Beam Dynamics Studies and RF Cavity Design for Proton Radiotherapy Machines

Matthew Southerby
School of Engineering
Lancaster University

A thesis submitted for the degree of
Doctor of Philosophy

February, 2024

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. A rough estimate of the word count is: 49850

Matthew Southerby

Beam Dynamics Studies and RF Cavity Design for Proton Radiotherapy Machines

Matthew Southerby
School of Engineering, Lancaster University
A thesis submitted for the degree of *Doctor of Philosophy*. February, 2024

Abstract

Abstract.

In recent years, there has been a rigorous effort to make proton radiotherapy a more viable and accessible source of cancer treatment. Due to the dose deposition rate of protons, there are an array of advantages of using protons over conventional x-rays. In order to treat deep seated tumours in all body sizes, protons of energy 250 MeV are required. As hospitals and treatment centers often have limited space, and budgets, proton machines must be compact and as affordable as possible. This thesis discusses a conceptual design of an S-band rf cavity for proton radiotherapy. The energy range is 150 - 250 MeV. The work is completed in conjunction with AVO and the Cockcroft Institute, in addition to Lancaster University. As the accelerator is the last stage of an all-linac machine, the beam emittance is relatively low, expanding the possible design space of solutions. For a limit on the cavity aperture of 2.5 mm, multiple cavity designs are explored, with respect to constraints such as cavity length, available rf power, and risk of rf breakdown. Firstly, a single cell is designed with the cell shape defined by splines, allowing for lossless increase in rf efficiency. The optimisation is completed using using multi-objective genetic algorithms and various visualisation techniques are explored to best represent the design space. The transverse beam dynamics are explored utilising a novel technique, and two newly developed focusing schemes are explored analytically. The schemes are solved such that the maximum cavity length for a given beam emittance is obtained. In order to quickly assess the performance of a cavity with respect to a given input power, a novel fast tracking algorithm is developed. The algorithm uses the on-axis electric field to approximate the energy gain of a particle over one rf cell assuming constant particle velocity. The tracking code is expanded to incorporate 6D phase space tracking, and is bench-marked relative to well-known tracking codes. With the limit on the cavity length, both Standing and Traveling Wave structures are explored with respect to input rf power. Cavity types are compared with respect to important design parameters, such as required power and fill time. The conceptual design is completed with the addition of matching cells and fine tuning of the structure, to allow for efficient power coupling. Finally, an approximation of the transmission is obtained using the fast tracking algorithm - confirming the high transmission required for linear proton machines.

Acknowledgements

I would like to firstly thank my Supervisor, Dr Robert Apsimon. Rob's continued support and friendship provided me with an environment in which I could thrive and allowed me to maintain my best effort throughout the last 3 years. I know that student-supervisor relationships can often be tested during post-graduate study, so I am eternally grateful to have had a supervisor who only strived to help me. It feels like only yesterday we were sipping beer on the Liverpool docks, ready for a busy week editing at LINAC'22.

I would also like to thank my second Supervisor, Prof. Graeme Burt, who was instrumental in my time at Lancaster. Whilst the second Supervisor is a fluid role, I greatly appreciated Graeme's extensive knowledge and assistance during my research. I would like to thank my industrial supervisor, Dr. Stefano Benedetti, for his continued advice, guidance, and interest, in the project.

I would like to thank fellow PhD student, Sam Smith, who has taught me everything I know about the CST software. In addition, for the brilliant week in Venice, for IPAC'23, one of my best memories during my time at Lancaster. I also want to extend my thanks to Dr. Sadiq Setiniyaz, for his unwavering encouragement and help over the last few years, in addition to our intense but rewarding week in Strasbourg.

I would like to thank my friends and family who have always been there for me. Specifically, I offer my deep thanks to my parents, who have always believed in me and my aspirations. I am grateful for my partner, Sophie, who I was able to spend my holidays with.

Lastly, I would like to thank the one person who kept me going throughout the entirety of my study, the only person who I wanted to keep me company during an afternoon of intense maths, Taylor Swift.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cancer Treatment with Radiation

1.1.1 Cancer Rates

In 2019, 18% of all deaths worldwide were attributed to cancer, with a total of 10 million deaths [1]. In approximately 50% of all cancer cases, radiotherapy will be used during the course of treatment [2]. Whilst high-income countries can provide radiotherapy treatment for all patients, low-income countries can offer the treatment to only 10% of patients [3], and the treatment is often palliative, as opposed to curative. Developments in Proton Beam Therapy (PBT) have been observed over the last 60 years due to the more advantageous dose deposition relative to conventional radiotherapy [4]. As the PBT technology improves, facilities can become more accessible to all patients worldwide, and provide patients with life-saving treatments.

1.1.2 Radiation to Treat Cancerous Cells

The X-ray was discovered by German scientist Wilhelm Conrad Röntgen in 1895 [5]. Soon after, doctors started using radiation to treat skin abnormalities due to disease such as lupus [6]. The first used of X-rays to treat cancer was in 1896, by French physician Victor Despeignes [7]. At the time, cancer was believed to be a parasitic infection, and ionising radiation could kill the bacteria causing the infection. Shortly after the discovery of radium in 1898 by Marie and Pierre Curie [8] the field of radium therapy was created [9].

Different radiation effects give rise to different death patterns for a given cell, there is no singular mechanism [10]. For example, apoptosis is a form of programmed cell death that can be initiated during radiotherapy. Ionising radiation induces water radiolysis which create reactive oxygen species and free radicals. Excessive free radicals result in apoptotic signaling pathways, due to IR radiation causing DNA double-strand breaks. As a result the

cell experiences extensive genomic instability which can lead to cell death. Other methods of cell death, for example necrosis, occurs due to the destruction of intracellular components directly, such as DNA. Fundamentally, all death mechanisms are a result of ionising radiation causing changes to cell DNA, and therefore non-cancerous cells are also affected.

As many different cell death mechanisms exist due to radiation, different cancerous cells are more resistant to certain mechanisms. In addition, a cells radiosensitivity is a function of the phase of the cell cycle. A cells resistance to radiation is described by its radiosensitivity [11]. Cancers like leukemia and lymphomas can die when exposed to low radiation doses. Other cancers, such as melanomas, require very large radiation doses, and are often treated with other methods. For some cancers, the cell death mechanism and molecular target are known, whilst for others the molecular target is unclear [11].

