BEAM HALO IN HIGH-INTENSITY
HADRON LINACS

vorgelegt von
Diplom-Ingenieur
Frank Gerigk

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr.-Ing. -

genehmigte Dissertation

Abstract

English:

This document aims to cover the most relevant mechanisms for the development of beam halo in high-intensity hadron linacs. The introduction will outline the various applications of high-intensity linacs and it will explain why, in the case of the CERN Superconducting Proton Linac (SPL) study a linac was chosen to provide a high-power beam, rather than a different kind of accelerator. The basic equations, needed for the understanding of halo development will be derived and employed to study the effects of initial and distributed mismatch on high-current beams. The basic concepts of the particle-core model, envelope modes, parametric resonances, the free-energy approach, and the idea of core-core resonances will be introduced and extended to study beams in realistic linac lattices. The approach taken is to study the behavior of beams not only in simplified theoretical focusing structures but to highlight the beam dynamics in realistic accelerators. All effects which are described and derived with simplified analytic models, are tested in realistic lattices and are thus related to observable effects in linear accelerators. This approach involves the use of high-performance particle tracking codes, which are needed to simulate the behavior of the outermost particles in distributions of up to 100 million macro particles. In the end a set of design rules will be established and their impact on the design of a typical high-intensity machine, the CERN SPL, will be shown. The examples given in this document refer to two different design evolutions of the SPL study: the first conceptual design report (SPL I) and the second conceptual design report (SPL II).

Deutsch:

Das Ziel dieser Arbeit ist, die relevantesten Mechanismen der Haloentwicklung für Teilchenstrahlen in Hochintensitätslinearbeschleunigern zu behandeln. In der Einleitung werden die vielfältigen Anwendungen dieser Linearbeschleuniger (kurz: Linac) vorgestellt. Es wird weiterhin erklärt warum im Falle der CERN Studie zur Konstruktion eines supraleitenden Protonenlinacs (SPL) ein Linac gewählt wurde um einen hochintensiven Protonenstrahl zu liefern, anstatt eines anderen Beschleunigertyps. Anschließend werden die grundlegenden Gleichungen abgeleitet, welche zum Verständnis der Haloentwicklung benötigt werden. Diese Gleichungen werden dann benutzt um den Einfluss von anfänglicher und statistisch verteilter Strahlfehlanpassung auf hochintensive Teilchenstrahlen zu untersuchen. Grundlegende Konzepte wie: das Teilchen-Kern Modell (particle-core model), Enveloppenmoden, parametrische Resonanzen, der "freie Energie" Ansatz und die Idee der Kern-Kern Resonanzen werden eingeführt und erweitert um Teilchenstrahlen in realistischen Fokussierungskanälen zu studieren. Eine Grundidee dieser Arbeit ist, das Strahlverhalten nicht nur in vereinfachten theoretischen Fokussierungsstrukturen zu beschreiben, sondern die Strahldynamik in realistischen Beschleunigern zu untersuchen. Alle Effekte welche mit vereinfachten analytischen Modellen abgeleiten werden, werden so mit beobachtbaren Effekten in Linearbeschleunigern in Zusammenhang gebracht. Dieser Ansatz bringt es mit sich, dass leistungsfähige Simulationsprogramme benutzt werden um die Trajektorien der äußersten Randteilchen einer Verteilung zu verfolgen, welche aus bis zu 100 Millionen Makropartikeln besteht. Gegen Ende der Arbeit wird eine Reihe von Regeln aufgestellt und es wird aufgezeigt, welchen Einfluss diese Regeln auf das Design eines typischen Linacs für hochintensive Teilchenstrahlen (den CERN SPL) hat. Die Beispiele in dieser Arbeit beziehen sich auf zwei Entwicklungsstadien dieses Linearbeschleunigers: den ersten konzeptionellen Designreport (SPL I) und den zweiten revidierten Report (SPL II).

iv

Symbols

s .... longitudinal position

x, y, z .... s-dependent single particle positions

a .... transverse r.m.s. beam envelope

$a_{x,y}$ .... transverse r.m.s. beam envelope in x,y

b .... longitudinal r.m.s. beam envelope

$\hat{a}, \hat{b}$ .... matched r.m.s. envelopes in the smooth approximation

$\kappa(s)$ .... focusing force depending on position

F_{sc} .... space-charge force term

I .... current

K_3 .... 3D space-charge parameter

K_1 .... 1D space-charge parameter

σ_0 .... zero-current phase advance per period

σ .... full-current phase advance per period

k_Q .... quadrupole focusing constant

k_0 ... zero-current phase advance per unit length (wave number for the transverse betatron oscillations)

k .... full-current phase advance per unit length

$\alpha_i, \beta_i, \gamma_i, (i = x,y,z)$ .... Twiss parameters

$\varepsilon_i, (i = x,y,z)$ .... r.m.s. emittance

$\varepsilon_{t,l}$ .... transverse, longitudinal r.m.s. emittance

q .... charge

G .... magnetic gradient

l .... length

l_Q .... quadrupole length

l_C .... cavity length

L .... length between two quadrupoles

L_p .... length of focusing period

m .... rest mass

c .... speed of light

v .... velocity

β .... normalised velocity (v/c)

γ .... relativistic mass factor

E_0 .... electric field gradient

T .... transit time factor

ϕ .... phase angle

λ .... RF wave length

q .... electric charge

$qn(r)$ .... radial electric charge density

ε_0 .... permittivity of free space

μ_0 .... permeability of free space

P .... power

W .... energy

A .... area

ABBREVIATIONS

AD	Antiproton Decelerator at CERN
ADS	Accelerator Driven Systems
AGS	Alternating Gradient Synchrotron at BNL
APT	Accelerator Production of Tritium
BNL	Brookhaven National Laboratory
BooNE	Booster Neutrino Experiment at FNAL
CCDTL	Coupled Cavity DTL
CCL	Coupled Cavity Linac
CERN	Centre Européenne pour la Recherche Nucléaire
CNGS	Cern Neutrinos to Gran Sasso
CONCERT	COmbined Neutron Centre for European Research and Technology
CW	Continuous Wave
DOE	Department of Energy (USA)
DTL	Drift Tube Linac
ESS	European Spallation Source
EURISOL	European Isotope Separation On-Line
FFAG	Fixed Field Alternating Gradient
FNAL	Fermi National Accelerator Laboratory
IFMIF	International Fusion Materials Irradiation Facility
IMPACT	Integrated Map and Particle ACcelerator Tracking Code
ISOLDE	Isotope Separation On-Line
JPARC	Japan Proton Accelerator Complex
KOMAC	KOrea Multi-purpose Accelerator Complex
linac	linear accelerator
LAMPF	Los Alamos Meson Physics Facility
LANL	Los Alamos National Laboratory
LANSCE	Los Alamos Neutron Science Center
LBNL	Lawrence Berkeley National Laboratory
LEBT	Low Energy Beam Transport
LEIR	Low Energy Ion Ring at CERN
LEP	Large Electron Proton Collider at CERN
LHC	Large Hadron Collider at CERN
MiniBooNE	first stage of the BooNE experiment
Nb/Cu	Niobium on Copper
NuMI	Neutrinos at the Main Injector at FNAL
PARMILA	Phase And Radial Motion In Linear Accelerators
PS	Proton Synchrotron at CERN
PSB	Proton Synchrotron Booster at CERN
PSI	Paul Scherrer Institute
PSR	Proton Storage Ring at LANL
RAL	Rutherford Appleton Laboratory
RCS	Rapid Cycling Synchrotron
RF	Radio Frequency

RFQ	RF Quadrupole
r.m.s.	root mean square
SC	SuperConducting
SDTL	Separated DTL
SPL (I/II)	Superconducting Proton Linac (I - design report 2000 [1], II - design report 2006 [2])
SPS	Super Proton Synchrotron at CERN
SNS	Spallation Neutron Source
TRASCO	TRAsmutazione SCOrie (design of an Accelerator Driven System for Nuclear Waste Transmutation)
WNR	Weapons Neutron Research Facility at LANL
XADS	eXperimental Accelerator Driven System

1. Introduction

1.1   High-power hadron linacs: machine types and their applications

High-power LINear ACcelerators (linacs) are under study for various applications since the early 70's. The first machine to actually achieve beam power in the megawatt range was the linac of the Los Alamos Meson Physics Facility (LAMPF), delivering its first 800 MeV beam at low duty cycle in 1972. After the construction of the proton storage ring (PSR) the neutron scattering community becomes the main user of the facility which in 1983 produces an average beam power of 1 MW using a pulsed beam. In 1993 the Department of Energy (DOE) drops the support for the LAMPF facility and the Los Alamos Neutron Science Center (LANSCE) is created. In 2005 the LANSCE machine remains the most powerful linear proton accelerator worldwide (Fig. 1.1).

Figure 1.1: The Los Alamos Neutron Science Center

Today LANSCE provides two different types of proton beams onto targets: a) a long-pulse proton beam (≈ 1 ms) with an average current of 1 mA for the Weapons Neutron Research Facility (WNR), and b) a short-pulse proton beam ($\approx 0.1\,\mu$s) with an average current of up to $125\,\mu$A for the Lujan Center which is produced by injecting an H$^-$ beam into the PSR where the beam is accumulated and compressed. The long-pulse beam is used either directly for irradiation experiments or to produce high-energy neutrons which complement those delivered by the Lujan Center, where "cold neutrons" are produced which allow

precise time-of-flight measurements used for general purpose material sciences. The Lujan Center is a national user facility for generally "non-military" basic and applied research in the fields of material science, engineering, condensed matter physics, polymer science, structural biology, chemistry, earth sciences, and neutron-nuclear-science research. It falls into the same science category as the ISIS facility at the Rutherford Appleton Laboratory (RAL) in the UK which, with an average beam power on target of 160 kW, is until today the most powerful pulsed spallation neutron source. In recent years neutron spallation sources became one of the most popular applications for high-intensity linacs. However, so far only the Spallation Neutron Source (SNS) project at Oakridge, Tennessee [3] was constructed and recently (spring 06) delivered the first beam on target. Similar projects are planned in Europe, India, China, and Korea.

1.1.1 H$^-$ injection and beam chopping

From the "accelerator point of view" long-pulse and short-pulse proton drivers have an important difference. Long-pulse proton drivers dump the beam from a linear accelerator directly onto the target, while short-pulse proton drivers need additional accumlator rings to compress the length of the linac pulses from the order of milliseconds to the order of microseconds. The size of the accumulator ring then determines the length of the compressed linac pulse as depicted in Fig. 1.2.
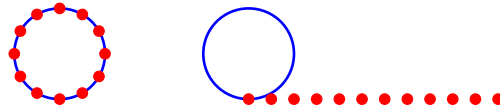

Figure 1.2: Compression of linac pulses with accumulator rings

In order to achieve pulse compression factors of up to 1000 (as in the case of SNS), linac bunches must be accumulated over a large number of injection turns (e.g. 1000). If one uses classical multi-turn injection with protons then the position of the circulating beam is always slightly shifted to make space for the injection of the incoming beam. With this technique the emittance or phase space area (i.e. the planes x/x', y/y', and z/z', respectively; see section 2.1) of the circulating beam increases with each turn approximately by the emittance of the injected beam. This principle is usually applied for up to a few tens of injection turns because one quickly reaches the acceptance limit of the circular machine which is basically given by the physical beam pipe aperture. This bottleneck can be removed by using H$^-$ charge-exchange injection which allows to inject particles over a large number of turns without increasing the emittance of the circulating beam. As shown in Fig. 1.3 the incoming H$^-$ beam and the circulating beam are deflected by the same dipole onto a common trajectory which makes it possible to inject repeatedly into the same phase space area. The stripping foil removes two electrons from the H$^-$ beam and the particles then continue on the same trajectory.

The maximum beam intensity that can be accumulated in a ring via H$^-$ injection is mainly limited by: a) the space-charge limit in the ring which is defined by the injection energy and the number of accumulated particles in the ring, b) heating of the stripper foil, c) repetition rate of the accumulation process, d) the existence of high duty-cycle, high-current H$^-$ sources for the linac, e) electron cloud and other instabilities in the rings.

To minimise losses when the H$^-$ beam is injected into a circular accelerator, one can employ low-energy beam choppers (see Section 7.2). These devices usually operate in an energy range of 2 to 3 MeV and are
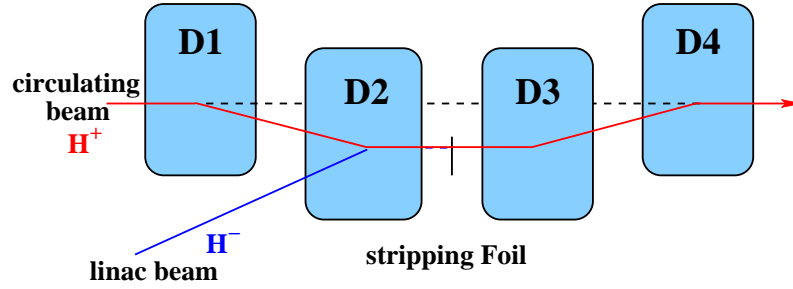
Figure 1.3: H^- multi-turn injection via charge exchange

located between Radio Frequency Quadrupole (RFQ) and Drift Tube Linac (DTL). The principle works as follows: when a continuous bunch train is injected into a ring RF system all bunches in the transition area between two ring RF buckets are either partially or completely lost as illustrated in Fig. 1.4. The effect is especially pronounced for non-accelerating RF buckets (as used in accumulator/compressor rings), where it can yield up to 30% of beam loss at injection. In case of synchrotrons the RF buckets are not fixed with respect to the location of the bunches and can thus still "collect" some of the particles in the transition areas. The purpose of low-energy beam choppers is to create gaps in the otherwise continuous bunch train to allow the transition from one ring RF bucket to the next without losing linac bunches. Depending on the acceptable losses the chopping ratio should be between 25 - 40% (see Fig. 1.4).
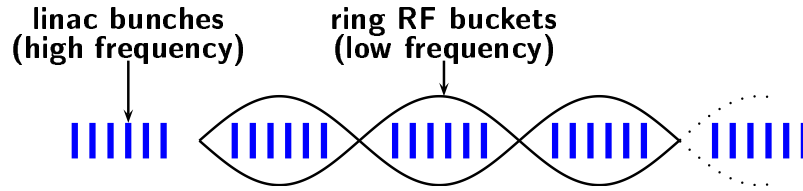

Figure 1.4: Low energy beam chopping

While Los Alamos was able to achieve one megawatt of beam power with its proton beam, only 80 kW could be produced with H^- which are injected into the PSR via charge-exchange injection. Many of the high-power hadron linacs which are under study or under construction today foresee H^- injection together with accumulator and compressor rings to produce short pulses of protons in the μs range containing short intense bunches down to the ns range. The two main applications for which this time structure is needed are short-pulse spallation targets [e.g. Spallation Neutron Source (SNS [3]), European Spallation Source (ESS study [4])] and neutrino factories [5] or super-beam facilities [6] (here only an accumulator ring is needed). Applications that use a direct proton beam on target can be partitioned into four main categories: a) long-pulse spallation targets, e.g: the WNR at LANSCE, the International Fusion Materials Irradiation Facility (IFMIF, under study), ESS, b) Accelerator Driven Systems (ADS, under study) for the transmutation of nuclear waste from conventional nuclear reactors or energy production with energy amplifiers (under study), c) radioactive ion beam facilities, e.g. On-Line Isotope Mass Separator (ISOLDE [7], in production), European Isotope Separation On-Line (EURISOL [8], under study), and d) the Accelerator Production of Tritium, e.g. APT [9], abandoned study, COmbined Neutron Centre for European Research and Technology (CONCERT [10], abandoned study). Table 1.2 lists the main characteristics of these linac-based projects.

1.1.2 Beam loss in linacs

The main concern in all high-power linac projects aiming at beams in the megawatt range is the maintainability of the machine. Beam loss leads to activation of accelerator components and may require long "cool-down" periods before hands-on repairs become possible. The currently accepted limit for beam loss that still allows hands-on maintenance is 1 W/m and stems from the experience at LANSCE (see Fig. 1.5)[1]. Even though this limit was already achieved by an accelerator that was conceived in the 70's one has to keep in mind that it took over 10 years to reach full beam power (with protons) under the condition of assuring hands-on maintenance for large parts of the machine. Until today only 80 kW of H^- can be accelerated at LANSCE and then compressed by the PSR. This means that low-loss acceleration in high-power proton linacs remains a "hot topic" and especially so for accelerators using H^- beams which suffer from the additional difficulties of: a) ionisation losses throughout the accelerator (e.g. via rest gas in the beam pipe, magnetic fields, or blackbody radiation of the beam pipe), b) generally worse beam quality out of H^- sources when compared to proton sources, c) deterioration of beam quality by the use of low-energy beam choppers, and d) injection into subsequent accumulator and compressor rings.
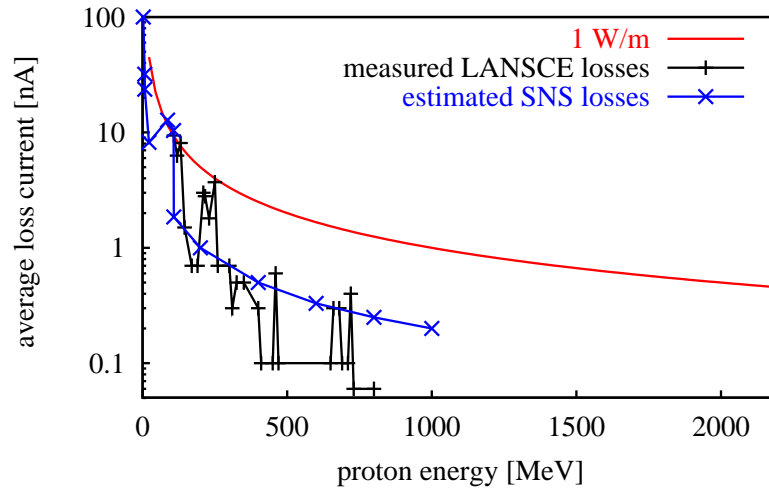

Figure 1.5: General beam loss budget for 1 W/m, measured losses at LANSCE, and estimated losses for SNS [11], [12]

With envisaged beam powers of up to 10 MW the loss limit of 1 W/m implies the need to control beam losses to a level of $10^{-6} - 10^{-7}$/m in order to avoid performance limitations due to machine activation. This poses not only a challenge to theoretical predictions of beam losses but also to the simulation tools, used to cross-check and quantify theoretical predictions. This challenge led to the development of multi-particle codes like IMPACT [13] which use parallelised space-charge routines. With these codes it became possible to routinely simulate 10^6 particles and to perform detailed halo studies with up to 10^8 particles within reasonable time scales. On the other hand the imperative for low-loss operation in many linac projects (APT, SNS, JPARC [14], CONCERT, TRASCO [15], ESS, SPL, etc; see also Section 1.3) led to a significant increase in the understanding of beam halo formation which is nowadays considered to be the dominant loss mechanism in high-intensity hadron linacs.

---

[1]the number 1 W/m assumes that the majority of the beam losses takes place inside of quadrupoles mainly consisting of iron, which provides a natural shielding for the radiation from the activated beam pipe.

Various mechanisms can contribute to the development of a low-density particle halo which surrounds the core and eventually results in beam loss. The most prominent are: a) parametric 2:1 resonances between the oscillations of a mismatched beam core and the movements of single particles, b) envelope-lattice resonances which occur between the beam envelopes and the elements of a periodic focusing structure, c) intra-beam scattering, and d) coherent space-charge coupling resonances. The envelope-lattice resonances can yield rapid r.m.s. emittance growth and halo development but are easily avoided by keeping the zero-current phase advance per period in all three planes below $90°$. Intra-beam scattering may be of importance in the Low-Energy Beam Transport (LEBT) section, where the ionisation of gas in the beam pipe is used to compensate the high space-charge forces at low energy. At higher energies, where the vacuum in the beam pipes is kept at much lower values than in the LEBT, this process is considered to be of minor importance (see e.g. [16], [17]). It will be shown in Chapter 6 that space-charge coupling resonances [18] are responsible for emittance exchange between the transverse and the longitudinal plane of the beam but that they do not (by themselves) contribute to the development of beam halo. This leaves the parametric resonances as the most important halo mechanism in high-intensity linacs.

1.2   High-intensity linacs versus other accelerator types

High-intensity linacs are not the only type of accelerator that is suited to produce high beam power. Promising candidates are cyclotrons and Rapid Cycling Synchrotrons (RCS). Another machine type that has been studied in recent years is the Fixed Field Alternating Gradient (FFAG) accelerator. So far, however, only small scale prototypes have been built.

The type of accelerator that is chosen for a particular high-power application depends to a large degree on the required beam pulse structure. It was already mentioned that short pulse operation (with pulses in the range of μs) necessitates the use of an accumulator ring. If the single bunches out of the accumulator ring need to be very short (ns range) then an additional compressor ring is needed to reduce the bunch length. A second decisive criterion for or against a certain accelerator type is the required output energy. For low-energy (≤ 500 MeV) high-power applications, cyclotrons appear as the most economical solution, while for high-energy (≥ 10 GeV) high-power applications, RCSs would clearly be the cheapest solution. In the following we compare the characteristics of a high-power linac with cyclotrons and RCSs and we review the arguments for a linac-based proton driver in case of the SPL project at CERN.

1.2.1   Requirements for a CERN-based proton driver

The most suitable machine type for a high-power proton driver at CERN needs to be matched to the requirements of potential high-power users at CERN. Presently, the most likely high-power proton applications to be located at CERN in the future are EURISOL and/or various types of neutrino facilities which are detailed in the following (compare also [2]). Apart from the high-power applications the SPL would also be beneficial for direct injection into the CERN Proton Synchrotron (PS), by-passing the PS booster (PSB), which is presently used to produce proton beams of 1.3 GeV. An increase of the injection energy together with the smaller emittances which can be obtained from a linac would result in higher brightness beams out of the PS which in turn would improve the beam quality out of the subsequent Super Proton Synchrotron (SPS) and the Large Hadron Collider (LHC) which is presently under construction at CERN. Figure 1.6 shows a diagram of the accelerator chain at CERN.
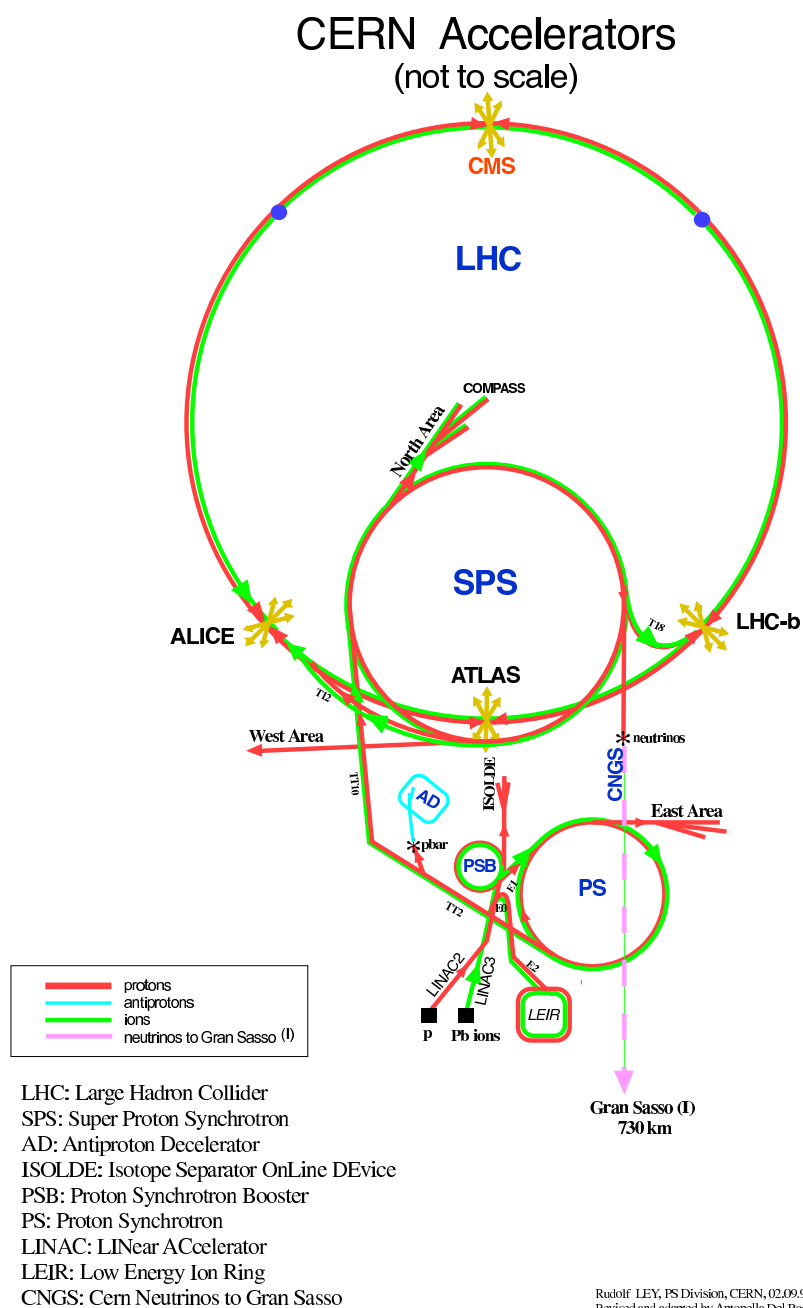
CERN Accelerators
(not to scale)

CMS

LHC

COMPASS

North Area

SPS

ALICE

ATLAS

LHC-b

T18

T2

West Area

T10

neutrinos

CNGS

East Area

AD

ISOLDE

pbar

PSB

PS

protons
antiprotons
ions
neutrinos to Gran Sasso (I)

T12

LINAC2

LINAC3

LEIR

p   Pb ions

Gran Sasso (I)
730 km

LHC: Large Hadron Collider
SPS: Super Proton Synchrotron
AD: Antiproton Decelerator
ISOLDE: Isotope Separator OnLine DEvice
PSB: Proton Synchrotron Booster
PS: Proton Synchrotron
LINAC: LINear ACcelerator
LEIR: Low Energy Ion Ring
CNGS: Cern Neutrinos to Gran Sasso

Rudolf LEY, PS Division, CERN, 02.09.96
Revised and adapted by Antonella Del Rosso, ETT Div.,
in collaboration with B. Desforges, SL Div., and
D. Manglunki, PS Div. CERN, 23.05.01

Figure 1.6: Accelerator chain at CERN

EURISOL

EURISOL is a study for a next generation radioactive ion beam facility which produces a wide range of exotic ions for users in the areas of nuclear physics, nuclear astrophysics, and material sciences. The facility would provide radioactive ion beams with intensities which are on average three orders of magnitude larger than current ISOL installations. Another recently discovered capability of EURISOL is the so-called beta-beam facility, which would make EURISOL an ideal partner user of a proton driver that delivers beam to a "super-beam" target (more details on beta-beams and super-beams in the follwing section). The "green field" EURISOL study foresees a continuous (CW) beam with 5 MW beam power at 1–2 GeV. The continuous operation as well as the target energy are chosen to minimise the impact of thermal stresses in the target. The proton beam will then be shared between one high-power target (max. 5 MW) and up to three low-power targets ($< 100\,\mathrm{kW}$ each). In terms of target stability and lifetime all targets would benefit from continuous beam operation which means that a pulsed beam from the SPL is certainly not the optimum choice. However, in order to mitigate the effects of thermal shocks and shock-waves created by a pulsed beam, the targets can be heated. Presently the minimum pulse length which is still considered suitable for the operation of EURISOL type targets is estimated to be in the range of ms. First studies on using pulsed beams (at energies above 1–2 GeV) from the SPL for EURISOL targets have recently started and it seems likely that a satisfactory operation can be achieved with pulsed operation [19] even if the lifetime of the targets will be reduced.

Neutrino facilities

In recent years neutrino physics has emerged as a new potential user of high-power accelerators and its supporters argue that the physics case is independent of the explorations at high-energy colliders [20, 21]. For neutrino facilities at CERN one has to distinguish three types of possible facilities which can be characterised as follows (see also [2]):

  i) super-beam: 4 MW proton beam + accumulator ring + pion production target,

 ii) beta-beam: $\approx 200\,\mathrm{kW}$ proton beam + ISOL type target + (existing) CERN Proton Synchrotron (PS) + (existing) CERN Super Proton Synchrotron (SPS) + decay ring,

iii) neutrino factory: 4 MW proton beam + accumulator and compressor ring + pion production target + muon cooling and capture channel + muon acceleration + muon decay ring.

All scenarios aim to study the topic of neutrino oscillations, which is currently being discussed as one of the major candidates for future high-energy physics studies at CERN and/or elsewhere in the world. The super-beam and neutrino factory schemes are based on a high-power (4 MW) proton beam hitting a pion-producing target. The pions decay within a few tens of metres into muons which then decay into neutrinos. In case of the neutrino factory the muons are captured and accelerated to energies of up to 20–50 GeV. Due to their short mean lifetime ($\approx 2.2\,\mathrm{ms}$ at rest) the muons must be accelerated quickly to relativistic speeds. The decay then takes place in a dedicated decay ring with long straight sections pointing to two detectors: one "near-detector" at a distance of several hundred kilometres and one "far-detector" at a distance of several thousand kilometres.

While the neutrino factory scheme offers the ultimate potential for neutrino physics, a combination of the super-beam and beta-beam scheme in the same energy range appears as a possible viable alternative. The beta-beam scheme would make use of the CERN infrastructure (PS and SPS) and it would use

the same target technology needed for EURISOL. The optimum energy for a CERN-based beta-beam plus super-beam facility is estimated to be around 3.5 GeV [22]. Simulations for pion production targets indicate an optimum energy range for the proton beam of 5–10 GeV, promising up to 50% higher transmission through an assumed muon front-end than at energies of 1–2 GeV [23, 24]. However, since these simulations rely on pion production models which are not yet experimentally verified it would be premature at this point in time to establish an optimum proton driver energy.

All three neutrino schemes can be based on a pulsed linac with a time structure which is dictated by the installations after the target. In the most challenging scheme, the neutrino factory, very short pulses in the order of μs are needed in combination with ultra short bunches in the ns range. The restriction in pulse length comes from the fact that in the final stage of all neutrino factory schemes, a muon beam decays in a dedicated ring into neutrinos. Assuming a circumference of ≈ 2 km for such a ring, the length of the muon pulse that can be injected within one turn is $\approx 6\,\mu$s. The short bunches are preferred in order to reduce the energy spread of the particles that are produced by the MW proton beam impinging on the pion production target. A small energy spread facilitates the task of capturing the pion beam (which quickly decays into a muon beam) and to accelerate the muons to energies between 20 and 50 GeV. The decay time of muons at 50 GeV is in the order of a few ms which imposes a timing restriction on the repetition rate with which one can re-fill the muon decay ring. This means that the repetition rate should be somewhere between 10 and 100 Hz. Considering different repetition rates while assuming the same average beam power one finds that the lowest repetition rate is limited by the space-charge forces of the accumulated beam in the accumulator/compressor rings. The highest repetition rate is limited by an increased power consumption of the accelerator components between the pion target and the muon decay ring. (This comes from the fact that for every beam pulse one has to add a certain time needed to fill all accelerating cavities with energy. For high repetition rates with relatively low currents the power consumption of the whole facility would be dominated by the power needed to fill the cavities before the arrival of the actual beam pulse.) The time structure of μs pulses and ns bunches can be achieved by combining an H$^-$ linac with an accumulator ring (see Section 1.1.1) and a compressor ring [25].

The synergy

For the time being, a pulsed beam at 3.5 GeV seems a viable compromise between the needs of EURISOL and possible neutrino facilities. At this energy both users can be supplied with high-power beams in a time-sharing operation mode. Another advantage of 3.5 GeV is that one can profit from an increased injection energy into the CERN PS, lowering the space-charge tune shift at injection and thus removing the present intensity limitation. While EURISOL (including beta-beams) could be operated with a proton front-end it is mandatory to foresee H$^-$ operation to drive a super-beam facility or a neutrino factory. The high-power long-pulse operation at or above 1 GeV needed for EURISOL can only be obtained from a linac and not from any other type of high-power accelerator as we will show in the following. Furthermore, a linac offers the unique opportunity to share the cost of the proton driver between two high-power users, both of which will certainly not operate for 12 months a year (due to maintenance, target exchanges, etc).

1.2.2 High-power cyclotrons

Average beam power in the MW range has already been achieved with cyclotrons, e.g. with the separated sector cyclotron at the Paul Scherrer Institute in Switzerland PSI [26]. A cyclotron makes use of the magnetic force on moving charges ($\boldsymbol{F} = q[\boldsymbol{v} \times \boldsymbol{B}]$) to bend particles into a circular path. The magnetic

field is supplied by large dipole magnets above and below the acceleration path (see Fig. 1.7). Between the two "dees" of magnetic field region an oscillating electric field accelerates the particles. After each acceleration step the radius of the circular path is increased due to the higher particle energy until finally the beam is extracted.

Cyclotrons are generally used in Continuous Wave (CW) operation with currents up to the order of mA. They are uneconomical for H^- acceleration above energies of 50 to 100 MeV because due to the high magnetic fields the extra electron of the H^- particles will be stripped. This means that very large magnets with low fields would be necessary to avoid H^- stripping. A general limitation for all particle species is the maximum energy which depends on the strength of the magnetic field and the maximum radius of the machine. So far proton energies in the order of ≈ 500 MeV have been realised with chains of cyclotrons but it seems unlikely that future machines will reach significantly higher values. For applications that need long pulses or CW beams of protons at low or medium energies (≤ 500 MeV), cyclotrons are a viable alternative to linacs, especially because of the compact size and because the RF power is used much more efficiently.
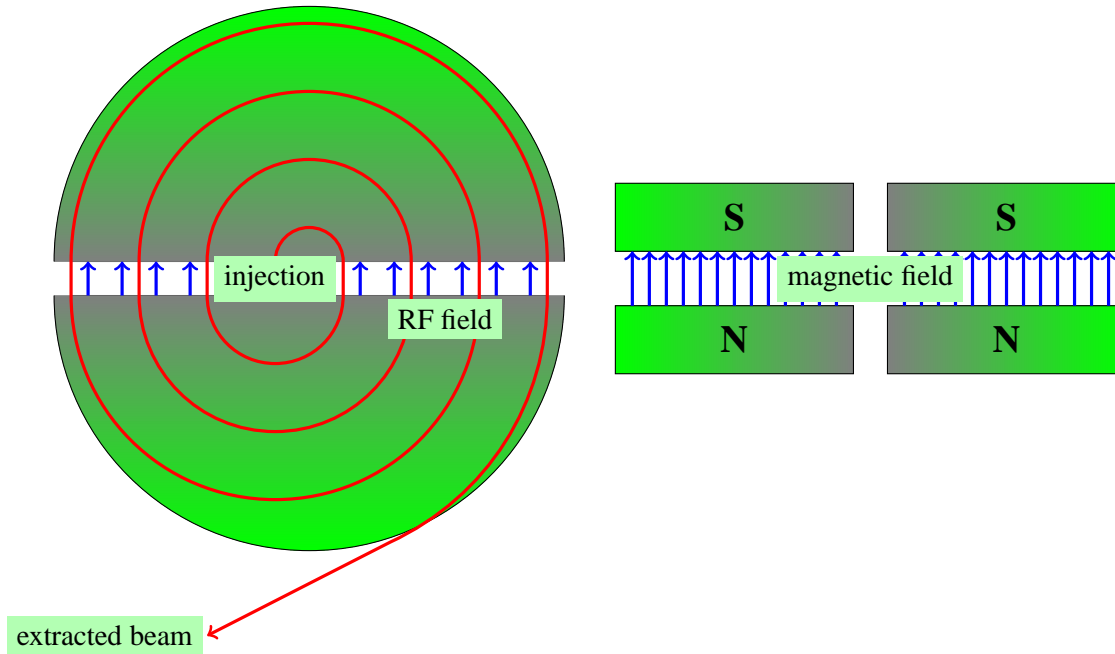

Figure 1.7: Cyclotron principle: (left) top view, (right) side view

1.2.3   Rapid cycling synchrotrons

The second type of circular accelerator which has already produced high-power beams is the rapid cycling synchrotron (RCS). In a synchrotron the beam is accelerated by one or more RF cavities, which can be matched in frequency to the revolution frequency of the particles. Contrary to the cyclotron the beam is kept on a constant-radius orbit by increasing the magnetic field in the dipoles, used to bend the beam, in synchronism with the increasing beam energy. Hence the name synchrotron. Within typically several thousand turns the beam is accelerated to its full energy and can be extracted. For an RCS the whole acceleration cycle only takes a few ms, which means one can have many cycles per second and

thus obtain a high average beam power. A review of existing high-power synchrotrons [27] in Table 1.1 shows that so far beam powers close to the MW range have only been achieved in high energy machines with low cycling rates. At lower beam energy the neutron spallation source ISIS remains since many years the RCS with the highest beam power (0.16 MW).

