

> ***Whenever possible, calculate the scattering angle in one go, possibly subdividing the integration into several parts as described above (10.2 - 3).***
>
> ***Adjust integration steps to be 1.6 times shorter than prior step.***
>
> ***When dealing with different layers or composite materials, always calculate a combined absorber using an average atomic number from the composites of the absorber.***

Nevertheless, there are situations where one cannot easily accommodate the above rules. One of these situations is when one wants to know the size of a beam somewhere in the middle of a fairly large scatterer. Here, the best procedure is to calculate the positions and angles of the particles always starting from the entry of the particle into the absorber. Never use intermediate results to continue the results of the scattering. Always restart from the beginning in one go to calculate effects of scattering at a later point.

10.2 SCATTERING SEEN THROUGH THE TWISS FUNCTIONS

The aim here is to incorporate multiple scattering in the standard Courant and Snyder beam optics theory. Once scattering can be parameterised and described by the phase-space ellipses, it can be included in lattice optics programs and even made part of matching routines. It is then possible, for example, to match to specified emittance values (providing of course that they are bigger than the original values), or to adjust the beam size as a function of the emittance and the betatron amplitude function combined, which is the situation within the patient's body for the spot size.

10.2.1 A thick scatterer with an uncoupled beam

The sigma matrix explained in Section 8.2 is a convenient way of describing the beam before and after the scattering process. Assuming that the beam is uncorrelated at the entry to the scatterer, the correspondence between the sigma matrix formalism and the Twiss formalism is given in Section 8.2.2 and is summarised by equation (8.13) that is repeated below.

$$\begin{matrix} \langle x^2 \rangle & \langle xx \rangle & 0 & 0 \\ \langle x\,x \rangle & \langle x^{\ 2} \rangle & 0 & 0 \\ 0 & 0 & \langle z^2 \rangle & \langle zz \rangle \\ 0 & 0 & \langle z\,z \rangle & \langle z^{\ 2} \rangle \end{matrix} \quad \frac{1}{} \begin{matrix} E_{\mathrm{x}\quad\mathrm{x}} & E_{\mathrm{x}\quad\mathrm{x}} & 0 & 0 \\ E_{\mathrm{x}\quad\mathrm{x}} & E_{\mathrm{x}\quad\mathrm{x}} & 0 & 0 \\ 0 & 0 & E_{\mathrm{z}\quad\mathrm{z}} & E_{\mathrm{z}\quad\mathrm{z}} \\ 0 & 0 & E_{\mathrm{z}\quad\mathrm{z}} & E_{\mathrm{z}\quad\mathrm{z}} \end{matrix} \ . \tag{8.13}$$

Bearing in mind that for hadron therapy the specification for the dose uniformity is typically 2%, the Highland formula [6,7] is sufficiently accurate to describe the multiple Coulomb scattering in a thick scatterer for the energies and materials normally used.

Although some aspects of scattering in an uncorrelated beam have been described earlier, this will now be redone using the sigma matrix formalism with Highland's formula. The analysis will start with an uncorrelated beam and will be extended later to the correlated case. The approach used here is more rigorous than that commonly found in the literature, since the correlation between angle and displacement, which occurs during scattering is taken into full account and thus provides a more general result. The earlier work made simplified assumptions which were well suited to thin scattering foils.

Consider a particle with the co-ordinates (y_1, y'_1). After passing through a scatterer, the particle will suffer a scattering angle θ_s and a displacement y and will assume the new co-ordinates (y_2, y'_2), The scattering is completely isotropic and unrelated to either y_1 or to y'_1, but there is a correlation between θ_s and y [9]. Three relationships can be established. Firstly, the change in divergence,

$$y_2 \quad y_1 \quad \theta_s \,. \tag{10.4}$$

By squaring (10.4) and then averaging over the whole beam, the increase in divergence can be related to the characteristic scattering angle of (10.1).

$$\langle y_2{}^2 \rangle \quad \langle y_1{}^2 \rangle \quad 2\langle y_1 \theta_s \rangle \quad \langle \theta_s^2 \rangle$$

where $\left\langle \theta_s^2 \right\rangle \quad \theta_0^2$ and $2\langle y_1 \theta_s \rangle \quad 0$ because y'_1 and θ_s are uncorrelated. Thus, the new divergence is given by,

$$\left\langle {y_2}^2 \right\rangle \quad \left\langle {y_1}^2 \right\rangle \quad \theta_0^2 \tag{10.5}$$

Similarly,

$$\left\langle y_2^2 \right\rangle \quad \left\langle y_1^2 \right\rangle \quad 2\langle y_1 \;\; y \rangle \quad \left\langle \;\; y^2 \right\rangle$$

$$\left\langle y_2^2 \right\rangle \quad \left\langle y_1^2 \right\rangle \qquad \frac{L}{\sqrt{3}}\theta_0 \quad {}^2 \tag{10.6}$$

where $\left\langle \;\; y^2 \right\rangle \qquad \frac{L}{\sqrt{3}} \;\; {}_0 \quad {}^2$ from θ[9]. Finally, the change in the cross-term yy' is given by,

$$\langle y_2 y_2 \rangle \quad \langle (y_1 \quad\;\; y)(y_1 \quad \theta_s) \rangle$$

$$(y_2 y_2) \quad \langle y_1 y_1 \rangle \quad \langle \theta_s \;\; y \rangle \quad \langle y_1 \theta_s \rangle \quad \langle y_1 \;\; y \rangle . \tag{10.7}$$

As mentioned earlier, the scattering is isotropic and has no correlation with y_1 or y_1', so the last two terms in (10.7) average to zero. The second term, however, makes a finite contribution, since large scattering angles tend to be associated with large displacements. This is intuitively obvious, since to have received a large displacement the trajectory angle must have been large on average while crossing the scatterer. If the angle is large on average while crossing the scatterer, it will on average be large at the exit of the scatterer. Reference 4 gives the correlation as 3/2. Using the definition of the correlation coefficient,

$$\langle \theta_0 \;\; y \rangle \qquad \sqrt{\left\langle \theta_0^2 \right\rangle \left\langle \;\; y^2 \right\rangle} \tag{10.8}$$

it follows that

$$\langle \theta_s \;\; y \rangle \quad \frac{L}{2} \;\; {}^2_0 . \qquad\qquad \theta \tag{10.9}$$

Thus (10.7) becomes,

$$\langle y_2 y_2 \rangle \quad \langle y_1 y_1 \rangle \quad \frac{L}{2} \;\; {}^2_0 . \qquad\qquad \theta \tag{10.10}$$

The changes due to scattering are described statistically by Equations (10.5), (10.6) and (10.10) and provide the three relationships needed to solve for the Twiss functions after scattering. Thus, using the correspondences in the matrix equation of (8.13),

Scattered ellipse for a thick scatterer;

$$E_2 \;\; {}_2 \quad E_1 \;\; {}_1 \qquad {}^2_0 \qquad\qquad \theta$$

$$E_2 \;\; {}_2 \quad E_1 \;\; {}_1 \qquad \frac{L}{\sqrt{3}}\theta_0 \quad {}^2 \tag{10.11}$$

$$E_2 \;\; {}_2 \qquad E_1 \;\; {}_1 \qquad \frac{L}{2}\theta_0^2$$

The right-hand sides of each equation in (10.11) are fully evaluated by the conditions at the entry to the scatterer and by the characteristics of the scatterer. Re-writing (10.11) with three constants E_2 $_2$ = A; E_2 $_2$ = B and E_2 $_2$ = C and substituting into the Twiss relation leads to the solution of the three equations by

$$\frac{1 \qquad ^{2}}{} \qquad \text{so that} \quad \frac{1 \quad (C/E_2)^2}{B/E_2} \quad \frac{A}{E_2} \quad \text{and} \quad E_2 \quad \sqrt{AB \quad C^2}\,. \qquad (10.12)$$

Consequently E_2 can be evaluated as

$$\boxed{\underline{\text{Scattered emittance}}\text{: } E_2 \quad E_1\sqrt{\quad 1 \quad \frac{\theta_0^2}{E_1} \quad 1 \quad \frac{L^2\theta_0{}^2}{3E_1} \qquad 1 \quad \frac{L}{2E_1}\theta_0^2 \quad ^{2}} \quad .(10.13)}$$

10.2.2 A thick scatterer with a coupled beam

This approach can now be generalised to coupled beams. The sigma matrix for a coupled beam has additional matrix elements:

$$\begin{matrix} \langle x^2\rangle & \langle xx'\rangle & \langle xz\rangle & \langle xz'\rangle \\ \langle x'x\rangle & \langle x'^2\rangle & \langle x'z\rangle & \langle x'z\ \rangle \\ \langle zx\rangle & \langle zx'\rangle & \langle z^2\rangle & \langle zz\ \rangle \\ \langle z'x\rangle & \langle z'x\ \rangle & \langle z'z\rangle & \langle z'^2\rangle \end{matrix} \qquad (10.14)$$

By adding the scattering in each term, squaring and then averaging over the beam, the overall effect of the scattering can be evaluated as before. The calculations for the elements not describing the coupling remain the same, since scattering is isotropic. The only additional work to be done is to deal with the three coupled terms in the matrix.

$$\langle x_2z_2\rangle \quad \langle(x_1 \qquad x)(z_1 \qquad z)\rangle$$

which can be transformed to

$$\langle x_2z_2\rangle \quad \langle x_1z_1\rangle \quad \langle x_1\ \ z\rangle \quad \langle z_1\ \ x\rangle \quad \langle\ \ x\ \ z\rangle.$$

Since scattering does not couple between the x and z planes all correlation products containing elements from two different planes equal to zero. Therefore,

$$\langle x_2z_2\rangle \quad \langle x_1z_1\rangle\,.$$

Similarly all other coupled terms stay unchanged after scattering and only the uncoupled matrix elements change.

$$\begin{aligned}\langle x_2 z'_2\rangle \quad & \langle (x_1 \quad x)(z'_1 \quad \theta_{z0})\rangle \\ & \langle x_1 z'_1\rangle \quad \langle x_1\theta_{z0}\rangle \quad \langle z'_1 \quad x\rangle \quad \langle \theta_{z0} \quad x\rangle \\ & \langle x_1 z'_1\rangle\end{aligned}$$

And finally,

$$\begin{aligned}\langle x'_2 \, z'_2\rangle \quad & \langle (x'_1 \quad \theta_{x0})(z'_1 \quad \theta_{z0})\rangle \\ & \langle x'_1 \, z'_1\rangle \quad \langle x'_1 \, \theta_{z0}\rangle \quad \langle z'_1 \, \theta_{x0}\rangle \quad \langle \theta_{x0}\theta_{z0}\rangle \\ & \langle x'_1 \, z'_1\rangle\end{aligned}$$

Thus, it is possible to treat the effect of a scatter on a coupled beam with the same methods as for an uncoupled beam, which is a welcome simplification although, at first, a little surprising.

10.2.3 Approximation for a thin scatterer

The above formulation often appears in an approximate form for thin scatterers. In a thin scatterer, it is assumed that the position of the particle is unchanged by its passage through the scatterer so that only the change in divergence needs to be taken into account. Neglecting the terms depending on *L* in (10.11) gives,

Scattered ellipse for a thin scatterer;

$$\begin{aligned} & E_{2 \;\; 2} \quad E_{1 \;\; 1} \quad {}^{2}_{0} \quad \theta \\ & E_{2 \;\; 2} \quad E_{1 \;\; 1} \\ & E_{2 \;\; 2} \quad E_{1 \;\; 1}\end{aligned} \qquad (10.15)$$

Equation (10.12) is still valid and yields,

$$E_2 \quad E_1\sqrt{1 \quad \frac{{}_1\theta_0^2}{E_1}} \, . \qquad (10.16)$$

The thin scatterer approximation is useful for such applications as stripping foils and vacuum windows.

10.2.4 Comparison with a simulation

A simple comparison with tracking shows the validity of Equation (10.11). An uncorrelated gaussian beam of 100’000 protons at 180 MeV with a momentum spread of 0.1% and a spatial cut-off at 2 was generated and tracked through a copper foil 3.67 mm thick. The absorber adds an average RMS angle of θ_0 = 20.4 mrad (using Highland’s formula (10.1)). The beam scatter plot in real space changes considerably as can be seen from Figures 10.3 and 10.4 that show one phase plane and the real plane before and after the scatterer. The uncorrelated beam is rectangular in real space, but elliptical in phase space, which can be seen from a careful study of Figure 10.3.

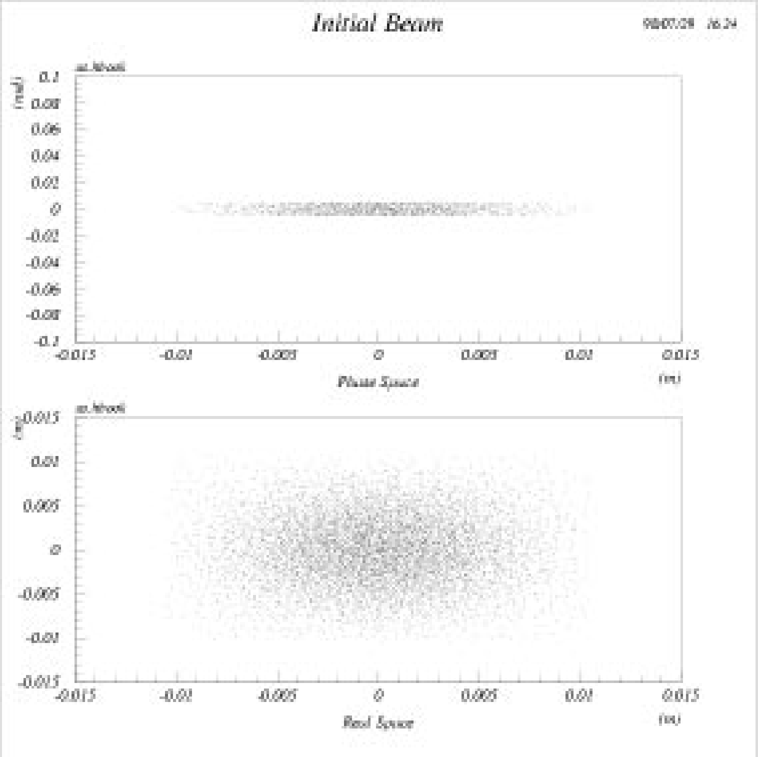


Figure 10.3 Phase and real space of the initial simulated beam

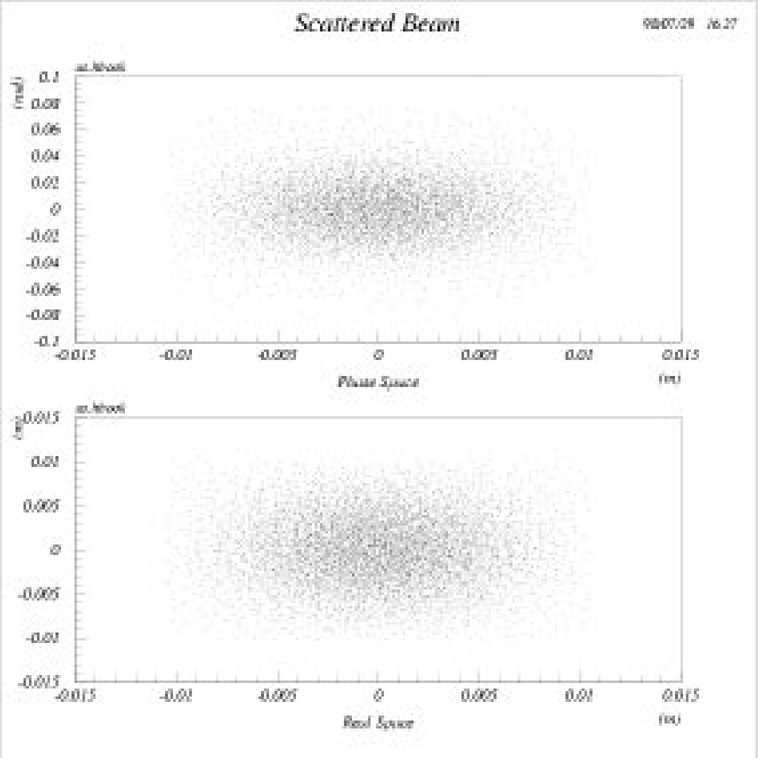


Figure 10.4 Phase and real space plots of the scattered beam

The numerical data corresponding to Figures 10.3 and 10.4 is collected in Table 10.1. The estimated Twiss parameters of the scattered beam obtained by statistical analysis of the distribution calculated by tracking agree extremely well with those calculated using the ‘ellipse’ equation (10.11).