Radiotherapy is a localised treatment, unlike treatments that are exposed to the entire body, such as Chemotherapy. Radiotherapy can be split into three branches, external, internal, and systemic radiation.

Today, the field of treatment with radiation is termed radiation therapy, with different branches, such as external beam radiotherapy. Particle therapy (or hadron therapy) is an area of external beam radiotherapy, that delivers energy using massive particles, as opposed to X-rays (conventional radiotherapy).

1.1.3 Proton Radiotherapy

Particles traversing matter deposit energy due to interactions. At different particle energies, different interactions occur more frequently, resulting in different loss mechanisms. In addition, the type of matter being traversed heavily influences the frequency of interactions that can take place. The stopping power is described as the retarding force that a particle experiences through a medium [12]. Protons, and other heavy ions, deposit energy as described by the Bethe-Bloch formula, which gives the rate at which a particle loses energy as a function of distance traversed. The formula is as follows, for a particle of velocity, $v = \beta c$, and electric charge z (in units of electron electric charge e)

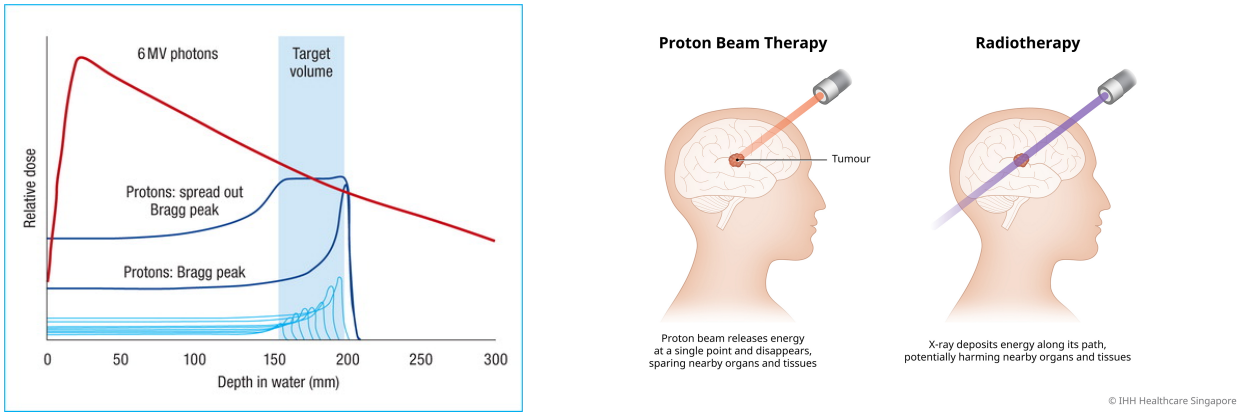
$$-\frac{dE}{dx} = \frac{4\pi}{m_e c^2}\frac{nz}{\beta^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\left[\ln\left(\frac{2m_e c^2\beta^2}{I(1-\beta^2)}\right) - \beta^2\right]. \tag{1.1}$$

Where m_e is the electron rest mass, c is the speed of light, n is the electron number density, ε_0 is the permittivity of free space, and I is the mean excitation energy (of the atoms in the medium). The energy loss mechanisms described with equation 1.1 are purely due to electronic (Coulomb) scattering. Correcting terms are often used to accurately describe the stopping power at lower energy levels, due to elastic scattering with atomic nuclei and electron capturing. The loss in particle energy as a function of depth, can be shown graphically with a Bragg curve, and describes the deposited dose as a function of depth into medium. A

common unit of dose is the Gray (Gy). One Gy is defined as 1 J deposited into 1 kg of material.

For protons at the minimum therapeutic energies ($\sim$ 70 MeV), $\beta^2 \sim 0.1$, and from Eqn. 1.1, $1/\beta^2$ is of the order 10. As the proton loses further kinetic energy, β approaches 0, creating a rapidly growing term in $1/\beta^2$, creating a Bragg peak [13]. Therefore for protons and other heavy ions, the Bragg peak occurs immediately before the particle comes to rest. At this depth, large amounts of energy is deposited into the medium. This demonstrates the advantage of PBT over conventional X-ray radiotherapy.

Figure 1.1 (a) shows the Bragg curve for different types of ionising radiation, including X-rays, single energy protons, and varied energy protons. The effect of combining protons of different energies (called range straggling [14]) produces a wider Bragg peak, called the spread out Bragg peak [15]. For a deep seated tumour, X-rays will deposit ionising radiation into healthy tissue both before and after the tumour. However, a proton beam will limit the dose to healthy tissue and maximise dose to the cancerous tumour, as shown in Fig. 1.1 (b). The position of the Bragg peak is a function of proton energy, and therefore must be tuned such that the peak occurs at the tumour site. Within a patient, metal implants and anatomical changes can lead to inaccuracies in the calculated position of the Bragg peak [16]. As the proton Bragg peak is very steep, a large dose can be deposited into potentially healthy tissue. This issue is not as damaging for conventional X-ray radiotherapy, due to the relatively flat Bragg curve at larger depths.



(a) (b)

Figure 1.1: (a) Bragg curve as a function of depth. The standard medium used is water [17]. (b) The stopping effect of PBT on a brain tumor relative to conventional radiotherapy [18].

Proton radiotherapy treatment requires protons energies to be in the range 70 - 250 MeV. This energy range is required in order to treat the entire range of potential tumours; from superifical tumours, to deep seated tumours - which are generally at depths of around 30 cm [14]. In order to produce protons of these energies, particle accelerators are used.

A brief history of particle accelerators is described below.

1.2 Short History of Particle Accelerators

In 1911 Rutherford published his theory on the structure of the atom, and in 1919 he induces a nuclear reaction with naturally occurring alpha particles [19]. The electrostatic machines that existed at the time could not produce high enough energy for Rutherford's research to continue. In 1932 Cockcroft and Walton produce a generator that produces 700 kV. At the same time, Van de Graaf invented another electrostatic generator, that could reach voltages of 1.5 MV [20]. The higher potential was reached mainly due to the suppression of sparking by placing the electrode system into a tank of dry nitrogen. Whilst the potential was higher, the Cockcroft and Walton generator could produce higher currents.