Table 1.1: Beam parameters of existing and proposed proton synchrotrons

machine	rep. rate [Hz]	energy [GeV]	power [MW]
RAL ISIS	50	0.8	0.16
BNL AGS	0.5	24	0.13
Fermilab MiniBooNE	7.5	8	0.05
Fermilab NuMI	0.5	120	0.3
CERN CNGS	0.17	400	0.5

Keeping the uncontrolled losses below the 1 W/m limit is one of the major design challenges for RCS accelerators. For example, a ring with 200 m circumference handling a 1 MW beam can accept a total fractional beam loss of $\approx 10^{-4}$ at its top energy (or 10^{-3} at injection energy assuming a typical energy multiplication factor of ≈ 10 in a synchrotron). In existing machines losses of around 10% or more occur mostly at injection. They are related to injection losses due to limited longitudinal acceptance (as explained in Section 1.1.1), a high space-charge tune shift during the accumulation phase, premature H^- stripping and injection foil scattering, magnet field errors and misalignments, various beam instabilities (e.g. head-tail instability, coupled bunch instability, negative mass and microwave instability, electron clouds, see also [28]), and accidental beam losses due to malfunctioning of single elements in the accelerator chain. The employment of low-energy beam choppers should help to drastically reduce the injection losses but with rising beam power requirements the space-charge forces at injection will rise and limit the current density in the ring. The losses during acceleration have to be controlled by dedicated beam collimation sections which scrape off the outermost particles before they get lost on the beam pipe.

One of the main cost drivers in an RCS are the main power converters for the fast cycling magnets, which become more expensive with rising repetition rates. Slower repetition rates, on the other hand, mean that the beam intensity at injection has to be very high in order to produce beams in the MW range. Although an RCS is generally considered to be more economic than a linac the beam dynamics and the technology becomes very challenging if beam energies of only a few GeV in combination with high-intensity beams are required. In the energy range of a few GeV existing machines, like ISIS are still an order of magnitude short of reaching MW beam power.

1.2.4 Accelerator choice for a CERN-based proton driver

We have already seen that the choice of a particular type of accelerator for a certain high-power proton application does not only depend on the beam power. The possible energy range as well as the needed time structure have to be taken into account along with possibly existing parts of an infrastructure that can be reused for the new facility.

The initial reason to choose a linear accelerator as proton driver for future high-intensity applications at CERN was triggered by the idea to re-use decommissioned RF equipment from the Large Electron Proton (LEP) Collider at CERN (compare also Section 7.3). The first idea was to put the super conducting (SC) LEP RF cavities in a straight line and to add a low-energy front-end. During subsequent design revisions

the actual percentage of recuperated material was drastically reduced and as a consequence also other accelerator types were considered as alternatives. The cyclotron idea was discarded because of its limited beam energy potential (especially for the acceleration of H^-) and the difficulties with high beam currents in such a machine. A rapid cycling synchrotron would be the most viable alternative, assuming that the stability problems for high-intensity beams at injection can be solved and that the uncontrolled losses can be kept below 1 W/m. Furthermore it would be extremely challenging to produce a MW beam at relatively low energies of only a few GeV which would be suitable for EURISOL.

In comparison to a synchrotron a linac offers a simpler beam dynamics which can be simulated in great detail with existing codes. In synchrotrons it is still difficult to simulate a complete acceleration cycle, and it is virtually impossible to make hundreds of simulations with different sets of statistical errors in order to test the machine lattice. This makes it difficult to judge wether new RCS designs can fulfil the need for low losses, which is required to ensure hands-on-maintenance. Another advantage of linacs is that there is a certain freedom to adapt the time structure of the pulses to the needs of various users. It is possible to deliver longer pulses with lower average currents while maintaining the same average beam power. In the case of the SPL at CERN, pulses in the ms range for EURISOL are extracted directly from the linac, while the more demanding time structure for neutrino applications is produced with a separate accumulator and compressor ring. This means that two completely different user requirements can be fulfilled with one proton front-end. A linac can deliver beam to several users within the same pulse simply by using a low-energy beam chopper to produce a gap (≈ 0.1 ms) in the bunch train, long enough to switch the high-power beam between different beam lines. This is important for instance in the case of EURISOL targets in order to avoid thermal stresses caused by the cooling of the target between beam pulses. Another option which still has to be verified experimentally is to share the beam within each pulse by using partial laser stripping at the linac output to convert a fraction of the beam from H^- to H_0. The beams can then be separated in a bending magnet and diverted to the respective users.

In the context of CERN a linac-based proton driver like the SPL offers the possibility for a staged approach towards neutrino physics and nuclear physics (EURISOL):

i) In the first stage the SPL would deliver a low-power beam (e.g. 0.5 MW) to a EURISOL / beta-beam facility. The targets for this facility basically exist or can be derived from the ISOLDE facility at CERN. In this stage it is not even necessary to have a high-duty H^- ion source but one can use existing proton sources in the front-end. Gradually one can then increase the beam power and start the testing of high-power targets first for EURISOL and then for a super-beam facility.

ii) Once the full beam power is reached (which will take several years) and once there is enough confidence in the short-pulse (μs) super-beam target technology, one can add an accumulator ring and use an H^- beam for charge-exchange injection into the ring. At this point one can supply the EURISOL low- and high-power targets, along with the low-power beta-beam targets plus the high-power super-beam target. At the same time one can start the prototyping and testing of the short-pulse neutrino factory target which must be able to withstand ns bunches of extremely high instantaneous power.

iii) The addition of a compressor ring now completes the installation to a complete neutrino factory proton driver and by this time it should also be clear if the energy of 3.5 GeV is sufficiently high. If necessary one could now still increase the final linac energy (and possibly also the beam power) to meet the required values. In parallel one should now also have enough confidence in neutrino factory targets and the necessary key technologies for the construction of a complete neutrino factory.

During all stages and even during the ramp-up in beam power at the very beginning, can the SPL deliver beam to the users including the regular CERN physics program which includes a luminosity upgrade for the LHC. Upgrades in terms of beam energy or beam power are achieved simply by adding more accelerating cavities at the end of the linac or by lengthening the beam pulses. This feature is important for the CERN scenario since it is not yet clear if 3.5 GeV will be sufficient for a neutrino factory. In the energy range of a few GeV and for beam powers of several MW (or higher) a linac offers more flexibility than circular accelerators. At this point in time FFAGs are not yet considered as a viable alternative, even though the technology has the potential to produce high-power beams. Future R&D work will have to show if FFAGs can indeed fulfil this potential.

1.3   High-power linac studies and the goal of this thesis

It was mentioned in the previous sections that the understanding of beam dynamics and halo development was boosted by the large number of high-power linac studies which have been undertaken in the last two decades (see Table 1.2). In order to simplify the problem, most of the analytical studies were performed using simplified accelerator lattices and beam geometries. One usually neglects acceleration and in the first stage one often uses two-dimensional beams, assuming un-bunched long cylinders of charged particles. To further simplify the analytic treatment all periodic transverse focusing forces are often averaged (smoothed) over the focusing periods. Without these simplifications it is in many cases impossible to study beam characteristics and one has to accept a drastic change of the original lattice. In the following chapter we will outline the analytic treatment of 3D beams with space-charge, which is used in most studies concerned with halo development. We will stress, which simplifications are made in order to derive the commonly used formulas.

The analytic findings of the last decades were usually verified with numerical multi-particle simulations. However, in most of these simulations the same simplifications as in the analytic treatment were used, meaning that 2D beams in continuous focusing lattices with artificially created particle distribution were used to test the analytic theories. In many cases the studies stopped at this point and the theories were not tested on realistic accelerator lattices including periodic focussing, acceleration, and transitions between different focussing lattices. This thesis was motivated by a different, maybe naive but entirely practical approach: during the design process of the SPL I [1] the beam dynamics simulations exhibited certain features which could not easily be related to well-known phenomena. In subsequent discussions with I. Hofmann, K. Bongardt and other colleagues, the relevant theory was revisited and amended if necessary. The first feature was the observation of emittance exchange between the transverse and longitudinal planes for certain settings of the phase advance. This phenomenon could be identified as a coherent space-charge resonance as described in [18]. The corresponding theory and the simulation results of this study are reported in Chapter 6.

After a number of tests with mismatched beams, the mismatch modes were used to systematically study halo development in the SPL (see Chapter 4). In an advanced stage of the design, when the SPL was simulated with statistical errors, features similar to the ones found for the mismatch modes were observed and were then related to the same theory (the particle-core model) that describes the basic effects of initial mismatch (see Chapter 5).

The work presented in this thesis is entirely based on phenomena that can be observed for realistic 3D beams in realistic linac lattices. These phenomena are then related to simplified analytic expressions (where possible), which will be derived in the following.

Table 1.2: High-power, high-energy linac studies and projects worldwide

machine	particle species	energy [GeV]	beam power [MW]	rep. rate [Hz]	peak current [mA]
existing:					
LANSCE (USA)	p/H$^-$	0.8	1/0.08	16.5/9.5	100/20
SNS* (USA)	H$^-$	1	1.4	60	38
under construction					
JPARC†	H$^-$	0.6	0.23	25	30
proposal					
SPL II (CERN, Switzerland)	H$^-$	3.5	5	50	64
ESS‡ (Europe)	H$^-$	1.33	10	50	114
CONCERT‡ (Europe)	H$^-$, p	1.33	25	50	114
KOMAC (Korea)	p	1	20	CW	20
TRASCO‡ (Italy)	p	1	30	CW	30
EURISOL (Europe)	p	1	5	CW	5
APT‡ (USA)	p	1	100	CW	100
FNAL (USA)	H$^-$, p	8	2	10	25
XADS (Europe)	p	0.6	6	CW	10
IFMIF	D$^+$	0.04	2x5	CW	2x125

* quoted are the goals, † if the final upgrade is approved, ‡ suspended study

2. Basic equations

Charged particle beams have velocity components which are perpendicular to the direction of propagation. These components arise for instance in the particle source where particles are extracted from a plasma with random thermal motions. After extraction space-charge forces will push particles away from the axis and without transverse focusing system the beam will soon start to diverge and be lost on the beam pipe. The acceleration of particles to energies beyond a few MeV is usually accomplished by radio frequency (RF) electric fields which partition the longitudinal axis in accelerating and decelerating areas, changing in time with the frequency of the RF system. For successful acceleration the particles must be confined longitudinally to the accelerating areas (RF buckets), meaning that the beam must be bunched in the longitudinal plane. Since the particles inside a bunch have different longitudinal momenta, a longitudinal focusing system must be put into place to complement the transverse focusing of the beam. In modern accelerators one often uses a periodic focusing lattice of alternating quadrupoles for transverse beam confinement, while longitudinal confinement is achieved by using off-crest acceleration, which ensures that faster particles are decelerated and that slower particles are accelerated with respect to the central (synchronous) particle.

In order to calculate the beam evolution and the stability properties of the complete ensemble of particles, it is useful to work with equations describing the r.m.s. beam envelopes. For effects that go beyond the regular oscillations of the beam one often uses the smoothed form of these equations, where one averages over the regular beam oscillations to study superimposed oscillations, e.g. due to mismatch.

In later chapters we will often use expressions that come from the smoothed approximation of Sacherers envelope equations [29]. Therefore this chapter will provide a derivation of these equations together with a set of useful formulae which are tailored for later employment.

2.1 3D envelope equations with space-charge

The smoothed envelope equations average over the rapid flutter of the beam envelopes which are caused by the focusing and defocusing forces, and provide relations for averaged r.m.s. beam sizes. The three main forces acting on the beam shall be considered, which are either external (quadrupoles and RF gaps), or internal (space-charge force) caused by the charged particles of the beam. For this purpose one assumes a uniform particle distribution in the first place, and can then, as Sacherer showed in [29], generalise the equations for various distributions without major changes.

Before deriving the actual envelope equations one needs to define some basic properties of a generic particle distribution. The variables used for the following derivations are:

- s is the independent variable denoting the longitudinal position on the beam axis,

- x, y, and z are the s-dependent single particle coordinates within a bunch at position s, which refer to the particle positions relative to the bunch center,

- a_{x}, a_{y}, and b are the s-dependent r.m.s. beam envelopes,

- r_{x}, r_{y}, and r_{z} are the semi-axes of an ellipsoidal bunch of particles, and

- $\hat{a}$ and $\hat{b}$ are the matched r.m.s. beam envelopes with space-charge which are in a simplified case (the smooth approximation) assumed constant over one focusing period, $\hat{a}_0$ and $\hat{b}_0$ are used for the zero-current case.

Assuming linear focusing forces all particle trajectories lie on ellipses in the the x/x' plane which is called trace-space and which represents a two-dimensional projection of the six-dimensional beam. Often one also refers to the x/x' plane as the unnormalized phase space projection (the $x/(p_x/mc)$ plane would be the normalized phase space projection). In the following we will simply refer to "phase space", regardless of the normalization of the vertical axis (see Fig. 2.1). For a matched beam the iso-density contours in phase space are concentric and geometrically similar to the trajectory ellipses. This statement remains valid even for mismatched beams as long as the focusing system remains linear. It is thus convenient to describe the properties of a linac beam with parameters that relate to r.m.s. ellipses in all three phase space planes (x/x', y/y', z/z'). For this purpose one can use the Courant-Snyder (or Twiss) parameters (α, β, γ) of Eq. (2.1) to describe the contour of a generic ellipse as shown in Fig. 2.1, which is centred at the origin and rotated by a certain angle.
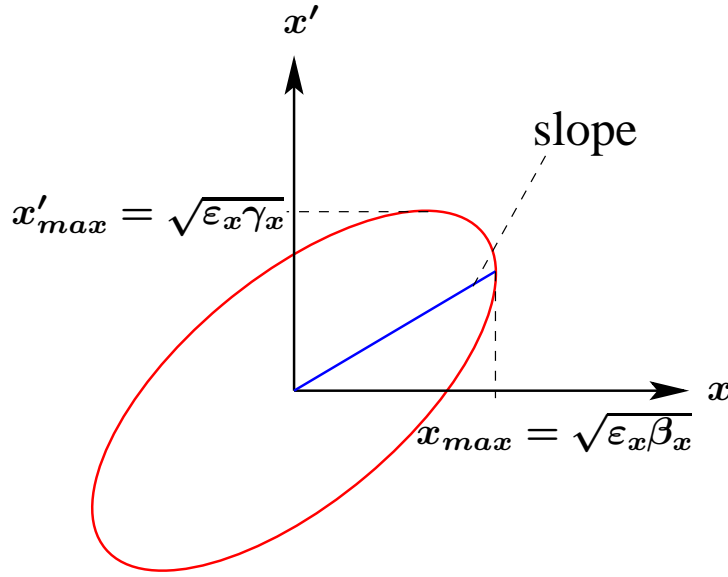

Figure 2.1: Generic r.m.s. ellipse for particle distributions in the x/x' phase space

$$\gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2 = \varepsilon_x \tag{2.1}$$

The Twiss parameters in each plane α_i, β_i, and γ_i are linked by

$$\gamma_i \beta_i - \alpha_i^2 = 1 \qquad i = \text{x,y,z} \tag{2.2}$$

and the area of the ellipse in each plane is given by $A_i = \pi \varepsilon_i$ which introduces the definition of the un-normalised r.m.s. emittance ε_i of an arbitrary particle distribution in phase space. When designing a linear accelerator one usually refers to the normalised emittance which remains constant during acceleration (provided that there is no emittance growth). It is calculated by multiplying the un-normalised emittance with the relativistic factors ($\beta\gamma$). In the following, however, we will continue to use the un-normalised r.m.s. emittance which relates to the physical properties of the beam.

Averaging over the particle distribution one can express the r.m.s. quantities of the beam which are only dependent on the longitudinal position s.

$$a_x^2 = \overline{x^2} = \beta_x \varepsilon_x \qquad a_x'^2 = \overline{x'^2} = \gamma_x \varepsilon_x \qquad \overline{xx'} = -\alpha_x \varepsilon_x \qquad \varepsilon_x = \sqrt{\overline{x^2}\,\overline{x'^2} - \overline{xx'}^2} \qquad (2.3)$$

The same definitions apply in the y-plane for a_y, ε_y and in the z-plane for b, ε_z. Using Eq. (2.3) one can derive the envelope equations starting from the transverse equation of motion for a single particle:

$$x'' + \kappa(s)x = 0 \qquad\qquad \textbf{transverse equation of motion,} \qquad (2.4)$$
$$\textbf{Hill's equation}$$

Hill's equation is a second-order linear differential equation with $\kappa(s)$ denoting the linear external focusing forces. In hadron linacs, one is usually interested in beam transport from low to high energy, which means one has to consider the space-charge force term F_{sc}, describing the forces between the charged particles which is most pronounced at low energy:

$$x'' + \kappa(s)x - F_{sc}(s) = 0 \qquad\qquad \textbf{transverse equation of motion} \qquad (2.5)$$
$$\textbf{with space-charge}$$

We note that in Eq. (2.5) the space-charge forces (in the term F_{sc}) acting on the single particle are also dependent on the transverse position of the particle with respect to the bunch centre. If we want to make the transition from the single particle equation to an envelope equation describing the r.m.s. quantities of all particles we will therefore, at some stage, need to average over F_{sc} in order to obtain a purely s-dependent quantity.

In order to derive the envelope equations one starts by averaging over the particle positions and the second moments of the distribution:

$$a_x = \sqrt{\overline{x^2}} \qquad a_x' = \frac{\overline{xx'}}{a_x} \qquad\qquad (2.6)$$

$$a_x'' = \frac{\overline{xx''}}{a_x} + \frac{\overline{x'^2}}{a_x} - \frac{\overline{xx'}^2}{a_x^3} = \frac{\overline{xx''}}{a_x} + \frac{\overline{x^2}\,\overline{x'^2}}{a_x^3} - \frac{\overline{xx'}^2}{a_x^3} \qquad\qquad (2.7)$$

Using the definition of the r.m.s. quatities in Eq. (2.3) one can simplify Eq. (2.7) to

$$a_x'' = \frac{\overline{xx''}}{a_x} + \frac{\varepsilon_x^2}{a_x^3} \qquad\qquad (2.8)$$

Differentiating $\overline{xx'}$ and using the equation of motion (2.5) yields

$$\overline{xx'}' = \overline{x'^2} + \overline{xx''} = \overline{x'^2} - \overline{x^2}\kappa(s) + \overline{xF_{sc}(s)} \qquad\qquad (2.9)$$

or

$$\overline{xx''} = -\overline{x^2}\kappa(s) + \overline{xF_{sc}(s)} \qquad\qquad (2.10)$$

Replacing the $\overline{xx''}$ term in Eq. (2.8) with Eq. (2.10) one arrives at the r.m.s. envelope equation which is similar to the equation of motion (2.5) but now contains an additional emittance term. After repeating

the same steps in the y and z plane one obtains:

$$
\begin{aligned}
a_{\mathrm{x}}'' + a_{\mathrm{x}}\kappa_{\mathrm{x}}(s) - \frac{\overline{xF_{\mathrm{sc,x}}(s)}}{a_{\mathrm{x}}} - \frac{\varepsilon_{\mathrm{x}}^2}{a_{\mathrm{x}}^3} &= 0 \\
a_{\mathrm{y}}'' + a_{\mathrm{y}}\kappa_{\mathrm{y}}(s) - \frac{\overline{yF_{\mathrm{sc,y}}(s)}}{a_{\mathrm{y}}} - \frac{\varepsilon_{\mathrm{y}}^2}{a_{\mathrm{y}}^3} &= 0 \\
b'' + b\kappa_{\mathrm{z}}(s) - \frac{\overline{zF_{\mathrm{sc,z}}(s)}}{b} - \frac{\varepsilon_{\mathrm{z}}^2}{b^3} &= 0
\end{aligned}
$$

r.m.s. envelope equations with space-charge (2.11)

These equations were first derived in 1959 by Kapchinskiy and Vladimirskiy [30] for a continuous (un-bunched) beam and in this context they are often referred to as the KV-envelope equations.

The term $\kappa(s)$ in the envelope equation takes on different forms depending on the type of focusing lattice. Generally one considers periodic focusing channels with period length L_{P} so that $\kappa(s) = \kappa(s + L_{\mathrm{P}})$. The emittance term in Eq. (2.11) is negative and acts as a defocusing force on the r.m.s. beam size.

One can simplify the envelope equations (2.11) by replacing the averaged space-charge terms in Eq. (2.11) using the results of Sacherer [29] and Lapostolle [31]. Sacherer showed that these terms depend very little on the actual particle distribution (uniform, Gaussian, hollow, parabolic), making the r.m.s. envelope equations a widely applicable tool to model the r.m.s. beam quantities.

In the next step we replace the space-charge terms by expressions relating to the particle mass and charge, their velocity and the aspect ratio of the bunches.

Lapostolle defines the electric potential for an ellipsoidal bunch in free space assuming similar dimensions for the semi-axes (r_{x}, r_{y}, r_{z}) of the bunch. The electric field components can then be written as [32]:

$$
\begin{aligned}
E_{\mathrm{x,sc}}(s) &= \frac{3I\lambda[1 - f(s)]x}{4\pi\varepsilon_0 c(r_{\mathrm{x}} + r_{\mathrm{y}})r_{\mathrm{z}}r_{\mathrm{x}}} \approx \frac{3I\lambda[1 - f(s)]x}{8\pi\varepsilon_0 cr_{\mathrm{x}}r_{\mathrm{y}}r_{\mathrm{z}}} \\
E_{\mathrm{y,sc}}(s) &= \frac{3I\lambda[1 - f(s)]y}{4\pi\varepsilon_0 c(r_{\mathrm{x}} + r_{\mathrm{y}})r_{\mathrm{z}}r_{\mathrm{y}}} \approx \frac{3I\lambda[1 - f(s)]y}{8\pi\varepsilon_0 cr_{\mathrm{x}}r_{\mathrm{y}}r_{\mathrm{z}}} \\
E_{\mathrm{z,sc}}(s) &= \frac{3I\lambda f(s)z}{4\pi\varepsilon_0 cr_{\mathrm{x}}r_{\mathrm{y}}r_{\mathrm{z}}}
\end{aligned}
\tag{2.12}
$$

where I is the averaged current over one RF period, λ is the (free space) RF wavelength, c is the speed of light and ε_0 is the permittivity of free space. The ellipsoidal form factor $f(s)$ depends on the aspect ratio of the bunch (which changes in a periodic focusing channel). Even though it will not be needed in the following we cite here Lapostolle's approximation [31] for $f(s)$

$$
f(s) \approx \frac{\sqrt{a_{\mathrm{x}}a_{\mathrm{y}}}}{3\gamma b}
\tag{2.13}
$$

In a three-dimensional uniform ellipsoid one can relate the semi-axes of the bunch to the r.m.s. value with (see [31]):

$$
r_i = \sqrt{5}a_i
\tag{2.14}
$$

The space-charge force terms can be replaced by their field components (see Appendix A.1)

$$
F_{i,\mathrm{sc}}(s) = \frac{qE_{i,\mathrm{sc}}(s)}{\beta^2\gamma^3 mc^2}
\tag{2.15}
$$

and one can express the space-charge terms in Eq. (2.11) as:

$$\frac{\overline{xF_{\text{x,sc}}(s)}}{a_{\text{x}}} = \frac{K_3[1-f(s)]}{2a_{\text{y}}b} \qquad \frac{\overline{yF_{\text{y,sc}}(s)}}{a_{\text{y}}} = \frac{K_3[1-f(s)]}{2a_{\text{x}}b} \qquad \frac{\overline{zF_{\text{z,sc}}(s)}}{b} = \frac{K_3f(s)}{a_{\text{x}}a_{\text{y}}} \qquad (2.16)$$

with K_3 being a 3-D space-charge parameter defined as:

$$K_3 = \frac{3qI\lambda}{20\sqrt{5}\pi\varepsilon_0 mc^3\beta^2\gamma^3} \qquad (2.17)$$

Using these expressions we can now write the envelope equations as:

$$
\begin{aligned}
a_{\text{x}}'' + a_{\text{x}}\kappa_{\text{x}}(s) - \frac{K_3[1-f(s)]}{2a_{\text{y}}b} - \frac{\varepsilon_{\text{x}}^2}{a_{\text{x}}^3} &= 0 \\
a_{\text{y}}'' + a_{\text{y}}\kappa_{\text{y}}(s) - \frac{K_3[1-f(s)]}{2a_{\text{x}}b} - \frac{\varepsilon_{\text{y}}^2}{a_{\text{y}}^3} &= 0 \\
b'' + b\kappa_{\text{z}}(s) - \frac{K_3f(s)}{a_{\text{x}}a_{\text{y}}} - \frac{\varepsilon_{\text{z}}^2}{b^3} &= 0
\end{aligned}
$$

3D envelope equations with space-charge (2.18)

In the following sections we will introduce some approximations to replace $\kappa(s)$ by averaged wavenumbers k for a periodic quadrupole channel with RF cavities.

2.2   The principle of smooth approximation

The smooth approximation is often used to simplify the envelope equations for practical usage. The idea is to average the focusing forces over one full focusing period of length L_{p}. For this purpose the s-dependent focusing term $\kappa(s)$ is replaced by the smoothed (averaged over one period) wave numbers k which are related to the phase advance per period σ by

$$
\begin{aligned}
\overline{\kappa_{\text{x}}(s)} \longrightarrow k_{\text{x,0}}^2 &= \left(\frac{\sigma_{\text{x,0}}}{L_{\text{P}}}\right)^2 \\
\overline{\kappa_{\text{y}}(s)} \longrightarrow k_{\text{y,0}}^2 &= \left(\frac{\sigma_{\text{y,0}}}{L_{\text{P}}}\right)^2 \\
\overline{\kappa_{\text{z}}(s)} \longrightarrow k_{\text{z,0}}^2 &= \left(\frac{\sigma_{\text{z,0}}}{L_{\text{P}}}\right)^2
\end{aligned}
$$

smooth approximation w/o space-charge (2.19)

When considering beams with space-charge the smooth approximation can be used in the following way:

$$
\begin{aligned}
\overline{\kappa_{\text{x}}(s) - \frac{K_3[1-f(s)]}{2a_{\text{x}}a_{\text{y}}b}} \longrightarrow k_{\text{x}}^2 &= \left(\frac{\sigma_{\text{x}}}{L_{\text{P}}}\right)^2 \\
\overline{\kappa_{\text{y}}(s) - \frac{K_3[1-f(s)]}{2a_{\text{x}}a_{\text{y}}b}} \longrightarrow k_{\text{y}}^2 &= \left(\frac{\sigma_{\text{y}}}{L_{\text{P}}}\right)^2 \\
\overline{\kappa_{\text{z}}(s) - \frac{K_3f(s)}{a_{\text{x}}a_{\text{y}}b}} \longrightarrow k_{\text{z}}^2 &= \left(\frac{\sigma_{\text{z}}}{L_{\text{P}}}\right)^2
\end{aligned}
$$

smooth approximation with space-charge (2.20)

The last step represents a strong simplification since it completely removes the dependence of the space-charge terms on the beam envelopes. From practical experience one finds that for weak or moderate space-charge the equations are still reasonably precise, while for very strong space-charge forces this approximation becomes inadequate.

In the next step one can relate the transverse and longitudinal wave numbers to actual properties of the focusing periods such as magnetic gradients, electric fields, period length, etc. For the transverse plane this can be done by comparing the transfer matrix of the thin lens approximation with a general periodic solution in matrix form (see Appendix A.2).

In the presence of RF cavities the situation becomes more complex since they not only provide acceleration and longitudinal focusing but also transverse defocusing. The transverse effects of RF cavities are again evaluated using the transfer matrix method as in the transverse plane. For the longitudinal effects one simplifies the longitudinal equation of motion without space-charge to specify the longitudinal wave-number $k_{l,0}$ (see Appendix A.4).

2.3   Quadrupole and RF focusing terms

Using the smooth approximation together with the thin lens approximations for quadrupoles and RF cavities one can replace the s-dependent focusing $\kappa(s)$ in the envelope equations (2.11).

In the longitudinal plane the longitudinal zero-current wave number $k_{l,0}$ can be derived as (see Appendix A.4)

$$\overline{\kappa_{z,0}(s)} \to k_{l,0}^2 = \frac{2\pi q E_0 T \sin(-\phi_s)}{mc^2 \lambda \beta^3 \gamma^3} \qquad \text{\textbf{smooth approximation for}} \atop \text{\textbf{longitudinal focusing}} \qquad (2.21)$$

For the quadrupole focusing forces one has to specify the type of focusing lattice. Here we will use a standard FODO lattice, consisting of focusing quadrupole, drift, defocusing quadrupole, and drift without acceleration to obtain (compare Appendix A.3):

$$\overline{\kappa_{x,Q}(s)} = \overline{\kappa_{y,Q}(s)} \longrightarrow k_Q^2 = \left(\frac{qGl_Q}{2mc\beta\gamma}\right)^2 \qquad \text{\textbf{smooth approximation for}} \atop \text{\textbf{transverse focusing in FODO}} \atop \text{\textbf{lattices w/o RF}} \qquad (2.22)$$

Adding the transverse defocusing from RF cavities located in the drifts, the transverse focusing constants become (compare Appendix A.4)

$$\begin{aligned} \overline{\kappa_x(s)} &= \overline{\kappa_y(s)} \longrightarrow k_{t,0}^2 \\ &\approx \left(\frac{qGl_Q}{2mc\beta\gamma}\right)^2 - \frac{\pi q E_0 T \sin(-\phi_s)}{mc^2 \lambda \beta^3 \gamma^3} \\ &= k_Q^2 - \frac{k_{l,0}^2}{2} \end{aligned} \qquad \text{\textbf{smooth approximation for}} \atop \text{\textbf{transverse focusing in FODO}} \atop \text{\textbf{lattices with RF}} \qquad (2.23)$$

2.4 Smooth approximation for a FODO channel with RF cavities

Starting with the 3D envelope equations with space-charge (same as Eq. 2.18)

$$
\begin{aligned}
a_{\mathrm{x}}'' + a_{\mathrm{x}} \kappa_{\mathrm{x}}(s) - \frac{K_3[1 - f(s)]}{2 a_{\mathrm{y}} b} - \frac{\varepsilon_{\mathrm{x}}^2}{a_{\mathrm{x}}^3} &= 0 \\
a_{\mathrm{y}}'' + a_{\mathrm{y}} \kappa_{\mathrm{y}}(s) - \frac{K_3[1 - f(s)]}{2 a_{\mathrm{x}} b} - \frac{\varepsilon_{\mathrm{y}}^2}{a_{\mathrm{y}}^3} &= 0 \\
b'' + b \kappa_{\mathrm{z}}(s) - \frac{K_3 f(s)}{a_{\mathrm{x}} a_{\mathrm{y}}} - \frac{\varepsilon_{\mathrm{z}}^2}{b^3} &= 0
\end{aligned}
$$

3D envelope equations with space-charge (2.24)

one can use the smooth approximation [Eq. (2.20)] to remove the s-dependence of the focusing forces.

$$
\begin{aligned}
a_{\mathrm{x}}'' + a_{\mathrm{x}} k_{\mathrm{t}}^2 - \frac{\varepsilon_{\mathrm{x}}^2}{a_{\mathrm{x}}^3} &= 0 \\
a_{\mathrm{y}}'' + a_{\mathrm{y}} k_{\mathrm{t}}^2 - \frac{\varepsilon_{\mathrm{y}}^2}{a_{\mathrm{y}}^3} &= 0 \\
b'' + b k_{\mathrm{l}}^2 - \frac{\varepsilon_{\mathrm{z}}^2}{b^3} &= 0
\end{aligned}
$$

smooth approximation for the 3D envelope equations with space-charge in a FODO lattice with RF (2.25)

The focusing constants are defined as [see Eq. (2.20)]

$$
\begin{aligned}
k_{\mathrm{t}}^2 &= k_{\mathrm{t,0}}^2 - \overline{\frac{K_3[1 - f(s)]}{2 a_{\mathrm{x}} a_{\mathrm{y}} b}} = \left(\frac{\sigma_{\mathrm{t}}}{L_{\mathrm{P}}} \right)^2 \\
k_{\mathrm{l}}^2 &= k_{\mathrm{l,0}}^2 - \overline{\frac{K_3 f(s)}{2 a_{\mathrm{x}} a_{\mathrm{y}} b}} = \left(\frac{\sigma_{\mathrm{l}}}{L_{\mathrm{P}}} \right)^2
\end{aligned}
$$

smooth approximation with space-charge (2.26)

As stated in the previous section the zero-current constants of a FODO quadrupole lattice are given by

$$
k_{\mathrm{l,0}}^2 = \frac{2\pi q E_0 T \sin(-\phi_{\mathrm{s}})}{m c^2 \lambda \beta^3 \gamma^3} \qquad k_{\mathrm{t,0}}^2 = \left(\frac{q G l_{\mathrm{Q}}}{2 m c \beta \gamma} \right)^2 - \frac{k_{\mathrm{l,0}}^2}{2}
$$

(2.27)

To calculate the matched r.m.s. beam envelopes one simply sets the second derivations to zero and obtains

$$
\begin{aligned}
\hat{a}_{\mathrm{x,y}} &= \sqrt{\frac{\varepsilon_{\mathrm{t}}}{k_{\mathrm{t}}}} = \sqrt{\frac{\varepsilon_{\mathrm{t}} L_{\mathrm{P}}}{\sigma_{\mathrm{t}}}} \\
\hat{b} &= \sqrt{\frac{\varepsilon_{\mathrm{l}}}{k_{\mathrm{l}}}} = \sqrt{\frac{\varepsilon_{\mathrm{l}} L_{\mathrm{P}}}{\sigma_{\mathrm{l}}}}
\end{aligned}
$$

smooth approximation for the matched full-current r.m.s. beam envelopes in a FODO lattice with RF (2.28)

In the zero-current case the space-charge terms vanish and the solution is found as

$$
\begin{aligned}
\hat{a}_{x,y,I=0} &= \sqrt{\frac{\varepsilon_t}{k_{t,0}}} = \sqrt{\frac{\varepsilon_t L_P}{\sigma_{t,0}}} \\
\hat{b}_{I=0} &= \sqrt{\frac{\varepsilon_l}{k_{l,0}}} = \sqrt{\frac{\varepsilon_l L_P}{\sigma_{l,0}}}
\end{aligned}
$$

smooth approximation for the matched zero-current r.m.s. beam envelopes in a FODO lattice with RF   (2.29)

We note that if the beam is matched Eq. (2.25) will describe a beam with constant beam envelopes in all three planes. Only in the case of mismatch will Eq. (2.25) describe a beam with changing r.m.s. envelopes. One can argue that this is a drastic simplification of the actual beam transport where a periodic lattice keeps the beam oscillating. On the other hand, having basically removed the lattice oscillations from the equations, one can now use them to determine the characteristics of mismatch oscillations in a relatively simple way. In the next chapter multi-particle simulations will be used to study the effects of mismatch and there we will plot the ratio of mismatched over matched r.m.s. beam envelopes to compare the results with the predictions from a simplified analytical model, the basis of which was derived in this chapter.

3. Multi-particle simulations with the IMPACT code

In the first stage of the SPL design, the standard beam dynamics code PARMILA [33] was used to establish a first layout [34, 35]. PARMILA is used since the 1960s for the design of linacs and was still used as the major design code for the SNS project [3]. PARMILA combines the creation of the accelerator layout with particle tracking and thus makes it difficult to isolate specific lattice characteristics. Furthermore, when the SPL study started, PARMILA only offered a 2D space-charge routine which assumes rotational symmetry of the beam. Another limitation of PARMILA was that only 10^5 particles could be used. While this number is perfectly adequate to calculate the r.m.s. quantities of a beam it is not enough to study detailed halo dynamics (see next section). In searching for a more precise particle tracking tool, subsequent simulations were done with IMPACT which is described in the following. IMPACT, however, was developed as a "verification code" rather than a "design code" which means that it lacks all features that make it easy for the user to build a linac with several acceleration sections. In order to use IMPACT for design improvements and simulations of the complete SPL linac, a number of tools had to be developed which are described in Section 3.2.