Comparison of calculated beam parameters and those obtained by tracking

	Input beam		**Beam after scattering**	
	Parameters for generated distribution	**Statistical estimation from generated distribution**	**Calculated by modified optics theory**	**Estimated statistically from tracking results**
$E_{\mathrm{RMS},x}$ [mm mrad]	2.30	2.30	77.3	76.5
$E_{\mathrm{RMS},z}$ [mm mrad]	2.30	2.30	77.3	76.6
$_x$ [m]	6.25	6.27	0.19	0.19
$_z$ [m]	6.25	6.25	0.19	0.19
$_x$	0.00	0.00	0	0.00
$_z$	0.00	0.00	0	0.00

Table 10.1 Comparison of calculated beam parameters and those obtained by tracking

10.3 MULTIPLE SCATTERING AND TRACKING

This topic will be discussed with reference to passive beam spreading systems that are based on specially designed scatterers that create a large (typically 20 20 cm^2) uniform (ideally < 2%) irradiation field. The design of such scatterers can be achieved using Monte Carlo tracking techniques, whereby a few 10^5 particles are tracked repeatedly through the scatterer optimisation loop that modifies the shape of the scatterer according to the particle density field that is created. Final runs with the order of 10^6 particles are needed to assure the specified accuracy to a high confidence level. An added complication in the design of such a system is the need to equalise the energy loss for all particles, which necessitates an additional scattering layer of adjustable thickness of lucite.

10.3.1 Energy loss from particles passing through matter

The energy loss for low energy particles passing through an absorber is dominated by ionisation. For heavy particles, i.e. particles heavier than a muon, this is well described by the Bethe-Bloch equation [7,12]

$$\frac{\mathrm{d}E}{\mathrm{d}x}\quad Kz_{\mathrm{inc}}{}^{2}\,\frac{Z}{A}\frac{1}{\;\;^{2}}\;\frac{1}{2}\ln\frac{2m_{\mathrm{e}}c^{2}\;\;^{2}\;\;^{2}T_{\max}}{I^{2}}\qquad{}^{2}\;, \tag{10.17}$$

where K equals to $4\;\;N_{\mathrm{A}}r_{\mathrm{e}}^{2}m_{\mathrm{e}}c^{2}$, Z is the atomic number, A the atomic mass of the medium, z_{inc} the number of elementary charges of the incident particle, is the velocity of the particle in units of the speed of light, I is the mean excitation energy in eV, m_{e} is the mass of the electron and r_{e} is the classical electron radius and finally, $T_{\max}$ is the maximum transferable kinetic energy that can be imparted to a free electron in a single collision.

The equation (10.17) is remarkable in two ways: firstly, it is strongly dependent on the velocity and, secondly, it is only very weakly dependent on the mass (via the maximum transferable kinetic energy T_{max}) of the incident particle. Since the ratio of *A/Z* is relatively constant over a large range of elements, it is possible to generate universal energy-loss curves by plotting the equation (10.17) versus the *area density* in g/cm^2. For ultra relativistic particles equation (10.17) can be extended to take into account shell effects and density corrections. For low-energy particles these correction factors can be neglected.

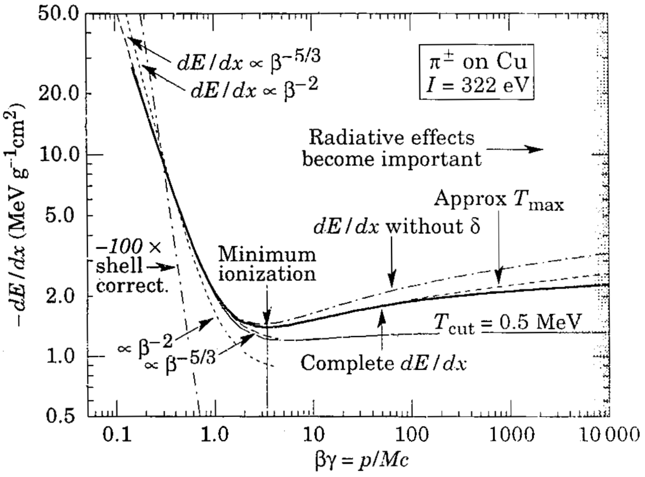


Fig 10.5 Bethe Bloch energy-loss curve

Figure 10.5 shows the d*E*/d*x* curve for charged pions in copper. Note that the energy loss per path-length is plotted versus momentum and not absorber thickness. In this way (10.17) is used for particle identification. The plot shows that the Bethe-Bloch equation can be cut conveniently into three regions. For low energies, which is the region of interest for medical purposes, the energy-loss is roughly proportional to $^{-5/3}$. The curve reaches a minimum at about 1 GeV/c, where a particle is called a minimum ionising particle (MIP). Finally, the energy loss rises again for ultra relativistic particles (relativistic rise).

The energy loss of a particle can be calculated by integrating (10.17):

$$E \quad \int_0^{x_0} \frac{dE}{dx} dx \tag{10.18a}$$

For particles with a kinetic energy larger than that of a MIP, it is usually sufficient to calculate the energy loss by evaluating:

$$E \quad \frac{dE}{dx} \quad x, \tag{10.18b}$$

where x is the absorber thickness. Unfortunately, particles for medical synchrotrons are well below the level of a MIP and therefore equation (10.18b) cannot be used. Instead the calculation of the energy loss, E, has to be done using numerical methods. One of these methods is the use of range tables, discussed below.

10.3.2 The range

Once the energy loss in an absorber can be calculated, it is natural to ask at what distance 50% of the particles will have lost all their kinetic energy T. This is known as the *range R* of a particle and can be found by integrating the inverse of (10.17), that is

$$R(T) \quad \int_{0}^{T_0} \frac{\mathrm{d}E}{\mathrm{d}x}^{\;\;1} \mathrm{d}E \tag{10.19}$$

The particle ranges are often plotted in tables and can be used to efficiently calculate the energy loss of particles passing through an absorber. For examples of such tables see reference [10, 11].

10.3.3 Energy straggling

The energy loss of a particle in matter is a statistical process. Equation (10.17) only describes the average energy loss of a particle, but does not describe the energy loss distribution. Calculating this distribution is mathematically complicated and is generally divided into two distinctive cases: thin absorbers and thick absorbers.

The case of thin absorbers it extremely difficult to calculate. The distribution, is asymmetric with long tails (see Figure 10.6). It was first described by Landau[12]. Part of the asymmetry is due to fact that a small energy loss is much more probable than a large one and that there is a cut-off for large energy losses given by the largest transferable kinetic energy T_{max}. The Landau theory has to be applied where the thickness of the absorber is only of a few hundred micrometers, e.g. for stripping foils.

Scatterers used for passive beam spreading on the other hand, can be treated as thick absorbers. For thick absorbers, where the number of collisions is large, the energy loss distribution can be shown to be gaussian. This follows directly from the Central Limit theorem in statistics. This very general theorem states that the sum of N random variables, all with the same distribution, approach a gaussian-distributed variable in the limit of N . The Central Limit Theorem therefore allows the calculation of the energy loss distribution without any subsequent knowledge of the distribution for a few-, or for single-scattering collisions. The width of this distribution was first calculated by Bohr (for non-relativistic particles)[12] for an absorber of thickness L:

$${}_{0}^{2} \quad 4 \;\; N_{\mathrm{A}} r_{\mathrm{e}}^{2} \left(m_{\mathrm{e}} c^{2}\right)^{2} \;\; \frac{Z}{A} L \quad 0.1569 \;\; \frac{Z}{A} L \left[\mathrm{MeV}^{2}\right], \tag{10.20}$$

which can easily be extended to relativistic particles via

$$ {}^{2} \quad \frac{1 \quad \frac{1}{2} \quad {}^{2}}{1 \quad {}^{2}} \quad {}^{2}_{0} \tag{10.21}$$

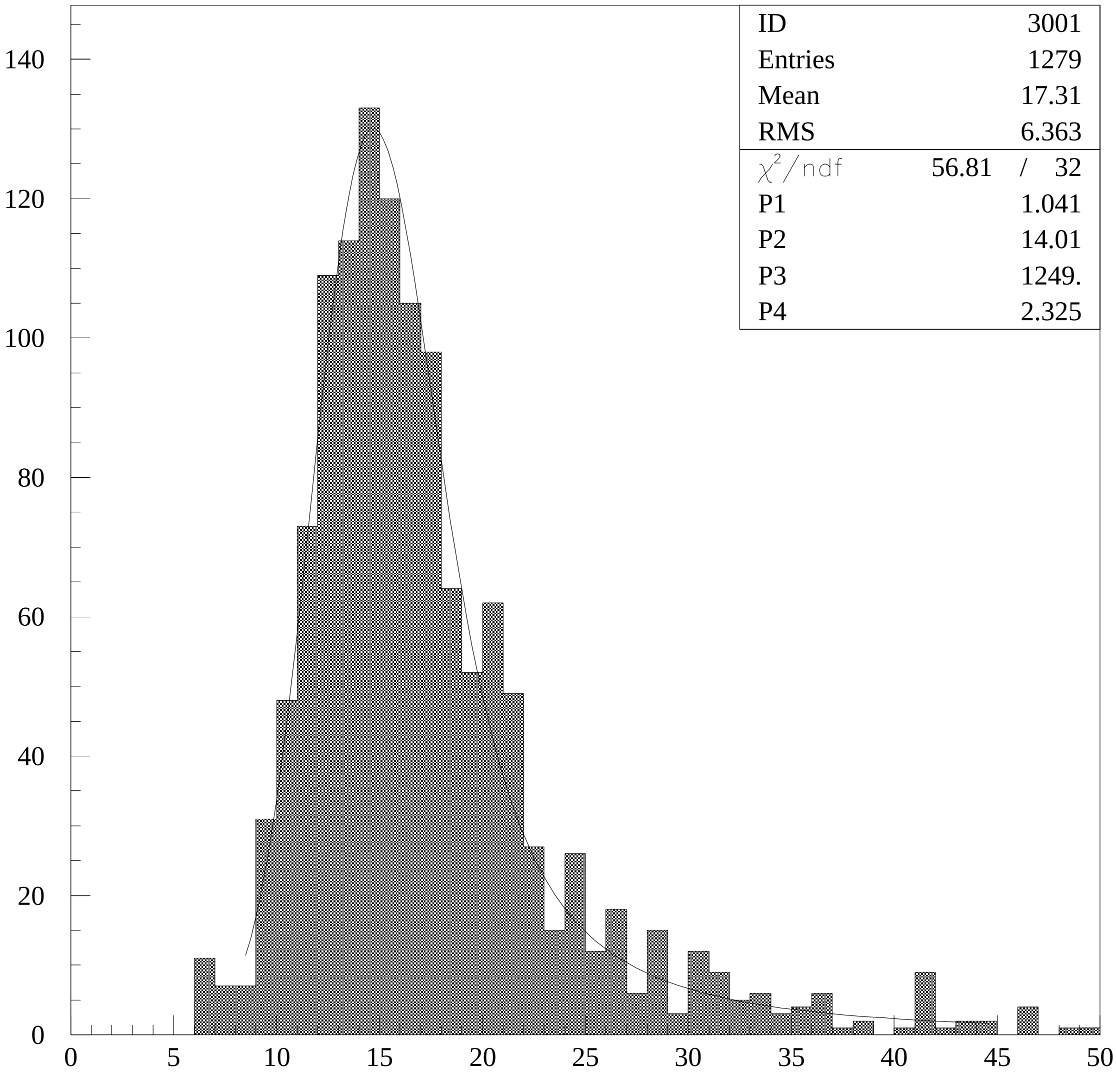


Figure 10.6 Landau plot for the energy straggling in a 300 μm thick silicon detector using a RU^{106} source

10.3.4 Energy loss calculations using range tables

When calculating the energy loss of a particle traversing an absorber, equation (10.17) is found to be numerically demanding and evaluating the integral (10.18a) may be prohibitive in the case of many simulated particles. It may therefore be more efficient to apply the following method.

A range table with a sufficient number of entries is used to calculate the range of a given particle with an initial energy E_0. Values of E_0 that are not included in the table have to be interpolated linearly. Knowing the value of R_0 at E_0 makes it possible to determine the average energy loss by subtracting the absorber thickness L from R_0, resulting in the 'leftover' range R_1.

$$R_1 \quad R_0 \quad L$$

Now the range table is inversely searched to find the energy E_1 that corresponds to a range of R_1. Subtracting E_1 from E_0 results in the average energy loss $E\ (\ E \quad E_0 \quad E_1)$. Again, values not found in the table have to be interpolated linearly. If the range table is populated densely enough, an accuracy of 0.13% can be achieved [10], well within experimental limits.

The computational cost of a 'look-up' table is small when using appropriate algorithms, such as a bisection search on an ordered table. Despite modern algorithms for numerical integration, such as Romberg's method[13], the use of range tables is far more efficient than the direct approach of evaluation of the integral in (10.18a).

10.3.5 Monte-Carlo recipe for calculating scattering in matter

Equation (10.1) gives the RMS value of the scattering angle when the three-dimensional scattered distribution is projected onto a plane. This can now be used to track a number of individual particles.

Consider a particle traversing an absorber of thickness x and being scattered along its path (see Figure 10.7). The projection of this path is $y \quad x \quad _{\text{plane}}$, where

$$_{\text{plane}} \quad \frac{1}{\sqrt{3}}\theta_{\text{plan}}. \tag{10.22}$$

A particle that has received a kick and exits with a non-zero value of θ_{plane} also has a greater chance of having a non-zero displacement y_{plane}. There is a correlation between θ_{plane} and y which is statistically expressed via a correlation coefficient $_{y\theta}$. It has a value [9] of

$$_{y} \quad \frac{\sqrt{3}}{2} \quad 0.87 \tag{10.23}$$

For a Monte-Carlo simulation it is probably most convenient to use two independent gaussian random numbers (z_1, z_2) with zero mean and variance one and then to set

$$\begin{aligned} y_{\text{plane}} \quad & \frac{z_1 x\theta_0\sqrt{1 \quad {}^{2}_{y}}}{\sqrt{3}} \quad \frac{z_2 \quad _{y}}{\sqrt{3}} \\ & \frac{z_1 x\theta_0}{\sqrt{12}} \quad \frac{z_2 x\theta_0}{2}; \qquad (10.24) \\ \theta_{\text{plane}} \quad & z_2\theta_0. \end{aligned}$$

This ensures that the correlation is properly taken into account. After the calculation of the scattering, one has to correct the energy absorbed by the scatterer. This can be done using the method described in Section 10.3.4. Of course the average energy loss should be randomised as well, using yet a third gaussian number z_3 and computing:

$$E^{\text{scattered}} \quad E_0 \quad z_3 \tag{10.21}$$

using the initial beam energy E_0 and the width of the energy straggling calculated according to (10.21).