In 1924, Ising had proposed accelerating particles using time varying fields, called 'resonant acceleration', as opposed to static fields that were used at the time. Four years later, Wideröe used 'resonant acceleration' to invent the drift tube linac (DTL) [21].

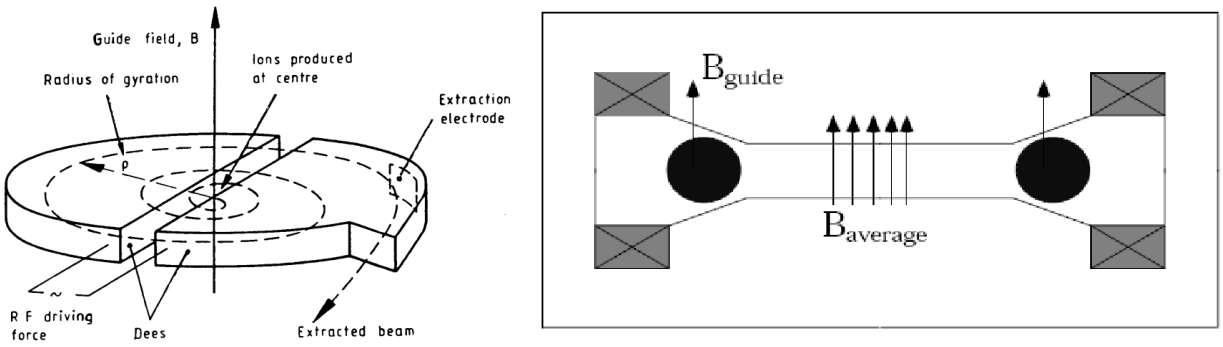
A circular accelerator is a machine that accelerates particles traveling along curved trajectories. One type of circular accelerator, called the betatron, was first built in 1940 by Kerst that accelerated electrons [22]. Similar to the cyclotron, the betatron confines electrons to a circular path, however the radius is fixed. The betatron is composed of a large electromagnet and a vacuum toroid. An increasing magnetic field will induce an electric field within the toroid, that accelerates injected electrons. The magnetic field increases at the same rate as the electron velocity, in order to keep the orbit radius constant. A schematic of the betatron accelerator is shown in Fig. 1.2 (b).

One year later, Lawrence invented the cyclotron concept, which was later proven by Livingston, who produced 80 keV protons [23]. By 1939, the University of California had produced a cyclotron capable of accelerating protons to 20 MeV. The cyclotron accelerates particles using a time varying voltage excited between electrodes. The particles travel in circular paths of increasing radius due to a constant transverse magnetic field. A schematic of the cyclotron accelerator is shown in Fig. 1.2 (a).

Whilst cyclotrons offer a compact and effective method to accelerate charged particles, the orbital frequency is assumed constant. The cyclotron frequency, f_{cyc}, in the non-relativistic limit is determined by equating the centripetal and Lorentz force,

$$f_{\text{cyc}} = \frac{eB}{2\pi m}. \tag{1.2}$$

Where B is the magnitude of the magnetic flux density perpendicular to the plane of particle motion, and m is the particle rest mass. Particles complete full rotations with a fixed frequency, provided the magnetic field is constant. When relativistic effects are incorporated,



(a) (b)

Figure 1.2: (a) Cyclotron accelerator Schematic [19]. (b) Betatron accelerator Schematic [24]. 1) guide magnetic field, 2) design orbit, 3) electromagnetic yoke and average magnetic field, 4) vacuum toroid.

Eqn. 1.2 changes to accommodate the Lorentz factor, γ_r,

$$f_{\text{cyc}} = \frac{eB}{2\pi\gamma_r m}. \tag{1.3}$$

The cyclotron frequency is thus inversely proportional to γ_r, and particles complete rotations more slowly as they increase in energy. As a result, particles are not continuously accelerated due to the constant frequency of the electric field between electrodes.

In order to overcome the Radio-frequency (rf) phase limitations of the cyclotron, the synchrocyclotron was invented, by Veksler and McMillan independently [25] [26]. The new invention provided phase stability by decreasing the frequency of the voltage source, so that the rf phase of the accelerating mechanism would remain synchronous with the beam. It can be shown that the change in frequency, for particles accelerated from $\gamma_{r0} \rightarrow \gamma_{r1}$, is as follows;

$$f_1 = f_0 + \frac{eB}{2\pi m} \ln\left(\frac{\gamma_{r0}}{\gamma_{r1}}\right). \tag{1.4}$$

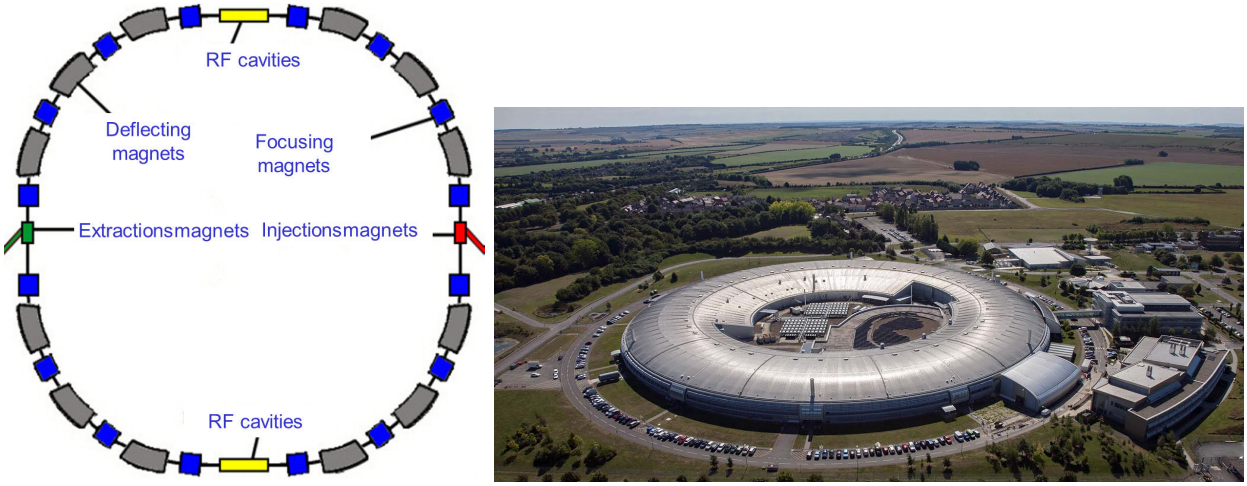
As a result, particles become accelerated as bunches. Another method that solves the phase limitation of a cyclotron is the isochronous cyclotron [27]. The isochronous cyclotron varies the magnetic field strength such that the cyclotron frequency is constant. The increase in magnetic field strength goes as

$$B_1 = B_0 \frac{\gamma_{r1}}{\gamma_{r0}}. \tag{1.5}$$

As the isochronous cyclotron is capable of accelerating higher beam currents, it is generally favoured over the synchrocyclotron.