3.1   The IMPACT code

The IMPACT [13] code was originally developed at the Los Alamos National Laboratory (LANL) by R. Ryne and J. Qiang and is now largely maintained by J. Qiang at the Lawrence Berkeley National Laboratory (LBNL). IMPACT stands for **I**ntegrated **M**ap and **P**article **AC**celerator **T**racking Code. The design goal was to provide the accelerator community with a fast and reliable multi-particle code, capable of using the resources of modern parallelised supercomputers. With up to several hundred processors working in parallel it becomes possible to use large (up to 256x256x256 in routine runs) space-charge grids together with a high number of particles (up 10^8 in routine runs) within reasonable time frames of typically a few hours.

One of the basic ideas in the code development was to use symplectic split-operator methods which allow to include space-charge effects in single-particle beam dynamics codes. The space-charge forces are then computed using time-efficient parallelised Particle In Cell (PIC) techniques which were largely developed by the plasma physics community. The single-particle Hamiltonian can be written as:

$$H = H_{\text{ext}} + H_{\text{sc}} \tag{3.1}$$

where H_{ext} refers to the external focusing forces of quadrupoles, RF cavities, etc. and H_{sc} stands for the space-charge forces of the beam. The Hamiltonians for standard beam line elements can be derived from standard magnetic optics and are listed in [13]. To calculate the space-charge potential, the charge is deposited on the grid and the potential is obtained by convolving the charge density with a Green's function.

Once the Hamiltonians are known, the mapping M_{ext} and M_{sc} can be computed, which corresponds to H_{ext} and H_{sc}. For 2nd order accuracy one can then use the following algorithm to advance the particles

$$M(\tau) = M_{ext}(\tau/2)M_{sc}(\tau)M_{ext}(\tau/2). \tag{3.2}$$

In the standard version of IMPACT τ equals the longitudinal coordinate z. The symplectic split operator method can easily be generalised to higher order but then needs more space-charge calculations per full step. Since the space-charge calculation dominates the execution time, this option should only be used for highly space-charge dominated beams.

A standard integration step thus involves: 1) transport of a numerical particle distribution through a half step based on M_{ext}, 2) solution of Poisson's equation based on the particle positions and performance of a space-charge kick based on M_{sc} which only affects the particle moments, and 3) transport through the 2nd half step. For intense space-charge this sequence can be used repeatedly on successive sections of one beam-line element. For weak space-charge forces, one can achieve good accuracy by including several beam line elements in one half step. The split operator method thus de-couples the rapid variation of external focusing forces from the slowly varying space-charge fields and allows to adapt the step size for each calculation separately.

Another feature of the IMPACT code is the use of on-axis RF field maps for all accelerating cells which allow a precise calculation of longitudinal focusing and acceleration as well as a good approximation of the transverse defocusing.

The justification for using large numbers of particles is illustrated in Figs. 3.1 to 3.3. A high number of particles provides a high resolution in the outermost areas of phase space, where halo formation takes place and where particles are lost on the beam pipe. It is quite obvious that a resolution of 10^{-6}, which is required for linacs with an output power in the MW range, can only be achieved with particle numbers of at least 10^{7} or even above. In Fig. 3.1 the transverse particle density at the end of a linac is shown for simulations with various particle numbers.
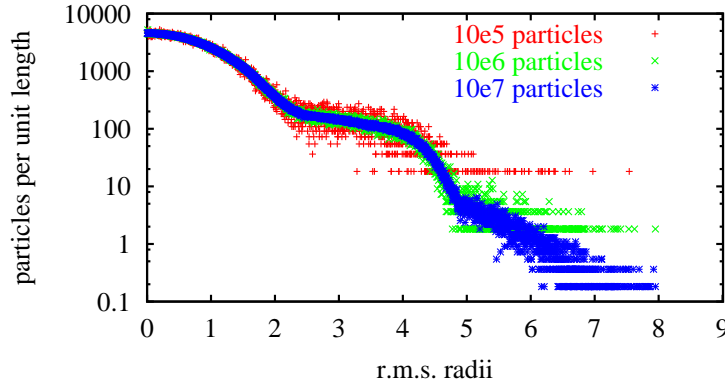
Figure 3.1: Radial particle density (in arbitrary units) at the end of a linac simulated with different particle numbers, IMPACT multi-particle simulation

One can see that the simulation with 10^{5} particles shows a double-Gaussian profile, which resembles the superposition of two Gaussian distributions. Using 10^{7} particles shows that a double-Gaussian does not correctly describe the tails in the distribution but that the outermost particles rather follow a third distribution type superimposed on the double-Gaussian.

Figure 3.2 shows an interesting example for the resolution that can be obtained with high numbers of particles [36]. The graph shows the halo distribution in longitudinal phase space for a mismatched beam. The distribution was calculated using 25 million particles on a 256^3 space-charge grid. Out of this distribution 10^5 particles were chosen randomly for the plot. In the core region (blue) the density changes from 1 to 1/10 at the edge of the blue area. In the subsequent green, yellow, and red areas the density changes from $1/10$ to $1/10^2$, to $1/10^3$, and to $1/10^4$. The figure also illustrates the need for suitable post-processing of the simulation results which can significantly facilitate the interpretation of the results.
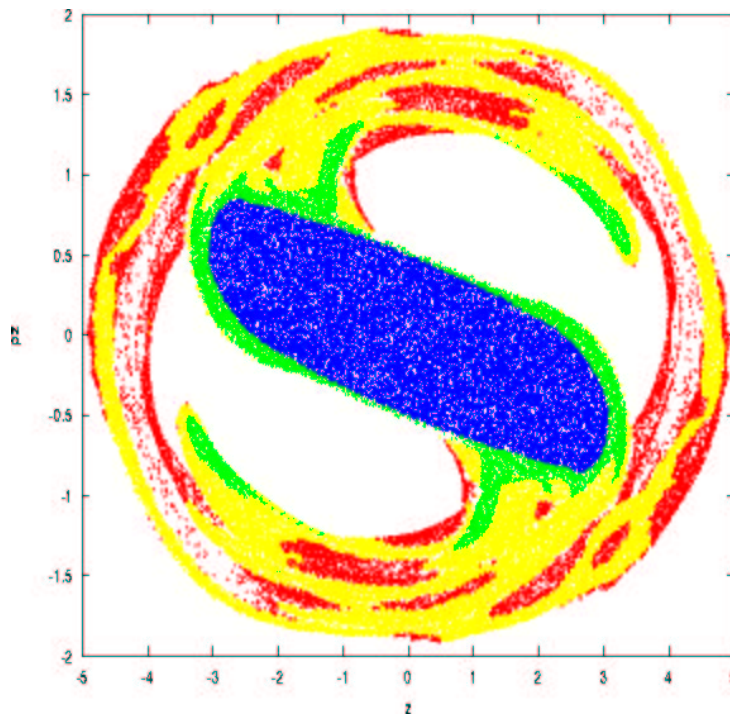
Figure 3.2: Post-processed data showing the longitudinal halo development for a mismatched beam, IM-PACT multi-particle simulation, source: R.D. Ryne

For the evaluation of statistical errors one rather uses many simulations with fewer particles in order to estimate the average particle loss due to errors. Also here one can profit from the short simulation times of IMPACT by using for instance 100 simulations with 10^5 particles rather than one simulation with 10^7 particles. As an example Fig. 3.3 shows the losses per beam line element in the drift tube linac of the ESS. The DTL is simulated with two different beam pipe over aperture ratios assuming a 1% randomly distributed quadrupole gradient error. Using a high number of error runs one can make very precise predictions on beam loss, assuming that the input distribution of the simulated lattice is known in sufficient detail. As demonstrated in Fig. 3.3 one can for instance make predictions on possible "hot spots" in a linac design.

Finally, Table 3.1 shows some typical simulation times for Linac4 and the SPL using a Linux PC cluster at Rutherford and an IBM supercomputer at NERSC. Linac4 has a length of 67 m and consists of 270 beam line elements while the SPL has a length of 580 m and consists of 460 elements.

Table 3.1: Simulation time for Linac4 and SPL on a linux PC cluster† or on the NERSC IBM SP RS/6000 supercomputer*

machine	space-charge grid	$N_{\text{particles}}$	$N_{\text{processors}}$	cpu time
linac4†	32^3	$5 \cdot 10^4$	2	25 min
linac4†	64^3	10^6	4	220 min
SPL†	128^3	10^7	8	1130 min
SPL*	64^3	10^6	64	40 min
SPL*	64^3	10^7	128	70 min
SPL*	128^3	10^7	64	430 min
SPL*	128^3	10^7	128	285 min
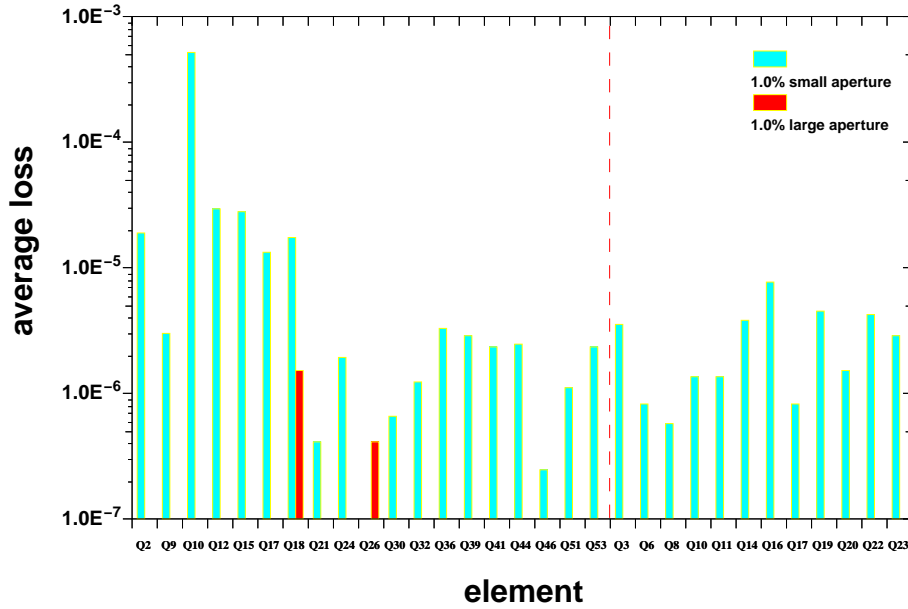
Figure 3.3: Average transverse losses in the ESS DTL for two different beam pipe over aperture ratios (4.5/6.5) assuming a 1% statistical quadrupole gradient error [37], IMPACT multi-particle simulation

3.2   Using IMPACT

In order to use IMPACT the input files have to be prepared by means of several "helper codes". Additionally it may be necessary to fine-tune the lattice that has been prepared in a first step with an envelope code like TRACE3D [38]. This is due to the fact that a beam that is matched with an envelope code is not necessarily matched for the use with a multi-particle code, especially if high space-charge forces are present. Different methods of transporting the beam through a lattice will results in slightly different matched beams and thus one may have to re-match the initial beam (and the lattice around structure transitions) when using IMPACT. For the simulations in this thesis a number of small scripts were written

to simplify the lattice set-up and the evaluation of results from IMPACT simulations. The idea was to keep the IMPACT code package (IMPACT, FIX3D, THETA) unchanged and to use scripts to start and evaluate the codes of the IMPACT package. This approach has the advantage that in case of changes in the IMPACT codes (bug fixes, implementation of new features), the scripts can be easily adapted to new formats of the input and output files.

The code asks for an element-by-element lattice description and does not use any graphical user interfaces. One of the interesting features of IMPACT is for example the exact integration over the on-axis electric fields in accelerating cavities, while most of the standard codes simplify this integration to a drift-kick-drift approach. In order to perform this integration correctly, however, the code needs as input the RF phase at the point when the synchronous particle enters the cavity (and the on-axis field map), rather than just the synchronous phase and field amplitude in the centre of the cavity. The first approach (by R.D. Ryne) to calculate this phase value was to use the fitting capabilities of the beam dynamics code MARYLIE [39]. Later on these preparatory calculations were simplified by J. Qiang who wrote the "THETA" code which calculates the correct phase settings for the different types of accelerating cavities and which uses the same lattice description format as IMPACT.

To improve the calculation of matched beams for IMPACT, the 3D envelope matching code FIX3D [40] was written (R.D. Ryne). It uses the same cavity field integration as IMPACT and thus produces a much better beam matching than TRACE3D. FIX3D also calculates the zero-current and full-current phase advance for the matched lattice period. This capability was then used by a script, written to calculate the phase advance evolution for a complete linac and to change the magnet settings in the lattice in order achieve a desired phase advance profile. This approach was used for the simulations in Section 6.1 where a certain ratio of longitudinal to transverse full-current phase advance is needed.

Together with R.D. Ryne a transition matching system was developed using again the fitting capabilities of MARYLIE and the FIX3D code. With this system the beam was propagated with FIX3D up to a certain lattice point before the transition. Then the matched beam was calculated for the first period after the transition and compared with the previously calculated beam that ended before the transition. Using the fitting routines of MARYLIE, the lattice before the transitions was changed until the difference between matched and propagated beam became minimal. With this approach IMPACT could be used to simulate complete linacs consisting of several different sections [41].

To run and evaluate a large number of simulations needed for the statistical studies presented in Chapter 5, several (python and bash) scripts were written and combined with fortran codes, to automatise this process and to perform error studies with only a few small input files. The input files specify the type of errors (quadrupole gradients, RF phases, etc.), their amplitude, their distribution type (uniform, Gaussian with or without cut-off at a certain multiple of the r.m.s value, parabolic), whether the errors occur in groups (e.g. all RF phases of one RF tank) or should be different for each element, and the number of simulations. A new error set is then created for each simulation and written into a new directory, from where the job is submitted to a linux cluster. After the completion of all runs (which is also monitored by a script) a small Fortran code is used to read the simulation output from each directory and to evaluate the results in terms of r.m.s. emittance growth, maximum envelope deviations, particle loss, energy and phase jitter, etc.

4. Initial mismatch

4.1   Space-charge and beam stability

In the design of linacs using hadron beams for high-intensity applications one has to take into account the effect of space-charge forces and all related beam instabilities. Some of these instabilities are known since many years and are well documented in text books on accelerator physics (e.g. [32]). One of the classic rules in hadron linac design is derived from perturbation studies of the envelope equations (e.g. [42]) which show that envelope instabilities (in beams with space-charge) occur for zero-current phase advance (σ_0) values above $90°$. As a consequence, lattices for high-intensity beams are generally designed to operate below $90°$ in all three planes. Below this classic threshold beam halo and emittance exchange are considered to be two of the most detrimental effects with respect to beam quality. The latter is triggered by coherent core-core resonances and will be treated in in detail in Chapter 6.

The development of a diffused beam halo is mainly triggered by single particle resonances, which can occur between single particles and either the periodicity of the lattice or the envelope oscillations of an r.m.s. mismatched beam core. For $\sigma_0 < 90°$, however, only the particle-envelope (particle-core) resonances are of real importance.

For both resonance types (core-core and particle-core) space-charge plays an important role. In case of the single particle resonances space-charge provides a large spread of single particle tunes and so increases the probability for single particles to enter into a resonance with the core. For core-core resonances space-charge increases the width of the stop bands, which finally, for very intense beams, yield a "sea of instability" in which no stable working point can be found for non-equipartitioned beams.

In the following sections and chapters the theoretical models for both resonances are discussed and tested with 3D tracking simulations of realistic linac lattices (3D in real space or 6D in phase space). The particle-core model is used to explain the basic mechanism for particle-core resonances. Then the three eigenfrequencies for a 3D beam envelope are approximated and used to excite mismatch oscillations for 3D tracking studies. The particle-core model is then extended to study statistical gradient errors (Chapter 5) and to show that even these random errors yield beam halo via particle-core resonances. Finally we consider the effect of core-core resonances and we discuss their occurrence in realistic designs (Chapter 6).

4.2   The particle-core model

The particle-core model is a simple means to study how single particles are influenced by oscillations of the beam core. A bunch of particles is treated as a "blob" of charge with distinct boundaries, which then interacts with a single particle crossing the core.

4.2.1 Introduction

The particle-core model described here follows the approach taken in [43] to derive maximum amplitudes for parametric beam halo. To study the principle of halo development it is sufficient to use the simplest description of the beam. An azimuthally symmetric continuous beam in a uniform focusing channel is described using the smooth approximation of the envelope equations. As outlined in Chapter 2 one averages over the external forces of a periodic channel and simplifies the space-charge calculation by assuming a uniform spatial distribution of the particles within the core. Using the derivations presented in Chapter 2 one can write the one-dimensional transverse equation of motion for the core-radius r_c as:

$$r_c'' + k_0^2 r_c - \frac{\varepsilon^2}{r_c^3} - \frac{K_1}{r_c} = 0 \tag{4.1}$$

where k_0 represents the transverse focusing forces as well as the zero-current phase advance per unit length. K_1 is a 1D space-charge constant containing mass, axial velocity, charge, and the number of core particles. Together with k_0 it defines the depressed phase advance per unit length ($k^2 = k_0^2 - \frac{K_1}{r_0^2}$). In this case ε is the *total* un-normalised emittance (as opposed to the un-normalised *r.m.s.* emittance used in Chapter 2). For a matched beam the core radius remains constant ($\Rightarrow r_c'' = 0$) and becomes:

$$r_c^2 = r_0^2 = \left(\frac{\varepsilon}{k_0}\right)\left[u + \sqrt{1 + u^2}\right] \text{ with } u = \frac{K_1}{2\varepsilon k_0} \tag{4.2}$$

For a single particle, moving transversely in the field of the core, the transverse equation of motion can be written as:

$$x'' + k_0^2 x - F_{sc} = 0 \tag{4.3}$$

where F_{sc} contains the space-charge forces acting on the single particle at the transverse position x. In contrast to the core radius r_c which is by definition symmetric in $(x, -x)$ and $(y, -y)$, the single particle particle will assume positive and negative values in the transverse plane, which is why the nomination x was chosen. Assuming the core as an infinitely long cylinder of uniform charge, particles experience linear forces inside the core and non-linear forces when leaving the core area:

$$F_{sc} = \begin{cases} K_1 x/r_c^2 & : & |x| < r_c \\ K_1/x & : & |x| \geq r_c \end{cases} \tag{4.4}$$

Outside the core, according to Gauss' law, the space-charge forces acting on the particles no longer depend on the actual size of the core. By using particles with different initial transverse positions in the range of $0 < x < n r_c$, the particle-core model is used to study the circumstances under which particles can be driven to large-amplitude oscillations. In all of the following simulations the initial transverse moments of the single particles are set to zero.

Compared to the equations derived in Chapter 2 we do not use r.m.s. quantities for calculations with the particle-core model. This is simply due to the fact that in Eq. (4.4) we need to establish when the particle is inside or outside the beam core, rather than when it is inside or outside the r.m.s. core radius. The variables used in this section are:

- $r_{\mathrm{c}} = x_{\mathrm{max}} = y_{\mathrm{max}}$ (100%) core radius or core envelope,

- r_0 (100%) matched core radius,

- x transverse single particle position,

- μ mismatch parameter: $\mu = r_{\mathrm{c}}/r_0$,

- τ tune depression: $\tau = k/k_0$,

- ε total (100%) emittance.

4.2.2   Initial mismatch

In the matched case the core radius as well as the amplitudes of the single particle oscillations remain constant. If the core is initially mismatched with $\mu = r_{\mathrm{c}}/r_0 \neq 1$, its radius starts to oscillate around its equilibrium value r_0. In the simple 1D model an initial envelope mismatch corresponds to the excitation of the breathing mode (fast mode), which is the only beam eigenmode for an azimuthally symmetric, continuous beam. Single particles traversing the core can obtain a net energy gain if the core radius is different for a particle when entering and leaving the core. As an example we integrate the equations of motion [Eq. (4.1), (4.3), and (4.4)] for a tune depression of $\tau = k/k_0 = 0.7$, a core mismatch of $\mu = 0.6$, and an initial single particle amplitude of $x(0) = 0.9r_0$. The parameters of the transport channel that is used in this section are listed in Table 4.1. For the integration the standard fourth-order Runge-Kutta algorithm with constant step size is used. The resulting core and single particle oscillations are shown in Fig. 4.1.
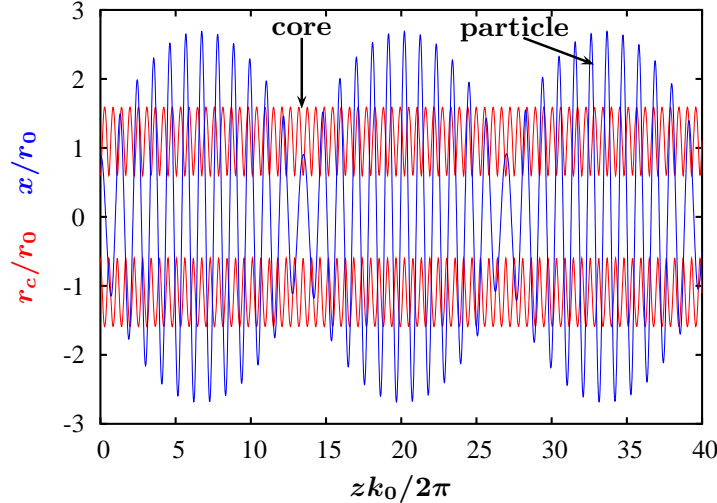
Figure 4.1: Particle and core envelopes for initial mismatch, $\mu = 0.6$, $x(0) = 0.9r_0$, $\tau = 0.7$

If the oscillation frequency of the single particle has a 1:2 parametric ratio with the oscillation frequency of the core, the particle gains energy and increases its oscillation amplitude until the resonant condition is no longer fulfilled. As a result one finds a certain maximum particle amplitude as well as so-called fixed-points in phase space around which halo particles start to conglomerate (compare stroboscopic

Table 4.1: Parameters for the simulated transport channels

tune depression	(τ)	0.5 - 0.9	
energy	(E)	11.4	MeV
focusing period	(l)	0.333	m
zero-current phase advance per period	(lk_0)	38.5	deg

plots in Fig. 4.2). The maximum halo extent is determined by the amount of initial mismatch, while the time constant for the development of beam halo is influenced by the tune depression. Wangler [43] also found that the position of the fixed-points is fairly insensitive to changes in tune depression or mismatch amplitude. In [44], however, it was shown that the fixed-point core distance is dependent on the equipartitioning ratio $(\varepsilon_t k_t / \varepsilon_l k_l)$. To visualise the fixed-points as well as the maximum halo amplitudes one can use stroboscopic plots showing the phase space position of the single particle once per core oscillation period. For this purpose we integrate Eqs. (4.1) and (4.3) and plot in Fig. 4.2 (a) and 4.2 (b) the transverse momentum and position of the particle every time the core oscillation reaches its minimum, which coincides with the occurrence of the maximum/minimum single particle amplitude (compare Fig. 4.1). To obtain a coherent resonance pattern a large number of focusing periods is needed (4000 in this example). Each dotted line represents a single particle with a certain initial amplitude and zero initial transverse momentum.
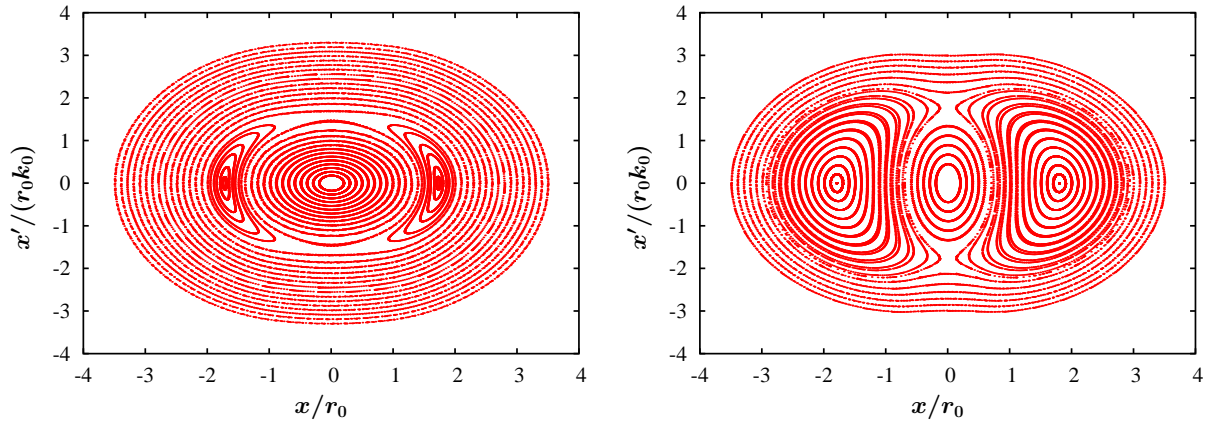
Figure 4.2: (a) Stroboscopic plot for $\mu = 0.95$, $0.2 \le x(0)/r_0 \le 3.5$, $\tau = 0.8$ (b) Stroboscopic plot for $\mu = 0.6$, $0.2 \le x(0)/r_0 \le 3.5$, $\tau = 0.8$

In the case of weak initial mismatch [e.g. 5%, Fig. 4.2 (a)] particles that start within the core ($|x|/r_0 \le 1$) remain basically undisturbed in their trajectory. They either need a non-zero initial momentum or an initial amplitude larger than the core radius to gain energy by means of a parametric resonance. In contrast, for strong initial mismatch [e.g. 40%, Fig. 4.2 (b)], the resonant regions surrounding the fixed-points become very large and the core area with undisturbed trajectories shrinks considerably. In both cases, for strong and weak mismatch particles with very large initial amplitudes (e.g. $|x|/r_0 \approx 3$), remain as unperturbed in their orbits as the inner core particles. This can be explained by the fact that in order to gain or to loose energy, the single particles must be in resonance (1:2 parametric ratio) with the core

oscillations. Due to the space-charge forces single particles with small amplitudes (within the core) have different oscillation frequencies than particles with larger amplitudes. Only the particles that can gain enough energy via a parametric resonance with the core will change their trajectories and will be "caught" by the resonance. The effect becomes stronger for high space-charge forces, providing larger "kicks" to the single particles. It becomes also stronger for larger amplitudes of the core oscillations. In that case the oscillation energy, stored in the core oscillations is large, meaning that the single particles can receive large "kicks" to change their orbits.

This classic 1D particle-core model predicts maximum halo amplitudes of $\approx 3r_0$. In terms of multiples of r.m.s. beam envelopes (in 3D) this corresponds to $\approx 7a$. It is worth noting that particles which are initially outside the core and get into resonance with the core oscillations have the same maximum halo amplitudes as particles starting inside the core. Studies with a more sophisticated 3D particle-core model [45] show that the maximum halo extent due to initial mismatch can amount to $> 5r_0$ if the longitudinal to transverse focusing ratio is larger than one.

4.3 3D envelope eigenmodes

In the previous section it was shown that initial mismatch yields the excitation of regular core oscillations with a fixed frequency. For a 3D bunched beam one can expect to find three such eigenmodes, which are likely to have different oscillation frequencies. Since the 3D envelope equations are coupled, one can also expect that each mode has different oscillation amplitudes in all three planes. Using the 3D envelope equations [Eq. (2.18)] and the smooth approximation for a FODO channel with RF, one can derive and analyse the eigenmodes of a 3D bunched beam. In the following the approach of Bongardt & Pabst (presented in [46] and [47]) is used to derive an approximation for the three eigenfrequencies of a mismatched beam. Similar, more exhaustive derivations are given in [48] and [49]. The derived formulae will then be used to excite the eigenmodes and to systematically study the effects of initial mismatch with a particle tracking code.

Starting point are the 3D envelope equations [see Eq. 2.18] which were derived in Chapter 2:

$$
\begin{aligned}
a_{\mathrm{x}}'' + a_{\mathrm{x}}\kappa_{\mathrm{x}}(s) - \frac{K_3[1-f(s)]}{2a_{\mathrm{y}}b} - \frac{\varepsilon_{\mathrm{x}}^2}{a_{\mathrm{x}}^3} &= 0 \\
a_{\mathrm{y}}'' + a_{\mathrm{y}}\kappa_{\mathrm{y}}(s) - \frac{K_3[1-f(s)]}{2a_{\mathrm{x}}b} - \frac{\varepsilon_{\mathrm{y}}^2}{a_{\mathrm{y}}^3} &= 0 \\
b'' + b\kappa_{\mathrm{z}}(s) - \frac{K_3 f(s)}{a_{\mathrm{x}}a_{\mathrm{y}}} - \frac{\varepsilon_{\mathrm{z}}^2}{b^3} &= 0
\end{aligned}
\qquad
\begin{array}{l}\textbf{3D envelope equations with}\\ \textbf{space-charge}\end{array}
\qquad (4.5)
$$

For these equations one can define a general matched s-dependent solution which has the same periodicity as the focusing system

$$
\begin{aligned}
\tilde{a}_{\mathrm{x,y}}(s + L_{\mathrm{P}}) &= \tilde{a}_{\mathrm{x,y}}(s) \\
\tilde{b}(s + L_{\mathrm{P}}) &= \tilde{b}(s)
\end{aligned}
\qquad \textbf{matched solution} \qquad (4.6)
$$

The mismatched solution will oscillate around the matched one with a periodicity different from the length L_{P} of the focusing periods (compare Fig. 4.1). Starting point of the derivation is the perturbation of the s-dependent matched beam envelopes:

$$
\begin{aligned}
a_{\mathrm{x,y}}(s) &= \tilde{a}_{\mathrm{x,y}}(s) + \Delta a_{\mathrm{x,y}}(s) \\
b(s) &= \tilde{b}(s) + \Delta b(s)
\end{aligned}
\qquad \textbf{mismatched solution} \qquad (4.7)
$$

The perturbed solution is then re-inserted into the envelope equations (4.5). For the sake of simplicity we will assume equal transverse emittances ($\varepsilon_{\mathrm{x}} = \varepsilon_{\mathrm{y}} = \varepsilon_{\mathrm{t}}$). Further simplification is achieved by using the smooth approximation for the zero-current focusing terms [$\overline{\kappa_{\mathrm{x,y}}(s)} \to k_{\mathrm{t,0}}^2$, $\overline{\kappa_{\mathrm{z}}(s)} \to k_{\mathrm{l,0}}^2$, compare (2.19)] and by ignoring the changing of the bunch shape along the transport channel [$f(s) \to \overline{f(s)}$]. We also use the smooth approximation to replace the s-dependent periodic solutions $\tilde{a}_{\mathrm{x,y}}(s)$ and $\tilde{b}(s)$ by the matched (constant) beam sizes $\hat{a}_{\mathrm{x}} = \hat{a}_{\mathrm{y}} = \hat{a}$ and $\hat{b}$. After some further manipulation (see Appendix C) one can derive three eigenmodes which are usually referred to as follows:

quadrupolar mode:

$$
\sigma_{\mathrm{env,Q}} = 2 \cdot \sigma_{\mathrm{t}} \qquad (4.8)
$$

with the eigensolution:

$$
\begin{aligned}
\frac{\Delta a_{\mathrm{x}}(s)}{\hat{a}} &= A_{\mathrm{m}} \cdot \cos\left(\sigma_{\mathrm{env,Q}} \cdot \frac{s}{L_{\mathrm{P}}} + \phi\right) \\
\frac{\Delta a_{\mathrm{y}}(s)}{\hat{a}} &= -A_{\mathrm{m}} \cdot \cos\left(\sigma_{\mathrm{env,Q}} \cdot \frac{s}{L_{\mathrm{P}}} + \phi\right) \\
\frac{\Delta b(s)}{\hat{b}} &= 0
\end{aligned}
\qquad (4.9)
$$

high-frequency mode or **breathing mode** or **fast mode:**

$$
\sigma_{\mathrm{env,H}} = \sigma_{\mathrm{t,0}}^2 + \sigma_{\mathrm{t}}^2 + \frac{1}{2}\sigma_{\mathrm{l,0}}^2 + \frac{3}{2}\sigma_{\mathrm{l}}^2 + \qquad (4.10)
$$
$$
\sqrt{\left(\sigma_{\mathrm{t,0}}^2 + \sigma_{\mathrm{t}}^2 - \frac{1}{2}\sigma_{\mathrm{l,0}}^2 - \frac{3}{2}\sigma_{\mathrm{l}}^2\right)^2 + 2\cdot\left(\sigma_{\mathrm{t,0}}^2 - \sigma_{\mathrm{t}}^2\right)\cdot\left(\sigma_{\mathrm{l,0}}^2 - \sigma_{\mathrm{l}}^2\right)}
$$

with the eigensolution:

$$
\begin{aligned}
\frac{\Delta a_{\mathrm{x,y}}(s)}{\hat{a}} &= A_{\mathrm{m}} \cdot \cos\left(\sigma_{\mathrm{env,H}} \cdot \frac{s}{L_{\mathrm{P}}} + \phi\right) \\
\frac{\Delta b(s)}{\hat{b}} &= \frac{A_{\mathrm{m}}}{g_{\mathrm{H}}} \cdot \cos\left(\sigma_{\mathrm{env,H}} \cdot \frac{s}{L_{\mathrm{P}}} + \phi\right) \\
g_{\mathrm{H}} &> 0
\end{aligned}
\qquad (4.11)
$$

low-frequency mode or **slow mode:**

$$\sigma_{\text{env,L}} = \sigma_{\text{t,0}}^2 + \sigma_{\text{t}}^2 + \frac{1}{2}\sigma_{\text{l,0}}^2 + \frac{3}{2}\sigma_{\text{l}}^2 - \tag{4.12}$$
$$\sqrt{\left(\sigma_{\text{t,0}}^2 + \sigma_{\text{t}}^2 - \frac{1}{2}\sigma_{\text{l,0}}^2 - \frac{3}{2}\sigma_{\text{l}}^2\right)^2 + 2 \cdot \left(\sigma_{\text{t,0}}^2 - \sigma_{\text{t}}^2\right) \cdot \left(\sigma_{\text{l,0}}^2 - \sigma_{\text{l}}^2\right)}$$

with the eigensolution:

$$\begin{aligned}
\frac{\Delta a_{\text{x,y}}(s)}{\hat{a}} &= A_{\text{m}} \cdot \cos\left(\sigma_{\text{env,L}} \cdot \frac{s}{L_{\text{P}}} + \phi\right) \\
\frac{\Delta b(s)}{\hat{b}} &= \frac{A_{\text{m}}}{g_{\text{L}}} \cdot \cos\left(\sigma_{\text{env,L}} \cdot \frac{s}{L_{\text{P}}} + \phi\right) \\
g_{\text{L}} &< 0
\end{aligned} \tag{4.13}$$

The form factors g_{H}, g_{L} are defined as

$$g_{\text{H,L}} = \frac{\sigma_{\text{t,0}}^2 - \sigma_{\text{t}}^2}{\sigma_{\text{env,H,L}}^2 - 2 \cdot \left(\sigma_{\text{t,0}}^2 + \sigma_{\text{t}}^2\right)} \tag{4.14}$$

A_m is the mismatch amplitude, and $\hat{a}$ and $\hat{b}$ are the matched beam envelopes in the smooth approximation. The nomenclature of the modes becomes clear when considering the characteristics of each mode:

- The quadrupolar mode consists of envelope oscillations of the transverse beam envelopes around their matched equilibrium with $180°$ phase difference between the planes. The longitudinal plane is unaffected (example in Fig. 4.3).

- The fast mode "breathes" in all three planes with the same phase but with different amplitudes in the transverse and longitudinal planes. From the eigenvalues one can see that its oscillation frequency is always higher than that of the slow and quadrupolar mode (example in Fig. 4.4).

- The slow mode also has different amplitudes in the transverse and longitudinal planes. The longitudinal and transverse envelopes oscillate with a phase difference of $180°$ (example in Fig. 4.5).

Since the three eigenmodes were derived with some strong simplifications, the solutions are only applicable within certain limits. For high space-charge forces or rapidly changing particle velocities, the derived excitation becomes less precise. Its quality can be judged by the smoothness of the ratio of mismatched over matched r.m.s. beam envelopes. An improvement can be achieved by measuring the envelope tunes in the simulation output and recalculating the mismatch excitation with these tunes. An even better excitation can be obtained by numerical computation of the envelope tunes.

Excitation of mismatch eigenmodes

Going back from the smoothed r.m.s. beam envelopes to the s-dependent envelopes ($\hat{a} \rightarrow \tilde{a}_{\text{x,y}}(s)$) one can excite the eigenmodes using the eigensolutions [Eqs. (4.9), (4.11), (4.13)] and the form factors

[Eq. (4.14)]. The matched Twiss parameters at the beginning of a focusing period ($s = 0$) are $\tilde{\alpha}_{x,y,z}(0)$ and $\tilde{\beta}_{x,y,z}(0)$. The goal is to modify α and β such that the mode envelope maximum is located at $s = 0$. The modified α and β values can then be used in a simulation code to study the mismatched beam behaviour.