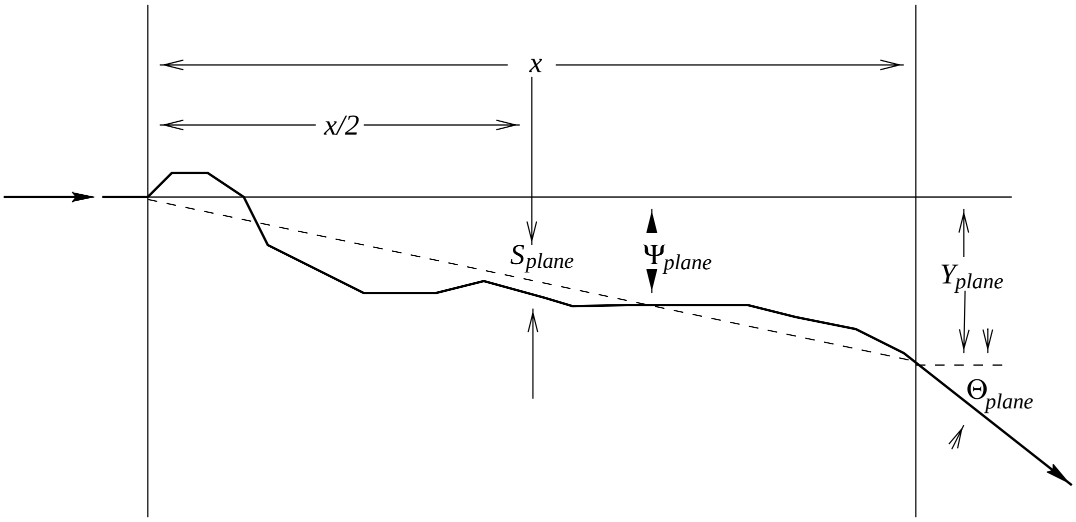


Figure 10.7 Schematic view of a particle traversing an absorber

References

[1] G. Molière, *Theorie der Streuung schneller geladener* Teilchen II, Z. Naturforschlg. 3a, (1948), p78-97.

[2] H.A. Bethe, *Molière's Theory of Multiple Scattering,* Phys. Rev. 89 (1953), p1256-66.

[3] U. Fano, *Inelastic Collisions and the Molière Theory of Multiple Scattering,* Phys. Rev. 93 (1954), p117-20.

[4] W.T. Scott, *Review of Modern Physics,* 35 (1963) p231-313.

[5] B. Gottschalk, A.M. Koehler, R.J. Schneider, J.M. Sisterson, M.S. Wagner, *Multiple coulomb scattering of 160 MeV protons,* Nucl. Instr. Meth.., B74, (1993), p467-90.

[6] V.L. Highland, *Some Practical Remarks on Multiple Scattering,* Nucl. Instr. and Meth. 129, (1975), p497-9 and Erratum Nucl. Instr. and Meth. 161, (1979), p171.

[7] Particle Data Group, *Review of particle properties,* Physics Letters, vol IIIB, (April 1982).

[8] G. Lynch, O.I. Dahl, *Approximations to multiple Coulomb scattering,* Nucl. Instr. Meth. B58 (1991), p6-10

[9] R.K. Bock, K. Bos, S. Brandt, J. Myrheim, M. Regler, editors, *Formulæ and methods in experimental data evaluation,* Volume I, (European Physical Society, 1983).

[10] H. Bichsel, *Charged-Particle Interactions,* in F.H. Attix, W.C. Roesch (Eds), *Radiation Dosimetry,* Vol 1 (1968), (Academic Press), p157-227.

[11] J.F. Janni, *Calculations of energy loss, range, pathlength, straggling, multiple scattering and the probability of inelastic nuclear collisions for 0.1 to 1000 MeV protons,* Air Force Weapons Laboratory Report AFWL-TR 65-150, (1996).

[12] W.R. Leo, *Techniques for Nuclear and Particle Physics Experiments,* (2nd Ed., Springer Verlag, 1994).

[13] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C,* (2nd Ed., Cambridge University Press, 1992).

* * *

I-11 ACTIVE SCANNING

So far this report has been concerned with the production of a smooth spill from the third-integer resonance, since this is a fundamental requirement for an efficient scanning system. The better the spill quality, the faster and more accurate the scan becomes. Fortunately, the scanning system itself is tolerant to high-frequency fluctuations and can be made to correct on-line for low-frequency fluctuations. The system's ability to compensate for a poor spill quality is the main subject of this section and is evaluated with reference to a target error of ± 2% in the dose distribution. The performance depends on the scanning technique applied and the technological limitations of the equipment. Should the spill irregularities overwhelm the capabilities of the system for compensation, then an alarm message must be issued and the beam be switch off.

There are three techniques for active scanning:

- Voxel (volume pixel),
- Mini-voxel
- Raster scanning.

Voxel scanning [1] was already discussed in Chapter 5 and will be briefly reviewed below. Mini-voxel scanning [2] resembles voxel scanning, except that the voxels overlap. Finally, true raster scanning uses a continuous motion of the beam spot that is modulated in velocity according to the spill intensity. Thus, mini-voxel scanning is an intermediate stage between the *static* voxel scanning and the *dynamic* raster scanning. In this chapter, the three techniques will be discussed and the practical limitations due to the discrete measurement intervals, the delays in the electronics and the response time of the power converters will be analysed.

11.1 MINI-VOXEL SCANNING

11.1.1 Comparison of voxel and mini-voxel scanning

It is useful at this stage to recall the main principles of voxel scanning. Let each voxel have a physical size *W* corresponding to the FWHM of the beam spot (see Figure 11.1(a)). The full cross-section of the tumour is covered by a mesh of distinct voxels, which are all irradiated separately. For each voxel, the scanning magnets are adjusted while the beam is turned off. The beam is then turned on until the desired dose for that voxel is delivered. After the beam is turned off, the magnets are readjusted for the next voxel. This technique is simple and the on-line dosimetry is relatively straightforward. However, it is sensitive to the alignment of the voxels with slow extracted beams. This is due to the near-rectangular distribution of the beam spot in one direction which will produce 'hot' and 'cold' spots with small misalignments. The treatments times are also increased by the regular switching off of the beam.

In mini-voxel scanning, the beam spot crosses from one 'main' voxel position to the next in *n* discrete steps of size *W*/*n* (see Figure 11.1(b)). The operational principle is to dwell at each intermediate position until the required fraction (1/n) of the full dose for a normal voxel is delivered and then to move on at the maximum scanning velocity to next intermediate position without switching off the beam. The particles that arrive after the decision to move are all attributed to the next mini-voxel position by the control system. This may appear strange, since an arbitrary sharing of say 50% in the last position and 50% in the next position would be closer to reality. However, the overall dose distribution is better managed, if this flux is attributed to the next position where there is still a dose deficit. In principle, a perfect correction can then be made at this position, except that it is misplaced by the mini-voxel step size. This procedure is repeated for the next move and so on. Intuitively, it appears that the smaller the mini-voxel step the better the smoothing, which suggests that true raster scanning could be the ideal technique. Mini-voxel scanning is quicker than voxel scanning and more efficient in its use of the beam.

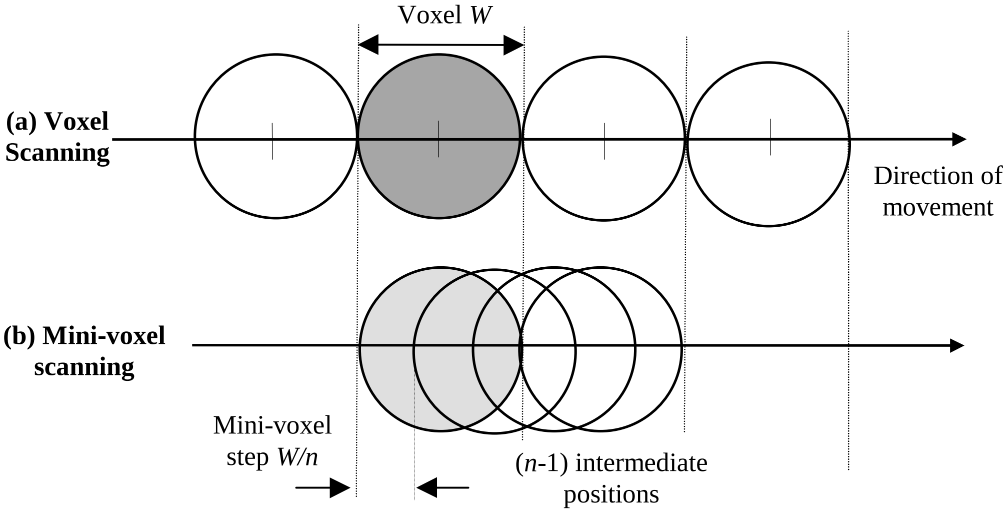


Figure 11.1 Schematic view of voxel and mini-voxel scanning

11.1.2 Influence of the beam distribution

In slow-extracted beams, the particle distribution is practically rectangular in one transverse direction and gaussian in the other. The rectangular distribution must be aligned with the principal direction of scanning. In this way, the tails of the gaussians overlap between adjacent scan lines and render the alignment of adjacent rows insensitive to small positional errors. In the scanning direction, the multiple overlap of the rectangular shape smoothes the effect of alignment errors mentioned earlier with reference to voxel scanning and prevents the formation of ‘hot’ and ‘cold’ spots.

11.1.3 Theory of mini-voxel scanning

With the rectangular beam distribution, it is easy to sum the dose at a given position as the beam spot steps past. In this case, the dwell time t_D at each mini-voxel will be proportional to required number of particles N_i and will be related to the full dose by,

$$t_D \propto N_i \approx \frac{N_{\text{Desired}}}{n} \tag{11.1}$$

where N_{Desired} is the number of particles that would correspond to the full dose desired at that position. The control system must keep a running total of the doses received at all mini-voxel positions with the goal of reaching the value of N_{Desired} specified by the treatment planner for each position. At any given position, the critical step is the last step the beam makes while covering that position. The decision to move will be based on the measured dose at that position reaching a specified threshold. This is illustrated in Figure 11.2.

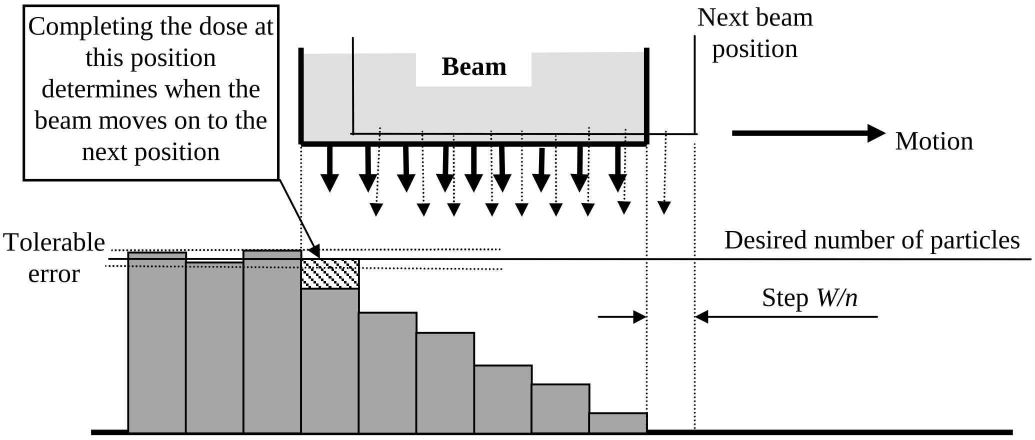

Completing the dose at this position determines when the beam moves on to the next position
Beam
Next beam position
Motion
Tolerable error
Desired number of particles
Step W/n

Figure 11.2 Working principle of the mini-voxel scanning

Figure 11.2 shows the expected situation. The last mini-voxel position at the trailing edge of the spot has approximately N_{Desired}/n deficit in its accumulated dose. Once it is full, the beam will step forward. The discrete measurement intervals of the on-line dosimetry system will cause small irregularities in the doses of the completed mini-voxels. In principle, the mini-voxel positions further downstream, that have yet to be completely irradiated, have larger deficits in their accumulated doses and it is not critical for them when the beam moves. Problems can arise in two situations:

- Firstly, if the dose delivered while the spot is moving exceeds the deficit at the next mini-voxel position (perhaps due to a beam spike), then there is no way to compensate and the scan stops with an out-of-tolerance point.
- Secondly, if the treatment planner asks for a dose reduction between adjacent mini-voxels that is greater than the dose deficit at the next position, then the dose is already exceeded before the beam has completely stepped past that position.

Thus, there is a limit on the rate at which the dose level can be changed that is further reduced by the quality of the beam spill. In theory, it is possible to make quicker changes and, for example, to create a sharp edge to the scan by shrinking the spot size, but this is a considerable operational difficulty. In the following, the analysis will be based on the case of constant spot size during a scan.

In order to evaluate the performance of a mini-voxel scanning system as described above, it is necessary to build a mathematical model that represents the distribution of the beam as closely as possible. The scan can be adequately represented in one dimension (see Figure 11.3). The number of particles delivered to the line element of length W/n is given by integration of the particle flux divided by number of steps during the crossing time. Let t_0 be the time at which the decision is taken to move from the position in which the beam does not touch chosen line element to the first position in which it covers the line element. Let this be the first step. Several steps later at time t_n, the decision will be taken to move from the last position in which the beam covers the line element to the position that is just beyond. Let this be the $(n+1)$th step. Let the power supply delay τ, be the time that the power converters need to start moving the beam. This delay is given by the internal switching frequency of the power supply. The delay of the electronics is small in comparison with the switching time of the power converter and can be considered as included. Finally, assume that the velocity of the movement is constant and it is equal to the maximum scanning velocity v_{max}.

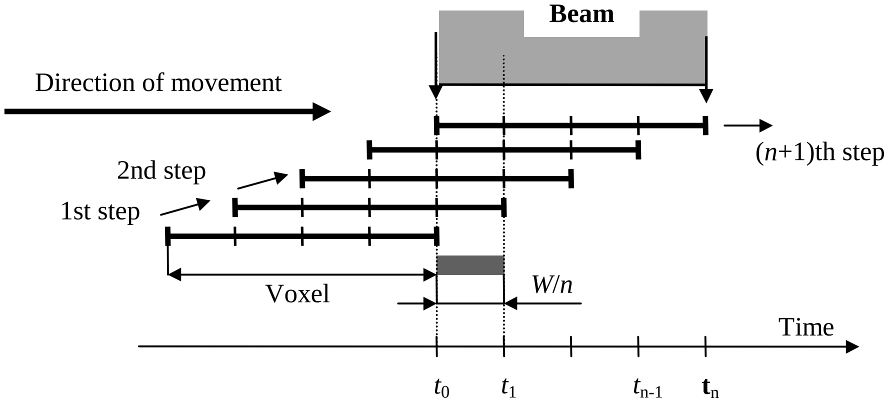


Figure 11.3 Model of mini-voxel scanning (crossing one spot width)

In this scenario, the number of particles, N_L, on the line element of length W/n will depend on the starting time t_0 and contributions from five integrals:

- $N_{\tau 1}(t_0, t_0+\tau)$ is the particle flux during the power converter delay in the first step.
- $N_{M1}(t_0+\tau, t_0+\tau+W/nv_{max})$ is the particle flux during the movement time to the next position.
- $N_C(t_0, t_n)$ is the main integral that quantifies the total particle flux during the time t_0 of the decision to move onto the line element up to the time t_n when the decision is made to finally leave the line element.
- $N_{\tau n}(t_n, t_n+\tau)$ is the particle flux during the power supply delay for the $(n+1)$th step.
- $N_{Mn}(t_n+\tau, t_n+\tau+W/nv_{max})$ is the particle flux during the movement time for the $(n+1)$th step.