One problem still remained, as there was no efficient way to focus the beam transversely, as only weak focusing was used. As a result, limits were placed on the size and energy

output of the machines. This problem was solved in 1952 by Courant, Livingston and Snyder, who proposed strong focusing [28]. This invention allowed synchrotron accelerators to become the dominant particle accelerator machine. Synchrotrons keep the beam confined to a circular orbit using dipole magnets, the beam is kept focused using strong focusing from quadrupole magnets. Higher order magnets, such as sextupoles are also used to correct chromatic aberrations [29]. Particles in synchrotrons lose energy via synchrotron radiation. The lost energy can be compensated with accelerating cavities, this is a feature of a storage ring. A synchrotron schematic is shown in Fig. 1.3 (a), and the UK Diamond Light Source is shown in Fig. 1.3 (b). The storage ring allowed for vast improvements in the available centre-of-mass energy of colliders [30].

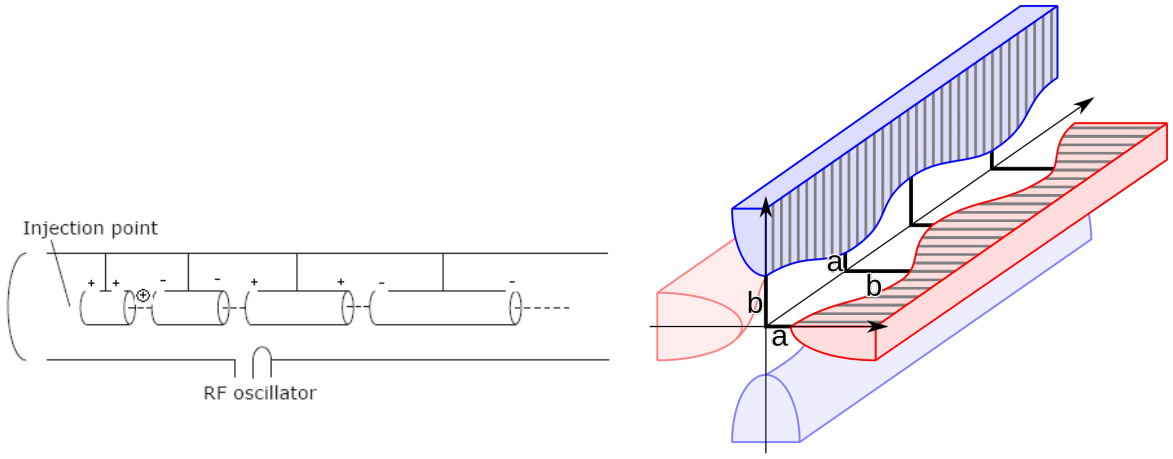


(a) (b)

Figure 1.3: Schematic of a simple synchrotron comprised of focusing magnets, bending magnets and rf cavities (a) [31]. Diamond Light Source, UK (b) [32].

Linear accelerators in the 1930s were limited by the increasing length of the drift spaces in order to keep the rf field synchronous with the beam. During world war II, research into radar provided higher frequency technology which allowed for linear machines to become more compact. The Alvarez accelerator, a DTL was built in 1946, which could produce protons at 32 MeV [33]. A schematic of the DTL is shown in Fig. 1.4 (a). In the same year, developments were made in linac technology for the acceleration of electrons [34]. The radio-frequency quadrupole (RFQ) was invented in 1970. It was increasingly used as in injector, over the commonly used Cockcroft-Walton injector. The RFQ provides both focusing and acceleration simultaneously, making it a highly desired accelerator at lower energies [35]. The accelerator has four electrodes (or 'vanes'), that when applied with a voltage, produce electric field lines that start and terminate on vanes of alternating polarity. If the shape of the vanes becomes sinusoidal with respect to the longitudinal dimension, a component

of the electric field acts along the direction of particle motion. As the polarity alternates the transverse focusing alternates between focusing/defocusing in the x/y transverse planes. The longitudinal component of the electric field will therefore also alternate between an accelerating and decelerating field. For a given rf phase and wavelength of longitudinal modulation, net acceleration can be achieved. A schematic of the RFQ electrodes is shown in Fig. 1.4 (b).



(a) (b)

Figure 1.4: Schematic of a DTL (a) [36] and the electrodes in a RFQ (b) [37].

Whilst RFQ and DTL are still used today for ion acceleration, the effective shunt impedance (rf efficiency) drops as a function of ion velocity. The modern rf linear accelerator is comprised of a set of coupled oscillators (or cells), that are excited in a particular normal mode and particular cavity mode (and are therefore called resonant structures). The cavity mode describes which field pattern is excited in a given cell. From Maxwells equations, only certain discrete electromagnetic field profiles can exist, called cavity modes. Each cavity mode can be excited in its own set of normal modes, describing the phase of cavity mode excitation relative to neighbouring cells. These accelerators are further divided into Standing Wave and Traveling Wave structures. Standing Wave accelerators are operated in the π mode (or the $\pi/2$ mode, see later) and the length of a cell is defined such that a particle observes a continuously accelerating field. In a Traveling Wave accelerator, the particle velocity is matched to the phase velocity, such that the observed field is continuously accelerating. An electron coupled cavity linac is shown in Fig. 1.5, the accelerator is a Standing Wave cavity operated in the $\pi/2$ mode. As an electron beam passes through the structure, it will absorb the electromagnetic energy that is stored in the individual cells. Normal conducting cavities are often machined out of a copper alloy, due to the high conductivity. The driving frequency of these resonant structures is an important design choice that effects nearly all aspects of the accelerator. For medical linacs, the most common frequency choices are around 3 and

5.7 GHz, which are known as S- and C-band structures, respectively. Radio-frequency cavities can also be made from niobium alloys (or more recently, niobium can be used as a thin-film coating), and at superconducting temperatures, these cavities have very low losses and thus high rf efficiency [38]. Superconductivity was introduced into the accelerating physics field in the 1960s.