For the excitation one considers the eigensolutions and their derivations

$$
\begin{aligned}
a_{x,y}(s) &= \tilde{a}_{x,y}(s) \cdot \left[1 \pm A_m \cos\left(\sigma_{env}\frac{s}{L_P} + \phi\right)\right] \\
\frac{d}{ds}a_{x,y}(s) &= \frac{d}{ds}\tilde{a}_{x,y}(s) \cdot [1 \pm A_m \cos\left(\sigma_{env}\frac{s}{L_P} + \phi\right)] \mp \\
&\quad \tilde{a}_{x,y}(s)A_m\sigma_{env} \sin\left(\sigma_{env}\frac{s}{L_P} + \phi\right)
\end{aligned}
\tag{4.15}
$$

with A_m being the amplitude of the mismatch oscillation. The maximum mismatch envelope value at $s = 0$ is obtained by setting $\phi = 0$:

$$
a_{x,y}(0) = \tilde{a}_{x,y}(0) \cdot [1 \pm A_m] \qquad \frac{d}{ds}a_{x,y}(0) = \frac{d}{ds}\tilde{a}_{x,y}(0) \cdot [1 \pm A_m]
\tag{4.16}
$$

Obviously the envelope value and the momentum have to be changed by the same factor, and since

$$
\alpha = -\frac{aa'}{\varepsilon} \qquad\qquad \beta = \frac{a^2}{\varepsilon}
\tag{4.17}
$$

the Twiss parameters α and β also have to be changed by the same factor, which in case of the quadrupole mode means:

$$
\begin{aligned}
\alpha_{m,x}(0) &= (1 + A_m)^2\tilde{\alpha}_x(0) & \alpha_{m,y} &= (1 - A_m)^2\tilde{\alpha}_y(0) \\
\beta_{m,x}(0) &= (1 + A_m)^2\tilde{\beta}_x(0) & \beta_{m,y} &= (1 - A_m)^2\tilde{\beta}_y(0) \\
\alpha_{m,z}(0) &= \tilde{\alpha}_z(0) & \beta_{m,z} &= \tilde{\beta}_z(0)
\end{aligned}
\tag{4.18}
$$

Figure 4.3 shows an example for the excitation of the quadrupole mode using the above formulae. The 30% envelope mismatch is excited at the beginning of two sections of the SPL I lattice (120 - 380 MeV) which has FODO quadrupole focusing and using a 6D waterbag distribution[1]. The oscillations in the x and y planes are of opposite phase and have the same amplitude. The longitudinal plane remains almost unperturbed (matched) and only small oscillations can be observed. Even though the excitation is not perfect, meaning that the oscillation maxima are not perfectly constant, the oscillations carry on almost unchanged. The lengthening of the oscillation period is due to the changing phase advance per metre, which decreases with increasing energy.

The same principle can be used to derive the excitation for the fast and slow mode and Fig. 4.4 and 4.5 show the resulting mismatch oscillations using the same lattice and using again a maximum envelope mismatch of 30%.

One can observe that also in case of the fast and slow mode, the oscillation characteristics fulfil the theoretical predictions. For both modes the transverse oscillations are in phase and of the same amplitude.

---

[1]The 6D waterbag distribution assumes a uniform distribution in all 6 phase space dimensions (compare Section 4.5).
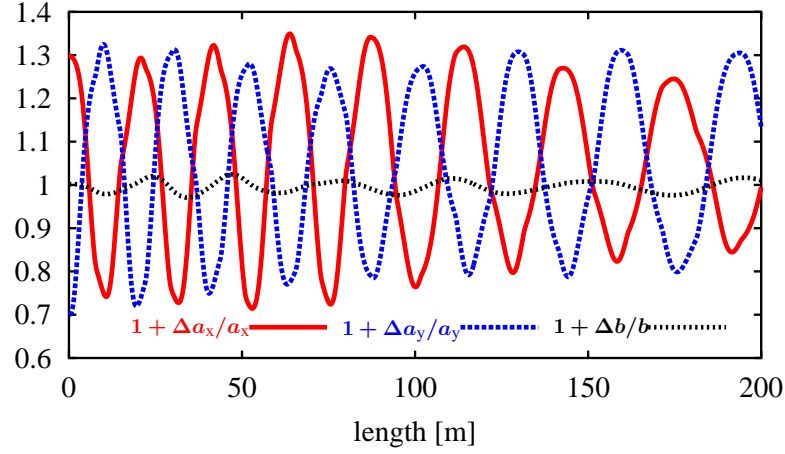
Figure 4.3: 30% quadrupole mode excitation with an initial 6D waterbag distribution in two sections of the SPL I SC lattice (120 - 380 MeV), plotted is the ratio of mismatched over matched r.m.s. beam envelopes, IMPACT multi-particle simulation
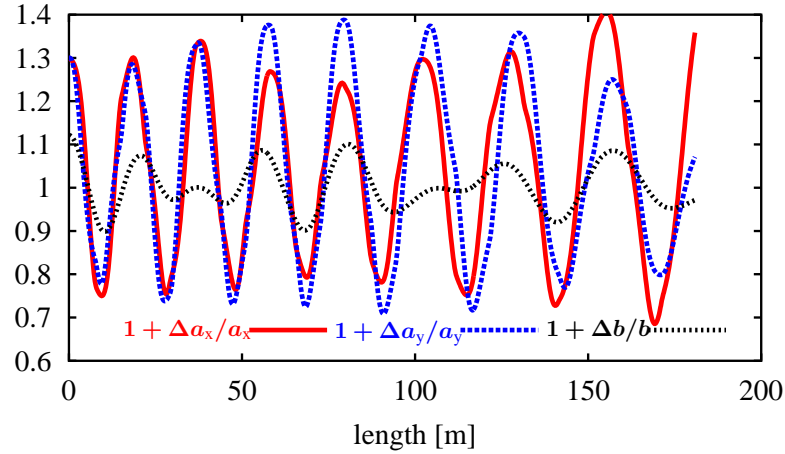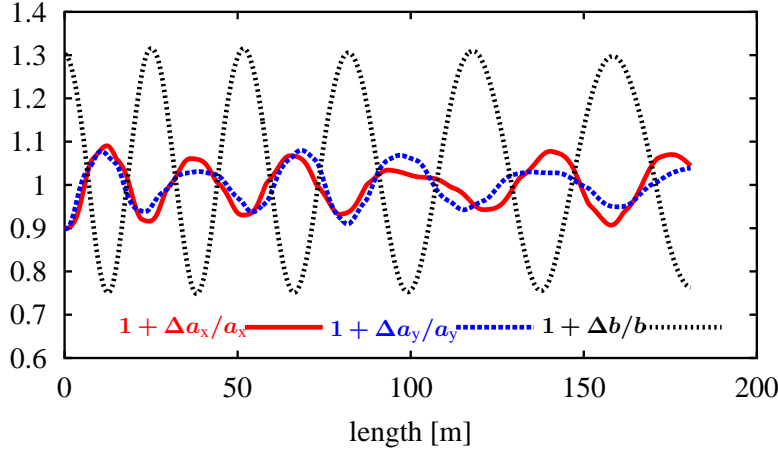

Figure 4.4: 30% fast mode excitation with an initial 6D waterbag distribution in two sections of the SPL I SC lattice (120 - 380 MeV), plotted is the ratio of mismatched over matched r.m.s. beam envelopes, IMPACT multi-particle simulation

The longitudinal oscillations have a different amplitude and in case of the slow mode also a different phase than in the transverse plane. Also here one can observe that the excitation is not perfect, especially in case of the fast mode. Despite the imperfection, the oscillations carry on basically undamped over the length of the lattice.

Figure 4.5: 30% slow mode excitation with an initial 6D waterbag distribution in two sections of the SPL I SC lattice (120 - 380 MeV), plotted is the ratio of mismatched over matched r.m.s. beam envelopes, IMPACT multi-particle simulation

4.4 The "free energy" limit for r.m.s. emittance growth

So far we have established that initial mismatch creates mismatch oscillations of the beam core and that these oscillations, via parametric resonances, are responsible for the development of beam halo and probably a certain r.m.s. emittance growth. The envelope equations assume a constant r.m.s. emittance and are thus not suited to describe the process of growing emittances. However, an estimation for r.m.s. emittance can be obtained from the "free energy" approach, which was introduced by Reiser [50]. Reiser uses a procedure that avoids the problem of frozen emittances in the envelope equations by stating that stationary (matched) and non-stationary (mismatched) beams have different average energies per particle. Borrowing an expression from thermodynamics he states that a matched beam is in its thermodynamic equilibrium which depends on the parameters of the beam, the focusing lattice and on the associated average particle energy W_i (which can be calculated out of the lattice and beam parameters). If the beam is mismatched, the associated particle energy changes to $W_f = W_i + \Delta W$ and the beam is no longer in its equilibrium. In this context the "excess energy" ΔW is called the "free energy" of the beam. He then assumes that the beam relaxes (or thermalises) into a new stationary distribution that corresponds to the increased average energy per particle W_f. By calculating the initial and final average particle energies (W_i and W_f) and by associating them to a beam described by the one-dimensional envelope equation (wave number k_0, emittance ε, space-charge constant K_1, beam radius r_c) it becomes possible to calculate the final emittance, circumventing the problem of constant emittances by excluding the transitional process from the calculation.

The thermalisation process can be triggered by non-linear space-charge forces, instabilities, or collisions, and causes emittance growth, increased beam radii, and possibly beam halo as the beam approaches a new stationary state at the higher average energy per particle. Reiser quotes three examples for non-stationary initial beams, which are: a) mismatch in the density profile, b) mismatch in the initial r.m.s. envelope, and c) off-centring.

In the following we will describe the three main steps of Reisers derivation to describe the emittance increase due to initial envelope mismatch. The complete calculation can be found in [42] or [50]. For the derivation Reiser uses an azimuthally symmetric beam in a uniform focusing channel using the same one-dimensional smoothed envelope equation [Eq. (4.1)] that was already used for the particle-core model in Section 4.2.

$$r_c'' + k_0^2 r_c - \frac{\varepsilon^2}{r_c^3} - \frac{K_1}{r_c} = 0 \tag{4.19}$$

If the initial beam is stationary the core radius is constant ($r_c = r_0$), the 2nd derivations vanish, and one can write Eq. (4.19) as

$$k_0^2 r_0 - \frac{\varepsilon^2}{r_0^3} - \frac{K_1}{r_0} = 0 \quad \text{or} \quad k^2 r_0 - \frac{\varepsilon^2}{r_0^3} = 0 \quad \text{or} \quad \varepsilon = k r_0^2 \tag{4.20}$$

where the relation

$$k^2 = k_0^2 - \frac{K_1}{r_0^2} \tag{4.21}$$

was employed.

For the subsequent derivation the following set of variables will be used:

- k_0 zero-current wave number,

- r_p beam pipe radius,

- r_0 general stationary beam radius,

- r_i, k_i, ε_i beam radius, full-current wave number, and emittance of an initially stationary beam,

- r_m beam radius of an initially mismatched beam,

- r_f, k_f, ε_f final beam radius, final full-current wave number, and final emittance after thermalisation.

Step I: average particle energy for a stationary beam

In the first step the average particle energy of a stationary beam is defined as the sum of its transverse kinetic energy W_k, its average potential energy due to the external focusing forces W_p, and the energy associated to its self fields W_s. For the latter it is necessary to introduce the radius r_p of the beam pipe. Summing up the three contributions Reiser derives the total energy per particle in a stationary beam as

$$W = \frac{\beta \gamma m c^2}{4} \left[r_0'^2 + k_0^2 r_0^2 + \frac{1}{2} K_1 \left(1 + 4 \ln \frac{r_p}{r_0} \right) \right] \tag{4.22}$$

Using Eq. (4.20) and (4.21) together with $r_0' = k_0 r_0^2$ one can write

$$W = \frac{\beta \gamma m c^2}{4} \left[(k^2 + k_0^2) r_0^2 + \frac{1}{2} (k_0^2 - k^2) r_0^2 \left(1 + 4 \ln \frac{r_p}{r_0} \right) \right] \tag{4.23}$$

The energy equation relating the final energy (index f), the initial energy (index i), and the "free energy" ($W_f = W_i + \Delta W$) can then be written as

$$
\begin{aligned}
\frac{\beta\gamma mc^2}{4} &\left[(k_f^2 + k_0^2)r_f^2 + \frac{1}{2}(k_0^2 - k_f^2)r_f^2 \left(1 + 4\ln\frac{r_p}{r_f}\right) \right] = \\
\frac{\beta\gamma mc^2}{4} &\left[(k_i^2 + k_0^2)r_i^2 + \frac{1}{2}(k_0^2 - k_i^2)r_i^2 \left(1 + 4\ln\frac{r_p}{r_i}\right) \right] + \Delta W
\end{aligned}
\tag{4.24}
$$

where the "free energy" ΔW can be introduced in the following convenient form

$$
\Delta W = \frac{1}{2}\beta\gamma mc^2 k_0^2 r_i^2 h
\tag{4.25}
$$

The parameter h is the "free energy" parameter which can be calculated for each of the three afore mentioned cases (distribution mismatch, envelope mismatch, off-centring).

Step II: for the emittance increase

From Eq. (4.21) we get

$$
k_i^2 = k_0^2 - \frac{K_1}{r_i^2} \quad \text{and} \quad k_f^2 = k_0^2 - \frac{K_1}{r_f^2} \quad \Rightarrow \quad k_f^2 = k_0^2 - \left(\frac{r_i}{r_f}\right)^2 (k_0^2 - k_i^2)
\tag{4.26}
$$

and together with Eq. (4.25) substituted into Eq. (4.24) one finds an expression that relates the change in beam radii to the "free energy" parameter h

$$
h = \left(\frac{r_f}{r_i}\right)^2 - 1 - \left(1 - \frac{k_i^2}{k_0^2}\right)\ln\frac{r_f}{r_i}
\tag{4.27}
$$

Using Eq. (4.20) one can write the emittance difference between the final and the initial beam as

$$
\Delta\varepsilon^2 = \varepsilon_f^2 - \varepsilon_i^2 = k_f^2 r_f^4 - k_i^2 r_i^4
\tag{4.28}
$$

With Eq. (4.26) and $\varepsilon_i^2 = k_i^2 r_i^4$ one can transform Eq. (4.28) into an expression for the final emittance

$$
\frac{\varepsilon_f}{\varepsilon_i} = \frac{r_f}{r_i}\sqrt{1 + \frac{k_0^2}{k_i^2}\left[\left(\frac{r_f}{r_i}\right)^2 - 1\right]}
\tag{4.29}
$$

Assuming that the mismatch is not too big ($r_f - r_i \ll r_i$) one can simplify Eq. (4.27) to

$$
\frac{r_f}{r_i} \approx 1 + \frac{h}{1 + (k_i/k_0)^2}
\tag{4.30}
$$

and get a first order approximation of the maximum emittance growth due to initial envelope mismatch.

$$
\frac{\varepsilon_f}{\varepsilon_i} = \sqrt{\left(1 + 2\frac{k_0^2}{k_i^2}h\right)}
\tag{4.31}
$$

Step III: the "free energy" parameter h for radial mismatch

A stationary (matched) beam in the smooth approximation can be represented by a circle in the (x/x') trace space. During the transport of the beam the circle rotates clockwise but the values for r_i and r'_i remain constant as shown in Fig. 4.6. If the beam is radially mismatched it has initially still the same emittance but an elliptic shape in trace space, denoted by r_m and r'_m in Fig. 4.6. Due to the clockwise rotation of the ellipse the beam radius r_m now oscillates between the depicted position of r_m and r_{max}.
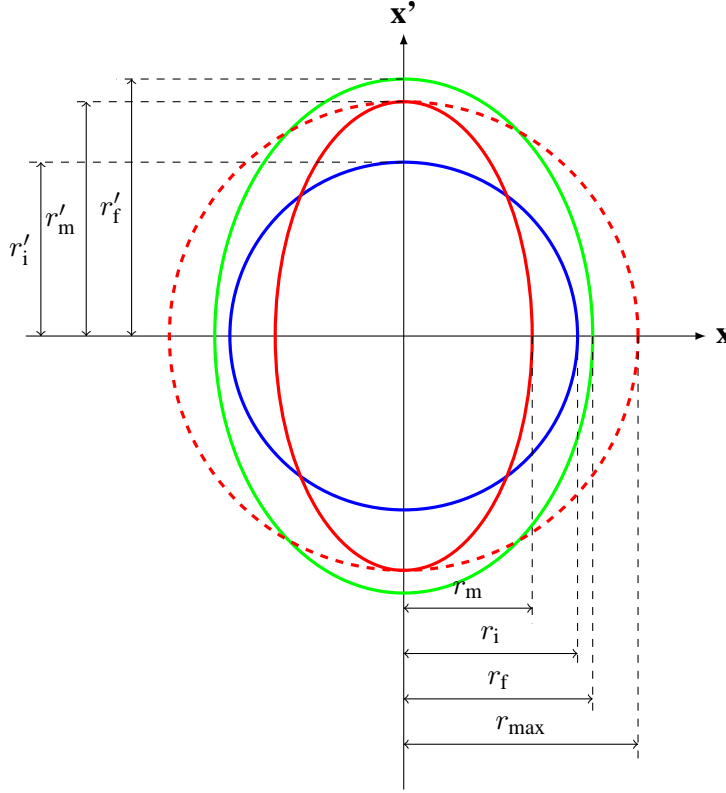
Figure 4.6: Upright trace-space ellipses for the initial mismatched beam (r_m, r'_m), the initial stationary beam (r_i, r'_i), and the final stationary beam (r_f, r'_f)

Using Eq. (4.22) we can calculate the energy difference ΔW between the initial mismatched beam (subscript m) and the initial matched beam (subscript i). Replacing K_1 with Eq. (4.26) then yields

$$\Delta W = \frac{\beta\gamma mc^2}{4} \left[r'^2_m - r'^2_i + k_0^2(r_m^2 - r_i^2) + 2r_i^2(k_0^2 - k_i^2)\ln\frac{r_i}{r_m} \right] \qquad (4.32)$$

Initially the emittances of the matched and mismatched beam are equal so that $r_m r'_m = r_i r'_i = \varepsilon_i$ (valid for upright emittances as shown in Fig. 4.6). Using

$$r'_i = k_i r_i \quad \text{and} \quad r'_m = r_i \frac{r'_i}{r_m} = k_i \frac{r_i^2}{r_m} \qquad (4.33)$$

and comparing Eq. (4.32) with Eq. (4.25) one can now express the "free energy" parameter h for a radially mismatched beam as

$$h = \frac{k_i^2}{2k_0^2}\left(\frac{r_i^2}{r_m^2} - 1\right) - \frac{1}{2}\left(1 - \frac{r_m^2}{r_i^2}\right) + \left(1 - \frac{k_i^2}{k_0^2}\right)\ln\left(\frac{r_i}{r_m}\right) \qquad (4.34)$$

where k_i/k_0 is the initial tune depression and r_m/r_i is the initial envelope mismatch.

In anisotropic 3D beams thermalisation not only occurs within one plane but can also occur between the planes (e.g. between the longitudinal and transverse planes) if certain resonance conditions are met (see Chapter 6). Generalising the "free energy" approach Hofmann suggested in [51] to relate the averaged r.m.s. emittance growth to the averaged mismatch in all three planes. Using Eq. (4.31) this yields an expression stating the theoretical upper limit for r.m.s. emittance growth in a 3D beam for an arbitrary initial mismatch assuming full thermalisation:

$$\frac{1}{3}\left(\frac{\Delta\varepsilon_{x,f}}{\varepsilon_{x,i}} + \frac{\Delta\varepsilon_{y,f}}{\varepsilon_{y,i}} + \frac{\Delta\varepsilon_{z,f}}{\varepsilon_{z,i}}\right) = \frac{1}{3}\sum_{n=x,y,z}\sqrt{1 + 2\frac{k_{n,0}^2}{k_{n,i}^2}h_n} \qquad (4.35)$$

The formulae assume a complete thermalisation of the free energy into r.m.s. emittance growth, while in practice one should expect slightly lower values.

4.5 Particle distributions for simulations in 6D phase space

Multi-particle simulations of realistic accelerating structures with space-charge suffer from the fact that there are no analytic solutions for stationary (self-consistent) particle distributions in 6D phase space. For simulations of very long bunches (or un-bunchd beams) one usually simplifies the problem to 4D phase space $(x/y/x'/y')$ where stationary solutions can be found for constant focusing lattices. In the following we will briefly describe the characteristics of the three most commonly used 4D distributions and then look at the 6D case which is needed for realistic simulations of high-intensity linacs.

4.5.1 KV, waterbag, and Gaussian distributions

A distribution often used for theoretical studies is the Kapchinskij-Vladimirskij distribution (KV) [30] which is the only self-consistent distribution where the external and internal forces (focusing system and space-charge) are linear and where the emittances are preserved. The KV distribtion is characterised by the fact that all particles have the same constant transvere energy, resulting in a distribution which covers the surface of 4-dimensional hyperellipsoid in 4D phase space. All 2D projections $(x/x', y/y', x/y, x'/y', x/y', x'/y)$ of this distribution are uniformly filled ellipses for beams with or without space-charge. In continuous focusing systems where the external nonlinearities are negligible and where the bunches are very long with respect to their transverse dimensions, the KV distribution is a useful tool to study space-charge related effects. However, it is clear that the KV distribution is not very realistic since all real beams will have a certain finite energy spread. Apart from that realistic accelerators hardly resemble the idealized beam transport systems (constant focusing, no acceleration) which are used in theoretical beam studies. It was also found that the KV distribution shows instabilities [52] which are not found in simulations with more realistic beams or in experiments. Another serious limitation is that there is no 6D solution for a KV-type beam which basically excludes its use for realistic linac simulations.

For these reasons one often uses the so-called waterbag distribution, where the particles are not mono-energetic but where particle energies are uniformly distributed between 0 and a certain maximum. Here

Table 4.2: Definition and properties of 4D particle distributions often used in multi-particle simulations

distribution	definition: $f(r_4)$ $r_4^2 = x^2 + x'^2 + y^2 + y'^2$	particle density in real space $r^2 = x^2 + y^2$
Kapchinskij-Vladimirskij (KV) $(r_c^2 = x_{max}^2 + y_{max}^2)$	$\frac{1}{2\pi^2 r_c^3}\delta(r_4 - r_c)$	$\frac{1}{\pi r_c^2}$
waterbag (WB) $(r_c^2 = x_{max}^2 + y_{max}^2)$	$\frac{2}{\pi^2 r_c^4}$ for $0 < r_4 < r_c$	$\frac{2}{\pi r_c^2}\left(1 - \frac{r^2}{r_c^2}\right)$
Gauss $(a^2 = \overline{x^2})$	$\frac{1}{4\pi^2 a^2}e^{-\frac{r_4^2}{2a^2}}$	$\frac{1}{2\pi a^2}e^{-\frac{r^2}{2a^2}}$

the 4D phase space is not just populated on the surface of a hyperellipsoid but it is represented by uniformly filled hyperspheres meaning that it is uniform in 4D phase space $(x/y/x'/y')$. In the zero-current case the 2D projections are again of elliptical shape with a density function in real space that depends linearly on the space-charge forces. For very high space-charge the 2D phase space projections $(x/x', y/y')$ become more and more square and the real space density becomes basically uniform (see Struckmeier [53] for detailed derivations and plots). However, even if the waterbag distribution is more realistic than the KV distribution it is still somewhat artificial since it imposes sharp beam boundaries within which all particles are confined. In realistic beams there will always be a "fuzzy" beam boundary and a certain amount of halo particles which are not adequately described with a waterbag distribution.

For a less sharp beam boundary one can use a Gaussian distribution which is defined by a Gaussian energy profile. For vanishing space-charge this distribution has a Gaussian density profile in real space while for high space-charge forces the density profile becomes uniform. Also this distribution is not completely realistic since it assumes infinite exponential tails which do not occur in realistic beams. In computer simulations with untruncated Gaussian distributions, the largest single particle amplitudes are basically related to the number of macroparticles generated. To obtain a better defined beam boundary one can cut off the Gaussian tails at a certain multiple of the r.m.s. radius.

In 4D simulations most computer codes generate the above mentioned particle distributions for the zero-current case which is summarized in Table 4.2 (from [54, 42]). One should note that for all three distribution types one can find a stationary solution for beams with space-charge in a continuous focusing channel. However, since most codes only generate the zero-current distribution (as listed in Table 4.2) the simulated beams are usually not stationary (apart from the KV case) and are subject to a certain redistribution of particles.

4.5.2 6D distributions

We have seen that even for a simplified 4D beam in a continuous focusing lattice it is a challenging task to find a realistic particle distribution that is stationary. In case of a 6D bunched beam in a realistic periodic focusing lattice including acceleration this task becomes impossible and one usually has to accept a certain initial emittance growth due the initial redistribution of particles. At very low energy, where

beams are usually space-charge dominated, one can show [42] and experimentally observe that beams develop a uniform density profile that is well approximated by a 6D waterbag distribution (which is uniform in 6D phase space $[x/y/z/x'/y'/z']$). At higher energy where the space-charge forces become very small beams usually develop a Gaussian profile, meaning that the equilibrium distributions in linacs will always be somewhere between a waterbag and a Gaussian profile. As we will show in the following Gaussian distributions generally yield more emittance growth than waterbag distributions meaning that one should use both distribution types to define upper and lower boundaries for the expected beam performance in a linac.

When studying initial mismatch in realistic lattices, the effects of radial mismatch are usually much more pronounced than the effects due to mismatch in the density profile which is the case for all multi-particle simulations presented here. If needed one can completely seperate the effects by tracking a beam over a certain number of equal periods without acceleration in order to provide enough time for the beam to find its equilibrium distribution. This distribution can then be used to specifically study effects due to radial (or other) mismatch without any influence of distribution mismatch.

For end-to-end simulations of a complete linac one is less interested in seperating different effects that yield emittance growth. In this case the focus should be on simulating a linac as realistically as possible and it is then sensible to use a simulated (or measured) output distribution of the RFQ which usually contains a certain amount of halo particles. The input distribution for the RFQ simulation itself is usually of the 4D waterbag type which becomes bunched and which develops some (usually Gaussian) tails during the passage of the RFQ.

4.5.3 Distributions and emittance definitions

The emittance definition used in this thesis is the one for the r.m.s. emittance introduced by Sacherer in [29]. Lapostolle suggested in [55] to use an emittance definition corresponding to four times the r.m.s. emittance which corresponds to the total emittance of a uniform continuous beam. Weiss has calculated [56] the ratio between the total and the r.m.s. emittance for beams which have a clearly defined border and which are uniformly filled in n-dimensional projections as

$$\frac{\varepsilon_{\text{total}}}{\varepsilon} = n + 2 \tag{4.36}$$

where ε is the r.m.s. emittance. The resulting ratios for waterbag and KV beams are listed in Table 4.3. In a 4D Gaussian beam which is truncated at p times the r.m.s. radius the ratio of total to r.m.s. emittance is approximately given by p^2 (assuming that $p > 4$).

The ratio of total to r.m.s. emittance is also a measure for the tails in a distribution and is thus important for the characterisation of halo development in a linac. However, using this ratio in multi-particle simu-

Table 4.3: Ratio of total to r.m.s. emittances for distributions with well defined boundaries which are uniform in n-dimensional hyperellipsoidal projections

distribution	n	$\varepsilon_{\text{total}}/\varepsilon_{\text{rms}}$	uniform in
KV	2	4	$(x/x', y/y', x/y, x'/y', x/y', x'/y)$
4D waterbag	4	6	$(x/y/x'/y')$
6D waterbag	6	8	$(x/y/z/x'/y'/z')$

lations is not very useful since it is determined by one single particle which happens to have the largest single particle emittance. For better statistics one often looks at fractional emittances which contain a certain percentage of all particles as for instance 95% or 99%. But also the fractional emittances do not offer a clear judgment on halo development since there is no straightforward algorithm to calculate the smallest emittance value which contains a certain percentage of particles. The usual (numerical) procedure to determine fractional emittances is to calculate the size and orientation of the r.m.s. ellipse [using Eq. (2.1) and (2.3)] and then to gradually increase this area until the desired percentage of particles is included, thereby maintaining the aspect ratio and orientation of the r.m.s. ellipse. Since halo development generally yields a fragmented S-shaped distribution in the 2D projections $(x/x', y/y', z/z')$ [as shown in Fig. 3.2)] this procedure becomes questionable and can only give an indication on the process of halo development. Another problem is that the tails have a different angular velocity than the core particles yielding oscillating values of the fractional emittances (an example of these oscillations can be appreciated in Figs. 6.6 and 6.7 where the 99%, 99.9%, and 99.99% fractional emittances during an emittance exchange are plotted). For this reason we recommend a halo characterisation that is based on the particle density versus beam radius (e.g. Fig. 4.13) or even more precise on the fraction of particles which is found outside multiples of the r.m.s. emittance (e.g. Fig. 5.10). As long as this curve is taken always at the same lattice position it provides a useful tool to compare the effect of different mismatch types on the development of beam halo.

4.6   Mismatch for realistic linac beams

In this section more realistic lattices and beams shall be considered, including acceleration and changing focusing elements. As an example the whole superconducting linac section (120 - 2200 MeV, simulation current: 40 mA, average tune depression: ≈ 0.7) of the SPL I study [1] is used to show in more detail the effect of the three eigenmodes on the beam. Furthermore an inhomogeneous 6D Gaussian beam density profile is used instead of the 6D waterbag distribution of the previous section. The lattice is simulated with the 3D multi-particle code IMPACT (see Chapter 3). The resulting mismatch oscillations for the three eigenmodes are plotted in Figs. 4.7 to 4.9.[2]
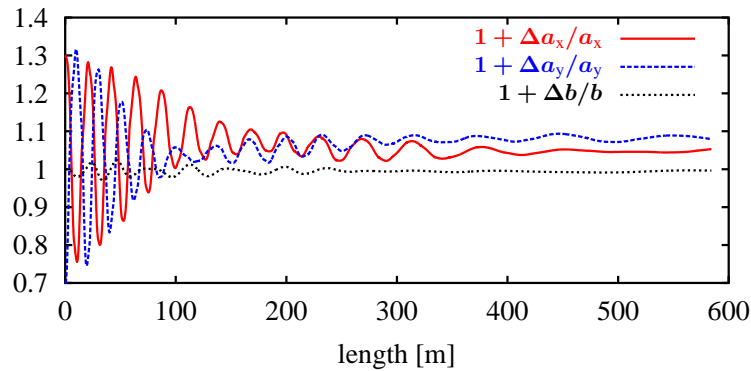
Figure 4.7: 30% quadrupolar mismatch in the SC section of SPL I, IMPACT multi-particle simulation

---

[2]The ratio of mismatched over matched r.m.s. envelopes is used to visualise the oscillations.
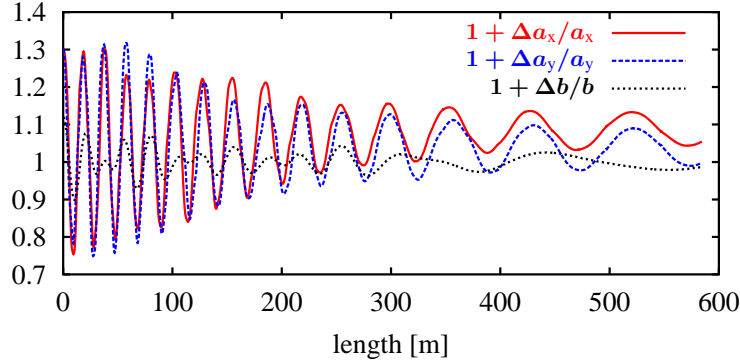
Figure 4.8: 30% fast mode mismatch in the SC section of the SPL I, IMPACT multi-particle simulation
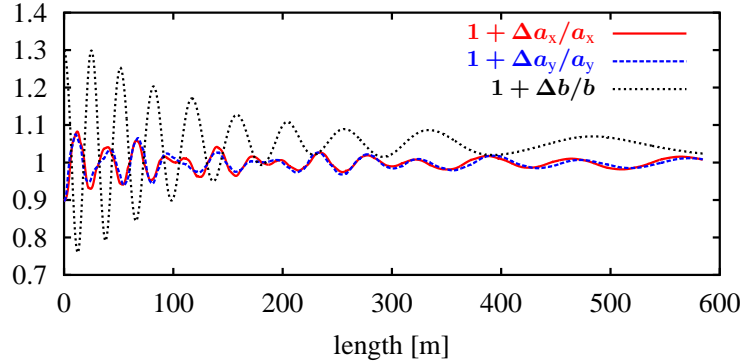
Figure 4.9: 30% slow mode mismatch in the SC section of the SPL I, IMPACT multi-particle simulation

One can observe that the phase and amplitude characteristics of all modes are basically the same as for the simulations with a 6D waterbag beam in the previous section. Although the tunes, and therefore also the oscillation frequencies, of the mismatch modes change considerably along the linac, and even though 3 lattice transitions are crossed, the oscillations remain remarkably stable. However, in contrast to the examples in Figs. 4.3 to 4.5, the oscillations now appear to be damped. This effect can be explained by stating that the "free energy" that has been introduced into the system via mismatch is transformed into beam halo and emittance growth. The emittance growth is visible in plots 4.7 to 4.9 through the shift of the oscillation average (1 at the beginning, and slightly higher at the end the simulations). At the end of the transformation the r.m.s. core has reached a new equilibrium and the oscillations basically stop. Figure 4.10 shows the r.m.s. emittance growth for the three modes and the predicted maxima from the free energy approach (Eq. 4.35).

In agreement with the transient behaviour in Figs. 4.7 to 4.9 the emittance growth for the quadrupolar mode has the shortest rise time, followed by the slow mode and the fast mode. None of the three modes surpasses the predicted maximum r.m.s. emittance growth from the free energy approach and only the slow mode does not reach its predicted maximum value. One can see that the agreement for the fast mode and quadrupolar mode are remarkably exact, meaning that the free energy limits seem to provide a useful estimate for maximum r.m.s. emittance growth. The simulations also suggest that a typical high-
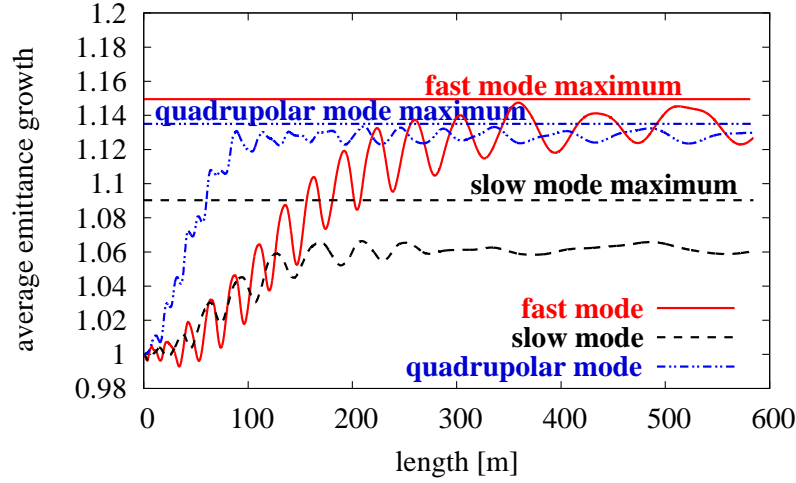
Figure 4.10: R.m.s. emittance growth for the 3 eigenmodes and predicted maxima from the free energy
approach, IMPACT multi-particle simulation

power linac is long enough for these maxima to be reached. Furthermore, one can conclude that if these
maxima are surpassed there must be more sources of mismatch than only the initial envelope mismatch
which was introduced here on purpose.

The difference in transformation speed between the example in Fig. 4.3 and those in Figs. 4.7 to 4.9 can
be explained as follows: in the first example the simulations with the 6D waterbag distribution result in
a very long rise time, while in the second example the Gaussian distribution provides enough "noise" to
quickly transform the mismatch oscillations into emittance growth. Reiser [42] suggests that the non-
uniformity of the beam is responsible for the quicker transformation and Hofmann [51] points out that
the Gaussian tails may be important for the transformation speed.