These integrals are combined to give the best approximation to the particle flux deposited on the line element,

$$\begin{aligned} N_L(t_0) &= -N_{\tau 1} - \frac{1}{2}\cdot N_{M1} + N_C + N_{\tau n} + \frac{1}{2}\cdot N_{Mn} \\ &= -\int_{t_0}^{t_0+\tau} \frac{\phi(t)}{n}\,dt - \frac{1}{2}\cdot \int_{t_0+\tau}^{t_0+\tau+\frac{W}{n\cdot v_{max}}} \frac{\phi(t)}{n}\,dt + \int_{t_0}^{t_n} \frac{\phi(t)}{n}\,dt + \int_{t_n}^{t_n+\tau} \frac{\phi(t)}{n}\,dt + \frac{1}{2}\cdot \int_{t_n+\tau}^{t_n+\tau+\frac{W}{n\cdot v_{max}}} \frac{\phi(t)}{n}\,dt. \end{aligned} \quad (11.2)$$

Note that during the two movements, half of the integrated flux is attributed to the previous position and half to the next position, which is exact in the case of a uniform spill. Note also that both N_{t1} and N_{M1} are subtracted, since the spot movement has not yet moved and hence these particles enter the previous line element. Similarly, the last two integrals are added, although they occur after the decision to move.

In (11.2), two types of error are shown explicitly. The first is a 'delay' due to the electronics and power supply, $\varepsilon_{\mathrm{Delay}}(t)$ and the second is the 'movement' error that will be called $\varepsilon_{\mathrm{Move}}(t)$. The definitions of these errors are,

$$\varepsilon_{\mathrm{Delay}}(t_0) = \frac{\displaystyle\int_{t_0}^{t_0+\tau} \frac{\phi(t)}{n}\,\mathrm{d}t}{\lambda_{\mathrm{Desired}} \cdot \dfrac{W}{n}} \qquad \varepsilon_{\mathrm{Move}}(t_0) = \frac{\dfrac{1}{2} \cdot \displaystyle\int_{t_0+\tau}^{t_0+\tau+\frac{W}{n\cdot v_{\max}}} \frac{\phi(t)}{n}\,\mathrm{d}t}{\lambda_{\mathrm{Desired}} \cdot \dfrac{W}{n}} \tag{11.3}$$

where $\lambda_{\mathrm{Desired}}$ is the desired line density [particle/m]. These errors appear with negative signs at the beginning (step 1) and with positive signs at the end (step n+1). In the ideal case of a uniform spill and a uniform scanning speed, these errors would cancel.

In a practical system, there will be an additional error that is caused by the 'quantisation' of the times t_0 and t_{n}. These times must be 'locked' to a clock that defines the measurement time interval, T_{bin} and the limits of the main integral will have to occur at integral numbers of T_{bin} at T_0 and T_{n}. There is no problem with t_0 and T_0 since this is a choice of origin, but in general t_{n} will be less than T_{n}, because the system will only learn that the desired dose has been achieved on the next clock cycle. An error for this 'clock' effect, similar to (11.3), can be defined as,

$$\varepsilon_{\mathrm{Clock}}(t_{\mathrm{n}}, T_{\mathrm{n}}) = \frac{\displaystyle\int_{t_{\mathrm{n}}}^{T_{\mathrm{n}}} \frac{\phi(t)}{n}\,\mathrm{d}t}{\lambda_{\mathrm{Desired}} \cdot \dfrac{W}{n}}\,. \tag{11.4}$$

A first indication of how a practical a scanning system would perform can be obtained by evaluating the influence of sinusoidal fluctuations in the spill. The fractional errors arising from the 'delay' and the 'movement' and the 'clock' for a given element of line of length W/n can be expressed as,

$$E(T_n \,|\, T_0) = \frac{1}{\phi_0 \cdot t_{\mathrm{n}}} \cdot \left(-\int_{T_0}^{T_0+\tau} \frac{\phi(t)}{n}\,\mathrm{d}t - \frac{1}{2} \cdot \int_{T_0+\tau}^{T_0+\tau+\frac{W}{n\cdot v_{\max}}} \frac{\phi(t)}{n}\,\mathrm{d}t + \int_{t_{\mathrm{n}}}^{T_n} \frac{\phi(t)}{n}\,\mathrm{d}t + \int_{T_{\mathrm{n}}}^{T_{\mathrm{n}}+\tau} \frac{\phi(t)}{n}\,\mathrm{d}t + \frac{1}{2} \cdot \int_{T_{\mathrm{n}}+\tau}^{T_{\mathrm{n}}+\tau+\frac{W}{n\cdot v_{\max}}} \frac{\phi(t)}{n}\,\mathrm{d}t \right). \tag{11.5}$$

Note that the 'delay' and 'movement' errors at the end of the irradiation have been modified to take account of the 'clock' time. If the particle flux is of the form $\phi(t) = \phi_0(1+A_{\mathrm{r}}\cos(\omega t))$, where A_{r} is the ripple amplitude, ϕ_0 is the nominal flux of particles and ω is ripple angular frequency, then

$$E(T_{\mathrm{n}} \mid T_0) = \frac{1}{2nt_{\mathrm{n}}\omega} \cdot [-2t_{\mathrm{n}}\omega + 2T_{\mathrm{n}}\omega + 2A_{\mathrm{r}}n\sin(t_0\omega) - 2A_{\mathrm{r}}\sin(t_{\mathrm{n}}\omega)$$

$$- A_{\mathrm{r}}\sin[\omega(t_0+\tau)] + A_{\mathrm{r}}\sin[\omega(T_{\mathrm{n}}+\tau)] \qquad (11.6)$$

$$- A_{\mathrm{r}}\sin\left[\omega\left(t_0 + \frac{W}{nv_{\max}} + \tau\right)\right] + A_{\mathrm{r}}\sin\left[\omega\left(t_{\mathrm{n}} + \frac{W}{nv_{\max}} + \tau\right)\right]].$$

The error can vary widely according to whether the delay is equal to the ripple period and so on, but for a given system, the variation depends only on the starting time. Figures 11.4, 11.5 and 11.6 show the general forms of these errors and their relative importance. The first graph shows the ‘clock’ error, the second shows the ‘delay’ and ‘movement’ errors combined, since these are identical in their effect and finally the third graph shows the behaviour of the total error. All graphs have been plotted against the starting time for the following conditions:

- Sinusoidal fluctuation in the spill of the form $\phi(t) = \phi_0(1+A_{\mathrm{r}}\cos(\omega t))$,
- Amplitude ripple modulation, $A_{\mathrm{r}} = 0.1$,
- Nominal particle flux, $\phi_0 = 2 \times 10^8$ particle/s,
- Ripple frequency $f = 100$ Hz,
- Sampling time bin, $T_{\mathrm{bin}} = 100$ µs,
- Delay time in electronics and power converters, $\tau = 100$ µs,
- Maximum scanning velocity, $v_{\max} = 10$ m/s,
- Beam spot size, $W = 10$ mm
- Number of steps to cross beam spot, $n = 10$.

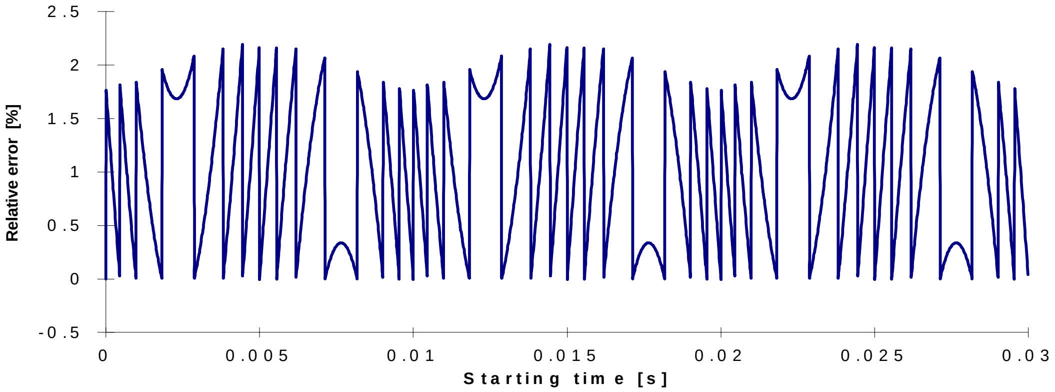


Figure 11.4 Dependence of the ‘clock’ error on the starting time

In Figure 11.4, only the ‘clock’ measurement error is shown. This is the principal error in voxel scanning, since the beam is switched off during the movement between voxels. For Figure 11.4, the discrimination level for the decision to move (or to switch off the beam in voxel scanning) was set to 100%. Thus the minimum measurement error is equal to 0%. The plot shows that the error just exceeds the specified limit of +2% for certain values of the starting time. This can be solved by including an off-set and setting the discrimination level for switching to somewhere between 98 and 99.5% of the desired dose.

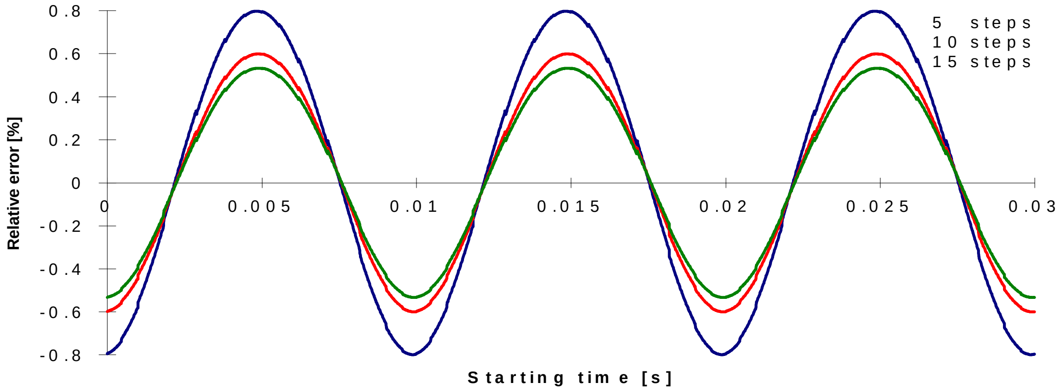


Figure 11.5 Dependence of the combined delay and movement errors on the starting time for 5, 10, 15 mini-voxels

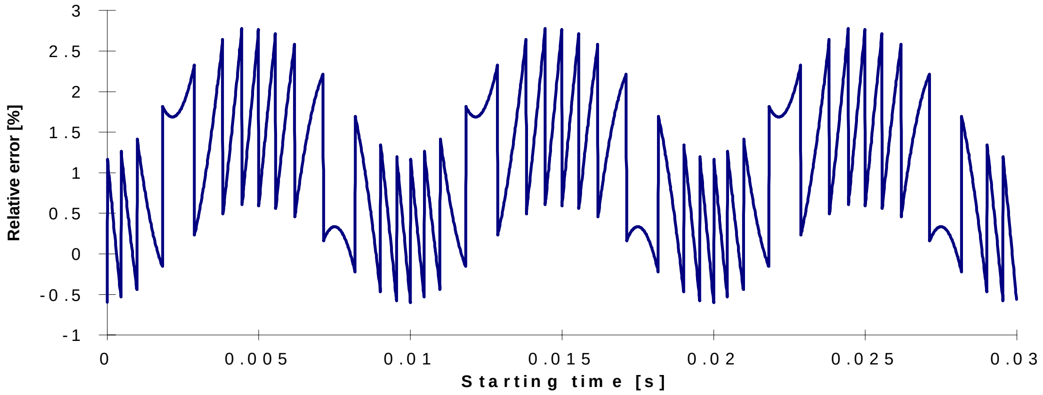


Figure 11.6 Dependence of the total error on the starting time

The total error shown in Figure 11.6 exhibits a total excursion of just over 3%. As with the ‘clock’ error in Figure 11.4, this can be brought within the tolerances by biasing the discrimination level down to say 98%.

11.1.4 Ripple specification for mini-voxel scanning

It is important to be able to use the analysis of the mini-voxel scanning to derive a specification for the maximum permissible ripple that the system can accept in the beam spill. This specification can then be used to determine the corresponding specifications for the power converters in the main ring. One possible approach is to form an inequality based on the expression (11.6) and to solve

$$E(T_\mathrm{n} \mid T_0) \le 0.04 \tag{11.7}$$

where 0.04 is the maximum error ±2 %. The main inconvenience is that before the equation can be solved, the time t_0 for which the error is maximum has to be found for each frequency, or an approximated maximum error has to be derived.

For a particle flux of the form $\phi(t) = \phi_0(1 + A_\mathrm{r}\cos(\omega t))$ the maximum ‘clock’ error can be found analytically. The maximum number of particles in one time bin will be obtained at the peak of the particle flux i.e. for time $t = 0$ in the case of the cosine function. Integration of the particle flux over interval -T$_{\text{bin}}$/2 to T_{bin}/2 gives the maximum 'clock' error.

$$\hat{\varepsilon}_{\text{Clock}} = \frac{\displaystyle\int_{T_{\text{bin}}/2}^{-T_{\text{bin}}/2} \frac{\phi(t)}{n}\,\mathrm{d}t}{\lambda_{\text{Desired}}\,\dfrac{W}{n}} = \frac{1}{\lambda_{\text{Desired}}\,W}\cdot\left[T_{\text{bin}}\ \phi_0 + \frac{2A_{\text{r}}\phi_0}{\omega}\sin\!\left(\frac{T_{\text{bin}}}{2}\,\omega\right)\right]. \tag{11.8}$$

An example of the variation of the maximum 'clock' error with ripple frequency is shown in Figure 11.7. The spill ripple is assumed to be a sinusoidal fluctuation of the form $\phi(t) = \phi_0(1 + A_{\text{r}}\cos(\omega t))$ with the amplitude modulation $A_{\text{r}} = 1$, the nominal particle flux $\phi_0 = 2 \times 10^8$ particle/s, the sampling time $T_{\text{bin}} = 100\ \mu\text{s}$, the electronic and power supply delay $\tau = 100\ \mu\text{s}$, the maximum scanning velocity $v_{\max} = 10$ m/s and the number of steps $n = 10$.