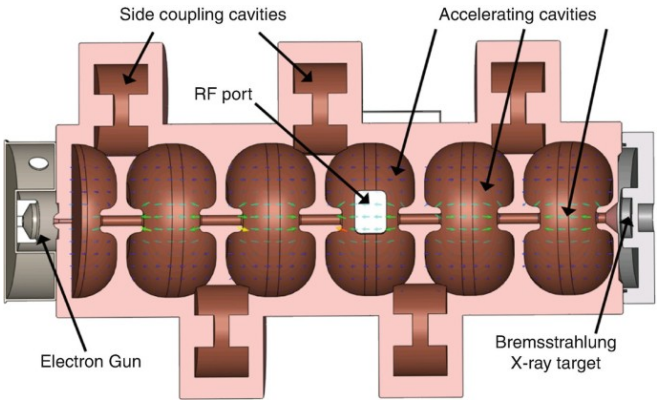


Figure 1.5: A coupled cavity electron linac excited in the $\pi/2$ standing wave mode. Electromagnetic energy is imported via the rf coupler [39].

1.3 Recent developments of Proton Medical Machines

In recent years, many proton accelerates have been designed for proton radiotherapy. The main limiting factor facing normal conducting cavities is the gradient achievable. As the accelerating gradient increases, the probability of rf breakdown increases (see Chapter 3). During breakdown, the cavity surface can be irreversibly damaged. After many events, an rf cavity may no longer be able to sustain electromagnetic field due to changes in the resonant frequency.

In 1991 a linac proposal for proton radiotherapy was developed [40]. The TERA foundation, together with CERN and INFN developed the first prototype of a 3 GHz proton linac in 1999 [41], called LIBO (LInac BOoster). In 2001, further tests were completed on the design [42]. LIBO was the linear accelerator for a cyclinac design (a cyclinac uses both cyclotron and linac accelerators in the same system), which used a cyclotron to accelerate protons to 62 MeV, before LIBO would further accelerate the protons to 74 MeV. TERA later began research into a single room facility called TULIP (TUrning LInac for Proton therapy) [43], another cyclinac design. In this machine, protons would be accelerated to 24 MeV with a cyclotron, with the remaining energy (24 - 230 MeV) coming from a 3 GHz side-coupled linac [44].

The first all-linac solution for a proton machine began with the ENEA group, in Italy. The lower energy levels would be accelerated using an RFQ into a DTL [45]. In 2014 CERN produced an efficient RFQ designed at accelerating protons to 5 MeV in 2 m [46]. This designed removed the requirement of cyclotron acceleration at the lower energy levels. The CERN spin-off company, A.D.A.M (Application of Detector and Accelerators to Medicine), began the design of a commercial all-linac solution, producing protons up to 250 MeV. In addition to protons machines, developments have been made towards a carbon ion linear accelerator [47, 44].

In the UK, there are currently two centres offering higher energy PBT. One centre is at the University College London Hospital (UCLH) in London, and the second is at the Christie NHS Foundation Trust, in Manchester. These centres were produced due to a government imitative, costing £250 M. Each centre has three treatment rooms, thus each treatment room costs approximately £40 M. Proton acceleration is achieved with the Varian ProBeam accelerator, an isochronous superconducting cyclotron [48]. Figure 1.6 shows the superconducting isochronous cyclotron capable of producing protons at 250 MeV (a) and a UCLH treatment room (b). There are lower energy PBT services available, such as the low energy proton machine (used to treat some eye cancers) in Clatterbridge, Liverpool.

Recently, Lancaster University, Manchester University, the Cockcroft Institute and the Christie hospital collaborated on a project that explored a linear booster concept for accelerating protons to 350 MeV. This project, called ProBE [51], developed a booster capable of accelerating protons from ~ 250 MeV (from the cyclotron) to 350 MeV, for medical imaging. The linac design was a high gradient S-band side-coupled SW cavity. A prototype has since been developed and tested at CERN.

Phase two of the ProBE project saw a collaboration with Advanced Oncotherapy (AVO), a specialist company that designed the LIGHT proton therapy system [52]. The LIGHT system is an all-linac solution, which has benefits over current systems due to the lower construction costs of the bunker system. As a result, the aim is to provide better accessibility to PBT.

This project specifically looks to design a high energy proton booster for the 150 - 250 MeV energy range. The lower energy levels (up to 5 MeV) will be completed by the 750 MHz RFQ designed and built at CERN [46]. The next acceleration stage will be completed with a side-coupled DTL, producing protons at 37.5 MeV. From 37.5 MeV to 250 MeV, the acceleration will be completed using a coupled cavity linac (CCL). For the CCL designed in this project, the minimum aperture limit was set to 2.5 mm. As the system is an all-linac solution, the beam emittance is kept small, relative to cyc-linac systems, and thus provides the ability to realise smaller apertures. The cavity is to be operated in S-band, at 2.9985 GHz. Initially, the cavity is expected to be a SW $\pi/2$ structure, however TW structures are also explored, due to developments in the work.

(a)

(b)

Figure 1.6: (a) Varian ProBeam superconducting cyclotron [49]. (b) Proton beam therapy treatment room at UCLH [50]

1.4 Thesis outline

Chapter 2 describes the theory of rf acceleration, including transverse and longitudinal beam dynamics. The aim of the chapter is to provide a basic understanding of the core concepts, such that a given reader can understand latter sections of the document.