To highlight the difference Fig. 4.11 shows the transverse mismatch oscillations due to fast mode exci-
tation in a periodic transport channel without acceleration for a 6D waterbag and a 6D Gaussian beam.
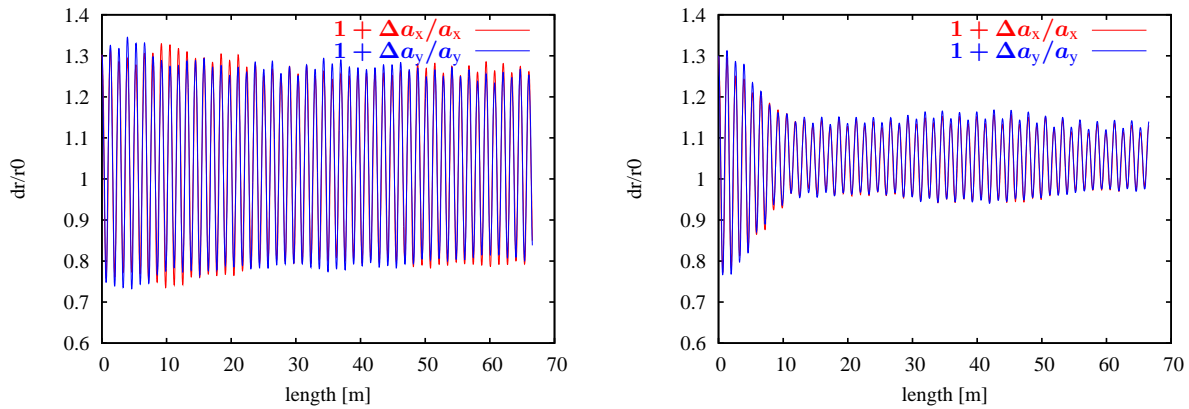
Figure 4.11: Mismatched over matched transverse r.m.s. envelopes for 6D waterbag (left) and Gauss
(right), 200 focusing periods, fast mode excitation, IMPACT multi-particle simulation

One can see that the transverse r.m.s. core-oscillations for a Gaussian beam with 30% fast mode excitation are rapidly damped from 30% to $\approx$ 12% where they remain almost constant. At the same time the waterbag beam loses only a fraction of its oscillation energy with the result of much smaller emittance growth (7% versus 27% for 6D Gauss) and halo development in a transport channel of equal length. Beams in realistic end-to-end simulations for high-intensity linacs usually develop particle distributions which are somewhere in between a 6D waterbag and a Gaussian distribution.

4.6.1 Particle redistribution

The effect of parametric resonances on the beam distribution can be illustrated by the redistribution of particles in a mismatched bunch. Since the three beam eigenmodes have three different frequencies, one can expect that each eigenmode affects particles with different oscillation frequencies (tunes). In order to plot a clear redistribution pattern a transport channel with 50 focusing periods (Gaussian input distribution) is simulated, exciting separately the three beam eigenmodes [Fig. 4.12 (a)].

We start with the quadrupolar mode, which oscillates only in the transverse plane. Its eigenfrequency is by definition [see Eq. (4.8)] twice as high as that of the transverse depressed tune. One can therefore expect a strong resonance with the core particles which also have the lowest tunes in the bunch. Looking at Fig. 4.12 (a) one can observe that indeed particles are removed from the core area and deposited at a higher radius. For the slow and fast modes the situation is more complicated since they both oscillate in all three planes. Nevertheless, it seems that one can obtain a consistent picture: the eigenfrequency of the slow mode is lower than for the quadrupolar mode, therefore the coupling with the core particles can only be very weak and consequently only few particles are removed from the core. Finally the fast mode seems to affect particles with higher tunes. The intuitive conclusion is that high-frequency core oscillations yield a large fixed point - core distance and thereby large halo amplitudes, whereas particles with lower tunes (from the core centre) are affected by the low-frequency core oscillations and end up around fixed points with a smaller fixed point - core distance. However, this pattern and the order of the fixed points should not be generalised since it is not only influenced by the mode frequencies but also by beam-lattice resonances and inherent redistribution patterns of the initial distribution (waterbag, Gauss, etc.). The only conclusion we can draw at this stage is that each eigenmode will generally trigger different redistribution patterns with different fixed point - core distances. It is also very likely that the fast mode will always yield the largest halo amplitudes, which makes it very suitable for halo studies.

For small mismatch amplitudes a general mismatch excitation yields a superposition of modes as shown in Fig. 4.12 (b). As an example we use a "$+ + +$" and "$+ - +$" mismatch which stands for +30% envelope mismatch in x, y, and z, or +30% in x and z and -30% in y, respectively.

Comparing Figs. 4.12 (a) and (b) we find that the redistribution for the "$+ - +$" excitation is dominated by the fixed point of the quadrupolar mode at $\approx 2r_{c,r.m.s.}$. This is consistent with theory since "$+-$" in x and y corresponds to the excitation of the quadrupolar mode. For the "$+ + +$" mismatch the quadrupolar mode is not excited at all and one obtains a mixture of the low and fast mode redistribution patterns. In this case the two fixed points are clearly visible, even if they are slightly shifted by the superposition.

We note that the partitioning of modes that are excited by a general mismatch is dependent on the emittances and tunes of the machine and can be completely different from the one in Fig. 4.12. When studying the effects of mismatch in a particular linac it is therefore advisable either to excite all three eigenmodes separately or to use at least a few different sets of general mismatches.
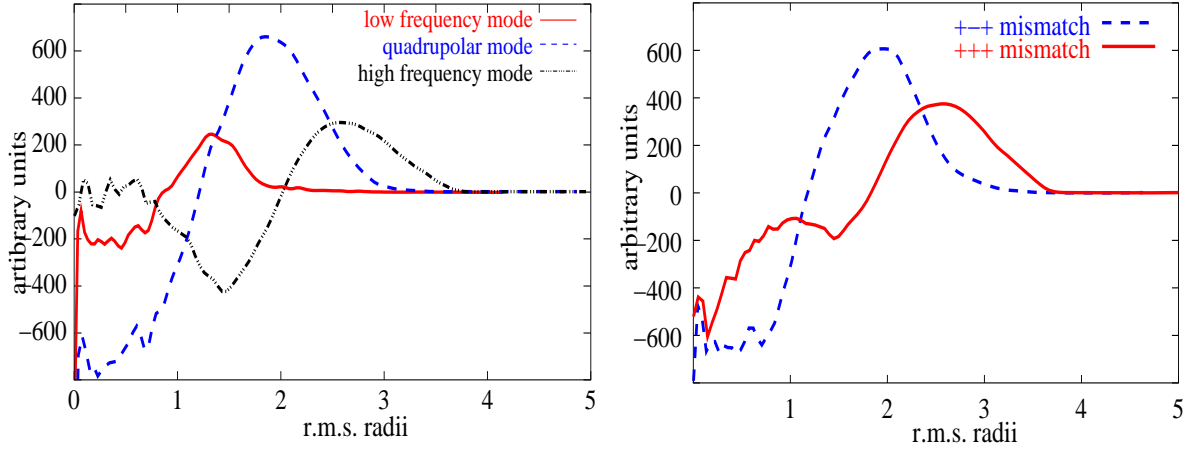
Figure 4.12: (a) Particle redistribution when excit-
ing the three eigenmodes in a transport
channel with 50 periods. Plotted is the
difference in transverse particle density
between the matched and mismatched
distribution at the end of the channel,
IMPACT multi-particle simulation.

(b) Particle redistribution when exciting a gen-
eral 30% envelope mismatch in a transport chan-
nel with 50 periods. Plotted is the difference
in transverse particle density between the mis-
matched and the matched distribution at the end
of the channel, IMPACT multi-particle simula-
tion.

4.6.2   Maximum halo extent

The previous section confirms the prediction from the simple 1D particle-core model that halo particles
gather around fixed points and that the distance of these fixed points from the core is given by the oscil-
lation frequency of the mismatched core and its 2:1 parametric ratio with the single-particle oscillations.
Beyond the orbits around these fixed points the particles cannot gain more energy from the core oscil-
lations, simply because the 2:1 resonance moves out of phase. Although it was found that fixed point -
core distances can theoretically go to infinity [57] one observes that maximum halo amplitudes for the
actual particles are limited to a certain threshold. Hofmann suggests in [58] that even if the oscillations
are still "in phase" for increasing fixed point - core distances, there is a decrease in the space-charge cou-
pling force, without which there can be no energy transfer from the core oscillations to the single-particle
orbits.

With a simple test one can verify the idea of a maximum halo extent: using again the superconduct-
ing section of SPL I, we excite a "+++" mismatch for a Gaussian input beam. Figure 4.13 shows the
transverse distribution at the end of the linac for different amplitudes of initial mismatch.

Even in simulations with very large numbers of particles (here: 10 million) one finds that the redistribu-
tion pattern as well as the maximum halo amplitude remain basically the same for the different mismatch
amplitudes. Only the particle density towards the maximum amplitude increases with increasing mis-
match. In this example the outermost particles extend to a maximum value of ≈ 8 times the r.m.s. enve-
lope ($8a$). However, due to imperfections in the lattice or transitions between sections, maximum values
of up to $\approx 12a$ are found in simulations. As an example we show in Fig. 4.14 the maximum transverse
halo extent for a strongly mismatched (40%) beam in a normal conducting linac (3 - 120 MeV), with two
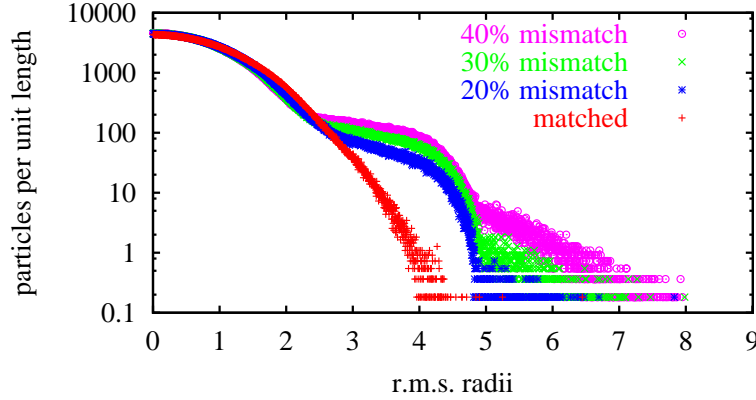lattice transitions (at 18 m and 41 m).

Figure 4.13: Transverse particle density (in arbitrary units) at the end of SPL I (120 – 2200 MeV, 40 mA simulation current, Gaussian input beam, 10^7 particles) when exciting a "+++" mismatch with different amplitudes, IMPACT multi-particle simulation
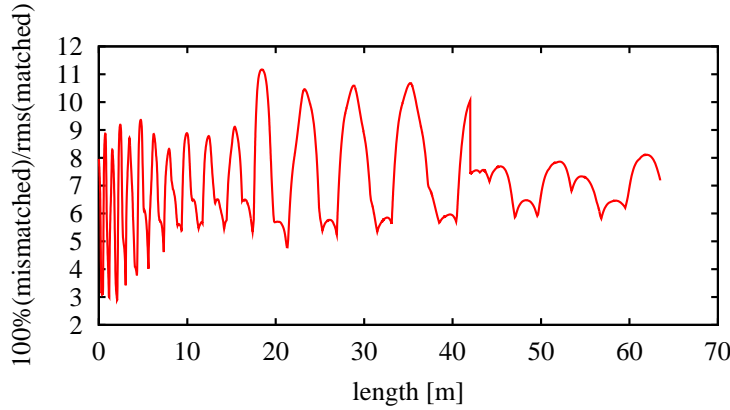

Figure 4.14: Transverse halo extent in multiples of the r.m.s. beam envelope for a normal conducting linac (3 - 120 MeV, 40 mA simulation current, 40% "+++" mismatch, Gaussian input beam), IMPACT multi-particle simulation

4.6.3 Beam collimation

An important aspect in the beam halo discussion is the question of the rise time for emittance growth or beam halo to develop. As discussed in the previous section, halo amplitudes between 8 and $12a$ are observed in simulations. This means that a simplistic approach to avoiding transverse beam loss would be to use apertures larger than 12 times the r.m.s. beam envelope. In normal conducting RF accelerating structures, however, the size of the aperture is directly linked to the power requirements of the cavities, and large bore radii unavoidably yield a poor RF efficiency. If the focusing quadrupoles are separated from the RF, as for instance in Coupled Cavity Drift Tube Linacs (CCDTL), Separated DTLs (SDTL), or Coupled Cavity Linacs (CCL), one can use quadrupoles with large bore radii without decreasing the RF efficiency[3]. Since the beam envelopes reach their highest values in the quadrupoles, enlarging their bore radius is an effective means to reduce beam loss. In a classical Alvarez DTL, where the quadrupoles are

---

[3]examples of the structures can be found in Appendix B

housed in the drift tubes, this trick cannot be applied and the apertures are typically in a range of 5 to $7a$, depending on the average beam power in the structure. Here, the only solution to reduce beam loss is the use of scrapers before entering the DTL. In that case the rise time for halo formation really becomes the crucial parameter for loss prediction in the machine. As an example Fig. 4.15 shows the evolution of the (longitudinally) most unstable particle that was found in the simulation of an ideal transport channel (the transverse behaviour can be assumed identical).
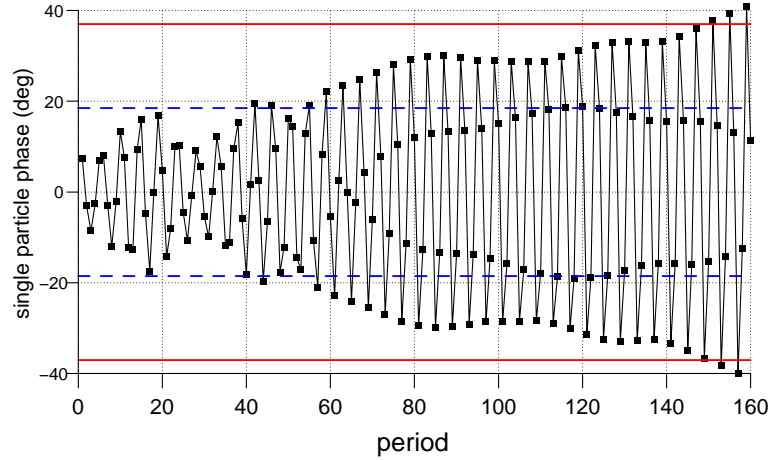
Figure 4.15: Evolution of the (longitudinally) most unstable particle in a transport channel of 160 periods with fast mode excitation and an initial 6D waterbag distribution. Plotted is one point per period. The dashed line indicates the maximum initial mismatched phase extent and the solid line indicates twice that value. Example from Bongardt & Pabst.

One can see that the most unstable particle after 160 periods is *not* the one that initially has the largest phase value. This means that even if halo scraping is employed at the low-energy stages of a linac, particles can still be expelled to large amplitude orbits. On the other hand the rise time in this particular example amounts to ≈ 60 focusing periods, which is already longer than many linacs.

Judging from the theory and the simulation results it is vital in every high-power linac design to avoid all sources of mismatch. These include not only initial mismatch but also mismatch between different sections. Once the beam is mismatched the only way to damp the core oscillations and to return to an equilibrium is the development of beam halo. This means that beam collimation and halo scraping can only help to control the effects of mismatch (beam loss) but they can certainly not correct the mismatch itself, nor will they prevent further halo formation.

This behaviour is illustrated in the following by placing collimators, long enough to remove halo particles from both planes of the x and y phase space, at 4 different positions in a FODO channel (Fig. 4.16). At all 4 positions particles beyond 3.2 r.m.s. amplitudes are removed from the beam.

From Fig. 4.17 one can appreciate that in all 4 cases the amount of beam loss in the scrapers is approximately the same, while the overall r.m.s. emittance increase and the maximum halo amplitude depend on the position of the collimators. Obviously, the scraping is most effective at position 4, where the core oscillations are already damped and where the beam has settled into a new equilibrium. From this position onward only little additional emittance growth takes place. Considering a realistic linac with several sources of mismatch (e.g. lattice transitions, statistical errors) this means that several collimators should be foreseen at positions sufficiently far away from the actual mismatch sources.
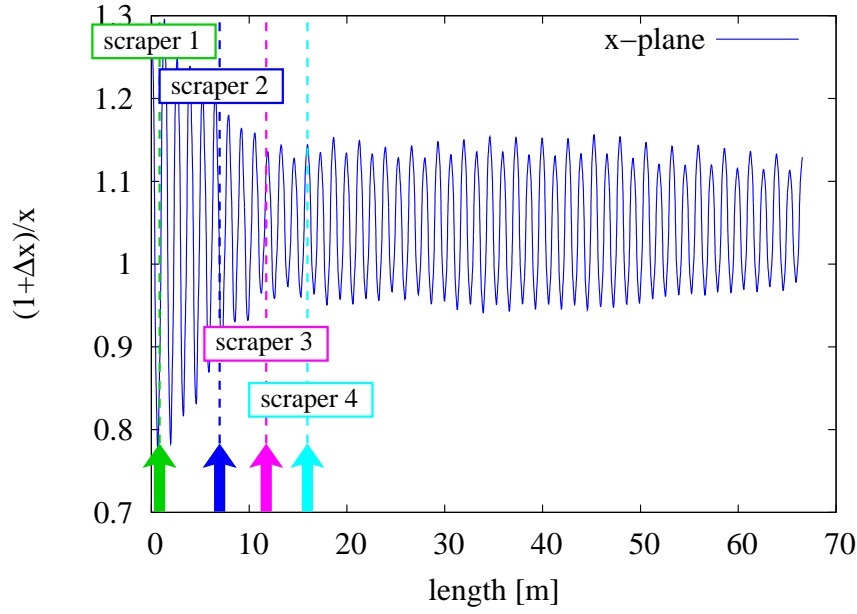
Figure 4.16: Mismatch oscillations due to a fast mode excitation with 30% initial amplitude mismatch in a FODO channel using a Gaussian distribution and showing 4 possible scraper positions (shown is the unscraped beam), IMPACT multi-particle simulation
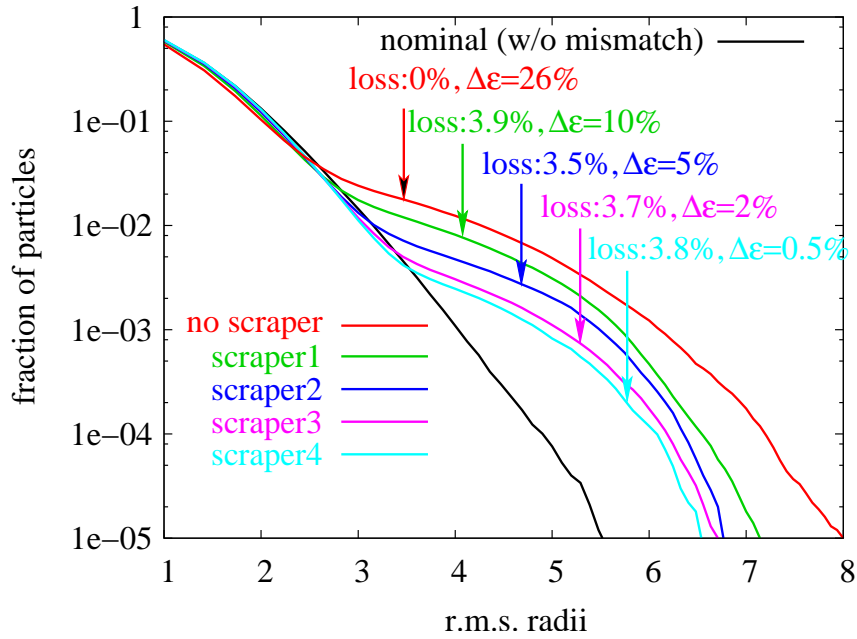

Figure 4.17: Final particle distribution and r.m.s. emittance growth for 4 scraper positions in a FODO channel, 30% fast mode excitation, Gaussian distribution, IMPACT multi-particle simulation

For linear accelerators which send their beam directly onto a target longitudinal halo is usually of minor interest. If, however, the linac beam enters a subsequent ring system for beam storage or further acceleration, the longitudinal behaviour becomes more important. Especially in accumulator and/or compressor rings with no acceleration, the longitudinal bunch shape has to be well confined to fit into the RF bucket of the ring system. In this case the bulk of the losses is triggered by the linac RF jitter, the uncertainty of the energy and phase level at the end of the linac which is caused by the finite tolerances of the RF system. Losses due to halo formation in the longitudinal plane will only have a minor contribution.

5. Distributed mismatch

Another contribution to the development of beam halo stems from statistical gradient errors which can be regarded as a multitude of individual mismatch sources. The particle-core model suggests that the main condition for halo development is the parametric 2:1 ratio between the core and single-particle oscillations. The 3D simulations (Fig. 4.7, 4.8, and 4.9) of the SPL I lattice show that once the core oscillations are excited they can remain remarkably stable throughout a complete linac even though several (matched) lattice transitions are crossed and even though the focusing constants change considerably. At some point the oscillations become damped by the transformation of the "free energy" (introduced via mismatch) into emittance growth and it is important to realize that this is the only process by which the oscillation amplitudes are reduced.

In 3D error simulations for Linac4 [59], the normal conducting front-end of the SPL, it was found that similar core oscillations can be excited by statistically distributed quadrupole gradient errors and that also there the resulting core oscillations remain remarkably stable throughout various lattice changes. Figure 5.1 shows an example of these core oscillations (triggered by statistical errors on the quadrupole gradients) for the Linac4 lattice, comprising a 3 MeV chopper line, followed by 3 DTL tanks and 37 CCDTL tanks.
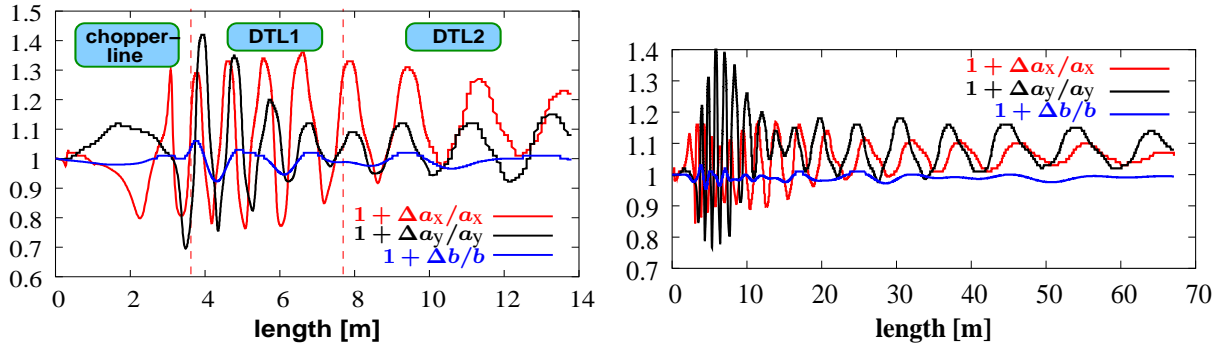
Figure 5.1: (a) Worst case envelope deviations for 1% (total) quadrupole gradient errors in the Linac4 front-end, IMPACT multi-particle simulation

(b) Worst case envelope deviations for Linac4 up to 120 MeV, IMPACT multi-particle simulation

One can see that for statistical gradient errors of only 1%, the worst case envelope deviations increase the regular beam amplitudes by up to 40% corresponding to a 40% initial mismatch!

In the following, the 1D particle-core model from Section 4.2 is extended to study if it can predict regular core oscillations caused by statistical gradient errors and if these core oscillations can be a source of halo formation. Furthermore systematic 3D tracking studies will be used to quantify the effects for a realistic lattice (see also [60]).

5.1 Particle-core model for statistical gradient errors

Starting point is again the 1D smoothed envelope equation [see Eq. (4.1)] and the same set of variables and conventions introduced in section 4.2. In every focusing period a small error is added to the focusing constant k_0, simulating the effect of statistically distributed gradient errors in a transport channel.

$$r_c'' + (k_0 + \Delta)^2 r_c - \frac{\varepsilon^2}{r_c^3} - \frac{K_1}{r_c} = 0 \tag{5.1}$$

The movements of the single particle in the space-charge field of the core are then given by [compare Eq. (4.3)]

$$x'' + (k_0 + \Delta)^2 x - F_{sc} = 0 \tag{5.2}$$

with the space the space-charge force term (of a continuous round beam) being defined as

$$F_{sc} = \begin{cases} K_1 x / r_c^2 & : \quad |x| < r_c \\ K_1 / x & : \quad |x| \geq r_c \end{cases} \tag{5.3}$$

The errors are applied using a Gaussian error distribution with a cut-off at twice the r.m.s. value. Figure 5.2 shows the result of integrating the core and single particle equations [Eq. (5.1) and Eq. (5.2)] assuming a randomly chosen error set with 1% (r.m.s.) focusing error. The initial single particle amplitude in this example is $|x(0)|/r_0 = 1.2$ with r_0 being the matched core radius for $\Delta = 0$. It is evident that once $\Delta \neq 0$, the beam will be mismatched in this period. With Δ changing from period to period the beam is subject to a different small mismatch in every period.

PSfrag replacements
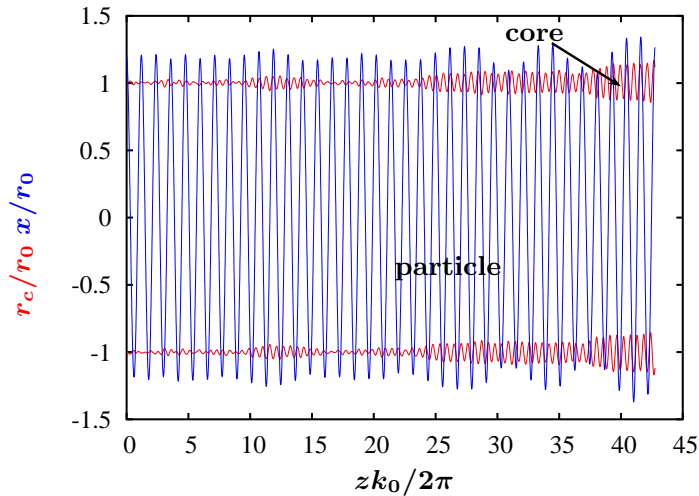
Figure 5.2: Particle oscillations & core envelope in case of a matched beam with 1% (r.m.s.) statistical focusing gradient errors (tune depression $\tau = 0.8$, $|x(0)|/r_0 = 1.2$)

After a certain number of periods the core starts oscillating in a similar manner as for initial mismatch. Due to the irregular excitation of the core, its oscillation amplitude is now subject to a slowly changing random modulation. The oscillation frequency, however, remains almost constant at the frequency of

the eigenmode and only changes within a few per cent around this reference value. Only for very small oscillations, as in the first periods of Fig. 5.2, does the oscillation frequency change by up to $\approx 20\%$.

Even though the errors are statistically distributed, the core oscillations can reach considerable amplitudes, a phenomenon that can be observed in realistic linac lattices as in Fig. 5.1. Since the core-oscillations are irregular as opposed to those resulting from initial mismatch (e.g. Fig. 4.1), it is unlikely for single particles to enter a stable parametric resonance for more than a few oscillation periods. Nevertheless, there are sections in the lattice, when the core oscillations maintain an almost constant amplitude (e.g. periods 25-40 in Fig. 5.2) and during these it seems that the single particle enters a similar oscillation pattern as for the initial mismatch in Fig. 4.1: the maximum amplitude of the single particle is rising and falling with a more or less sinusoidal modulation. This pattern suggests that even a few oscillation periods which are more or less 'in phase' with the core oscillations suffice to transfer energy from the core to the single particle trajectories. Comparing once more with the case of initial mismatch in Fig. 4.1 this explanation seems very likely, since also there the energy transfer practically starts from the first few core oscillations onwards, implying that the mechanism to transfer energy is the same as for initial mismatch. In the case of initial mismatch the core immediately oscillates with a large amplitude yielding an increase of single particle amplitudes by a factor of 2-3 within a few periods. For statistical gradient errors, however, the particle-core model suggests that several hundred periods may be necessary to achieve core oscillation amplitudes that are large and stable enough to trigger a significant increase in single particle amplitudes.

Plotting the maximum and minimum values of core and single particle amplitudes along the channel [using again Eq. (5.1) and Eq. (5.2)] for different initial single particle amplitudes (Fig. 5.3) confirms that the single particle envelopes are oscillating in a manner similar to initial mismatch. It also shows that particles from within the core boundaries are only weakly affected, whereas particles starting outside of the core boundaries show a clear increase in their oscillation amplitudes.
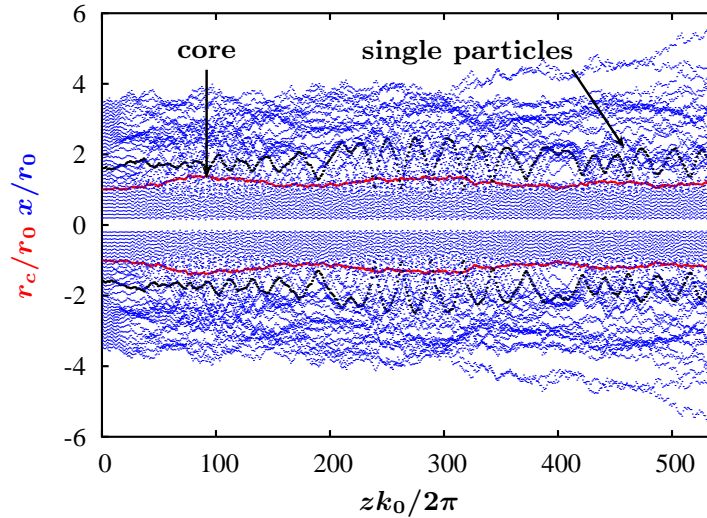

Figure 5.3: Maximum particle/core amplitude values in case of a matched beam with 1% r.m.s. statistical focusing gradient errors ($\tau = 0.8$, $0.2 \leq |x(0)|/r_0 \leq 3.5$)

5.1.1 Average effects and evidence for a resonant process

In order to show which particles experience on average the largest amplitude growth, Figs. 5.4 (a), (b), and 5.5 show the growth factors for single particle amplitudes as a function of their initial values. Each of the plotted points represents the average of maximum amplitude values found in 1000 simulations [using Eq. (5.1) and Eq. (5.2)] with different error sets. The curves connecting the points are interpolated with cubic splines.
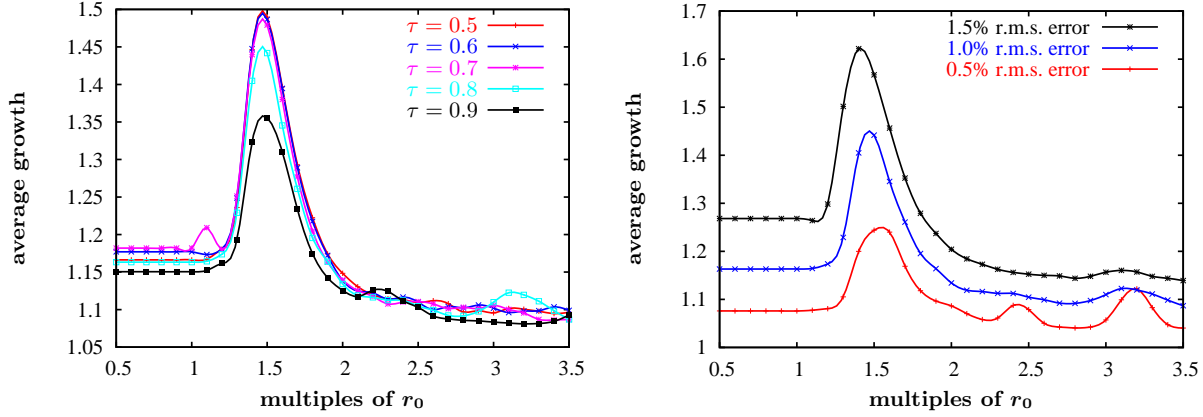
Figure 5.4: (a) Maximum amplitude growth for single particles over 100 zero-current betatron periods versus their initial amplitude and tune depression τ. Each dot represents the averaged results of 1000 runs with different error sets with 1% r.m.s. error.

(b) Maximum amplitude growth for single particles over 100 zero-current betatron periods as a function of their initial amplitude and of the r.m.s. error amplitude. Each dot represents the averaged results of 1000 runs with different error sets for a tune depression of $\tau = 0.8$.

Figure 5.4 (a) shows that for an emittance dominated beam ($0.71 < \tau < 1.0$) the maximum growth factors (or halo amplitudes) clearly depend on the tune depression of the beam, while they remain basically constant for space-charge dominated beams ($0.0 < \tau < 0.71$). The same observation was made by Wangler [43] who studied halo development due to initial mismatch using the particle core model.

Figure 5.4 (b) shows, not surprisingly, that the maximum oscillation amplitudes depend on the amplitude of the statistical variation that is applied to the focusing forces, and finally, Fig. 5.5 explores how the maximum amplitude growth depends on the length of the simulated transport channel.

In all three cases (Figs. 5.4 (a), (b), and 5.5) the maximum amplitude growth is found for particles with initial amplitudes around $1.5\,r_0$, suggesting that these particles have the highest probability of entering a parametric resonance with the core oscillation.

In order to show that the underlying mechanism is indeed based on a resonance, Fig. 5.6 shows the wave numbers (tunes per metre) for the core (k_{core}) and the single particle ($2k_{\mathrm{particle}}$) oscillations, assuming one randomly chosen error set with different initial amplitudes for the single particles. The data is smoothed by approximating the raw data with bezier curves of n'th order (n - number of data points) which connect the end points. For simplicity we use the bezier smoothing which is available in gnuplot [61] and which reduces the amplitudes of the rapidly changing wave numbers while making it easier to distinguish the trends in the raw data.
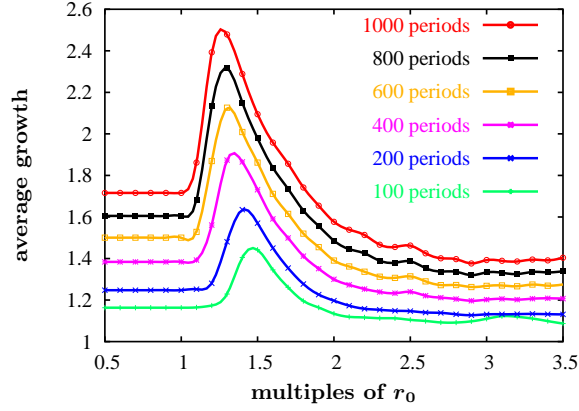
Figure 5.5: Maximum amplitude growth for single particles as a function of their initial amplitude and the length of the transport channel (in zero-current betatron periods). Each dot represents the averaged results of 1000 runs with different error sets for a tune depression of $\tau = 0.8$ and an r.m.s. error amplitude of 1%.
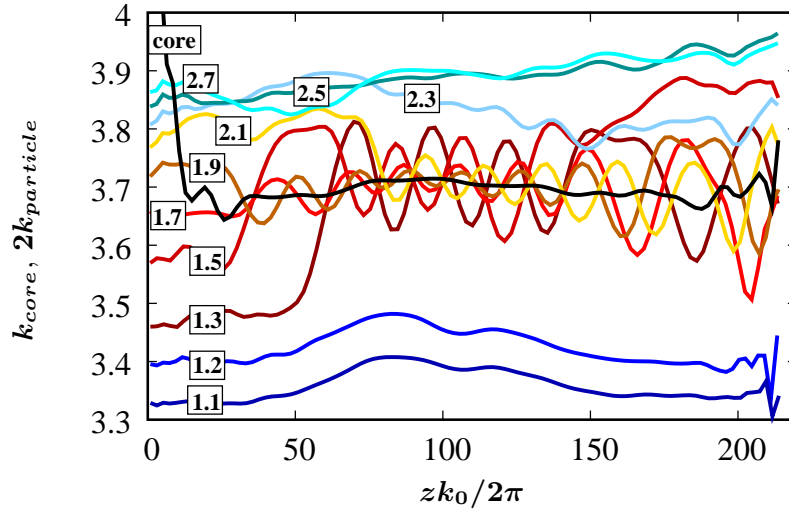

Figure 5.6: Wave numbers (arbitrary units) for core and single particle oscillations. Single particle wave numbers are multiplied by 2 and depict particles starting with different initial amplitudes (1.1 - 2.7 times the matched core radius), $\tau = 0.8$, r.m.s. error amplitude: 1%.

After 20 - 30 zero-current betatron periods the core wave number assumes a relatively stable value of ≈ 3.7 (arbitrary units). A few periods later the wave numbers of particles, that start with initial amplitudes between 1.3 and 2.1 times the matched core radius, begin to oscillate around an average that is half as big as the core wave number, meaning that they fulfil the condition for a parametric 1:2 resonance with the core. It is interesting to note that particles whose initial wave numbers are very close to 0.5 times the core wave number start very quickly to resonate with the core, while it takes more time until particles with larger or smaller wave numbers are affected. Particles that start with initial amplitudes of $|x(0)| \leq 1.2\, r_0$ or $|x(0)| \geq 2.3\, r_0$ have either too small or too high a wave number to enter a resonance within the length of the calculated transport channel.