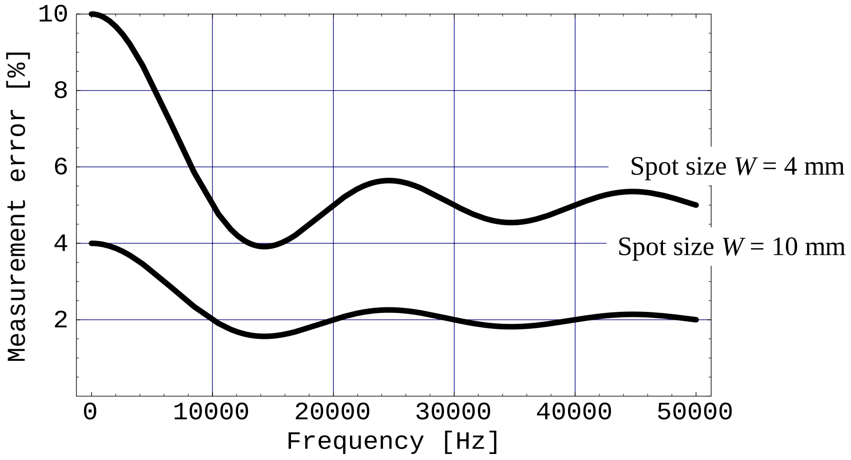


Figure 11.7 Variation of the maximum 'clock' error with the ripple frequency

The maximum 'clock' error is bigger for the smaller spot size because the same particle flux was assumed in both cases. The dwell time on each mini step is then shorter and this causes a bigger error when the sampling time, T_{bin}, is kept constant. Thus, the 'clock' error is proportional to the spot size and decreases with frequency in the region 1 to 10 kHz. For higher frequencies, it oscillates around a non-zero value that depends on the spot size.

The movement error is maximum and positive when the ripple is at its minimum value during the initial movement at T_0 and at its maximum during the final movement at T_{n}, see Figure 11.8. The maximum error is then given by the shaded area as,

$$E_{\max} = \frac{1}{\lambda_{\text{Desired}} W}\left|\frac{1}{\pi f}[2A_r\phi_0 \sin \pi f\left(\frac{W}{v_{\max} n} + \tau\right)\right| \tag{11.9}$$

where $\omega = 2\pi f$. The delay error has a similar form to the movement error and can be included in the same expression.

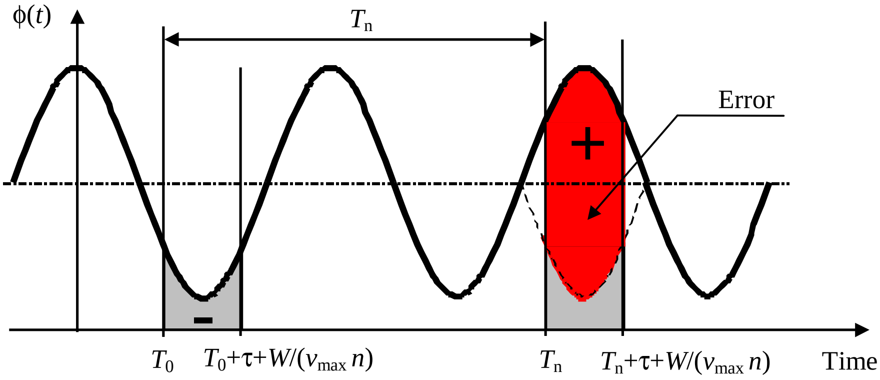


Figure 11.8 Schematic view of the maximum movement error

The variation with ripple frequency of the combined maximum delay and movement error is shown in Figure 11.9 for an amplitude modulation $A_r = 1$, the nominal particle flux $\phi_0 = 2 \times 10^8$ particle/s, the sampling time $T_{bin} = 100\ \mu s$, the electronic and power supply delay $\tau = 100\ \mu s$, the maximum scanning velocity $v_{max} = 10$ m/s and the number of steps $n = 10$.

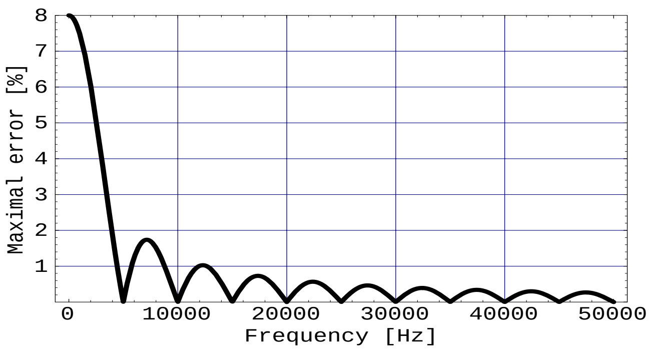


Figure 11.9 Variation of the combined maximum delay and movement error with the ripple frequency

The combined maximum delay and movement error is largest at low frequency and tends to zero at high frequencies. This behaviour does not correspond exactly to the original model. The expression (11.9) is fully valid for high frequencies or low flux intensities where the time of integration t_n is longer than the half period of the ripple. Under the conditions mentioned above ($\phi_0 = 2 \times 10^8$ particle/s), the average value for t_n is 5 ms, which corresponds to a frequency of 200 Hz i.e. that for frequencies lower than 100 Hz the expression (11.9) is not valid because the time t_n is shorter than time between minimum and maximum of the ripple oscillation and the maximum error can not be reached. The correct 'low-frequency' formula has the form,

$$\hat{E}_{Low} = \frac{1}{\lambda_{Desired} W} \frac{A_r \phi_0}{2\pi f}\left[-\sin(2\pi f T_n) + \sin\left[\pi f\left(\frac{W}{v_{max} n} + \tau\right)\right] + \sin\left[2\pi f\left(T_n + \frac{W}{v_{max} n} + \tau\right)\right]\right]. \quad (11.10)$$

The low-frequency dependence of the combined maximum delay and movement error under the chosen conditions (i.e. $T_n \approx 50\ T_{bin} = 5$ ms) is shown in Figure 11.10.

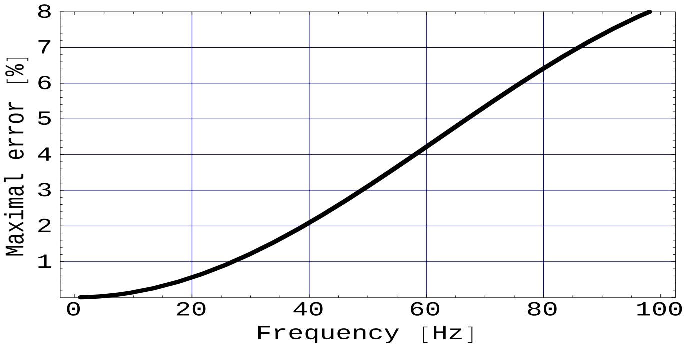


Figure 11.10 Variation of the combined maximum delay and movement error with ripple frequency computed according to the ‘low-frequency’ formula (11.10)

Combining the ‘high’ and ‘low’ frequency formulæ (11.9) and (11.10) gives an accurate picture of the variation of the combined maximum delay and movement error with the ripple frequency. This is shown on a log scale in the Figure 11.11.

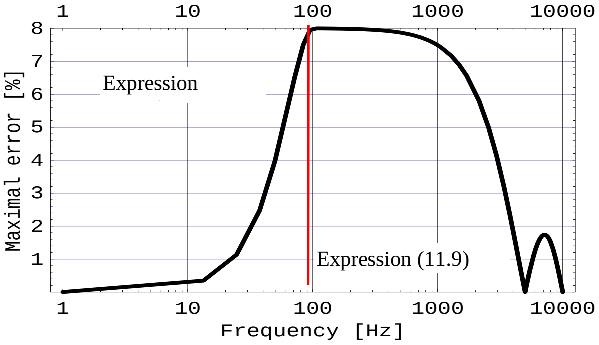


Figure 11.11 The combined maximum delay and movement errors versus ripple frequency

Now that the maximum error functions have been established for the ‘clock’ and the combined delay and movement errors, tolerances can be determined for the permissible ripple over the full frequency range by solving (11.8), (11.9) and (11.10), taking care to apply the low-frequency and high-frequency formulæ. This result is shown in Figure 11.12.

The maximum permissible ripple modulation depends strongly on the spot size. The requirements on the medical machine are to change spot size continuously between 4 and 10 mm (FWHH). Figure 11.12 shows that the maximum permissible modulation in the spill diminishes as the spot size is reduced. For the minimum spot size of 4 mm the equations do not have reasonable solutions, which means that for the conditions chosen the dose precision of ± 2% cannot be reached for the 4 mm spot. This is caused in particular by the measurement error that is greater than 4% at all frequencies. In other words, the dwell time at one mini-step position is too short in comparison with sampling time. The possible solutions are to:

- Decrease particle flux (to increase the irradiation time)
- Shorten the sampling time (to reduce the 'clock' error)
- Reduce the number of mini-steps (to increase the dwell time).

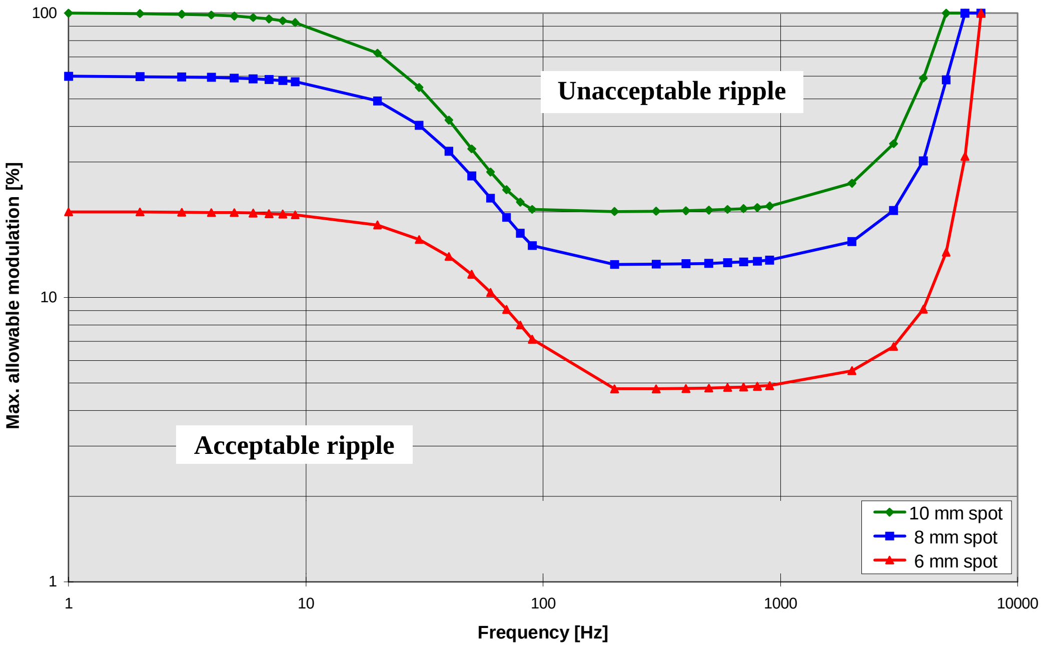


Figure 11.12 Maximum permissible ripple amplitude with ripple frequency to obtain ±2% dose uniformity

[Conditions: Sinusoidal fluctuation in the spill of the form $\phi(t) = \phi_0(1+A_r\cos(\omega t))$, nominal particle flux $\phi_0 = 2 \times 10^8$ particle/s, sampling time $T_{\mathrm{bin}} = 100$ µs, electronic and power supply delay $\tau = 100$ µs, maximum scanning velocity $v_{\max} = 10$ m/s and number of steps $n = 10$]

In all cases, the aim is to increase the ratio of the dwell time to the sampling time to reach a total measurement error of less than 4%.

The plot in Figure 11.12 can be conveniently divided into three regions in the frequency domain:

- 1 to 10 Hz where the maximum permissible ripple amplitude depends mainly on the 'clock' error and is relatively high.
- 10 to 1000 Hz where the maximum permissible ripple amplitude declines to a minimum because of the growing importance of the delay and movement errors.
- 1 to 10 KHz where the maximum permissible amplitude tends to 100% as all errors, except the 'clock' error, tend to zero.

11.2 RASTER SCANNING

In true raster scanning, the beam is moved continuously with a scanning velocity that is controlled by a feed-forward loop using the measured intensity of the incoming particle beam. For higher intensities than nominal, the scanning velocity is increased and, for lower intensities, it is decreased to obtain the desired dose. Three levels of approximation will be considered in the theory present in this section.

11.2.1 Instantaneous measurement and velocity changes (1st approximation)

If a ‘point’ beam is moving continuously along a line, then the line density of the particles, $\lambda(s)$, deposited by the beam will be,

$$\lambda(s) = \frac{\mathrm{d}N}{\mathrm{d}s} = \frac{\delta N}{\delta t}\,\frac{\delta t}{\delta s} = \phi(t)\cdot\frac{1}{v_{\mathrm{scan}}(t)} \tag{11.11}$$

where N is the number of particles, δt and δs correspond to small steps in time and distance, $\phi(t)$ is the particle flux and $v_{\mathrm{scan}}(t)$ is the speed of the moving beam spot. This is illustrated in Figure 11.13.

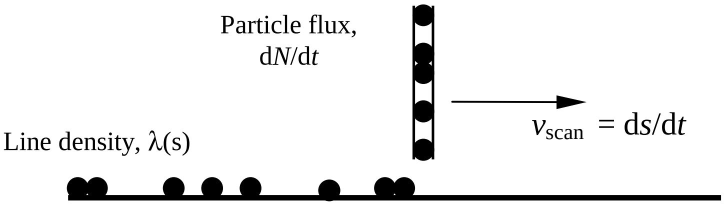


Figure 11.13 Particle distribution by a ‘point’ source

Under ideal conditions these parameters would have exactly their nominal time independent values and the desired dose would be achieved by adjusting only the scanning velocity by a dose factor $\Re$,

$$\lambda(s)_{\mathrm{Desired}} = \Re\,\phi_0\,\frac{1}{v_0}, \quad \text{so that} \quad v_{\mathrm{scan}} = \frac{1}{\Re}v_0. \tag{11.12}$$

Under practical conditions, the particle flux ϕ varies with the time, but for sufficiently slow variations, a reasonable result should be obtained by adjusting the scan speed according to the particle flux at some slightly earlier time (t-τ),

$$v_{\mathrm{scan}}(\mathrm{t}) = \underbrace{\frac{1}{\Re}v_0}_{\substack{\text{Set by}\\ \text{treatment}\\ \text{planning}}} \times \underbrace{\frac{\phi(t-\tau)}{\phi_0}}_{\substack{\text{Correction for}\\ \text{spill}\\ \text{fluctuations}}} \tag{11.13}$$

where τ is the lag before the velocity correction (assumed in this model as being instantaneous) is made and includes the delays needed for the electronics, the computation and the power supplies. Thus each measurement bin is compared to the reference flux level ϕ_0. This is the simplest way of correcting the spill fluctuations. The substitution of (11.13) into (11.11) gives the actual line density deposited and if this is divided by (11.12), the relative error is obtained,

$$\frac{\lambda}{\lambda_{\mathrm{Desired}}}(t) = \phi(t)\,\frac{1}{\phi(t-\tau)}. \tag{11.14}$$

In principle, the shorter the delay between measurement and correction and the more frequently the measurements are made, the better the compensation of irregularities in the spill intensity $\phi(t)$ becomes. This type of correction is known as a *simple compensation* and is limited to frequencies much lower than $1/\tau$.

11.2.2 Instantaneous response but finite measuring times (2nd approximation)

In this model, the quantisation and finite length of the measurement time is taken into account. This filters out the frequencies above the sampling frequency.