Chapter 3 describes the design procedure of the single cell. This chapter discusses the optimisation methods implemented, such as Multi-Objective Genetic Algorithms, for solving Multi-Objective problems. A novel method was used to select a single cell geometry design, that utilised different visualisation techniques to best portray the design space of the single cell. Once a single cell was selected, the Standing Wave side-coupled cells and Traveling Wave coupling slots were designed, with respect to various objectives.

Chapter 4 introduces a novel self-consistent framework that incorporates longitudinal acceleration into transverse beam dynamics. Two new focusing schemes are analytically explored, the FODO-like focusing scheme, and the Minimum Aperture Scheme (MAS). The FODO-like scheme is solved analytically, such that the maximum beam size is minimised at the cavity entrance and exit. For a given beam emittance, the maximum cavity length is thus determinable. The MAS considers the focusing scheme by which the acceptance ellipse of the cavity is matched to the beam ellipse, providing the maximum cavity length feasible for a given beam emittance.

Chapter 5 describes a novel Fast Cell to Cell Tracking (FC2CT) algorithm. The developed algorithm tracks the 6D phase space of a particle beam by analytically solving for the momentum change once per rf cell, in order to reduce computation time. The method approximates constant particle velocity over the integration region and requires the cell longitudinal electric field component, E_z, determined by electromagnetic solver codes. FC2CT is developed for tracking through both Standing and Traveling Wave structures.

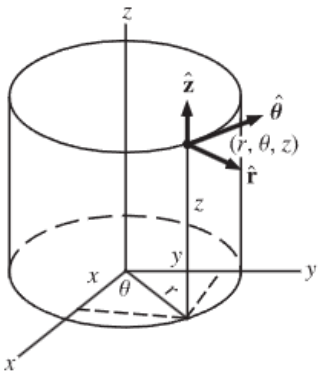
Chapter 6 uses FC2CT, in addition to results from Chapter 3, to approximate the energy gain for a given cavity type, with both the input power and cavity length being variable. This allows the exploration and comparison of potential linac solutions for the 150 - 250 MeV proton booster.

Lastly, Chapter 7 discusses the final electromagnetic design of the chosen cavity, including the matching cell coupler design. The chapter also displays results obtained from tracking (with FC2CT) a realistic beam through the chosen linac design. The initial beam ellipse is determined analytically, from results obtained in Chapters 4 and 6.

Chapter 2

Theory of Particle Acceleration

2.1 Solving Maxwell's Equations in a Cylindrical Pillbox



(a)

Figure 2.1: Pillbox cavity and relationship between cylindrical and cartesian coordinates [53].

Cylindrical pillbox are used as vessels for electromagnetic energy that produce longitudinal acceleration of a charged particle, due to the properties of the first principal mode, TM_{010}. The wave equation [54] in cylindrical coordinates (see Fig. 2.1) is as follows,

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}. \tag{2.1}$$

Where u is a scalar wave function. The first step to solving Eqn. 2.1 is to separate u into multiple functions of one variable,

$$u(r, \theta, z, t) = R(r)\Theta(\theta)Z(z)T(t), \tag{2.2}$$

substituting Eqn. 2.2 into Eqn. 2.1 and dividing by u before simplifying produces the following;

$$\frac{1}{c^2}\frac{T''}{T} = \left(R'' + \frac{R'}{r}\right)\frac{1}{R} + \frac{1}{r^2}\frac{\Theta''}{\Theta} + \frac{Z''}{Z}, \tag{2.3}$$

where $\frac{\partial^2 X}{\partial x^2} = X''$. Firstly, it can be noted that the terms $\frac{1}{c^2}\frac{T''}{T}$ and $\frac{Z''}{Z}$ are purely functions of one variable, and therefore are constants with respect to changes in other variables. Defining

$$\frac{1}{c^2}\frac{T''}{T} = -\frac{\omega^2}{c^2} = -k^2, \tag{2.4}$$

and

$$\frac{Z''}{Z} = -k_z^2. \tag{2.5}$$

Multiplying Eqn. 2.3 by r^2, the term $\frac{1}{r^2}\frac{\Theta''}{\Theta}$ becomes independant of other variables and is also a constant,

$$\frac{\Theta''}{\Theta} = -\alpha^2. \tag{2.6}$$

Substituting the constant terms into Eqn. 2.3 and simplifying,

$$0 = \left(R'' + \frac{R'}{r}\right)\frac{r^2}{R} - \alpha^2 - k_z^2 r^2 = -k^2 r^2. \tag{2.7}$$

Defining $k^2 - k_z^2 = k_r^2$ and multiplying by R produces the final form of the differential equation

$$r^2 R'' + rR' + (k_r^2 - \alpha^2)R = 0. \tag{2.8}$$

Equation 2.8 is called the Bessel's differential equation and the solutions are the Bessel functions of the first kind [55]. Bessel functions of the second kind Y_α, are also solutions to Bessel's equation, however are valid for singularities at $r = 0$, which is not the case here. Bessel functions of the first kind can be described as follows

$$J_\alpha(r) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(m+\alpha+1)}\left(\frac{r}{2}\right)^{2m+\alpha}. \tag{2.9}$$

Where Γ is the Gamma function. The solution in Eqn. 2.9 for $\alpha = 0$ describes the longitudinal electric field, E_z, as a function of radial displacement, r. This mode is called a Transverse-Magnetic (TM) mode, as there is no longitudinal magnetic field component. The fundamental mode has azimuthal symmetry and is called the TM_{010} mode. The general result for TM modes are described with TM_{mnp} (and similarly transverse electric modes are TE_{mnp}). The subscript m describes the number of full wave variations around theta, n describes the number of half wave variations along the diameter and p describes the number of half wave variations along the longitudinal direction. Only a discrete set of solutions can be produced within a

cylindrical pillbox, each described with different values of m, n and p. Each mode is a different cavity mode. For the TM_{010} mode, the only other non-zero field component is B_θ, describing the component of the magnetic field azimuthally. The scalar field can be determined from E_z using Ampere's law,

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}, \tag{2.10}$$

where μ_0 is the permeability of free space and J is the current density (not to be confused with the Bessel Function, J_α). The fields are as follows;

$$E_z = E_0 J_0(k_r r) \cos(\omega t), \tag{2.11}$$

$$B_\theta = -\frac{E_0}{c} J_1(k_r r) \sin(\omega t). \tag{2.12}$$

Where E_0 is the maximum longitudinal electrical field. The field profiles for the TM_{010} mode are shown in Fig. 2.2.