One can expect that the range of particles likely to interact with the core oscillations will be larger in a more realistic 3D beam with altogether 3 different core-eigenfrequencies. Furthermore the maximum amplitude values and their associated time constants will certainly be different for more realistic beams with non-uniform distributions and non-linear space-charge forces within the core area.

Even though one could imagine that for statistical gradient errors, the particles experience on average as much negative energy kicks as positive kicks, Fig. 5.5 clearly shows that statistical gradient errors have a cumulative effect. This can be understood by interpreting these errors as a continuous supply of 'free energy' which can only increase the transverse beam temperature (or energy) but never yield a decrease (compare [50, 62]). In the case of initial mismatch the beam obtains one large energy kick at the beginning and the particles as well as the core perform regular oscillations around their equilibrium. Without further disturbance these oscillations are maintained indefinitely in the particle-core model, while a continuous supply of 'energy' via statistical gradient errors yields ever increasing particle amplitudes. We note at this point that the core oscillations of a realistic 3D beam, caused by initial mismatch, are eventually damped by the energy transfer from the core to the single particle oscillations. This transformation from mismatch to beam halo and r.m.s. emittance growth has been found to develop much more rapidly for Gaussian beams than for the idealised 6D waterbag beams (compare Fig. 4.11 and also [51]).

5.1.2 Halo development

While the average effects give a good indication of the general particle behaviour due to statistical gradient errors, they provide little information about the development of beam halo and its maximum extent in phase space. To illustrate the formation of halo driven by statistical gradient errors we use the same kind of stroboscopic plot that showed the resonance pattern for initial mismatch (e.g. Fig. 4.2). Figure 5.7 depicts the pattern of one particle for one particular error set, which has a larger than average amplitude growth. One can see that initially the particle maintains a phase space trajectory close to the core boundaries. At some point the particle receives a kick that, within a few periods, rapidly increases its amplitude from $\approx 2\,r_0$ to $\approx 3.5\,r_0$, where it remains for another ≈ 150 zero-current betatron periods.
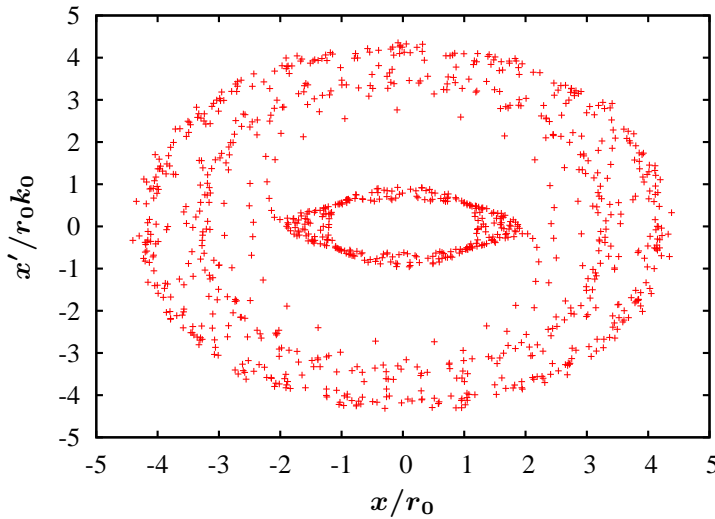

PSfrag replacements

Figure 5.7: Stroboscopic plot for single particle with: $|x(0)|/r_0 = 1.2$, 1% r.m.s. error, 1000 zero-current betatron periods

Later on another kick expels the particle even further to amplitudes $> 4\,r_0$. Contrary to initial mismatch, particles which are subject to statistical gradient errors seem to increase their amplitudes further and further, provided the length of transport channel is long enough.

To assess more generally the potential for halo development we evaluate in Fig. 5.8 the probability for particles to reach large amplitudes as a function of their initial amplitude and the length of the transport channel.
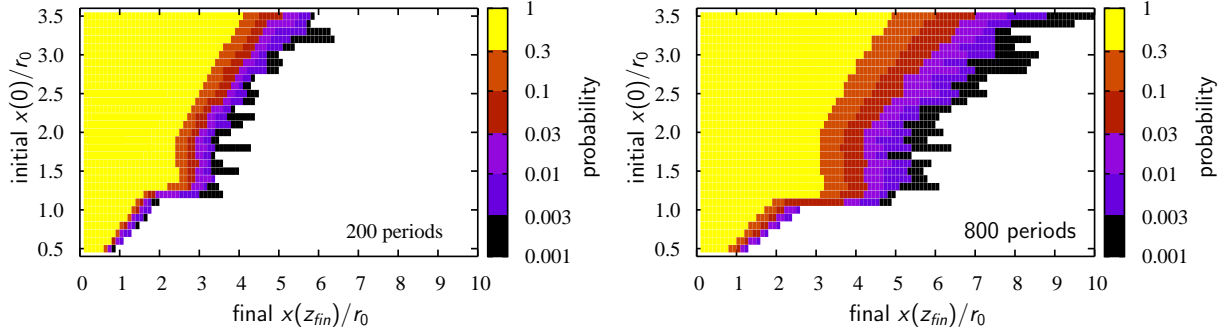

Figure 5.8: Probability for single particles to reach large amplitudes; 200 and 800 zero-current betatron periods. The black bars represent one out of a total of 1000 simulations.

Two features in Fig. 5.8 are worth noting: a) particles that start at $1.2 < |x(0)|/r_0 < 2.1$ show an almost equal probability of reaching certain large amplitudes, which can be explained by the enhanced amplitude growth that was observed earlier for particles with initial amplitudes around $\approx 1.5\,r_0$. b) Particles starting from within the core area ($x(0)/r_0 < 1$) show extremely low probability to actually transgress the core boundaries and reach larger amplitudes. As soon as the initial particle amplitudes are slightly larger than the matched core radius ($x(0)/r_0 > 1$) the probability for amplitude growth increases considerably.

5.2 Limitations of and conclusions from the particle-core model

A continuous, azimuthally symmetric, un-bunched beam in a constant focusing channel is a very crude approximation for a realistic 3D linac beam propagating through a periodic, accelerating lattice which is subject to the influence of non-linear forces. However, as a proof-of-principle, the model shows that halo development can be triggered by statistical gradient errors and that the process of expelling single particles to large amplitudes is based on a parametric resonance. It also suggests that in a space-charge dominated beam the maximum halo amplitudes are independent of the tune depression. Furthermore it shows that the maximum halo extent depends on the error amplitudes and on the length of the transport channel. Contrary to initial mismatch there seems to be no boundary for the extent of the halo produced, suggesting that particles will inevitably get lost even if one uses accelerating lattices that allow large bore radii (e.g. superconducting cavities). The model cannot predict r.m.s. emittance growth and only allows very limited conclusions about the time constants for halo development along with the associated maximum amplitudes. While the time constants for single particles to become halo particles might seem very long, the situation certainly changes when considering the large number of particles in a realistic beam. There, even a tiny fraction of particles ($10^{-5} - 10^{-7}$) acquiring large amplitudes, may cause losses that can impose limits on the maximum beam power for the operation of the machine.

5.3   3D particle tracking

To study the effects of statistical gradient errors on a realistic 3D beam we use a periodic focusing channel without acceleration and with the same basic properties as used for the particle-core model (see Table 4.1). For each case considered in the following, 500 randomly created error sets (Gaussian error distribution with cut-off at twice the r.m.s. value) are simulated with the 3D tracking code IMPACT [13], using 10^5 particles.

The particle-core model predicts (see Fig. 5.4) that the maximum amplitude growth for single particles only depends on the tune depression if the beam is in the emittance dominated regime ($0.7 < \tau < 1$). To verify this finding for a realistic 3D beam, IMPACT is used to calculate the r.m.s. emittance growth and the maximum beam amplitudes for a Gaussian particle distribution. The first plot in Fig. 5.9 shows the r.m.s. emittance growth versus tune depression assuming 0.5% and 1.0% r.m.s. errors for the quadrupoles of the periodic focusing channel.
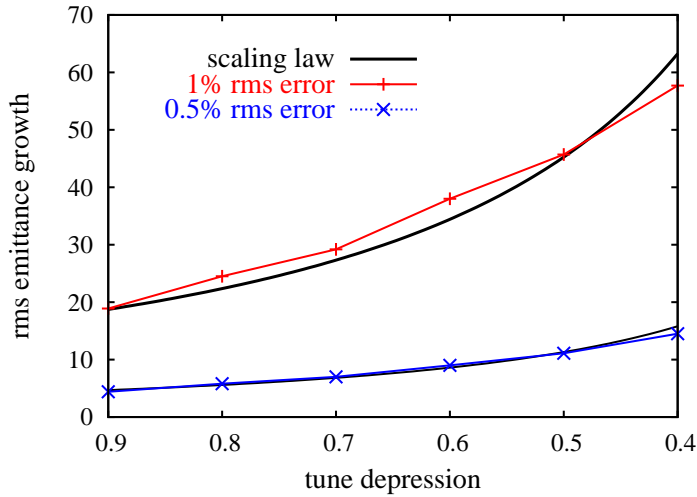

Figure 5.9: R.m.s. emittance growth for 0.5% and 1% r.m.s. quadrupole errors versus tune depression, scaling law: $\sim (k/k_0)^{-3/2}$ (200 focusing periods, 6D Gauss, 500 simulations per data point, IMPACT multi-particle simulation)

One can see that the emittance growth rises with stronger tune depression and also with increasing errors on the quadrupoles. The increase in r.m.s. emittance growth in the simulated region ($0.4 < \tau < 0.9$) scales approximately with: $(k/k_0)^{-3/2}$ and, in contrast to the predictions from the particle-core model, no difference in the growth pattern can be seen for emittance dominated or space-charge dominated beams. In order to judge if the maximum halo amplitude is also sensitive to changes in tune depression for space-charge dominated beams Fig. 5.10 can be used. It shows the average fraction of particles exceeding a certain multiple of the (average) transverse r.m.s. *input* emittance ($\varepsilon_t = (\varepsilon_x + \varepsilon_y)/2$).

Again, in contrast to the prediction from the particle-core model, the maximum halo amplitude increases with stronger tune depression in the emittance-dominated *and* space-charge dominated regime. However, in agreement with the particle-core model, the halo amplitudes increase with increasing amplitudes of the r.m.s. quadrupole errors.
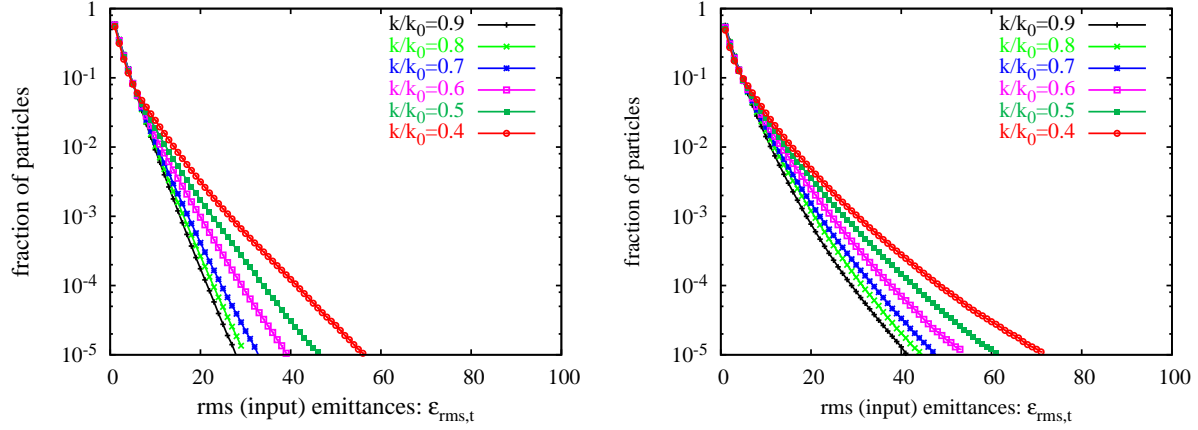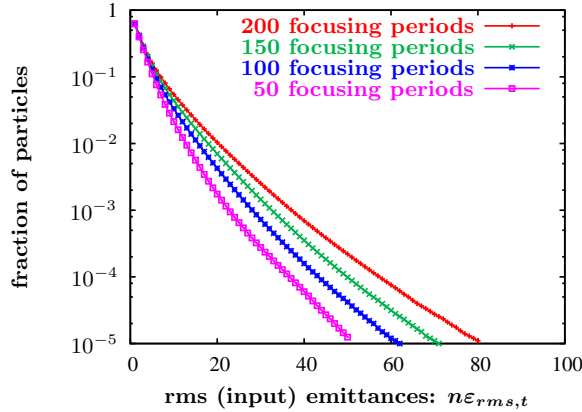
Figure 5.10: Averaged fraction of particles exceeding multiples of the r.m.s. input emittance ($\varepsilon_{r.m.s.,t}$) for transport channels with different tune depressions, left: 0.5% r.m.s. quadrupole error, right: 1.0% r.m.s. quadrupole error (200 focusing periods, 6D Gauss, 500 simulations per curve, IMPACT multi-particle simulation)

The next comparison with the particle-core model is shown in Fig. 5.11 which characterises the output distribution for transport channels of different lengths.

Figure 5.11: Averaged fraction of particles exceeding multiples of the r.m.s. input emittance ($\varepsilon_{r.m.s.,t}$) for transport channels of different length (100 focusing periods correspond to ≈ 10.7 zero-current betatron periods), 6D Gauss with 1% (r.m.s.) quadrupole gradient errors, 500 simulations, normalised to transverse *input* emittance, IMPACT multi-particle simulation

The fraction of particles exceeding the *input* distribution increases approximately linearly with the length of the transport channel which coincides with the results from the particle-core model in Fig. 5.5. In terms of the 'free energy approach' one can argue that statistical gradient errors represent a continuous feed of 'free energy' into the system, which is transformed into r.m.s. emittance growth and beam halo.

As predicted by the particle-core simulations, the maximum halo amplitudes in the case of statistical gradient errors can reach significantly higher values than in the case of initial mismatch. The probability, however, of reaching halo amplitudes in excess of 10 times the r.m.s. envelope amplitudes, seems very

low. Nevertheless, the simulations suggest that the effects of statistical gradient errors can be seen in linear accelerators, which have a high number of focusing elements in their low-energy sections. Below 150 MeV one finds approximately between 5 and 10 zero-current betatron periods in a typical normal conducting linac ($\approx$ 200 quadrupoles plus 300 RF gaps, not counting the RFQ), which, depending on the lattice characteristics, may already be long enough to yield significant losses due to statistical errors. Using for instance only 5 zero-current betatron periods with a 1% r.m.s. gradient error (Fig. 5.11), one already finds a fraction of 10^{-5} of the particles beyond 50 r.m.s. input emittances which corresponds to $\approx$ 7 r.m.s. envelope amplitudes ($7a$, an aperture limitation that often used on normal conducting accelerating structures).

In synchrotrons or storage rings for space-charge dominated beams, where the bunches are transported through a large number of lattice periods, statistical gradient errors may well account for the development of a substantial parametric beam halo that has to be scraped by dedicated beam collimation systems. In order to show that statistical errors not only increase the r.m.s. emittance but do in fact produce a low-density beam halo we plot in Fig. 5.12 the averaged fraction of particles exceeding certain multiples of the transverse r.m.s. *output* emittance for three different r.m.s. error amplitudes. By normalising each curve by its output r.m.s. emittance one basically removes the contribution of the r.m.s. emittance growth from the plots. The plot also compares the halo development with respect to an initial mismatch, using a 30% fast mode excitation at the beginning of the simulation.
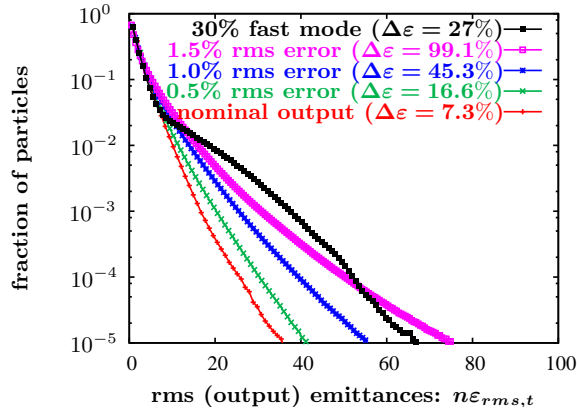

Figure 5.12: Averaged fraction of particles exceeding multiples or the respective r.m.s. output emittance $\varepsilon_{\text{r.m.s.,t}}$ for statistical gradient errors. (500 simulations per curve, 6D Gauss, 200 focusing periods, IMPACT multi-particle simulation, $\Delta\varepsilon \rightarrow$ r.m.s. emittance growth.)

In all cases one can observe a clear halo development. However, compared to the fast mode initial mismatch we find that, on average, the effects of statistical errors are much less dramatic. For the fast mode excitation one can observe a certain 'hump' in the output distribution which is likely to be a result of the redistribution pattern of the mismatched beam core (compare particle redistribution due to initial mismatch in Fig. 4.12a). For statistical gradient errors the resonant conditions are changing very rapidly and thus the output distributions become very smooth.

It was stated earlier (page 46, Fig. 4.11) that in case of initial mismatch beams with Gaussian distributions are much quicker to develop beam halo than beams with waterbag distributions. For statistical errors we find that there is not such a distinct difference between the two distributions. Figure 5.13 shows the averaged fraction of particles exceeding a certain multiple of the r.m.s. *output* emittance for simulations with initial waterbag and Gaussian distributions using initial mismatch and/or statistical errors.
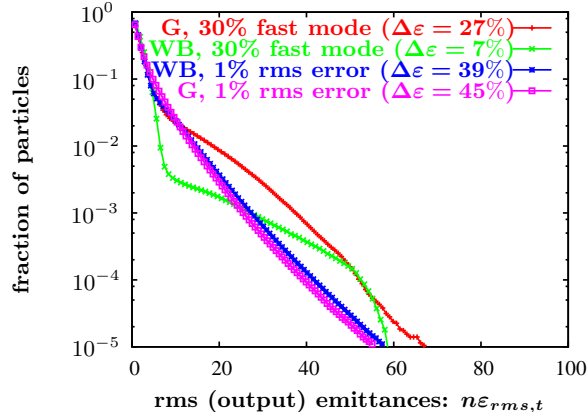
PSfrag replacements

Figure 5.13: Averaged fraction of particles exceeding multiples of the r.m.s. output emittance $\varepsilon_{\mathrm{r.m.s.,t}}$ for statistical error runs. (500 simulations, 6D Gauss and waterbag, 200 focusing periods, IMPACT multi-particle simulation, $\Delta\varepsilon \rightarrow$ r.m.s. emittance growth)

While in the case of initial mismatch the difference between the Gaussian and the waterbag beam (27% r.m.s. emittance growth versus 7%) is clearly visible, only a very small difference can be observed in case of statistical errors (45% versus 39%). This is indeed quite surprising since the particle-core model seems to suggest that those particles mainly affected by statistical errors are initially outside the core in which case the Gaussian distribution should be affected much stronger than the waterbag distribution.

5.4 Conclusions on statistical gradient errors

Using a simple particle-core model one can show that statistical gradient errors can excite oscillations of the beam core which can then, via parametric resonances, transfer energy to single particles. 3D simulations for a periodic transport channel show that this mechanism not only results in r.m.s. emittance growth but that it also initiates the development of a low-density beam halo surrounding the core. The process scales more or less linearly with the error amplitude (see Fig. 5.12) and the length of the transport channel (see Fig. 5.11) but seems to be almost insensitive to the type of particle distribution that is chosen in 3D simulations (Fig. 5.13). The particle-core model predicts that for emittance dominated beams the maximum amplitudes depend on the level of tune depression and reach their maximum when entering the space-charge dominated regime ($\tau < 0.7$) where no further dependency on the tune depression is observed. More realistic 3D simulations, however, show that also in the space-charge dominated regime the r.m.s. emittance growth as well as the maximum halo amplitudes rise with stronger tune depression.

For most simulations a 1% r.m.s. gradient error was applied, which may seem high in view of the specifications for new linac projects like the Spallation Neutron Source (SNS) (0.14% r.m.s. quadrupole gradient error [63]). Nevertheless, most of todays operating accelerators do not fulfil such tight margins and one should keep in mind that quadrupole gradient errors are only one error type that has to be added to a whole range of error sources. For high-power linacs, which are planned to operate with a beam power of 10 MW or more (e.g. linacs as fusion drivers or for waste transmutation), even much smaller error margins may not be sufficient to prevent the development of unwanted and harmful beam halo. As a rule of thumb one finds that the effect of a 1% r.m.s. quadrupole gradient error in terms of emittance growth and halo development, is comparable to the effect of a 20% initial fast mode excitation (assum-

ing a transport channel of $\approx$ 15 zero-current betatron periods and a space-charge dominated beam). In circular accelerators, statistical gradient errors provide a continuous source for halo development as long as the beam is in a regime with relevant space-charge forces. Here, they may at least partly explain the characteristic 'halo shoulders' that can be observed in synchrotrons and that have to be controlled by beam collimation systems.

For the time being the author is not aware of any significant studies on the subject of halo development for off-centred beams, which are caused by the misalignment of lattice elements. The resulting oscillations of the beam core will be introduced in the next chapter as 1st order eigenmodes of the beam. Since in this case the whole beam including the outermost particles oscillates around the nominal beam axis there is no obvious possibility for exciting a resonance between the particles and the core. It is therefore likely that any potential halo development will only occur as a 2nd order effect, and will certainly be much smaller than halo caused by statistical gradient errors.

6. Core-core resonances

For a long time it was believed that equipartitioning is a necessary feature of high-current linacs to avoid emittance exchange due to space-charge effects. A beam is called equipartitioned in all three planes if the following equation holds

$$\sigma_x \varepsilon_x = \sigma_y \varepsilon_y = \sigma_z \varepsilon_z \qquad \textbf{equipartitioning condition} \qquad (6.1)$$

where σ is the phase advance (or tune) per focusing period and ε the r.m.s. emittance. The expressions in Eq. (6.1) are usually understood as energies ($W_i = \sigma_i \varepsilon_i$) in the different planes and the idea is that anisotropic (or non-equipartitioned) beams may exchange a certain amount of "energy" so that the beam becomes equipartitioned. This effect was first identified as a coherent space-charge instability in [64] where a number of coherent eigenmodes were derived analytically for anisotropic two-dimensional beams. Jameson stressed the importance of these instabilities [65] and continued to recommend their avoidance in the design of high-intensity linacs [66] The main difference with respect to the eigenmode analysis for bunched beams in Section 4.3 is the requirement to incorporate non-elliptical and changing bunch shapes as well as changing emittances. Therefore the analytical treatment is based on an integration of the Vlasov equation and is limited to the evaluation of a KV distribution in two dimensions (see [18]). Until now no three-dimensional treatment has been found. Details of the 2D derivation are found in [18] and [67]. The analysis provides a set of 2D eigenmodes which can be characterised as in Fig. 6.1.
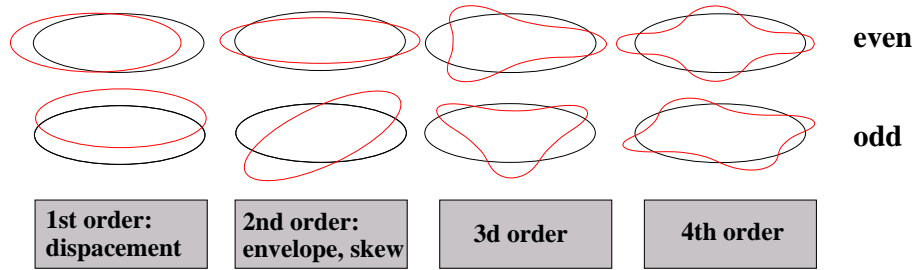

Figure 6.1: Characterisation of core-core eigenmodes

Since the analytic treatment could not be taken any further, computer simulations were employed to evaluate the effect of the identified eigenmodes on the performance of simulated multi-particle beams [68, 67]. It was found that ideal KV beams are affected by oscillating and non-oscillating modes, while waterbag beams are only sensitive to non-oscillatory (purely growing) instabilities. In both cases the r.m.s. emittances are only affected if they have different starting values in the different planes. In this context one needs to explain the terms "oscillatory" and "non-oscillatory": an eigenmode can be characterised by the eigenfrequency ω with which a beam distribution changes in time ($e^{-j\omega t}$). If the real part of the eigenmode frequency is non-zero ($Re(\omega) \neq 0$) one speaks of an oscillatory mode, while non-oscillatory modes are characterised by $Re(\omega) = 0$. Strictly speaking the non-oscillatory modes should not be labelled as eigenmodes but rather as coherent space-charge instabilities.

From Fig. 6.1 one can see that the even modes are symmetric with respect to the horizontal axis, while the odd modes have no such symmetry. Interpreting the two planes as *x* and *y*, the odd symmetry corresponds to a lack of rotational symmetry around the longitudinal axis. It should be noted that this type of mode cannot be found with *r-z* simulation codes which assume azimuthal beam symmetry.

First-order modes represent the trivial case of displacement from the beam axis, and the oscillation frequencies are just the betatron frequencies without space-charge in each direction. The second-order modes correspond to the beam eigenmodes that were derived in the Section 4.3. However, since this is a 2D analysis we would find only the two eigenmodes for a DC beam (quadrupolar and breathing mode). The real novelty of the Vlasov analysis are the third- and forth-order modes and, looking at the oscillation shape, it becomes immediately clear why these modes are found only if the mathematical tools allow non-elliptical bunch shapes. The areas where these coherent space-charge instabilities (due to the third and fourth-order modes) affect the beam can then be visualised in "stability charts" (see Fig. 6.2 or Appendix D, [67]).

Up to this stage the theory was developed using 2D beams in idealised constant focusing structures without acceleration. Even though the theory was developed already in 1979/80 it was never systematically applied for the design of high-intensity linacs, and still in 1995 Jameson [69] writes with respect to the coherent space-charge instablities: "Contemporary RF linac design,... has evolved empirically to usually avoid the unstable regions." A common approach was to simply choose an equipartitioned design, which puts severe constraints on the design of the machine. In the following section we report the results of a systematic study of emittance exchange for the superconducting part of the SPL I project that was initally performed without knowledge of the stability charts. In subsequent studies the results were compared with the predictions of the charts and first published in [70].

6.1 Application of stability charts

We now make the transition from the simplified 2D theory to a realistic linac structure with acceleration using 3D bunched beams. Although the mathematical model is derived for anisotropy between the two transverse planes of a DC beam, Hofmann suggested in [18] to apply the results for anisotropy between the transverse and the longitudinal plane. This idea was tested and verified with 3D multi-particle simulations using a simple constant-focusing lattice (as opposed to a periodic focusing lattice which is found in real accelerators) without acceleration [71]. This approach was used for the superconducting part of the SPL I project [70]. Figure 6.2 shows the stability chart for the SPL I emittance ratio of $\varepsilon_l/\varepsilon_t = 2$. The shaded areas of the chart indicate where emittance exchange between the longitudinal and the transverse plane is to be expected (the degree of shading indicates the speed of the process). For strong tune depression ($k_x/k_{x0} \leq 0.4$) one obtains emittance exchange for all tune ratios, while for moderate values stable beam operation seems possible in certain areas. The dashed line indicates the condition for an equipartitioned beam [see Eq. (6.1)].

The application of the charts for a 3D beam with anisotropy between the longitudinal and transverse plane is demonstrated with the following example: in two linac sections of SPL I (120 - 383 MeV, 40 mA simulation current, waterbag distribution), the quadrupole gradients are modified to create three different focusing lattices that fall into different areas of the stability chart in Fig. 6.2. The original SPL I and the "case 1" lattice should both be stable, while "case 2" is clearly located in an unstable area. The result of the simulations is plotted in Figs. 6.3 to 6.5.

In case of the SPL I and case 1 there is no significant change in the longitudinal/transverse r.m.s. emit-
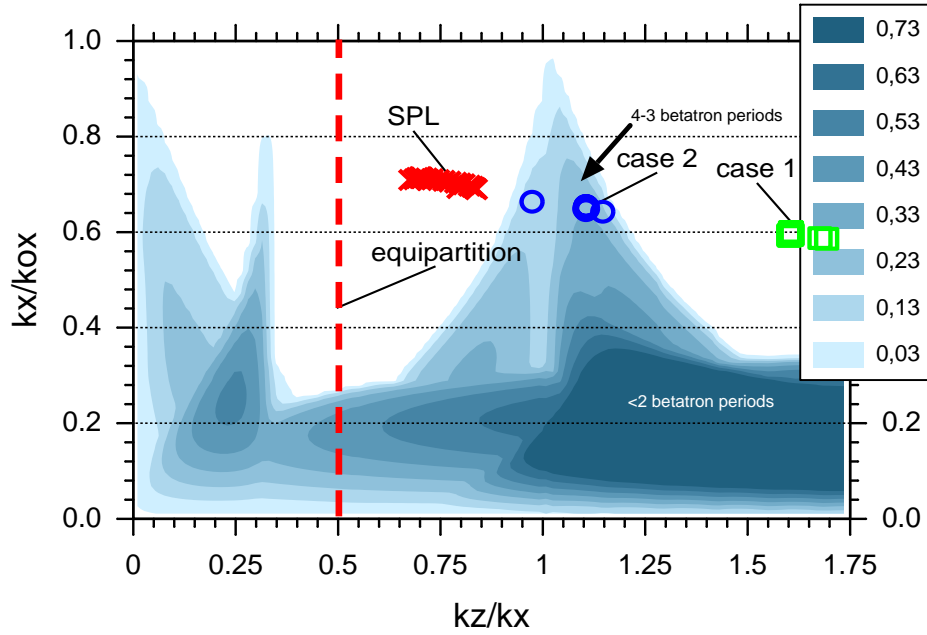
Figure 6.2: Stability chart for the SPL I emittance ratio of $\varepsilon_l/\varepsilon_t = 2$, source: Ingo Hofmann

tance values, while for case 2 one can observe a substantial emittance exchange from the longitudinal to the transverse plane. We note that in this case there is one "hot" plane, the longitudinal one, which is feeding the two "cold" transverse planes. For this reason there seems to be much more longitudinal decrease than transverse increase. In an actual linac one should try to avoid a design where the transverse emittance is higher than the longitudinal one, because in case of an exchange two "hot" planes would feed one "cold" plane and a high longitudinal emittance increase would be the consequence.
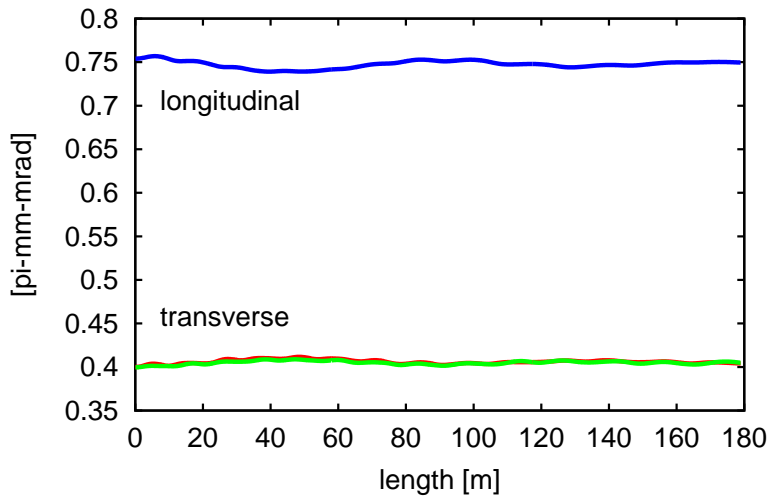

Figure 6.3: R.m.s. emittance evolution for the SPL I lattice, IMPACT multi-particle simulation
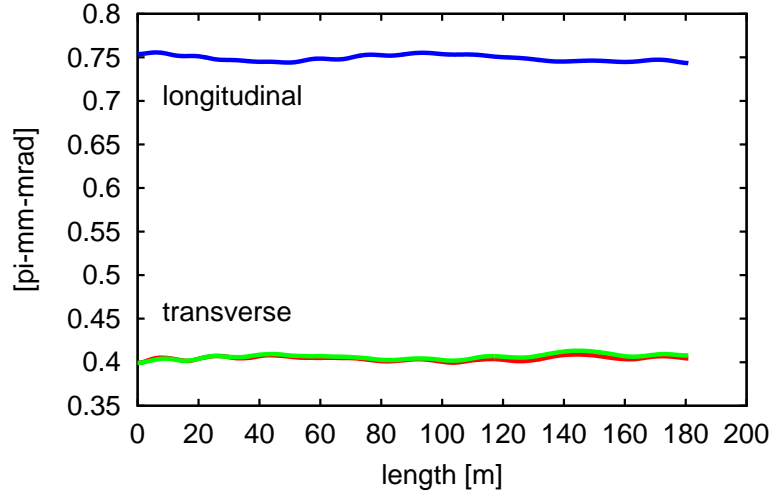
Figure 6.4: R.m.s. emittance evolution for case 1, IMPACT multi-particle simulation
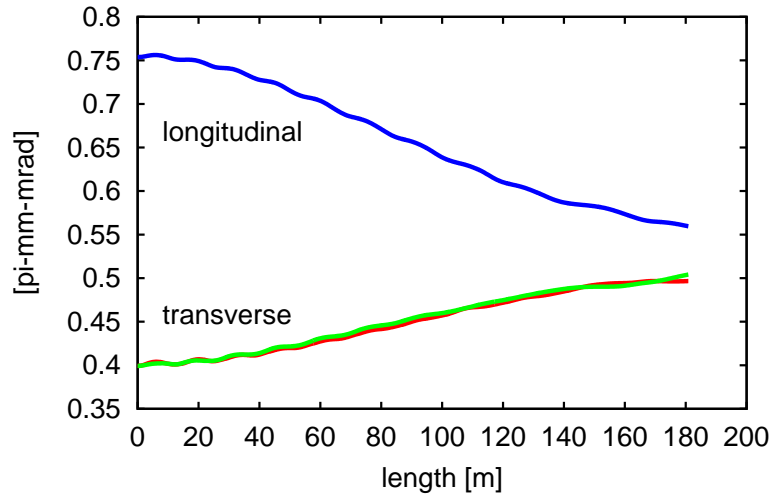

Figure 6.5: R.m.s. emittance evolution for case 2, IMPACT multi-particle simulation

In recent years the stability charts have been successfully applied in various high-intensity linac projects and should be regarded as a new tool in the design of linac lattices. Looking at the various stability charts for different emittance ratios in Appendix D, one can see that small emittance ratios close to 1.0 yield large stable areas in the charts and vice versa. Nevertheless, equipartitioning does not appear to be an obligatory design feature as was often assumed in the past. Even though emittance ratios close to 1.0 provide larger stable areas, the demand for equipartitioning puts severe restrictions on the lattice design which seems no longer justified. The emittance exchange itself takes place only if we have a combination of beam anisotropy plus a certain tune ratio plus a minimum tune depression. Up to now the understanding is that the most harmful resonance in the charts is the fourth-order even mode, which is always located around a tune ratio of 1.0.

6.2 Core-core resonances & beam halo

The obvious question to ask is if core-core resonances contribute to the development of beam halo and a simple answer can be given by looking at the fractional emittances (99%, 99.9%, etc.) for case 2 of the previous section.

In Fig. 6.6 we see that the outermost particles seem to be unaffected by the decreasing longitudinal r.m.s. emittance, while the rising transverse r.m.s. emittance in Fig. 6.7 yields a slight increase in higher fractional emittances. Altogether the ratio between the 99.99% emittances and the r.m.s. values increases by $\approx 25\%$, which can hardly be referred to as halo development. This behaviour confirms the model, which describes this kind of instability as a space-charge driven core-core resonance, which is an altogether different process than the particle-core resonances previously discussed.
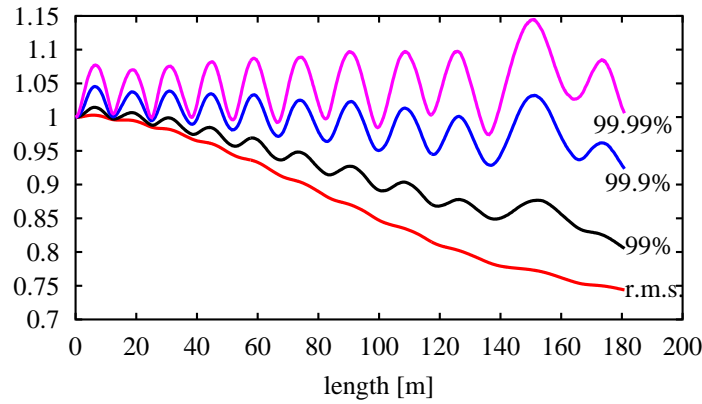

Figure 6.6: Evolution of fractional longitudinal emittances for case 2. All emittances are normalised to an initial value of 1.0, IMPACT multi-particle simulation.
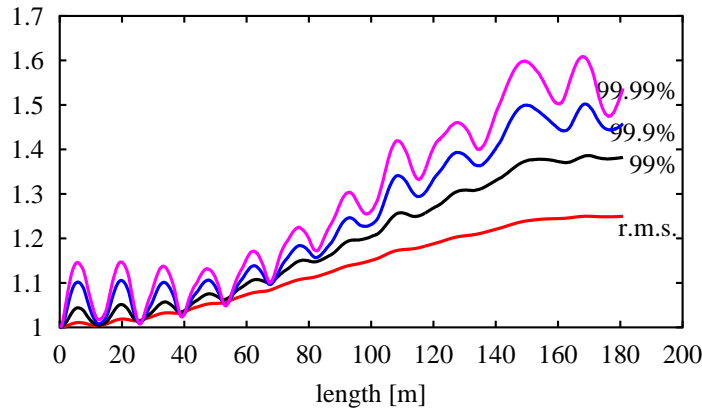

Figure 6.7: Evolution of fractional transverse emittances for case 2. All emittances are normalised to an initial value of 1.0, IMPACT multi-particle simulation.