The flux will be measured in bins T_{bin}, so equations (11.13) and (11.14) should be more accurately written as,

Scan velocity:

$$v_{\text{scan}}(t) = \frac{1}{\Re} v_0 \quad \frac{1}{\phi_0 T_{\text{bin}}} \int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t')\,\mathrm{d}t' \tag{11.15}$$

Fractional error with zero spot

$$\frac{\lambda}{\lambda_{\text{Desired}}}(t) = \phi(t) \frac{T_{\text{bin}}}{\int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t')\,\mathrm{d}t'}. \tag{11.16}$$

This situation is illustrated in Figure 11.14, where at time t the velocity is set according to the flux measured at a time τ earlier using (11.15) and at time $t+T_{\text{bin}}$ the velocity is set according to flux at a time τ-T_{bin} earlier. The error at the time t can be calculated from (11.16). The time lag includes the time to measure one bin, to read and treat the data and to change the power converters.

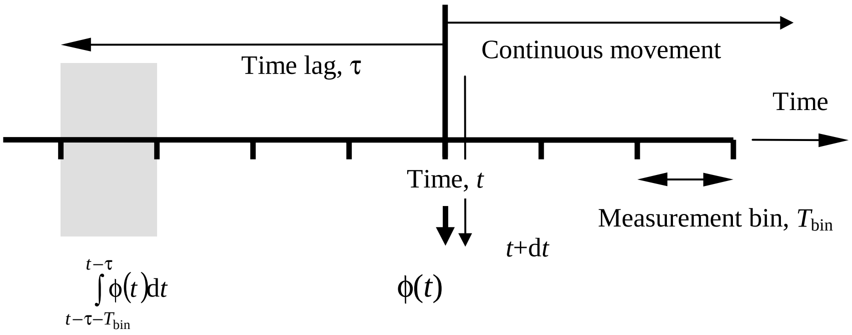
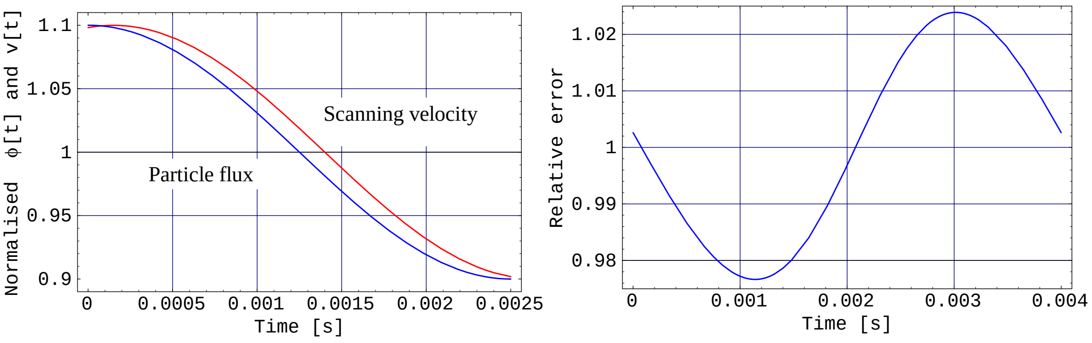

Time lag, τ
Continuous movement
Time
Time, t
Measurement bin, T_{bin}
t+dt
$\int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t)\mathrm{d}t$
$\phi(t)$

Figure 11.14 Measurement of particle flux

According to equation (11.15), the velocity follows the particle flux with time lag τ and then the error oscillates according to (11.16). This is illustrated in the Figure 11.15.



(a) (b)

Figure 11.15 (a) Dependence of the normalised particle flux and normalised scanning velocity on the time, (b) Dependence of the relative error of the dose on the time

Equations (11.12) to (11.16) all describe a scanning system that has zero spot size. The inclusion of a finite spot size helps to smooth the fluctuations, providing that the time needed to traverse one spot width is larger than the delay, τ, and larger than the ripple period. The scanning velocity remains unchanged, but the error equation is modified slightly. Consider a point on the track of the spot. This point will 'see' the beam for the time that the spot takes to traverse its own width, see Figure 11.16.

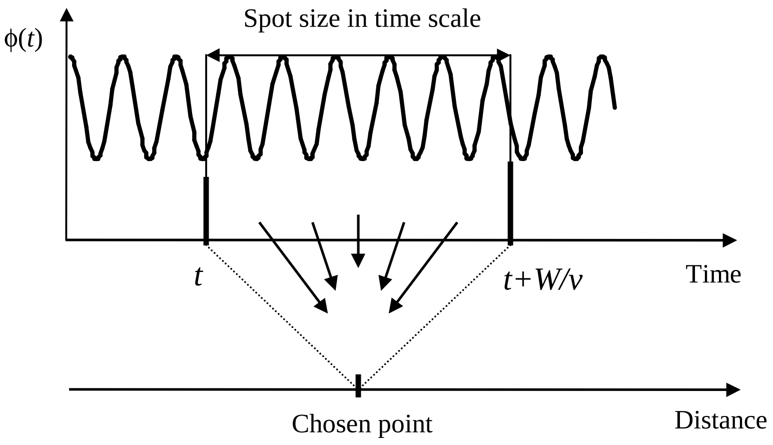


Figure 11.16 Schematic view of the spot smoothing

If the spot width is W, then,

Fractional error with finite spot size

$$\left(\frac{\lambda}{\lambda_{\text{Desired}}}\right)_{\text{average}} = \frac{v_0}{W}\int_{t-\frac{W}{v_0}}^{t} \frac{\phi(t)\cdot T_{\text{bin}}}{\int_{t-\tau-T_{\text{bin}}}^{t-\tau}\phi(t')\,\mathrm{d}t'}\,\mathrm{d}t \qquad (11.17)$$

In order to simplify (11.17), the lower limit of the time span for the integration has been set to W/v_0. Providing v_{scan} does not vary widely this should be a reasonable approximation. Consider a sinusoidal ripple fluctuation in the spill of the form $\phi(t) = \phi_0(1+A_r \cos(\omega t))$, then the expression for the scan velocity according to (11.15) is

$$v_{\mathrm{scan}}(t) = \frac{v_0}{T_{\mathrm{bin}}\,\omega}\left(T_{\mathrm{bin}}\,\omega + A_{\mathrm{r}}\sin[(t-\tau)\omega] - A_{\mathrm{r}}\sin[(t-T_{\mathrm{bin}}-\tau)\cdot\omega]\right) \qquad (11.18)$$

and the fractional error with zero spot size according (11.16)

$$\frac{\lambda}{\lambda_{\mathrm{Desired}}}(t) = \frac{T_{\mathrm{bin}}\,\omega[1 + A_{\mathrm{r}}\cos(t\cdot\omega)]}{T_{\mathrm{bin}}\,\omega + A_{\mathrm{r}}\sin[(t-\tau)\omega] - A_{\mathrm{r}}\sin[(t-T_{\mathrm{bin}}-\tau)\cdot\omega]} \qquad (11.19)$$

and fractional error with a finite spot size *W* according to (11.17) is written below with the use of (11.19).

$$\left(\frac{\lambda}{\lambda_{\mathrm{Desired}}}\right)_{\mathrm{average}} = \frac{v_0}{W}\int_{t-\frac{W}{v_0}}^{t} \frac{T_{\mathrm{bin}}\,\omega[1 + A_{\mathrm{r}}\cos(t\cdot\omega)]}{T_{\mathrm{bin}}\,\omega + A_{\mathrm{r}}\sin[(t-\tau)\omega] - A_{\mathrm{r}}\sin[(t-T_{\mathrm{bin}}-\tau)\cdot\omega]}\,\mathrm{d}t\,. \qquad (11.20)$$

An analytic solution to (11.20) can be obtained with a symbolic algebra program, but the result is too complicated to be useful. The analytic solution depends on the parameters of the system (sampling time T_{bin}, delay τ and nominal velocity v_0), chosen parameters (time t_0 and spot size *W*) and ripple parameters (modulation A_{r}, frequency *f* and $\omega = 2\pi f$) as expected.

As for the mini-voxel scanning system, the ripple frequency and the chosen spot size determine the error of the dose. The dependence of this error on the frequency and spot size for modulation $A_{\mathrm{r}} = 0.2$, sampling time $T_{\mathrm{bin}} = 100$ µs, delay $\tau = 100$ µs, nominal velocity $v_0 = 2$ m/s and starting time $t_0 = 0$ is shown in the Figure 11.17.

The error is maximum for the smallest spot size and for the frequency between 1 and 3 kHz. The exact value of the frequency for which the error is maximum varies with the starting time t_0 and level of ripple modulation. For the low frequencies up to 100 Hz the compensation by controlling the velocity is effective. At frequencies above 100 Hz, the system becomes too slow to follow fast changes in the spill. For frequencies above 1 kHz, the spot size starts to smooth ripple and this becomes effective above 4 kHz ripple. The spot size smoothing plays a significant role in the raster scanning process. The critical frequencies are between 1 and 3 kHz, where the compensation is too slow and the number of oscillations in the spot is too small to compensate fluctuations by the spot-size smoothing.

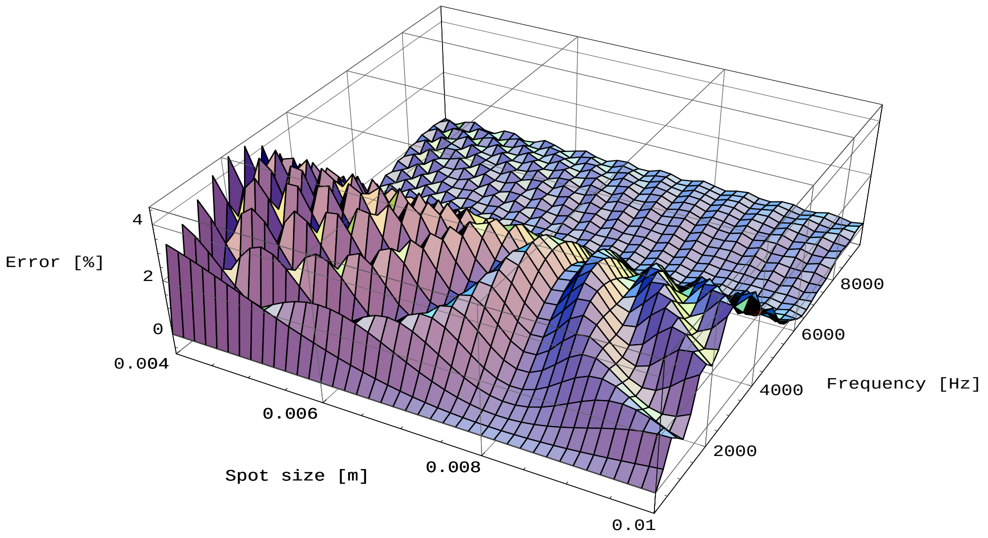


Figure 11.17 Dependence of the fractional error on the spot size and frequency

The errors in Figure 11.18 for a given t_0, but different spot sizes, have similar variations with frequency. The system over-doses in the low to medium frequency range up to 5.8 kHz, under-doses for higher frequencies up to 8 kHz and finally for still higher frequencies the error tends to zero. The phenomenon of the overdosing and under-dosing is caused by the velocity compensation. The major difference between the curve for a 4 mm spot and a 10 mm spot is the amplitude of the residual error oscillations that is naturally smaller in the case of 10 mm spot because of the better spot smoothing.

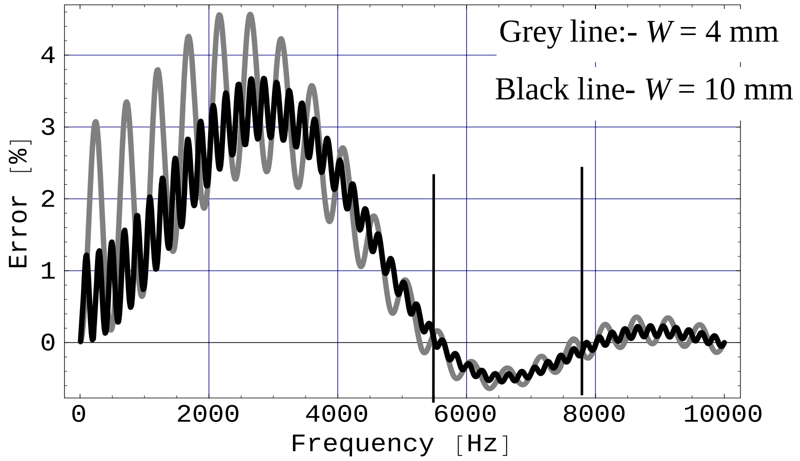


Figure 11.18 Dependence of the fractional error on ripple frequency

11.2.3 Ripple specification for raster scanning (2^{nd} approximation)

The first stage is to determine the maximum error at each frequency and then to convert this to a specification for the maximum permissible ripple over the frequency range. The maximum error for a given frequency depends on the starting time t_0, which defines the phase advance (the dependence of the error on the t_0 is periodic). The situation for modulation $A_r = 0.5$, the sampling time $T_{bin} = 100$ μs, the delay $\tau = 100$ μs, the nominal velocity $v_0 = 2$ m/s, the frequency 150 Hz and the spot size $W = 10$ mm is illustrated in Figure 11.19.

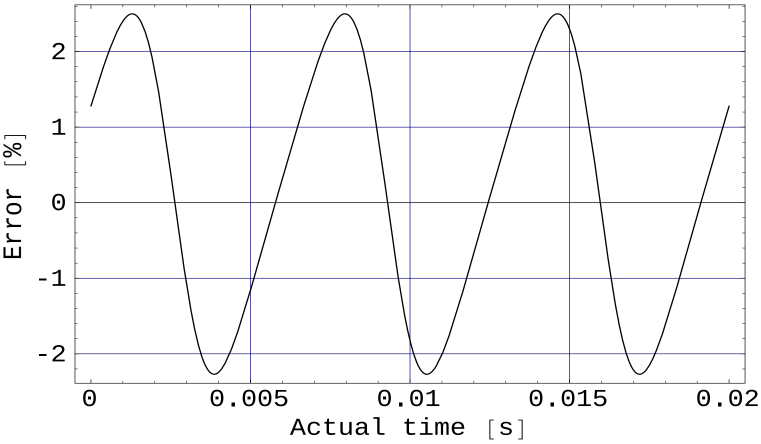


Figure 11.19 Dependence of the fractional error on the starting time t_0

The error curve is slightly off-set above zero because the frequency of 150 Hz is located in the over-dosed region. The maximum fractional error also depend on the ripple amplitude (see Figure 11.20) which means that for each frequency and each ripple amplitude the maximum error has to be found.