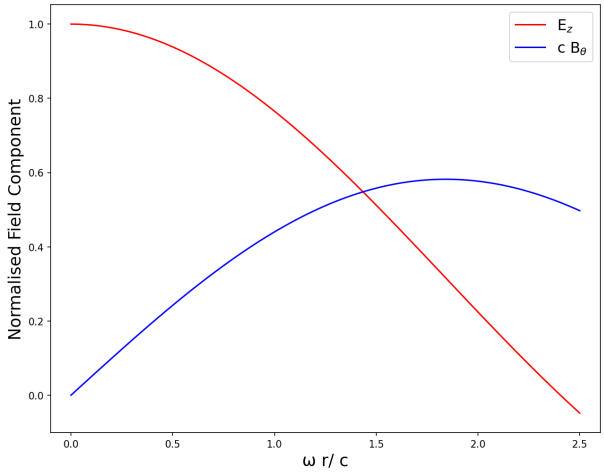


Figure 2.2: First order Bessel functions describing field components E_z (J_0) and cB_θ (J_1).

The boundary condition states a perfect conductor can not have a tangential electric field component or a normal magnetic field component at the surface. Therefore, E_z must terminate on the circumference of the cylindrical pillbox, occurring at $r \approx \frac{2.405\omega}{c}$. Therefore, for a cylindrical pillbox of radius R_{cyl}, the frequency of the mode, ω_{010}, is $\frac{2.405c}{R_{cyl}}$. The stored energy in the cavity mode is

$$U = \frac{\pi \varepsilon_0 l_{cyl} R_{cyl}}{2} E_0^2 J_1^2(2.405), \tag{2.13}$$

where ε_0 is the vacuum permeability and l_{cyl} is the length of the cylinder.

Note the above discussion has assumed a perfectly cylindrical pillbox, in which E_z is constant with longitudinal variable, z. In reality, an aperture is required, on-axis, in order to excite the cavity mode in addition to providing a path for particle to traverse. With the addition of the aperture, E_z is no longer a constant with z. This effect also produces a non-zero E_r component for the TM_{010} mode, due to Gauss' Law,

$$\frac{1}{r}\frac{\partial(rE_r)}{\partial r} + \frac{1}{r}\frac{\partial E_\theta}{\partial \theta} + \frac{\partial E_z}{\partial z} = \frac{\rho}{\varepsilon_0}. \tag{2.14}$$

Where ρ is the charge density and is zero in the absence of free charges. This effect is discussed more rigorously in Section 5.2 when deriving the field components of a TM_{010} mode given the $E_z(r, z)$ field map.

In general, solving Maxwell's equations analytically for single cell geometries is not possible. As a result, computer codes (such as CST [56]) are used to solve Maxwells equations using the Maxwell's Grid Equations [57] using finite difference techniques.

Slater's Perturbation Theory

A cavity mode operating at resonant frequency will observe the stored energy oscillating between the electric and magnetic field evenly, $U = U_{E,\mathrm{max}} = U_{H,\mathrm{max}}$. For a small change in the geometry of the cylindrical waveguide, the electric and magnetic field lines will shift until the maximum stored energy in both fields is equal, changing in the resonant frequency. The change in the frequency can be approximated using Slater's Perturbation theory, which states for a volume change ΔV,

$$\frac{\Delta\omega_0}{\omega_0} = \frac{\Delta U_H - \Delta U_E}{U} = \frac{1}{4U}\int_{\Delta V}(\mu_0 H^2 - \varepsilon_0 E^2)dV. \tag{2.15}$$

U_H and U_E describe the (time averaged) energy removed due to the change ΔV. This can be visualised as follows. Suppose removing a concentric shell from a cylindrical pillbox excited in the TM_{010} mode, so that the resulting pillbox has a smaller radius. As E_z is close to zero in the removed region (and $E_r = E_\theta = 0$), $U_E \approx 0$. As B_θ is non-zero in this region, a portion of magnetic field energy has been removed ($U_H > 0$), thus from Eqn. 2.15, $\Delta\omega_0$ is positive and the resonant frequency increases, as expected.

2.2 RF Cell Figures of Merit

When designing rf accelerator structures a single cell is often designed before multiple cells are coupled together to produce coupled cavity structures. The are multiple figure of merits for a single cell and will be defined now. The first figure of merit is the quality factor, Q. Defined as

$$Q = \frac{\omega U}{P}, \tag{2.16}$$

Table 2.1: Frequency scaling of multiple cavity parameters.

Cavity Parameter	Normal Conducting	Super Conducting
Surface Resistance, R_s	$f^{1/2}$	f^2
Power Loss, P	$f^{-1/2}$	f
Shunt Impedance, Z	$f^{1/2}$	f^{-1}
Quality Factor, Q	$f^{-1/2}$	f^{-2}
Z/Q	f	f

where ω is $2\pi f$, U is the stored energy and P is the power loss. The Q factor describes the ratio of the stored energy to the energy dissipated in one oscillation for the described system.

The efficiency of particle acceleration over a single cell can be described with the shunt impedance, R,

$$R = \frac{V_0^2}{P}. \tag{2.17}$$

Where V_0 is the voltage gain over the cell, and is given

$$V_0 = \int_0^L E_z(z)dz. \tag{2.18}$$

It is convenient to normalise the shunt impedance to one cell length to give the shunt impedance per unit length, R',

$$R' = \frac{V_0^2}{L_{cell}P}. \tag{2.19}$$

In practice, the accelerating axial field varies with time as the particle traverses the cell ($E_z(z) \rightarrow E_z(z,t)$) and the voltage gain is modulated by the transit time factor, T, to account for the varying field. The effective shunt impedance per unit length is given;

$$Z = \frac{(V_0T)^2}{L_{cell}P}. \tag{2.20}$$

Henceforth, any reference to the shunt impedance is a reference to the effective shunt impedance per unit length. Dividing the shunt impedance by the quality factor cancels the power loss and is independent of material properties

$$\frac{Z}{Q} = \frac{(V_0T)^2P}{L_{cell}P\omega U} = \frac{(V_0T)^2}{\omega U L_{cell}}. \tag{2.21}$$

The peak field values are also vital figures of merit for single cell design.