7. Practical linac design

7.1   general rules

An important choice for each linac lattice concerns the evolution of the transverse and longitudinal phase advance. High electric fields at low energy yield strong longitudinal focusing forces [see Eq. (2.27)] and the limitation of keeping the longitudinal phase advance $< 90°$ can impose severe constraints on the maximum electric fields to be used in the first stages of acceleration. If beam loss has to be kept as low as possible then this relation enforces the use of short focusing periods and clashes with the idea of using superconducting cavities at low energy for the following reasons: the use of SC technology is most efficient in case of long drift spaces between quadrupoles, so that there is enough space to fit a cryostat with preferably several accelerating gaps. Long cryostats with several gaps, however often push the phase advance too high, so that the electric field has to be kept lower than technologically feasible and thus one of the advantages of SC technology is lost. For this reason the most common low-energy ($2 - 50\,\text{MeV}$) structure for high-intensity linacs is still the classic Alvarez Drift Tube Linac (DTL). New approaches for SC linacs use transverse focusing elements within the cryostat and thus try to overcome the space limitations.

The longitudinal focusing of the accelerating cavities decreases with increasing beam energy (for the same accelerating gradient) and if the linac aims to accelerate beyond a few hundred MeV then the longitudinal phase advance has to be limited even further at low energy: we have seen that the stability charts (see Appendix D) offer more than one stable area for linac operation. The two largest areas are either close to equipartitioning with $k_l < k_t$ or on the opposite side of the 4th order instability (which is located around $k_t \approx k_l$) at $k_l > k_t$. If the final linac energy is above a few hundred MeV and if one wants to avoid crossing the unstable areas then the area close to equipartitioning is most suitable for accelerating the beam and the appropriate tune ratio must be maintained throughout the machine. This means that in order to stay in the stable area we have to impose a certain maximum tune ratio which depends on the space-charge tune depression (e.g. $k_l/k_t < 0.8$). Restricting the maximum transverse zero-current phase advance to $85°$ per period and assuming that the zero-current tune ratio approximately equals the full-current ratio we obtain a limit of $\approx 65 - 70°$ for the maximum longitudinal zero-current phase advance per period. Using these limits one can determine the maximum energy gain per period that respects this limit which can be translated into a certain number of accelerating gaps or a maximum number (n_max) of $\beta\lambda$ per period[1]. With Eq. (2.21) one obtains a simple relation to determine n_max:

$$n_\text{max} \leq \frac{65°\pi}{180°} \sqrt{\frac{mc^2\beta\gamma^3}{2\pi\lambda\Delta W_\text{p}\tan\phi_\text{s}}} \tag{7.1}$$

where ΔW_p denotes the energy gain per focusing period. Depending on the actual tune depression one can replace the $65°$ in Eq. (7.1) with an appropriate value. The choice of focusing periods for SPL I which was chosen [72, 73] according to Eq. (7.1) is plotted in Fig. 7.1. For SPL II the same rule was applied to define the maximum electric gradients and the maximum number of accelerating cells per period [2].

---

[1]in a DTL (0-mode structure) one $\beta\lambda$ corresponds to one accelerating gap, while in a CCL ($\pi/2$-mode structure) one $\beta\lambda$ corresponds to two accelerating gaps.
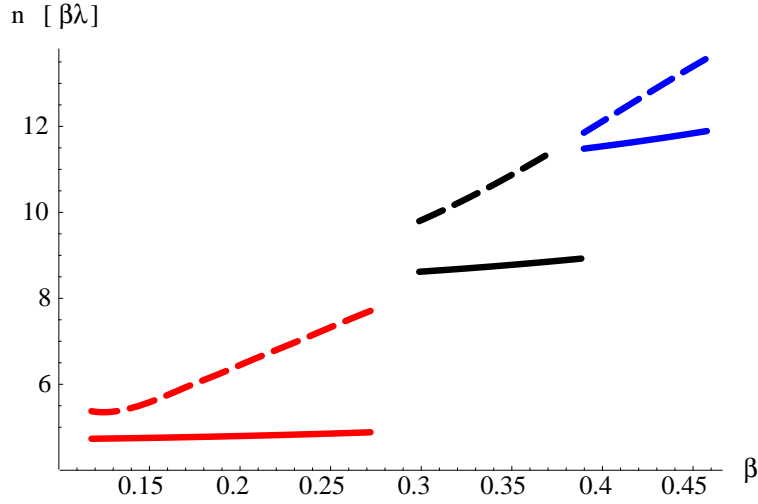
Figure 7.1: Dashed lines → maximum number of $\beta\lambda$ per focusing period to keep $\sigma_{l,0} < 65°$, solid lines → actual number of $\beta\lambda$ per period between 3 and 120 MeV. Lattice structures from left to right: DTL, CCDTL with 3 gaps per cavity, CCDTL with 4 gaps per cavity.

In low-energy machines (e.g. < 200 MeV) resonance crossing can be avoided and it can be considered to operate the complete linac in a regime where $k_l > k_t$. In this case SC cavities would be preferable in order to keep the longitudinal phase advance high throughout the machine.

Another design goal is to minimise the mismatch created by transitions from one accelerating structure to the next. Different accelerating structures have different RF efficiencies depending on the particle velocity and the RF frequency used. As one can see from Fig. 7.1 structure transitions often entail an increase in the number of accelerating gaps per focusing period yielding an increase in the "real estate" electric gradient. The most suitable way to minimise transition mismatch is to keep the phase advance per metre smooth across the transition. This measure equals the matched beam size on both sides (see Eq. [2.28]) even for beam currents different from the design current. Thus the transition becomes less sensitive to changes in the nominal beam conditions and the additionally created mismatch is minimised.

7.2   Low-energy beam chopper

In many high-intensity linac designs low-energy beam choppers are used to minimise losses at injection of the linac beam into the RF buckets of a subsequent ring system (see Section 1.1). To minimise the deflecting voltage of the chopper plates while keeping the space-charge forces at an acceptable level, these devices are usually located at energies of 2-3 MeV. Since the lattice of these transport lines has to provide enough space to house 1–2 deflecting plates of typically 0.5 m length, they demand a considerable change in the length of the focusing period. Figure 7.3 shows the impact of such a transition on the phase advance per metre (in case of the ESS chopper line which is shown in Fig. 7.2) [37] and Fig. 7.4 shows the corresponding change in transverse and longitudinal beam size.

3.13 m
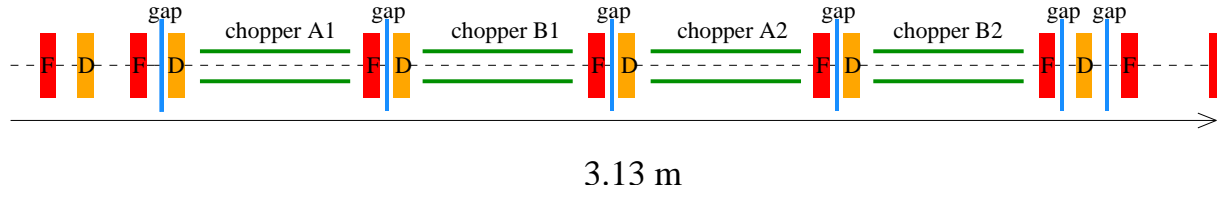
Figure 7.2: Layout of the ESS chopper line
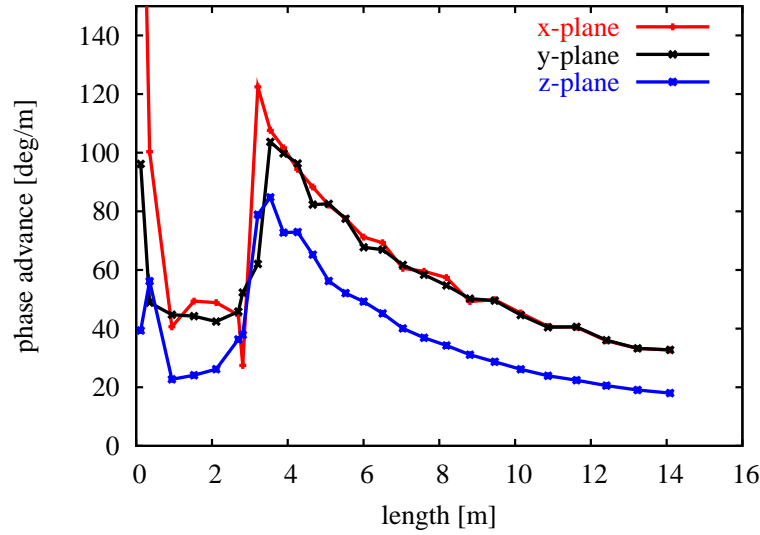
Figure 7.3: Full current phase advance per metre along chopper line and two DTL tanks in the ESS,
          IMPACT multi-particle simulation
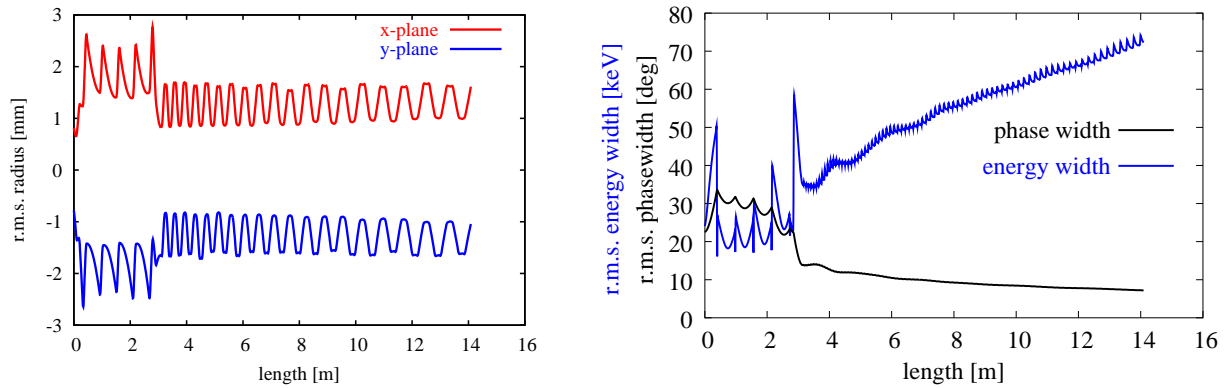
Figure 7.4: **left:** r.m.s. transverse beam envelopes along chopper line and two DTL tanks in the ESS;
          right: total r.m.s. phase and energy width, IMPACT multi-particle simulation

To accommodate the chopper line, the phase advance per metre has to be changed drastically, and, as a
consequence the otherwise smooth focusing is severely disturbed making the matching into and out of

the chopper line not only difficult but also very sensitive to errors.

For successful operation the deflecting field between the chopper plates has to rise between two successive bunches in order to avoid partly chopped bunches. For this purpose the chopper amplifiers have to provide rise times in the nanosecond range for electric fields of 0.5 – 3 kV (depending on the design of the line). Two main approaches can be taken to achieve clean chopping: a) Minimising the voltage on the chopper plates eases the task of the amplifier but implies the use of long chopper lines, so that the small deflection angle is sufficient to deviate the beam. In this case it is mandatory to optically enhance the deflection by using a deflecting quadrupole between the chopper plates and the dump which collects the deflected beam. b) Minimising the length of the chopper line reduces the emittance growth but increases the demands for high voltage at fast rise times from the amplifier. Both solutions are currently under study at RAL [74] and CERN [75] with the goal of building two test stands to investigate the technical feasibility of each solution. The RAL approach is a revised version of the ESS chopper line and aims at providing a front-end for future upgrades of the ISIS injector linac [76].

7.3 The SPL project at CERN

7.3.1 Introduction

After the shut-down of the Large Electron Proton (LEP) Collider at CERN a large number of superconducting RF cavities (including the cryogenic infrastructure) became available. It was before the actual shut-down that the idea was conceived to recuperate these RF cavities for a superconducting proton linac (SPL) at CERN [77, 78]. The first conceptual design report [1] focused on re-using the obsolete cavities, which were produced with the CERN-specific technique of sputtering a thin layer of niobium onto welded copper cavities (Nb/Cu). The main application of this linac was a CERN-based neutrino factory which dictated the timing scheme for the linac. Soon after this design report it became clear that the use of newly developed medium-β Nb/Cu cavities would be more efficient than the recuperation of the old LEP cavities [79]. In 2005 the progress in superconducting bulk niobium cavities led to abandoning the old Nb/Cu technology and to a further revision of the design, making the linac 40% shorter while increasing the energy by 60% to 3.5 GeV. The revised design of the SPL is referred to as SPL II and is published in a 2nd conceptual design report [2].

There are two potential high-power users of the SPL II: 1) EURISOL and 2) neutrino production facilities. The beam energy of 3.5 GeV was chosen as a compromise between the preferred scenarios of the EURISOL design study ([8], 1 GeV, 5 MW, CW) and the requirements of a neutrino production target (energy range 5-10 GeV, 4 MW, pulsed). For SPL II a relatively short linac pulse length was chosen in order to minimise the number of injection turns into subsequent circular machines as well as to limit the pulse length of the H$^-$ source. This means that SPL II has to use a relatively high peak current of up to 64 mA (output current) to produce a 4-5 MW beam. Table 7.1 lists the main SPL II parameters for "EURISOL" and "neutrino" operation and also lists the design differences to the first conceptual design report.

The most dramatic change for the beam dynamics is the increase in peak current from 18.4 mA to 64 mA. The resulting higher space-charge forces make it much more difficult to control the r.m.s. emittance growth and to limit the development of beam halo.

Table 7.1: Main linac parameters and changes from SPL I to SPL II

	SPL I	SPL II (neutrino)	SPL II (EURISOL)	
energy	2.2	3.5	3.5	MeV
average beam power	4	4	5	MW
length	690	430	430	m
repetition rate	75	50	50	Hz
beam pulse length	2.2	0.57	$0.71 + 0.014$	ms
average pulse current (after chopping)	11	40	40	mA
peak bunch current (after 3 MeV)	18.4	64	40	mA
beam duty cycle (after chopping)	16.5	2.9	3.6	%
injection turns (into ISR)	660	176	–	
peak RF power	32	163	163	MW
no. of 352.2 MHz tetrodes (0.1 MW)	79	3	3	
no. of 352.2 MHz klystrons (1 MW)	44	14	14	
no. of 704.4 MHz klystrons (5 MW)	–	44	44	
cryo temperature	4.5	2	2	K

7.3.2   Layout and design

SPL II comprises the following sections: H$^-$ ion source, Low Energy Beam Transport (LEBT), RFQ, DTL, Coupled Cavity DTL (CCDTL) (see Fig. B.2), Side Coupled Linac (SCL) (see Fig. B.3), and a Superconducting (SC) linac that accelerates the beam from 180 MeV to 3.5 GeV. Figure 7.5 shows a basic layout of SPL II. A detailed parameter list for this layout is given in Table 7.2.
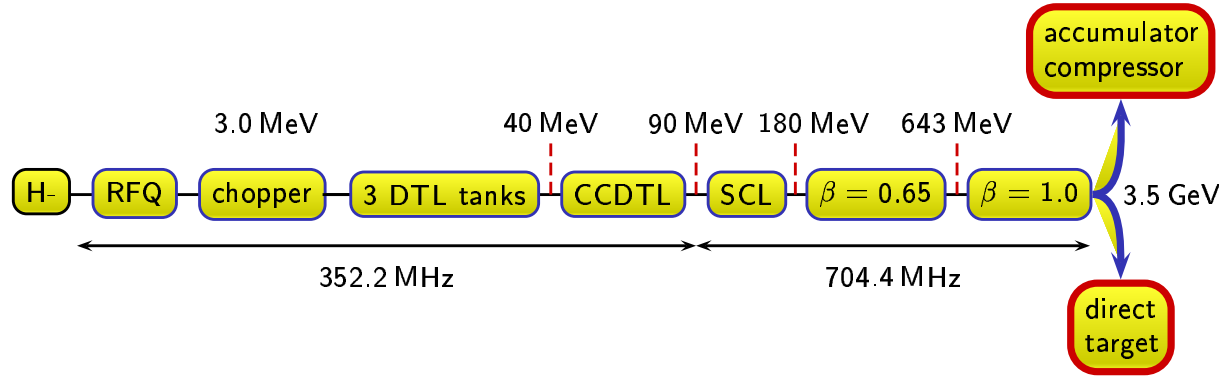
Figure 7.5: Basic layout of SPL II [2]

The H$^-$ source is necessary to make use of H$^-$ charge-exchange injection into a subsequent accumulator/compressor ring. After a magnetic LEBT with with 2 solenoids, the RFQ accelerates the beam from 95 kV to 3 MeV. Following the RFQ a beam chopper allows to create gaps in the bunch train corresponding to the transitions between RF buckets of a subsequent circular machine, reducing the injection losses to acceptable levels. A performance test of H$^-$ source, LEBT, RFQ and chopper is foreseen in a dedicated 3 MeV test stand at CERN [75]. The RF frequency of 352.2 MHz for the normal conducting part

Table 7.2: Linac layout of the SPL II at CERN

section	energy range [MeV]	no. of cavities	no. of cells	peak RF power [MW]	no. of LEP* klystrons	no. of 704.4 MHz† klystrons	length [m]
source, LEBT	– 0.095	–	–	–	–	–	3
RFQ	0.095 – 3	1	560	1.0	1	–	6
chopper line	3	3	3	0.1	–	–	3.7
DTL	3 – 40	3	85	3.8	5	–	13.4
CCDTL	40 – 90	24	72	6.4	8	–	25.2
SCL	90 – 180	24	264	15.1	–	5	34.4
$\beta = 0.65$	180 – 643	42	210	18.5	–	7	86
$\beta = 1.0$	643 – 3560	136	680	116.7	–	32	256
total				161.6	14	44	

* 352.2 MHz, 1 MW, † 4-5 MW

of the linac is determined by the existing LEP klystrons. From 90 MeV onwards 704.4 MHz klystrons are foreseen to power the normal conducting SCL which accelerates the beam up to 180 MeV. From there 704.4 MHz superconducting elliptical multi-cell cavities raise the beam energy up to its final value of 3.5 GeV. The higher frequency reduces the size and cost of the structures and it allows to use higher RF gradients in the SCL. On the other hand the frequency transition introduces the difficulty of matching a 352.2 MHz beam out of the CCDTL into the 704.4 MHz buckets of the SCL. After the chopper line a classic Alvarez DTL is used to accelerate the beam to an energy of 40 MeV. Due to the demands of high current and low losses, short focusing periods are mandatory in this energy range. Having its maximum shunt impedance at around 20 MeV, a DTL is considered the most suitable structure for this section. After 40 MeV longer focusing periods can be accepted and one can separate the quadrupoles from the RF structure. This measure simplifies the construction and eases considerably the alignment of the quadrupoles, which is why the CCDTL structure was chosen to cover the energy range from 40 to 90 MeV. Separating the quadrupoles from the RF structure has the additional advantage that larger bore diameters can be used in the quadrupoles without compromising the RF efficiency of the accelerating structure. With the increasing length of the drift tubes at higher energies it becomes more efficient to use RF structures with a gap distance of $\beta\lambda/2$ rather than $\beta\lambda$ as used in the DTL-type structures. At 704.4 MHz the chosen SCL structure can be machined out of solid copper at a reasonable price and is used to raise the beam energy up to 180 MeV. From this energy onwards, two families of bulk-niobium elliptical cavities with geometrical lengths adapted to particle velocities of $\beta = 0.65$ and $\beta = 1.0$ accelerate the beam to its final energy of 3.5 GeV. This approach reduces the amount of R&D and the production costs compared to options that foresee more families of SC cavities, which may result in a slightly shorter linac length.

The basic beam dynamics design of SPL II follows the design rules described in the beginning of this chapter. The zero-current phase advance in all three planes is below 90° per period, and the ratio between the longitudinal and transverse full-current phase advance is kept such that it avoids unstable areas in the stability charts. Furthermore the phase advance per metre is kept smooth across all lattice transitions (with the exception of RFQ/chopper and chopper/DTL). Table 7.3 shows the emittance growth per section

of the nominal beam in all three planes assuming a Gaussian beam which is matched at each structure transition.

Table 7.3: Emittance growth and losses in SPL II for an initially Gaussian distribution [2]

	RFQ	chopper	DTL	CCDTL	SCL	SC	total
energy [MeV]	3	3	40	90	180	3500	3–3500
$\Delta\varepsilon_x$ [%]	8.5	29.7	-2.4	0.7	0.75	0.3	40
$\Delta\varepsilon_y$ [%]	10.6	1.1	20.7	1.5	8.8	0.4	49
$\Delta\varepsilon_z$ [%]	–	9	11.9	0.3	0.9	4	25
transmission [%]	99.6	91.1	99.9	100	100	100	90.7
length [%]	6	3.7	13.4	25.2	34.2	341.8	425.5

In case of the CERN SPL, the chopper line also acts as a beam collimator which removes halo particles. Nevertheless one can observe the largest emittance growth in the chopper line and after the transition from chopper to DTL. As explained earlier, this behaviour is triggered by the extreme change in the focusing constants (or phase advances per metre) around the transitions into and out of the chopper line.

Extensive error simulations with TRACE_WIN [80] have been used to define limits for the statistical errors in the SPL II lattice [2]. The chosen values, which are listed in Table 7.4, ensure that no additional beam loss is caused by the statistical errors. The additional r.m.s. emittance growth due to these errors amounts to 10%/12%/17% in the x/y/z planes of the normal conducting part and to 42%/44%/17% in the x/y/z planes of the superconducting section. The working assumption is that these values can be reduced or at least maintained for the real machine with the help of an orbit correction system. In case of larger than expected emittance growth and halo development dedicated beam collimators will be needed to localise any beam loss.

Table 7.4: Total acceptable error amplitudes in the normal and superconducting sections of the SPL II

quadrupole gradient	± 0.5	± 0.5	%
quadrupole displacement	± 0.1	± 0.5	mm
quadrupole rotation (x,y)	± 0.5	± 0.25	deg
quadrupole rotation (z)	± 0.2	± 0.5	deg
cavity field phase	± 1.0	± 1.0	deg
cavity field amplitude	± 1.0	± 1.0	%

Acknowledgements

Bibliography

[1] (Editor) M. Vretenar, Conceptual design of the SPL, a high-power superconducting H^- linac at CERN, CERN 2000-012.

[2] (Editor) F. Gerigk, Conceptual design of the SPL II, a high-power superconducting H^- linac at CERN, CERN 2006-006.

[3] Spallation Neutron Source, Oakridge, USA, `http://www.sns.gov`

[4] ESS Accelerator Team, The ESS Project, technical report, ISBN 3-89336-303-3 **3** (2002).

[5] P. Gruber *et al.* Study of a European neutrino factory complex, CERN-PS-2002-080-PP (2002).

[6] A. Blondel *et al.*, Superbeam studies at CERN, CERN-Nufact-Note-95, 2001.

[7] `http://isolde.web.CERN.ch/ISOLDE`

[8] European Isotope Separation On-Line, `http://www.eurisol.org`

[9] J.D. Schneider, APT accelerator technology, Proc. Linear Accelerator Conf., Geneva, Switzerland, 1996.

[10] J.M. Lagniel, High power proton linac for a multi-user facility, Proc. European Particle Accelerator Conf., Vienna, Austria, 2000.

[11] Ed. N. Catalan-Lasheras, SNS/AP technical note 07.

[12] S. Schriber, NC/SC proton linacs, seminar, CERN, 2001.

[13] J. Qiang, R.D. Ryne, S. Habib and V. Decyk, An object-oriented parallel particle-in-cell code for beam dynamics simulation in linear accelerators, *J. Comp. Phys.* **163** (2000) 1-18.

[14] Japan proton accelerator complex, `http://j-parc.jp`

[15] D. Barni *et al.*, Status of the high current proton accelerator for the TRASCO program, Proc. European Particle Accelerator Conf., Paris, France, 2002.

[16] R.L. Gluckstern and A.V. Fedotov, Coulomb scattering within a spherical beam bunch in a high-current linear accelerator, Proc. Particle Accelerator Conf., New York, USA, 1999.

[17] N. Pichoff, Intrabeam scattering on halo formation, Proc. Particle Accelerator Conf., New York, USA, 1999.

[18] I. Hofmann, Stability of anisotropic beams with space charge, *Phys. Rev. E* **57** (1998) 4713.

[19] E. Noah, F. Gerigk, J. Lettry, M. Lindroos and T. Stora, EURISOL target stations operation and implication on its proton driver beam, Proc. European Particle Accelerator Conf., Edinburgh, United Kingdom, 2006 [CERN-AB-2006-055].

[20] International Scoping Study of a Future Neutrino Factory and Super-beam Facility, `http://www.hep.ph.ic.ac.uk/iss/`.

[21] BENE Steering Group, Beams for European neutrino experiments (BENE), Midterm scientific report, CARE-Report-06-009-BENE (2006) [CERN-2006-005, ECFA/06/242] .

[22] J.E. Campagne and A. Caze, The Θ_{13} and δ_{CP} sensitivities of the SPS-Fréjus project revisited, CERN-NUFACT-NOTE 142 (2004) [LAL-2004-102].

[23] S.J. Brooks and K.A. Walaron, Computed pion yields from a tantalum rod target: comparing MARS15 and GEANT4 across proton energies, Proc. 7th International Workshop on Neutrino Factories and Superbeams, Frascati, Italy, 2005.

[24] K.A. Walaron, Simulations of pion production in a tantalum rod target using GEANT4 with comparison to MARS, UKNF-NOTE-30 (2005).

[25] F. Gerigk and R. Garoby, Operational flexibility of the SPL as proton driver for neutrino and other applications, ICFA workshop on High-Brightness High-Intensity Beams, Tsukuba, Japan, 2006 [CERN-AB-2006-020].

[26] Paul Scherrer Institute, Villigen, Switzerland, `http://www.psi.ch`

[27] W. Chou and J. Wei, Report of the Snowmass M6 working group on high intensity proton sources, 2001.

[28] J. Wei, Synchrotrons and accumulators or high intensity protons: issues and experiences, Proc. European Particle Accelerator Conf., Vienna, Austria, 2000.

[29] F.J. Sacherer, R.m.s. envelope equations with space charge, CERN/SI/Int. DL/70-12, 1970 [*IEEE Trans. Nucl. Sci.* NS-**18**, 1105 (1971)].

[30] I. Kapchinskyi and V. Vladimirskiy, Limitations of proton beam current in a strong focusing linear accelerator associated with beam space charge, Proc. 2nd Conf. High Energy Accel., CERN, Geneva, 1959.

[31] P.M. Lapostolle, Effets de la charge d'espace dans un accélérateur linéaire à protons, CERN/AR/INT SG/65-15, 1965.

[32] T.P. Wangler, *RF Linear Accelerators*, Wiley, ISBN 0-471-16814-9 (1998).

[33] D. Swenson, D.E. Young and B. Austin, Comparison of particle motions as calculated by two different dynamics programs, Proc. Linear Accelerator Conference, Los Alamos, USA 1966.

[34] F. Gerigk, Beam dynamics in the superconducting section of the SPL (120 MeV - 2.2 GeV), CERN-NUFACT-NOTE 24 (2000).

[35] F. Gerigk and M. Vretenar, Design of the 120 MeV drift tube linac for the SPL, CERN-NUFACT-NOTE 37 (2000).

[36] The DOE grand challenge in computational accelerator physics, `http://www.nersc.gov/news/greenbook/nersc3/node15.html`, 2005.

[37] F. Gerigk, Revised ESS front-end (2.5 - 20 MeV), ESS 03-138-A, ISSN 1433-559X, 2003.

[38] L. Singleton and C.Y. Yao, TRACE3D. Interactive beam-dynamics program, ESTSC-000380SUN0000 FNAL Technical note 1993.

[39] A. Dragt, F. Neri, G. Rangarajan and J. van Zeijts, Release of Marylie 3.0, Proc. Particle Acclerator Conf., Chicago, USA 1989.

[40] R.D. Ryne, Finding matched rms envelopes in rf linacs: a Hamiltonian approach, LA-UR-95-391 (1995).

[41] F. Gerigk, Revised beam dynamics and layout for the superconducting section of the SPL, CERN-NUFACT-NOTE 62 (2001).

[42] M. Reiser, *Theory and Design of Charged Particle Beams*, J. Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore, 1994.

[43] T.P. Wangler, K.R. Crandall, R. Ryne and T.S. Wang, Particle-core model for transverse dynamics of beam halo, *Phys. Rev. Special Topics Acc. Beams* **1** (1998) 084201.

[44] G. Franchetti, I. Hofmann and D. Jeon, Anisotropic free-energy limit of halos in high-intensity accelerators, *Phys. Rev. Lett.*, **88** (2002) 254802.

[45] J. Qiang and R.D. Ryne, Beam halo studies using a three-dimensional particle-core model, *Phys. Rev. Special Topics Acc. Beams*, **3** (2000) 064201.

[46] K. Bongardt and M. Pabst, Analytical approximation of the three mismatch modes for mismatched bunched beams, ESS Note 97-95-L, 1997.

[47] A. Letchford, K. Bongardt and M. Pabst. Halo formation of bunched beams in periodic focusing systems, Proc. Particle Accelerator Conf., New York, USA, 1999.

[48] N. Pichoff, Envelope modes of a mismatched bunched beam, DAPNIA/SEA 98/44, 1998.

[49] C.K. Allen, Envelope oscillations and halo formation in bunched beams: a perturbational approach, Proc. ICFA Workshop on The Physics of High Brightness Beams, Los Angeles, USA, 1999.

[50] M. Reiser, Free energy and emittance growth in nonstationary charged particle beams, *J. Appl. Phys.*, **70** (1991) 1919–1923.

[51] I. Hofmann, G. Franchetti, J. Qiang, R. Ryne, F. Gerigk and N. Pichoff, Review of beam dynamics and space charge resonances in high intensity linacs, Proc. European Particle Accelerator Conf., Paris, France, 2002.

[52] I. Hofmann, L.J. Laslett, L. Smith and I. Haber, Stability of the Kapchinskij-Vladimirskij (K-V) distribution in long periodic transport systems, *Particle Accelerators* **13** (1983) 145.

[53] J. Struckmeier and I. Hofmann, The problem of self-consistent particle phase space distributions for periodic focusing channels, **39** (1992) 219.

[54] J. Struckmeier, Selbstkonsistente und Nichtselbskonsistente Phasenraumverteilungen intensiver Ionenstrahlen, PhD thesis (1985) [GSI-85-14].

[55] P.M. Lapostolle, Possible emittance increase through filamentation due to space charge in continuous beams, *IEEE Trans. Nucl. Sci.* NS-**18** (1971) 1101.

[56] P.M. Lapostolle and M. Weiss, Formulae and procedures useful for the design of linear accelerators, CERN-PS-2000-001-DR, 1999.

[57] I. Hofmann, J. Qiang and R.D. Ryne. Cross-plane resonance: a mechanism for very large amplitude halo formation, Proc. Particle Accelerator Conf., Chicago, USA, 2001.

[58] G. Franchetti, I. Hofmann and D. Jeon, Anisotropic free-energy limit on halos in high-intensity accelerators, *Phys. Rev. Lett.*, **88** (2002) 25.

[59] F. Gerigk, Halo studies for the ESS and Linac4 (CERN) front-ends, Proc. Workshop HALO 03, Montauk, USA, 2003, [AIP Conf. Proc., **693** (2003) 61-64].

[60] F. Gerigk, Beam halo in high-intensity hadron accelerators caused by statistical gradient errors, *Phys. Rev. Special Topics Acc. Beams*, **7** (2004) 064202.

[61] Gnuplot, a portable comman-line driven interactive data plotting utility: http://www.gnuplot.info

[62] G. Franchetti and I. Hofmann, Effect of lattice random errors on a space charge dominated beam, Proc. European Particle Accelerator Conf., Paris, France, 2002.

[63] J. Galambos, I. Anderson, N. Holtkamp and C. Strawbridge, SNS parameter list, report no. SNS 100000000-pl0001-r09, 2003.

[64] I. Hofmann, Coherent space charge instability of a two-dimensional beam, Proc. HIF workshop, Berkeley, 1979.

[65] R.A. Jameson, Beam-intensity limitations in linear accelerators, Proc. Particle Accelerator Conf., Washington DC, USA, 1981.

[66] R.A. Jameson, On scaling & optimization of high-intensity, low-beam-loss RF linacs for neutron source drivers, Proc. Workshop on Advanced Accelerator Concepts, Port Jefferson, USA, 1992 [AIP Conf. Proc. 279:969-998, 1993].

[67] I. Hofmann and O. Boine-Frankenheim, Resonant emittance transfer driven by space charge, *Phys. Rev. Lett.*, **87/3** (2001) 034802.

[68] I. Hofmann, Emittance growth of beams close to the space charge limit, IEEE Trans. Nucl. Sci, **NS-28**, 3 (1981).

[69] R.A. Jameson, Beam losses and beam halo in accelerators for new energy sources, Int. Symposium on Heavy Ion Fusion, 1995 [LA-UR-96-175].

[70] F. Gerigk and I. Hofmann, Beam dynamics of non-equipartitioned beams in the case of the SPL project at CERN, Proc. Particle Accelerator Conf., Chicago, USA, 2001.

[71] I. Hofmann, J. Qiang and R.D. Ryne, Collective resonance model of energy exchange in 3D non-equipartitioned beams, *Phys. Rev. Lett.*, **86** (2001) 2313-2316.

[72] F. Gerigk, M. Vretenar, Design choices for the SPL normal conducting front-end (3-120 MeV), CERN-Nufact-Note NF110, 2002.

[73] F. Gerigk and M. Vretenar, Design of a 120 MeV H- Linac for CERN high intensity applications, Proc. Linear Accelerator Conf., Gyeongju, Korea, 2002.

[74] M.A. Clarke-Gayther, D.C. Faircloth, D.J.S. Findlay, F. Gerigk, P.F. Harrison, A.P. Letchford and K.R. Long, The RAL front-end test stand, Proc. Workshop NuFact 04, 2005, [Nuclear Physics B, **149**, 323-325].

[75] C. Rossi *et al.*, The SPL front-end: a 3 MeV H$^-$ test stand at CERN, Proc. Linear Accelerator Conf., Lübeck, Germany, 2004.

[76] F. Gerigk, A new 180 MeV H$^-$ linac for upgrades of ISIS, Proc. European Accelerator Conf., Lucerne, Switzerland, 2004.

[77] R. Garoby and M. Vretenar, Proposal for a 2 GeV linac injector for the CERN PS, CERN PS/RF/Note 96-27, 1996.

[78] R. Garoby, H. Haseroth, C.E. Hill, A.M. Lombardi, P.N. Ostrumov, J.M. Tessier and M. Vretenar, Feasibility study of a 2 GeV superconducting H$^-$ linac as injector for the CERN PS, Proc. Linear Accelerator Conf., Chicago, USA, 1998.

[79] F. Gerigk, Revised beam dynamics and layout for the superconducting section of the SPL, CERN PS/RF/Note 2001-003, 2001.

[80] N. Pichoff and D. Uriot, TRACE_WIN, DSM/DAPNIA/CEA 2000/45.

[81] E. Wilson, An Introduction to Particle Accelerators, Oxford University Press, 2001.

[82] K.L. Brown and R.V. Servranckx, First-and second-order charged particle optics, Proc. Physics of High Energy Particle Optics, BNL/SUNY Summer School, 1983 [AIP Conf. Proc., **127** (1985) 102-103].