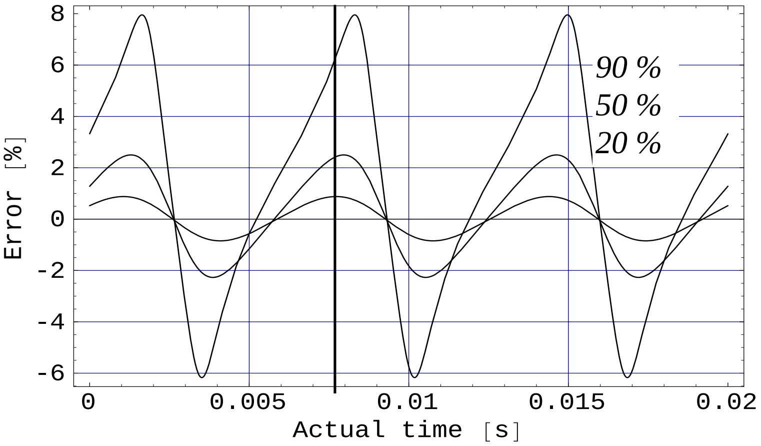


Figure 11.20 Dependence of the fractional error on the starting time t_0 for different amplitude modulations

The maximum permissible ripple amplitude is given by solving inequalities that follow from (11.20).

$$\left|\left(\frac{\lambda}{\lambda_{\text{Desired}}}\right)_{\max}\right| \le 0.02 \quad \text{and} \quad \left|\left(\frac{\lambda}{\lambda_{\text{Desired}}}\right)_{\min}\right| \le 0.02 \tag{11.21}$$

Numeric solution of the inequalities (11.21) for the conditions of sinusoidal fluctuations in the spill, the sampling time $T_{bin} = 100$ μs, the electronic and power supply delay $\tau = 100$ μs, the maximum scanning velocity $v_{max} = 2$ m/s and spot sizes of 4 to 10 mm is shown in Figure 11.21 and the detailed zoom in Figure 11.22.

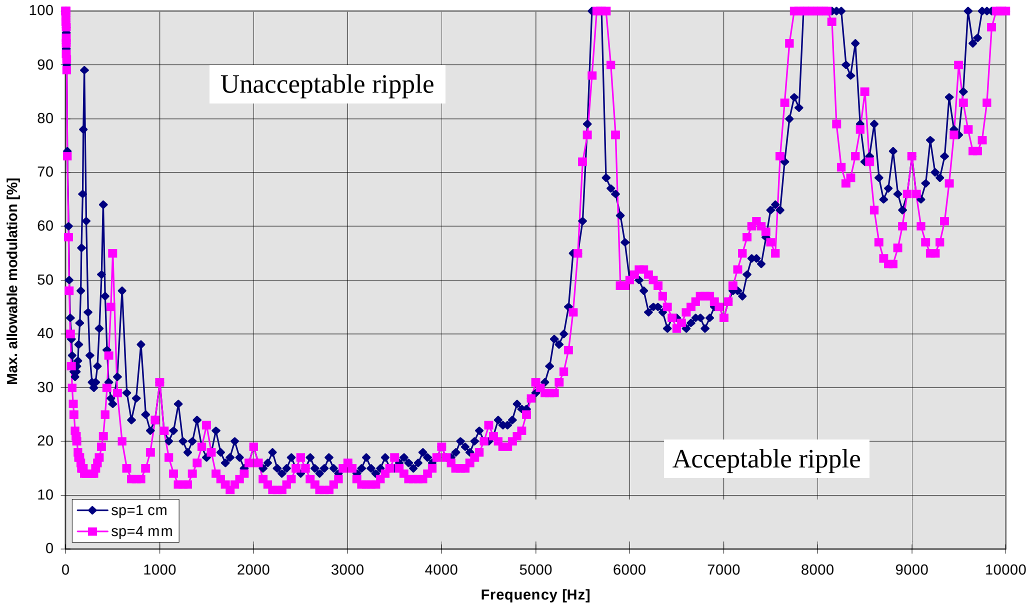


Figure 11.21 Dependence of the maximum permissible ripple amplitude against frequency

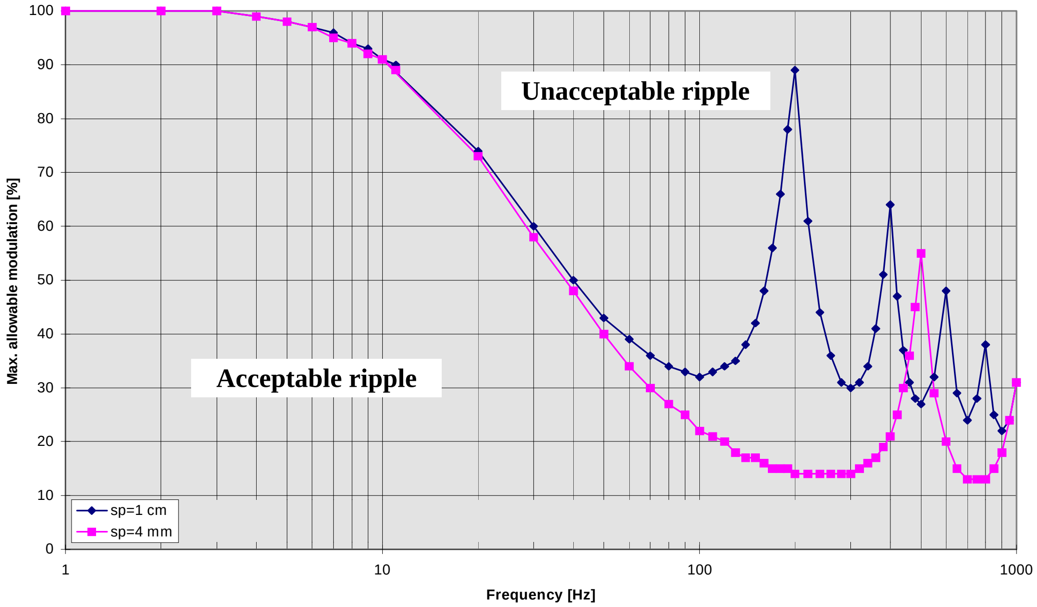


Figure 11.22 Dependence of the maximum permissible ripple amplitude against frequency in the range1to1000 Hz.

Figures 11.21 and 11.22 show that for low frequencies between 1 and 10 Hz the velocity compensation is very effective over the full range of spot sizes and modulation between 90 and 100% is permissible. With increasing frequency, the permissible modulation drops because of declining efficiency of the velocity compensation. At low ripple frequencies the spot smoothing is negligible.

The curve for the 10 mm spot has minima at odd multiples of 100 Hz and maxima at even multiples 100 Hz, whereas the curve for the 4 mm spot that has minima and maxima shifted by 50 Hz. The maxima correspond to the situation where the spot size contains an integral number of ripple periods (minimum error) assuming the average scanning velocity equals 2 m/s. The number of wavelengths integrated by the spot increases and the difference between maxima and minima decreases.

For the practical interpretation of the results, the bottom envelope of each curve has to be taken. Figures 11.21 and 11.22 show that the frequency interval for which the ripple amplitude is small (between 10 and 20%) is very wide (0.12 to 4.75 kHz for the 4 mm spot and 1.1 to 4.5 kHz for the 10 mm spot). Both curves permit 100% ripple modulation at 5.8 kHz, 8 kHz and 10 kHz where the system passes from the over-dosing to under-dosing and vice versa.

11.2.4 Finite times for response and measurement (3rd approximation)

The final level of approximation simulates the real system in which the system gives the measured flux every T_{bin} and the scanning velocity is changed in steps corresponding to T_{bin}. Once the velocity is set, the value is kept constant over the next measurement bin, when a new value of the velocity can be computed. A schematic model of the relation between the flux measurement and the velocity control is shown in Figure 11.23. The only disharmony between the model and the real world is the instantaneous change of velocity from one value to the next.

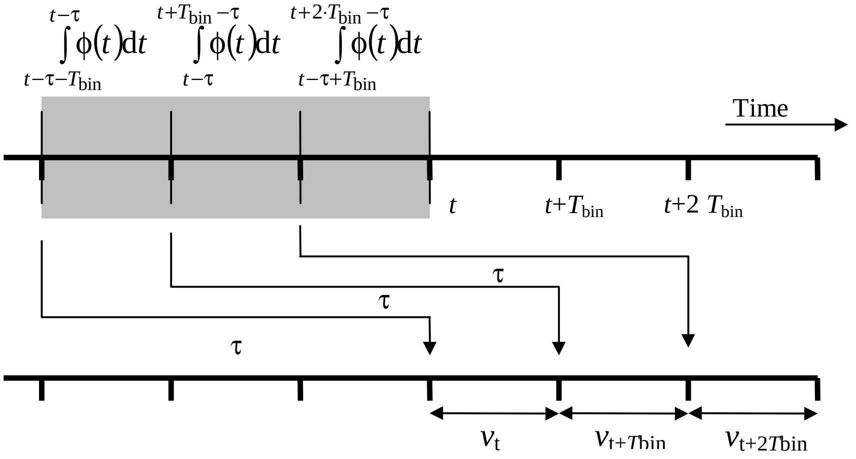

$\int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t)\mathrm{d}t$
$\int_{t-\tau}^{t+T_{\text{bin}}-\tau} \phi(t)\mathrm{d}t$
$\int_{t-\tau+T_{\text{bin}}}^{t+2\cdot T_{\text{bin}}-\tau} \phi(t)\mathrm{d}t$
Time
t
$t+T_{\text{bin}}$
$t+2\ T_{\text{bin}}$
τ
τ
τ
v_t
$v_{t+T\text{bin}}$
$v_{t+2T\text{bin}}$

Figure 11.23 Measurement of particle flux and control of the velocity

The expression for the scanning velocity (11.15) has then to be adjusted by unit step function Θ to the form

$$v_{\text{scan}}\left[t, t+T_{\text{bin}}\right] = \frac{1}{\Re} v_0 \, \frac{1}{\phi_0 T_{\text{bin}}} \, \left[\Theta(t) - \Theta(t+T_{\text{bin}})\right] \int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t')\mathrm{d}t' . \qquad (11.22)$$

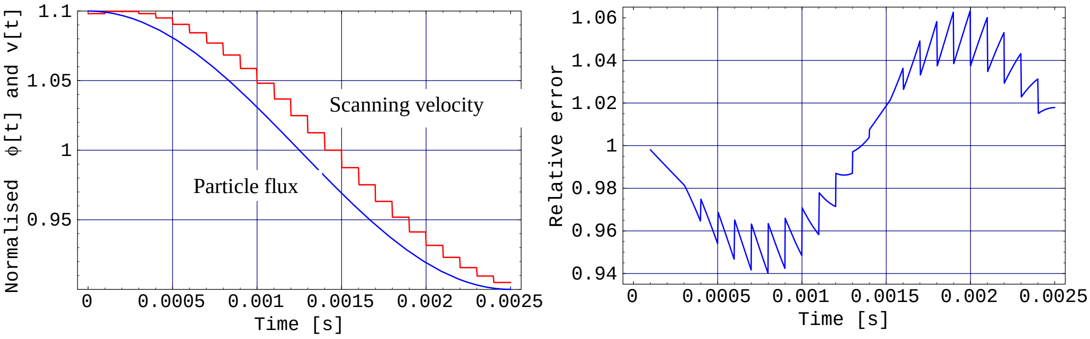
Using expression (11.22) in equation (11.16) for fractional error with zero spot size of previous model one obtains

$$\frac{\lambda}{\lambda_{\text{Desired}}}(t) = \phi(t) \cdot \frac{T_{\text{bin}}}{\left[\Theta(t) - \Theta(t+T_{\text{bin}})\right] \cdot \int_{t-\tau-T_{\text{bin}}}^{t-\tau} \phi(t')\mathrm{d}t'} . \qquad (11.23)$$

and for the fractional error with finite spot size W

$$\left(\frac{\lambda}{\lambda_{\mathrm{Desired}}}\right)_{\mathrm{average}} = \frac{v_0}{\Lambda\Re}\int\limits_{t-\frac{W}{v_0}}^{t} \frac{\phi(t')\mathrm{d}t'}{\left[\Theta(t)-\Theta(t+T_{\mathrm{bin}})\right]\cdot \int\limits_{t-\tau-T_{\mathrm{bin}}}^{t-\tau}\phi(t')\mathrm{d}t'}\mathrm{d}t\,. \qquad (11.24)$$

The scanning velocity changing in steps controlled by particle flux according to (11.22) and the fractional error for zero spot size computed according to (11.23) are shown on Figure 11.24.



a) b)

Figure 11.24 (a) Dependence of the normalised particle flux and normalised scanning velocity on the time, (b) Dependence of the relative error of the dose on the time

The Figure 11.24 shows that the improved model has caused a deterioration in the efficiency of the system. For example, at the frequency of 200 Hz, only 6% of ripple is acceptable compared to 10% in the previous model shown in Figure 11.14. Furthermore, the ability of the system to correct fluctuations in the spill drops with the increasing frequency of the ripple. For high frequencies, the maximum error with a zero spot size can be even higher than the ripple modulation. In practice, this situation will not occur because the finite spot size will smooth out this effect.

The behaviour of a raster scanning system with a simple velocity compensation can be summed up as:

- At low ripple frequencies, the velocity compensation is effective for smoothing.
- At high frequencies, the finite spot size is effective for smoothing,
- But, at high frequencies, the velocity compensation becomes counter-productive and deteriorates the otherwise near-perfect smoothing from the finite spot size.

This is the reason in Figures 11.21 and 11.25 for the second dip after 5 kHz. The overall performance of the raster scanning system with simple velocity compensation is shown in Figures 11.25 and 11.26. The dependence of the maximum allowable ripple amplitude is similar to that obtained with the previous model (see Figures 11.21 and 11.22), except that the general performance is less good. The differences are due to the velocity control regime.

11.3 SUMMARY

Two dynamic active scanning methods have been investigated for comparison with the static voxel scanning method. The initial conditions and requirements used were those of the PIMMS machine. To simplify the problem only sinusoidal ripple was assumed and only single frequencies were assumed to be active in the spill at any one time. This approach should be sufficient for comparing the methods, but the specification of the a practical system will need to be tightened to account for the mixing of several ripple frequencies.