A. R.m.s. envelope equations and the smooth approximation

A.1   Space-charge force term

In relativistic beams the expressions for fields and momenta need to take account of the relativistic transformations. From the relativistic momentum

$$\boldsymbol{p} = \gamma m \boldsymbol{v} \quad \text{with} \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad \text{and} \quad \beta = \frac{v}{c} \tag{A.1}$$

one can derive Newton's equation for the relativistic case

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \gamma m \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} + m \boldsymbol{v} \frac{\mathrm{d}\gamma}{\mathrm{d}t} \tag{A.2}$$

For longitudinal forces, parallel to the direction of beam propagation, this relation can be written as

$$F_\parallel = \frac{\mathrm{d}p_\parallel}{\mathrm{d}t} = \gamma m \frac{\mathrm{d}v_\parallel}{\mathrm{d}t} + m v_\parallel \frac{\mathrm{d}\gamma}{\mathrm{d}t} \tag{A.3}$$

and using $\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \beta \gamma^3 \frac{\mathrm{d}\beta}{\mathrm{d}t}$ one can write the (relativistic) Newton's equation for the longitudinal plane as

$$F_\parallel = \gamma^3 m \frac{\mathrm{d}v_\parallel}{\mathrm{d}t} \qquad \textbf{Newton's equation} \atop \textbf{(longitudinal)} \tag{A.4}$$

For transverse motion, perpendicular to direction of beam propagation, the 2$^{\text{nd}}$ term on the right hand side of Eq. (A.2) vanishes and one obtains

$$F_\perp = \gamma m \frac{\mathrm{d}v_\perp}{\mathrm{d}t} \qquad \textbf{Newton's equation} \atop \textbf{(transverse)} \tag{A.5}$$

In the next step we can now relate Eq. (A.4) and (A.5) to the forces created by the space-charge field of the beam. A single particle of charge q which is moving with the beam experiences the effects of the Lorentz Force

$$\boldsymbol{F} = q \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} \right) \tag{A.6}$$

which means that the radial (transverse) forces are given by

$$F_\mathrm{r} = q \left(E_\mathrm{r} - v_\mathrm{z} B_\phi \right) = q E_\mathrm{r} \left(1 - \beta^2 \right) = \frac{q E_\mathrm{r}}{\gamma^2} \tag{A.7}$$

and the longitudinal forces are given by

$$F_z = qE_z \tag{A.8}$$

Combining Eq. (A.4) with (A.8) and Eq. (A.5) with (A.7) we find that the longitudinal and transverse acceleration can be written identically as

$$\frac{dv_i}{dt} = \frac{qE_i}{\gamma^3 m} \quad \text{with} \quad i = \text{x,y,z} \tag{A.9}$$

If one now makes the transformation from time derivatives to space derivatives using

$$\frac{dv_x}{dt} = \frac{d^2 x}{dt^2} = \beta^2 c^2 \frac{d^2 x}{ds^2} = \beta^2 c^2 x'' \tag{A.10}$$

one express the space-charge force term $F_{\text{sc},i}$ in the envelope equations as

$$F_{\text{sc},i} = \frac{qE_i}{\beta^2 \gamma^3 m c^2} \quad \text{with} \quad i = \text{x,y,z} \tag{A.11}$$

A.2   Thin lens approximation

The thin lens model is a first step to derive the transverse focusing constants for the smooth approximation. Here one assumes constant forces along each element with transitions being treated as hard-edged. This approach allows to treat focusing elements as magnetic lenses of zero length, hence the name *thin lens approximation.* Starting point is the transverse equation of motion *without* space-charge:

$$x'' + \kappa(s)x = 0 \qquad\qquad \textbf{equation of motion w/o space-charge (Hills equation)} \tag{A.12}$$

The general solution of Eq. (A.12) can be written in matrix form:

$$\begin{bmatrix} x \\ x' \end{bmatrix} = M \cdot \begin{bmatrix} x_0 \\ x_0' \end{bmatrix} \qquad\qquad \textbf{transfer matrix solution to Hills Equation} \tag{A.13}$$

When transporting a beam through a series of lattice elements 1, 2, 3, ..., n, the total transport matrix M is given by the multiplication of all single element matrices $M = M_n \cdot \ldots \cdot M_3 \cdot M_2 \cdot M_1$. The matrices for focusing and drift elements have the form:

$$M_{\text{drift}} = \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} \qquad\qquad M_{\text{lens}} = \begin{bmatrix} 1 & 0 \\ \pm\frac{1}{f} & 1 \end{bmatrix} \tag{A.14}$$

with l being the length of a drift and f being the focal length of a focusing element with length $l = 0$.

Once the total transfer matrix for one focusing period is calculated, one can relate the (known) focusing constants f of the single lattice elements to the average (unknown) focusing constants $k = \sigma/L_p$ of the whole period by comparing the transfer matrix elements with the general transverse periodic solution in matrix form:

$$M_p = \begin{bmatrix} \cos\sigma_{\text{x,y},0} + \alpha_{\text{x,y}} \sin\sigma_{\text{x,y},0} & \beta_{\text{x,y}} \sin\sigma_{\text{x,y},0} \\ -\gamma_{\text{x,y}} \sin\sigma_{\text{x,y},0} & \cos\sigma_{\text{x,y},0} - \alpha_{\text{x,y}} \sin\sigma_{\text{x,y},0} \end{bmatrix} \tag{A.15}$$

The matrix in Eq. (A.15) can be derived from Hills equation (A.12) assuming that the beam is transported through a periodic lattice. The derivation can be found in various text books (e.g. [81]) and is attributed to [82].

A.3 Quadrupole focusing in the smooth approximation

Assuming magnetic gradients of equal strength in x and y ($G = \frac{\partial B_\mathrm{x}}{\partial y} = -\frac{\partial B_\mathrm{y}}{\partial x}$) a particle of charge q will experience opposite Lorentz forces in x and y inside the quadrupole so that one can write the s-dependent magnetic focusing components as:

$$\kappa_{\mathrm{Q,x}}(s) = k_\mathrm{Q}^2(s) = \frac{|qG(s)|}{mc\beta\gamma} \qquad \kappa_{\mathrm{Q,y}}(s) = -k_\mathrm{Q}^2(s) \tag{A.16}$$

The focal length of a quadrupole in the thin lens approximation is given by

$$\frac{1}{f_\mathrm{Q}} = |\kappa|l_\mathrm{Q} = \left|\frac{qGl_\mathrm{Q}}{mc\beta\gamma}\right| \qquad \textbf{quadrupole thin lens} \atop \textbf{approximation} \tag{A.17}$$

In order to replace the s-dependent quadrupole wave number $k_\mathrm{Q}(s)$ which is only valid inside a quadrupole by a constant wave number k_Q for a full lattice period we use the smooth approximation. Starting in the middle of a focusing quadrupole the thin lens transfer matrix for one FODO period is:

$$\boldsymbol{M}_\mathrm{p} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f_\mathrm{Q}} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \frac{1}{f} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f_Q} & 1 \end{bmatrix} \tag{A.18}$$

$$= \begin{bmatrix} 1 - \frac{l^2}{2f_\mathrm{Q}^2} & 2l\left(1 + \frac{1}{2f_\mathrm{Q}}\right) \\ -\frac{l}{2f_\mathrm{Q}^2}\left(1 - \frac{l}{2f_\mathrm{Q}}\right) & 1 - \frac{l^2}{2f_\mathrm{Q}^2} \end{bmatrix} \tag{A.19}$$

Equating (A.15) and (A.19) with $L_\mathrm{P} = 2l$ yields an expression for the transverse phase advance per period

$$\frac{l}{2f_\mathrm{Q}} = \frac{L_\mathrm{P}}{4f_\mathrm{Q}} = \sin\left(\frac{\sigma_\mathrm{Q}}{2}\right) \approx \frac{\sigma_\mathrm{Q}}{2} \tag{A.20}$$

which can be used with Eq. (A.17) and $k = \sigma/L_\mathrm{P}$ to define the quadrupole wave number in the smooth approximation. The s-dependent focusing $\kappa(s)$ in the envelope equation can thus be replaced by

$$\overline{\kappa_{\mathrm{Q,x}}(s)} = \overline{\kappa_{\mathrm{Q,y}}(s)} \longrightarrow k_\mathrm{Q}^2 = \left(\frac{qGl_\mathrm{Q}}{2mc\beta\gamma}\right)^2 \qquad \textbf{smooth approximation for} \atop \textbf{transverse focusing in FODO} \atop \textbf{lattices w/o RF} \tag{A.21}$$

Please note the difference in power and sign between the expressions for $k_\mathrm{Q}^2(s)$ in Eq. (A.16) describing a beam inside a quadrupole and k_Q^2 in Eq. (A.21) describing a FODO quadrupole channel in the smooth approximation.

A.4   RF focusing in the smooth approximation

The forces of the RF system on the beam have to be divided into a longitudinally focusing part and a transversely defocusing part. The longitudinal part is based on the phase and energy difference between the particles of a bunch and the synchronous particle. The transverse component is based on the radial kick of the electric fields in the gap: due to the time-dependent field rise in the gap most particles experience a higher field in the second half of the gap than in the first half, resulting in a net defocusing force. Effects due to different radial positions of the particles and velocity change are less important for ion linacs.

A.4.1   Longitudinal focusing

Following [42] in the case of small acceleration rate one can write the linearised longitudinal equation of motion as

$$(\phi - \phi_{\mathrm{s}})'' = \frac{2\pi q E_0 T}{mc^2 \lambda \beta^3 \gamma^3} \sin \phi_{\mathrm{s}}(\phi - \phi_{\mathrm{s}}) \tag{A.22}$$

where ϕ_{s} denotes the synchronous phase and where $E_0 T$ is given by the transit time factor integral over the accelerating gap:

$$E_0 T = \int\limits_{-L/2}^{L/2} E(t=0, s) \cos\left(\frac{\omega s}{\beta c_0}\right) \mathrm{d}s \tag{A.23}$$

For small amplitude oscillations and ignoring the non-linear parts, Eq. (A.22) can be simplified to

$$(\phi - \phi_{\mathrm{s}})'' + k_{\mathrm{l},0}^2 (\phi - \phi_{\mathrm{s}}) = 0 \tag{A.24}$$

or, using $\phi - \phi_{\mathrm{s}} = -\frac{\omega}{\beta c}(s - s_{\mathrm{s}}) = -\frac{\omega}{\beta c} z$ one can transform (A.22) into the z frame:

$$z'' + k_{\mathrm{l},0}^2 z = 0 \qquad \text{\textbf{longitudinal equation of}} \atop \text{\textbf{motion w/o space-charge}} \tag{A.25}$$

with the longitudinal zero-current wave number $k_{\mathrm{l},0}$ being defined as

$$\overline{\kappa_{\mathrm{l},0}(s)} \rightarrow k_{\mathrm{l},0}^2 = \frac{2\pi q E_0 T \sin(-\phi_{\mathrm{s}})}{mc^2 \lambda \beta^3 \gamma^3} \qquad \text{\textbf{smooth approximation for}} \atop \text{\textbf{longitudinal focusing}} \tag{A.26}$$

A.4.2   Transverse defocusing

The electric field in an RF gap rises during the passage of a bunch yielding a net radially defocusing force. From $\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B})$ we can write the radial Lorentz Force as:

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{\mathrm{r}} = q(E_{\mathrm{r}} - \beta c B_{\Phi}) \tag{A.27}$$

The corresponding field components of an RF gap as seen by a particle of phase ϕ are [32]:

$$E_{\mathrm{r}} = -\gamma_s E_0 T I_1 \left(\frac{2\pi r}{\gamma_{\mathrm{s}} \beta_{\mathrm{s}} \lambda} \right) \cos(\phi) \tag{A.28}$$

$$B_{\Phi} = -\frac{\gamma_{\mathrm{s}} \beta_{\mathrm{s}}}{c} E_0 T I_1 \left(\frac{2\pi r}{\gamma_{\mathrm{s}} \beta_{\mathrm{s}} \lambda} \right) \sin(\phi) \tag{A.29}$$

with I_0 and I_1 being modified Bessel functions and β_{s} and γ_{s} being the relativistic factors for the synchronous particle. Assuming small acceleration and $\beta_{\mathrm{s}} \approx \beta$ one can express the radially defocusing force as:

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{\mathrm{r}} = -q\gamma E_0 T I_1 \left(\frac{2\pi r}{\gamma\beta\lambda} \right) \sin(\phi)(1 - \beta^2) \tag{A.30}$$

a form that nicely shows that this force vanishes for particle velocities approaching the speed of light. Substituting

$$p_{\mathrm{r}} = mc\gamma\beta r' \qquad \mathrm{d}s = \beta c\,\mathrm{d}t \tag{A.31}$$

and approximating the modified Bessel function by $I_1(z) \approx z/2$ one can write a smoothed equation of motion (assuming a small acceleration rate and a small velocity difference between the synchronous particle and the particles of the bunch)

$$\frac{1}{\beta\gamma} \frac{\mathrm{d}}{\mathrm{d}s} \beta\gamma r' - \frac{k_{1,0}^2}{2} r = 0 \tag{A.32}$$

with $k_{1,0}^2$ being the longitudinal zero-current wave number defined in Eq. (2.21). Integrating Eq. (A.32) one can then derive the thin lens approximation for an accelerating gap:

$$\frac{1}{f_{\mathrm{g}}} = \frac{\pi q E_0 T \sin(-\phi) l_{\mathrm{g}}}{mc^2 \lambda (\beta\gamma)^3} \qquad \textbf{gap thin lens approximation} \tag{A.33}$$

Now one can replace the drifts in the FODO lattice by gaps of length l_{g} and calculate the transfer matrix for a FODO period with RF gaps. Starting in the centre of the focusing quadrupole one obtains

$$
M_{\mathrm{P}} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f_{\mathrm{Q}}} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l_{\mathrm{g}} \\ \frac{1}{f_{\mathrm{g}}} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \frac{1}{f_{\mathrm{Q}}} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & l_{\mathrm{g}} \\ \frac{1}{f_{\mathrm{g}}} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -\frac{1}{2f_{\mathrm{Q}}} & 1 \end{bmatrix} \tag{A.34}
$$

$$
= \begin{bmatrix} 1 + \frac{l_{\mathrm{g}}}{f_{\mathrm{g}}} - \frac{l_{\mathrm{g}}^2}{2f_{\mathrm{Q}}^2} & 2l_{\mathrm{g}} + \frac{l_{\mathrm{g}}^2}{f_{\mathrm{Q}}} \\ \frac{2}{f_{\mathrm{g}}} - \frac{l_{\mathrm{g}}}{2f_{\mathrm{Q}}^2} - \frac{l_{\mathrm{g}}}{f_{\mathrm{Q}}f_{\mathrm{g}}} + \frac{l_{\mathrm{g}}^2}{4f_{\mathrm{Q}}^3} & 1 + \frac{l_{\mathrm{g}}}{f_{\mathrm{g}}} - \frac{l_{\mathrm{g}}^2}{2f_{\mathrm{Q}}^2} \end{bmatrix} \tag{A.35}
$$

Equating (A.15) and (A.35) yields an approximate expression for the transverse phase advance per period

$$\sigma_{\mathrm{t},0}^2 \approx \left(\frac{l_{\mathrm{g}}}{f_{\mathrm{Q}}} \right)^2 - \left(\frac{4l_{\mathrm{g}}}{f_{\mathrm{g}}} \right) \tag{A.36}$$

which one can use to define the transverse wave number for a FODO channel including RF focusing in the smooth approximation [using Eqs. ((A.17), (2.21), (2.22)]:

$$
\begin{aligned}
\overline{\kappa_{\mathrm{x}}(s)} &= \overline{\kappa_{\mathrm{y}}(s)} \longrightarrow k_{\mathrm{t},0}^2 \\
k_{\mathrm{t},0}^2 &= \left(\frac{qGl_{\mathrm{Q}}}{2mc\beta\gamma} \right)^2 - \frac{\pi q E_0 T \sin(-\phi_{\mathrm{s}})}{mc^2 \lambda \beta^3 \gamma^3} \\
&= k_{\mathrm{Q}}^2 - \frac{k_{\mathrm{l},0}^2}{2}
\end{aligned}
$$

smooth approximation for transverse focusing in FODO lattices with RF (A.37)

It is important to note that the RF defocusing term in Eq. (A.37) remains constant for different lattice types while the quadrupole term depends on the lattice structure (e.g. FDO, or FOFODODO).

B. Accelerating structures
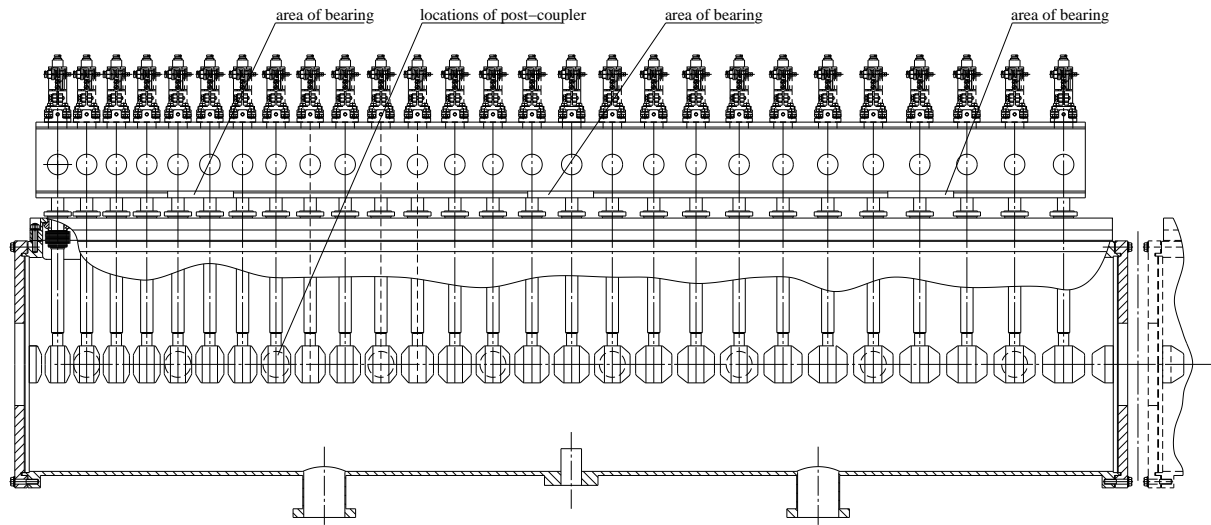
Figure B.1: Proposed 1st DTL tank for Linac4 with external girder for the alignment of the drift tubes, drawing from VNIIEF, Sarov
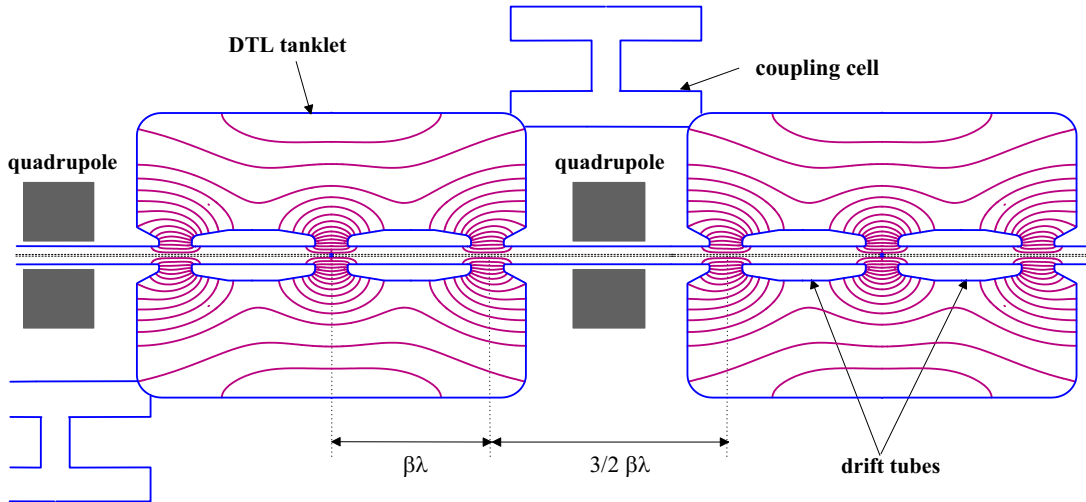
Figure B.2: Basic Coupled Cavity Drift Tube Linac (CCDTL) structure with 3 gaps per tank and quadrupoles between the tanks
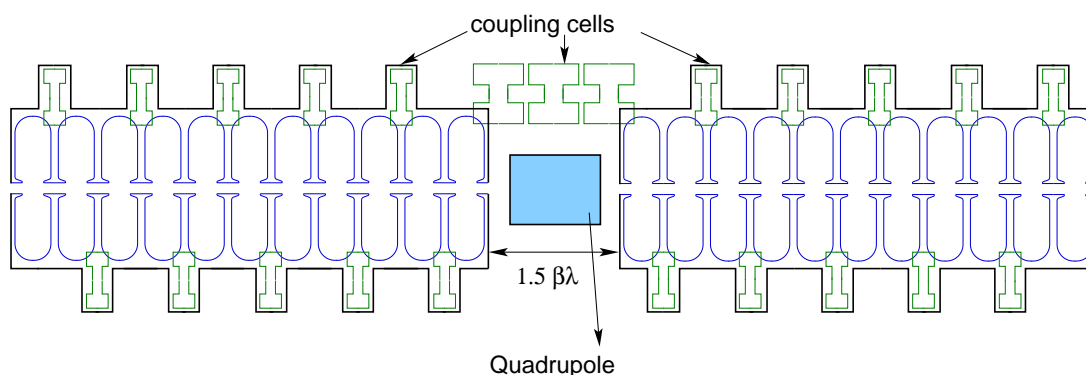
coupling cells

1.5 βλ

Quadrupole

Figure B.3: Basic Side Coupled Linac (SCL) structure

C. Derivation of envelope eigenmodes

Equations (4.5) together with (4.7) are linearised using Taylor expansions around the points $\Delta a_x = 0$, $\Delta a_y = 0$, $\Delta b = 0$

$$\frac{\varepsilon_t^2}{(\tilde{a}_{x,y} + \Delta a_{x,y})^3} \approx \frac{\varepsilon_t^2}{\tilde{a}_{x,y}^3} - \frac{3\varepsilon_t^2}{\tilde{a}_{x,y}^4} \cdot \Delta a_{x,y} \tag{C.1}$$

$$\frac{\varepsilon_l^2}{(\tilde{b} + \Delta b)^3} \approx \frac{\varepsilon_l^2}{\tilde{b}^3} - \frac{3\varepsilon_l^2}{\tilde{b}^4} \cdot \Delta b \tag{C.2}$$

$$\frac{K_3[1 - f(s)]}{(\tilde{a}_{x,y} + \Delta a_{x,y})(\tilde{b} + \Delta b)} \approx K_3[1 - \overline{f(s)}] \cdot \left(\frac{1}{\tilde{a}_{x,y}\tilde{b}} - \frac{\Delta a_{x,y}}{\tilde{a}_{x,y}^2\tilde{b}} - \frac{\Delta b}{\tilde{a}\tilde{b}^2} \right) \tag{C.3}$$

$$\frac{K_3 f(s)}{(\tilde{a}_x + \Delta a_y)(\tilde{a}_y + \Delta a_y)} \approx K_3\overline{f(s)} \cdot \left(\frac{1}{\tilde{a}_x\tilde{a}_y} - \frac{\Delta a_x}{\tilde{a}_x^2\tilde{a}_y} - \frac{\Delta a_y}{\tilde{a}_x\tilde{a}_y^2} \right) \tag{C.4}$$

neglecting all non-linear terms. After this modification we obtain a system of three coupled linear differential equations of Hill's type for the perturbation:

$$\frac{d^2}{ds^2} \begin{pmatrix} \Delta a_x \\ \Delta a_y \\ \Delta b \end{pmatrix} + \underbrace{\begin{pmatrix} \left(k_{t,0}^2 + \frac{3\varepsilon_t^2}{\tilde{a}_x^4} \right) & \frac{K_3[1 - \overline{f(s)}]}{\tilde{a}_y^2\tilde{b}} & \frac{K_3[1 - \overline{f(s)}]}{\tilde{a}_y\tilde{b}^2} \\ \frac{K_3[1 - \overline{f(s)}]}{\tilde{a}_x^2\tilde{b}} & \left(k_{t,0}^2 + \frac{3\varepsilon_t^2}{\tilde{a}_y^4} \right) & \frac{K_3[1 - \overline{f(s)}]}{\tilde{a}_x\tilde{b}^2} \\ \frac{K_3\overline{f(s)}}{\tilde{a}_x^2\tilde{a}_y} & \frac{K_3\overline{f(s)}}{\tilde{a}_x\tilde{a}_y^2} & \left(k_{l,0}^2 + \frac{3\varepsilon_l^2}{\tilde{b}^4} \right) \end{pmatrix}}_{\mathbf{M}} \cdot \begin{pmatrix} \Delta a_x \\ \Delta a_y \\ \Delta b \end{pmatrix} = \vec{0} \tag{C.5}$$

Employing once more the equations of the smooth approximation (2.19) - (2.20) one can replace the matrix elements in $\mathbf{M}$ with expressions for the phase advances per period. Furthermore one can now, after having removed the s-dependency of the matched envelopes, assume the transverse envelopes to be equal ($\tilde{a}_x, \tilde{a}_y \to \hat{a}_x = \hat{a}_y = \hat{a}$, $\tilde{b} \to \hat{b}$), which at this point corresponds to assuming equal quadrupole strengths in x and y or to a channel with solenoid focusing. With Eqs. (2.28) and (2.29) the previous system of equations can be re-written as

$$\frac{d^2}{ds^2} \begin{pmatrix} \frac{\Delta a_x}{\hat{a}} \\ \frac{\Delta a_y}{\hat{a}} \\ \frac{\Delta b}{\hat{b}} \end{pmatrix} + \frac{1}{L_P^2} \begin{pmatrix} \sigma_{t,0}^2 + 3\sigma_t^2 & \sigma_{t,0}^2 - \sigma_t^2 & \sigma_{t,0}^2 - \sigma_t^2 \\ \sigma_{t,0}^2 + 3\sigma_t^2 & \sigma_{t,0}^2 - \sigma_t^2 & \sigma_{t,0}^2 - \sigma_t^2 \\ \sigma_{l,0}^2 + 3\sigma_l^2 & \sigma_{l,0}^2 - \sigma_l^2 & \sigma_{l,0}^2 - \sigma_l^2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\Delta a_x}{\hat{a}} \\ \frac{\Delta a_y}{\hat{a}} \\ \frac{\Delta b}{\hat{b}} \end{pmatrix} = \vec{0} \tag{C.6}$$

Using the ansatz

$$\begin{pmatrix} \frac{\Delta a_x}{\hat{a}} \\ \frac{\Delta a_y}{\hat{a}} \\ \frac{\Delta b}{\hat{b}} \end{pmatrix} = e^{i\left(\frac{\sigma_{env}}{L_P}\right)s} \cdot \begin{pmatrix} A \\ B \\ C \end{pmatrix} \tag{C.7}$$

we obtain a homogeneous system of linear equations for the envelope tunes (σ_{env}) of the eigenmodes:

$$\frac{1}{L_{\text{P}}^2} \begin{pmatrix} \sigma_{\text{t},0}^2 + 3\sigma_{\text{t}}^2 - \sigma_{\text{env}}^2 & \sigma_{\text{t},0}^2 - \sigma_{\text{t}}^2 & \sigma_{\text{t},0}^2 - \sigma_{\text{t}}^2 \\ \sigma_{\text{t},0}^2 - \sigma_{\text{t}}^2 & \sigma_{\text{t},0}^2 + 3\sigma_{\text{t}}^2 - \sigma_{\text{env}}^2 & \sigma_{\text{t},0}^2 - \sigma_{\text{t}}^2 \\ \sigma_{\text{l},0}^2 - \sigma_{\text{l}}^2 & \sigma_{\text{l},0}^2 - \sigma_{\text{l}}^2 & \sigma_{\text{l},0}^2 + 3\sigma_{\text{l}}^2 - \sigma_{\text{env}}^2 \end{pmatrix} \cdot \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \vec{0} \qquad \text{(C.8)}$$

By setting the determinant to zero one finds three non-trivial solutions for σ_{env}.
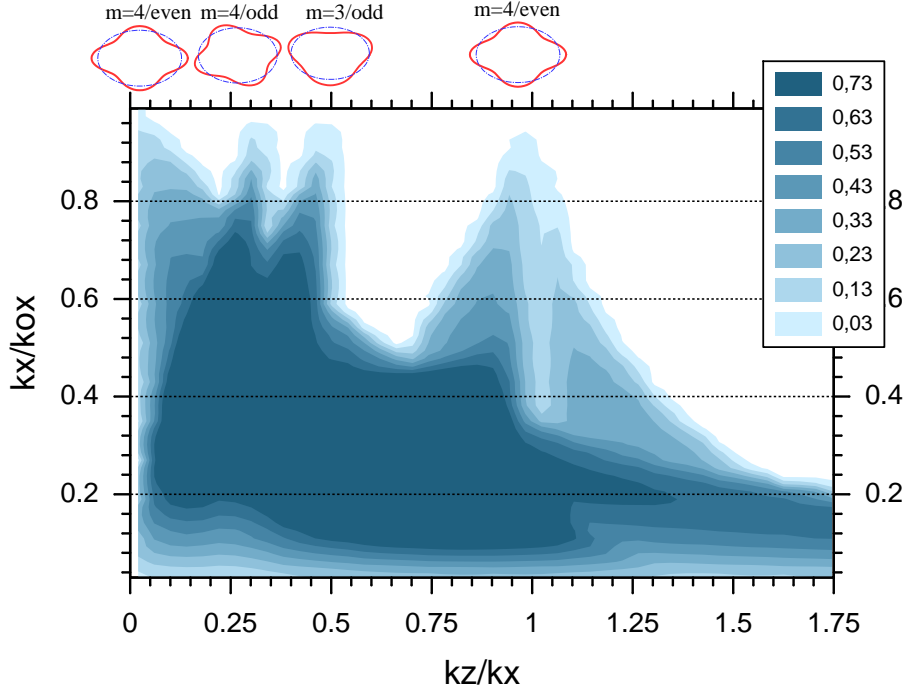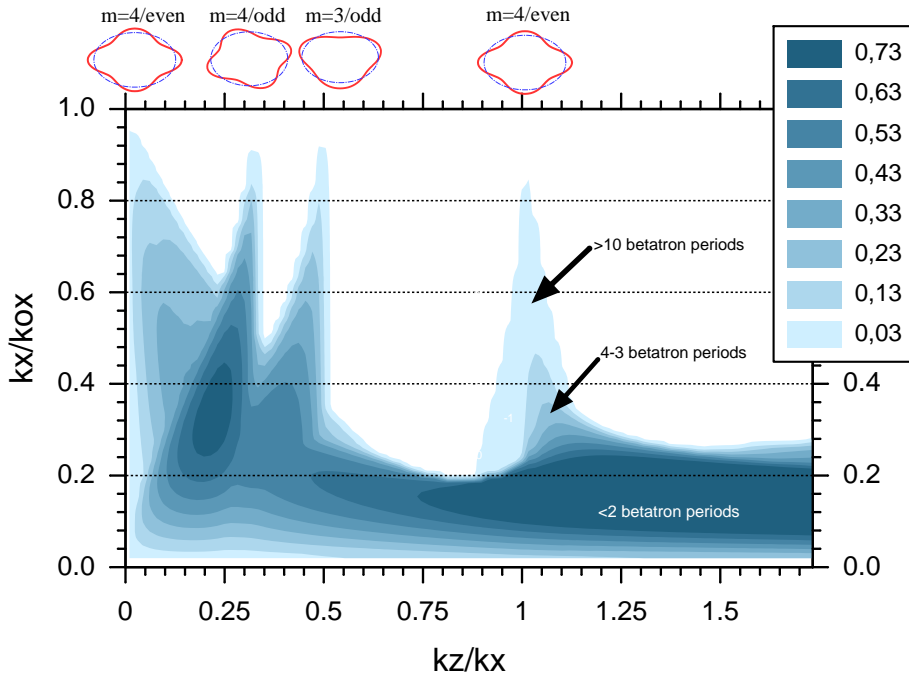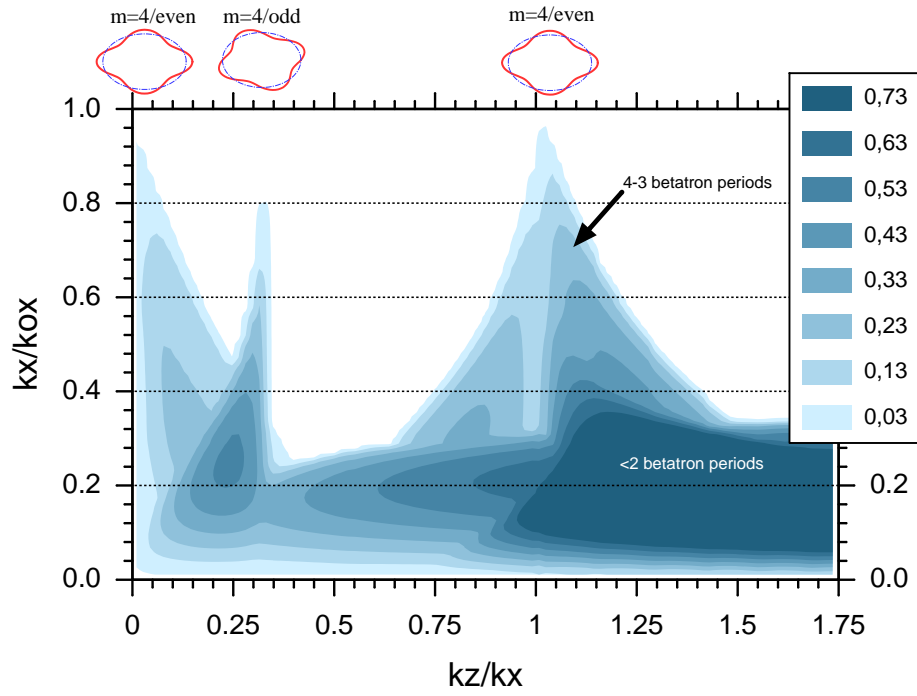
D. Stability charts

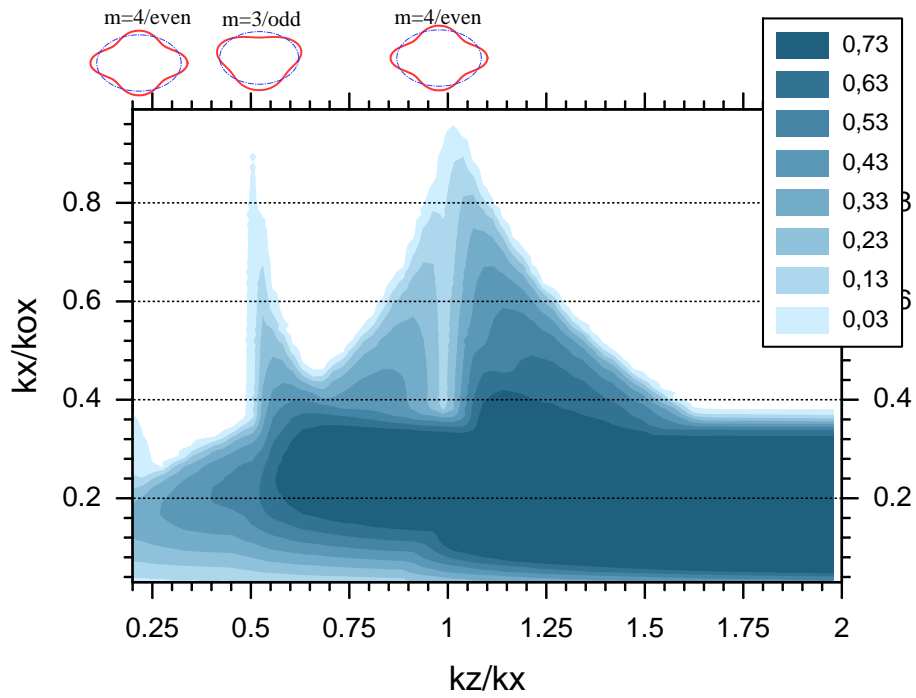Figure D.1: Stability chart for $\varepsilon_l/\varepsilon_t = 0.5$

Figure D.2: Stability chart for $\varepsilon_l/\varepsilon_t = 1.2$

Figure D.3: Stability chart for $\varepsilon_{\mathrm{l}}/\varepsilon_{\mathrm{t}} = 2.0$

Figure D.4: Stability chart for $\varepsilon_{\mathrm{l}}/\varepsilon_{\mathrm{t}} = 3.0$