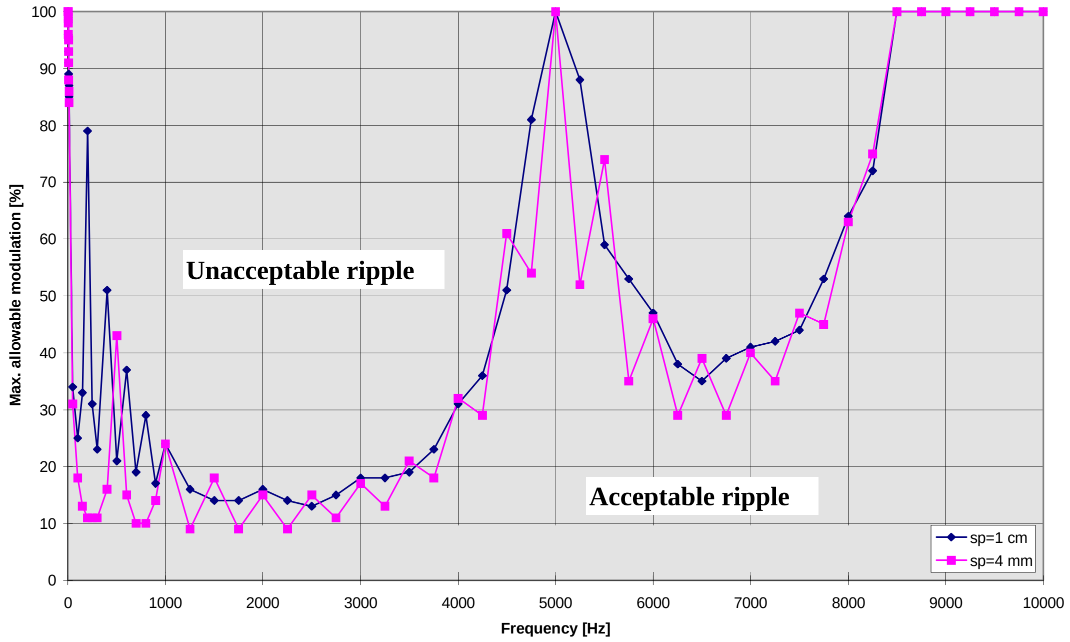


Figure 11.25 Dependence of the maximum allowable ripple amplitude with ripple frequency

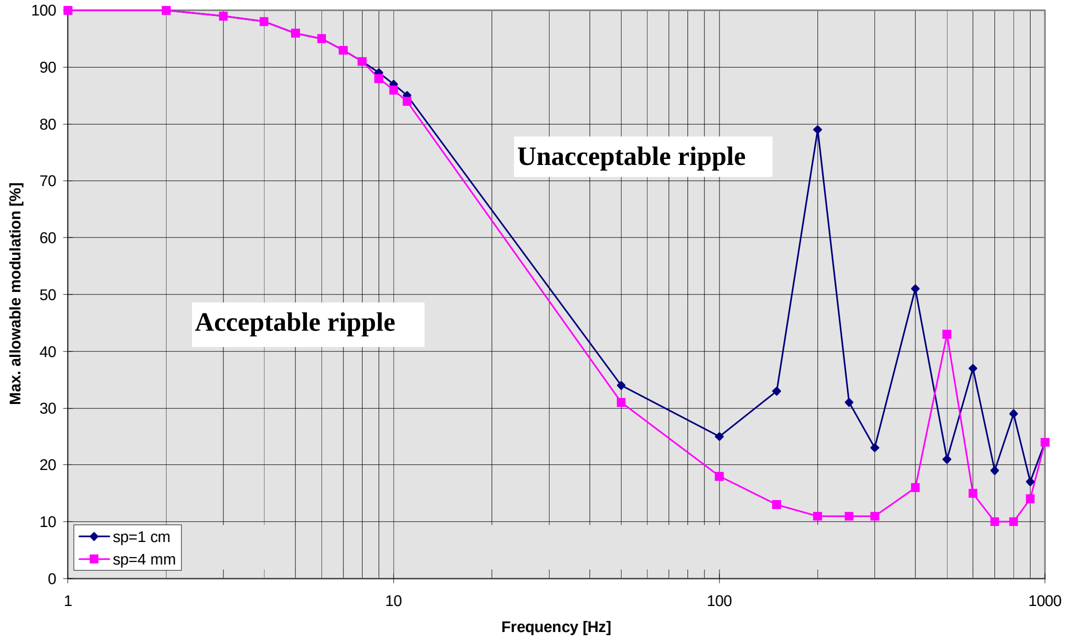


Figure 11.26 Dependence of the maximum allowable ripple amplitude with ripple frequency in the reduced range 1-1000 Hz

The same equipment (power supplies, detectors, extraction lines, magnets) can be used for all three methods. The only differences reside in the on-line flux measurement system and the control system (software) that has to work according to a different philosophy for each method.

Both dynamic methods pose similar demands on the spill quality for the largest spot size of 10 mm. The compensation of spill fluctuations is very good for the low frequencies between 1 and 10 Hz (using simple velocity compensation) and for frequencies higher than 8 kHz (using the automatic smoothing from the finite spot size). For the intermediate frequencies the maximum allowable ripple modulation drops and, in the case of mini voxel scanning, it stays on the level 20% in the region 100 to 1000 Hz. In the case of raster scanning, the critical frequencies are between 1 and 3 kHz where the level of acceptable modulation drops to 15%.

The smaller the spot size, the more marked the differences between mini-voxel and raster scanning. Up to 10 Hz the raster scanning can accept all spot sizes with almost equal and high efficiency, whereas mini-voxel scanning has problems so severe that 4 mm spots cannot be accepted for the test parameters. Between 100 and 3000 Hz, the difference among spot sizes becomes more noticeable in the case of raster scan but the maximum allowable modulation is never lower than 8%.

The efficiency of both methods can be enhanced by decreasing the particle flux, which for mini-voxel scanning increases the dwell time and for raster scanning slows the scanning velocity. This would attenuate the ‘clock’ error that plays the major role in the mini-voxel and voxel methods. However, this means an increased time for treatment.

Another way of tolerating larger ripple amplitudes is to increase the sampling frequency and hence to reduce the ‘clock’ error. This approach would be more beneficial for mini-voxel scanning than raster scanning where the delay and movement errors play a more important role. Decreasing the delay and movement errors is a technological problem that depends on the internal switching frequency of the power converter. The raster scanning method stands to gain most from advances in this domain.

The raster scanning could be improved, if the simple velocity compensation could be ‘filtered’ so as to act on low frequencies only. This would remove the negative impact of this technique at high frequencies where it introduces more error than it corrects and reduces the efficiency of the spot smoothing. It is also possible that the useful frequency range of the velocity compensation technique could be extended by powerful, on-line, real-time computing to frequency analyse the signal and to predict the underlying wave form. Some simple algorithms were tried, but no spectacular gains were achieved. However, it would seem possible that an improvement could be made in this way.

References

[1] E. Pedroni, H. Blattmann, T. Böhringer, A. Coray, S. Lin, S. Scheib, U. Schneider, *Voxel scanning for proton therapy*, Proc. of the NIRS International Workshop on Heavy Charged Particle Therapy and Related Subjects, Chiba, Japan, (July 1991), 94-109.

[2] T. Haberer, W. Becher, D.Schardt, G. Kraft, *Magnetic scanning for heavy ion therapy*, Nucl. Instr. and Meth. In Physics Research, A330, (1993), p296-305.

* * *

APPENDIX IA

POISSON STATISTICS

Basic problem

Consider N boxes, in which M balls have been randomly hidden. Make $M<N$, so that some boxes must be empty and make the boxes so small that they can contain only a single ball. The probability that a box contains a ball will be,

$$q = \frac{M}{N} \tag{A-1}$$

and the probability that it is empty will be,

$$p = 1 - q\,. \tag{A-2}$$

If n boxes are randomly selected, the probability of finding some specified sequence of m balls and $(n\text{-}m)$ empty boxes is

$$q^m p^{(n-m)}. \tag{A-3}$$

There will be several ways of specifying the sequence of balls and empty boxes and the overall probability of finding m balls, independent of the sequence, will be,

$$w_{\mathrm{n}}(m) = \binom{n}{m} p^{(n-m)} q^m\,. \tag{A-4}$$

Of course, the sum of all the probabilities of finding no balls, one ball, two balls, three balls etc. up to n balls will be unity,

$$\sum_{\mathrm{m=0}}^{\mathrm{m=n}} w_{\mathrm{n}}(m) = 1\,. \tag{A-5}$$

Consider now the identity

$$(p+qy)^n = \sum_{\mathrm{m=0}}^{\mathrm{m=n}} w_{\mathrm{n}}(m) y^m\,, \tag{A-6}$$

where y is a variable and differentiate with respect to y,

$$n(p+qy)^{n-1} q = \sum_{\mathrm{m=0}}^{\mathrm{m=n}} m w_{\mathrm{n}}(m) y^{m-1}\,.$$

If y = 1, then

$$\overline{m} = nq\,. \tag{A-7}$$

where $\overline{m}$ is the average number of balls that one would expect to find in n boxes and this equals the probability of finding a ball, times the number of boxes that are searched.

Now differentiate (A-6) a second time with respect to *y*,

$$n(n-1)(p+qy)^{n-2}q^2 = \sum_{\mathrm{m=0}}^{\mathrm{m=n}} m(m-1)w_{\mathrm{n}}(m)y^{m-2}\,.$$

If *y* is again made equal to unity, then

$$n(n-1)q^2 = \left\{\sum_{\mathrm{m=0}}^{\mathrm{m=n}} m^2 w_n(m) - mw_n(m)\right\}$$

$$n(n-1)q^2 = \overline{m^2} - \overline{m}$$

From (A-7), $\overline{m} - (\overline{m})^2 = nq(1-nq)$, which when added to the previous result gives,

$$n(n-1)q^2 + nq(1-nq) = \overline{m^2} - \overline{m} + \overline{m} - (\overline{m})^2$$

$$nq(1-nq) = \overline{m^2} - (\overline{m})^2$$

$$\sigma^2 = \overline{m^2} - (\overline{m})^2 = npq\,. \qquad \text{(A-8)}$$

Statistics of the ‘waiting beam’

The above model corresponds to the beam ‘waiting’ for extraction. The particles within this beam will be randomly distributed by rf noise between two limiting frequencies (momenta). The boxes correspond to the infinitesimally small phase-space volume needed to contain a particle. Thus $N>>M$ and $p \approx 1$ and *q* will be very small, so that equation (A-8) becomes,

$$\overline{m} = \sigma^2 = nq\,. \qquad \text{(A-9)}$$

If now the beam is moved slowly into the resonance and *n* ‘boxes’ are sampled at regular time intervals, then the mean number of particles in each time bin and the variance of these samples will be given by equation (A-9).

This is known as Poisson statistics. There are two features in the above which characterise this type of statistics. Firstly, there can only be no more than one ball in a box. This means that it is not possible to force a second particle into exactly the same phase space as another particle. Secondly, the very large number of boxes ensures that $p \approx 1$ and *q* is small. This reflects the very small ‘granular size’ of a beam, which is indeed mostly space.

Example calculation

The overall beam gives the probability of finding, or not finding, a particle at given moment (i.e. in a ‘box’. Let there be 10^8 particles in the beam and let *N* be the number of ‘boxes’, which is very large, but need not be specified. The probability of finding a particle is then $q = 10^8/N$ and of not finding a particle, $p = (1-q)$. Now, sample the spill from the beam in time bins of $n = 10^{-5}N$ boxes. Suppose that the whole spill last 1 s (N boxes) and a bin last 10 μs (n boxes). From (A-9),

$$\overline{m} = \sigma^2 = nq = 10^{-5} N \frac{M}{N} = 10^{-5} 10^8 = 10^3$$

Thus it is expected that the spill will have on average 10^3 particle every 10 μs and the distribution in the time bins will have a standard deviation of $\sigma \approx 30$.

$$\overline{m} = 1000 \pm 30 (\pm \sigma)$$

* * *

INDEX

Acceleration-driven extraction 27-29
Acceptance 15
Achromatic transfer between septa 47-51
Active scanning 1,5,159-160,195-216
Amplitude-momentum selection extraction 27-29.69-76,80-82,105
Amplitude-selection extraction 27-29,79
Aperture, schematic 44
Band spill 68-76
‘Bar-of-charge’ 4,161
Beam delivery 102-104,195-216
Beam delivery, influence of distribution 196
Beam intensity, extracted 4
Beam shaping 107,112-114
Bethe Bloch energy loss curve 189
Betatron core 2,105-106
Bolus 1
Bragg peak 1
Channelling rf bucket 3,123-133
Characteristic scattering angle 179-182
Chopper 169-171
Chromaticity 17,26
‘Clock’ error 199-205
Closed-dispersion bend 174
Conformal treatment 1
Coulomb scattering (see Scattering)
Coupled beam, with scattering 185-186
Courant and Snyder and sigma formalisms 147-148
Customised lattice 137-138
Delay error 199-201
Diffusion constant 111-112
Diffusion equation 108-111
Doublet 139-140
Driving term 20
Duty factor 87,111,131-132
Dynamic aperture 15
Eddy current damping 100-101
Electrostatic septum 29-31,32-34,43-44
Emittance 35
Emittance of extracted beam 82-85
Emittance, asymmetry 3-4
Emittance with scattering 183-186
Energy straggling 190-191
EULIMA 136
Extracted beam, distribution 26
Extracted beam, fitting an ellipse 166-169
Extracted beam, maximum emittance 84-85
Extraction methods, comparison for ripple 76-78
Extraction, methods 26-29,107
Feeding the resonance 105-134
FODO 139-140
FOFDOD 139-140
‘Front-end’ acceleration 123-133
Gantry 2,143-160
Hardt condition 2,37-43
Hardt condition equation 38
Hardt condition, choosing dispersion 42-43

Hardt condition, with zero chromaticity	51
Highland’s formula	179-182
Highland’s formula, generalised	180
Highland’s formula, improved	180
Horizontal beam size control	3-4
‘Hot’ and ‘cold’ spots	1,195
Intrinsic smoothing	3
Iso-centric gantry	143
Kobayashi Hamiltonian	9,14-16
Kobayashi Hamiltonian, generalised	16-18
Kobayashi Hamiltonian, shifted	53-54
KV distribution	144
Lattice	135-141
Magnetic septum	31-32,43-44
Mini-voxel scanning	195-205
Momentum-selection extraction	27-29
Monte Carlo tracking with scattering	188-193
Movement error	199-201
Multiple scattering (see Scattering)	
‘One-to-minus one’-module	161
‘One-to-one’ module	161
Open dispersion bend	174-175
Over-modulation of the spill	88-90
Passive spreading	1,159-160
Performance parameters	5-6
Phase displacement acceleration	141-120
Phase shifter	171-172,173
Phase-space map	14,18-19
Poisson statistics in a spill	90,217-219
Quadrupole-driven extraction	27-29
Ramping power converters	98-100
Range of charged particles in an absorber	190,191-192
Raster scanning	205-214
Regular-cell lattice	135
Resonant extraction	2,25-52
RF bucket	115
RF-knockout-driven extraction	27-29
Riesenrad gantry	5,163,175-176
Ripple in the spill	87-91
Rotator	3,152-155
Rotator design	153-155
Rotator method	144,152-153
Rotator, beam sizes	156-158
Round-beam method	144,150-152
Scaling	155-156
Scatterers, subdivision	182
Scattering, in absorbers	179-193
Scattering, in phase displacement acceleration	118
Scattering in the Twiss formalism	182-186
Scattering in ‘thick’ absorbers	183-185
Scattering in ‘thin’ absorbers	186
Separatrices, resonance	15,19-20,32-34
Separatrix, rf	115
Sextupole	9-11
Sextupole, chromatic effects	21-23
Sextupole, closed-orbit distortion	18
Sextupole, nomenclature, F and D type	11
Sextupole, normalised gradient	10
Sextupole, normalised strength	10

Sextupole, resonance excitation 20--21
Sextupole, sign convention 11
Sextupole, virtual 20
Sextupole-driven extraction 27-29
Sigma-matrix formalism 144-147
Sinusoidal modulation of the spill 88
Slow extraction (see resonant extraction)
Source to surface distance 159-160
Spill 9,53-85
Spill quality factor 87
Spill specification, mini-voxel scanning 201-205
Spill specification, voxel scanning 102-104
Spiral step and kick 13,29-31
Split FODO 139-140
'Square-ring' lattice 136
Stable triangle 15
Stationary bucket 115
Steinbach diagram 25-26
Stepper 161,172-173
Stochastic noise 107-114
Stochastic resonant extraction 107
Stopband 25-26,36
Strip spill 61-68
Strip spill length 65-67
Strip spill spike 62,67
Strip spill tail 62-67
Symmetric beam method 144,148-150
Synchrotron 2
Synchrotron frequency 115
Telescope module 161,162-163
Third-order resonance 9-23
Transfer between septa 44-51
Transit time, storage time 53,55-61
Triplet 139-140
Tune distance, modified tune distance 12
Tune ripple 76-82,92-97
Tune ripple, defence against 97-101
Uncoupled beam 146-147
Uncoupled beam with scattering 183-185
Unstacking 121-122
Vertical beam size control 4
Voxel scanning 102-104,195

* * *