Table 2.1 describes the frequency dependence of various cavity parameters. The shunt impedance increases as $f^{1/2}$ for normal conducting cavities. One may think that the higher

the design frequency the better, as the rf efficiency improves. However the Q-factor decreases with the same dependence. In the next section it is shown that higher Q suppresses the power flow droop and power flow phase shift.

2.3 Dispersion Relation and Coupled Oscillators

A smooth uniform waveguide excited in the TM_{010} mode will have a dispersion relation, that relates the frequency of the normal mode oscillations to the wave-vector, k, [35]

$$\omega^2 = \omega_c^2 + (kc)^2. \tag{2.22}$$

Where ω_c is the cutoff frequency, and is a function of the waveguide geometry. $k = \frac{2\pi}{\lambda_G}$ for guide wavelength λ_G. The phase velocity is

$$v_{phase} = \frac{\omega}{k} = \frac{c}{\sqrt{1 - \omega_c^2/\omega^2}} > c, \tag{2.23}$$

and greater than c. A particle can not be efficiently accelerated as the synchronicity condition is not met. Inserting conducting walls at periodic positions (with period p) along the waveguide (often called 'disk loaded') is a method used to slow down the wave phase velocity. The act of inserting periodic conducting walls introduces modulation to the longitudinal electric field magnitude with the same period as the walls. The axial electric field now takes the form

$$E_z(r, z, t) = E_p(r, z)e^{i(\omega t \pm kz)}. \tag{2.24}$$

Where $E_p(r, z)$ is periodic over the distance p. $E_z(r, z, t)$ obeys Floquet theorem, which requires the field value at any two points separated by an integer number of periods must differ by a complex number. This is the same as;

$$E_p(r, z + np) = E_p(r, z)e^{\pm inkp}. \tag{2.25}$$

as $E_p(r, z)$ is periodic over some distance p it can be represented by a Fourier series;

$$E_p(r, z) = \sum_{n=-\infty}^{\infty} c_n(r)e^{-i2n\pi\frac{z}{p}}. \tag{2.26}$$

Inserting Eqns. 2.24 and 2.26 into the wave equation in cylindrical coordinates;

$$\frac{\partial^2 E_z(r, z, t)}{\partial r^2} + \frac{1}{r}\frac{\partial E_z(r, z, t)}{\partial r} + \frac{\partial^2 E_z(r, z, t)}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E_z(r, z, t)}{\partial t^2} = 0. \tag{2.27}$$

Thus;

$$c_n''(r) + \frac{1}{r}c_n'(r) + K_n^2 c_n(r) = 0, \tag{2.28}$$

where a prime refers to a partial derivative with respect to r, and

$$K_n^2 = \frac{\omega^2}{c^2} - k_n^2, \tag{2.29}$$

where $k_n = k + \frac{2n\pi}{p}$. For real values of K_n the solution for the propagating axial field is

$$E_z(r, z, t) = \sum_{n=-\infty}^{\infty} E_n J_0(K_n r) e^{i(\omega t - k_n z)}. \tag{2.30}$$

Where $J_0(k_n r)$ is the zeroth order Bessel function of the first kind, and describes the radial effect on $E_z(r)$.

Equation 2.30 demonstrates the axial electric field can be described as an infinite sum of traveling waves called spatial harmonics. Each spatial harmonic travels at its own phase velocity

$$v_{ph,n} = \frac{\omega}{k_n c} = \frac{v_{ph,0}}{1 + (n v_{ph,0} \lambda / p)}, \tag{2.31}$$

where $v_{ph,0}$ is the phase velocity of the principal spatial harmonic, with $n = 0$. Selecting a spatial harmonic of high enough index will produce a phase velocity that is suitable for particle acceleration. It can be shown the effect of all other spatial harmonics on a particle energy cancel, leaving only the contribution from the chosen harmonic. An infinite disk/iris loaded waveguide can also be described as a set of identical coupled oscillators. For a given cavity mode, like the TM_{010} mode, there exists an infinite number of normal modes of oscillations.

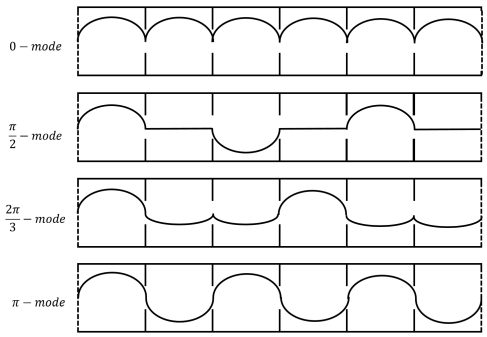


Figure 2.3: Normal modes excited in a periodic chain of oscillators.

For a given normal mode, each cell oscillates at the same frequency, given by the dispersion relation. From one cell to the next there exists a phase advance, from Floquet's theorem. The phase advance per cell defines which normal mode is being excited, as shown in Fig. 2.3. For a finite set of coupled oscillators, the number of normal modes is equal to the number of oscillators. As a result, the dispersion relation goes from a continuous to a discrete curve, with

each point representing a normal mode of oscillation. The π mode describes the normal mode by which each cell oscillates π radians out of phase. One can visualise a particle traveling at such a velocity (v) that it only observes a positive E_z field as it traverses multiple cells. This would require the particle crosses one cell in the time it takes the field to oscillate by half a period, $T/2$,

$$\frac{L_{cell}}{v} = \frac{T}{2} = \frac{\pi}{\omega}. \tag{2.32}$$

An arbitrary dispersion curve is shown in Fig. 2.4 for an infinite set of oscillators (so that the curve is continuous). As the field is comprised of an infinite number of spatial harmonics of increasing k_n, the curve is indefinitely periodic. The π mode is displayed, showing phase velocity of each spatial harmonic. Due to the periodicity, a dispersion curve is shown with only one branch. Thus, a given dispersion curve is actually demonstrating whichever spatial harmonic provides the correct phase velocity for the application. Due to the periodicity, the group velocity - calculated as the gradient of the tangent to a given normal mode - is constant for all spatial harmonics of the same normal mode.

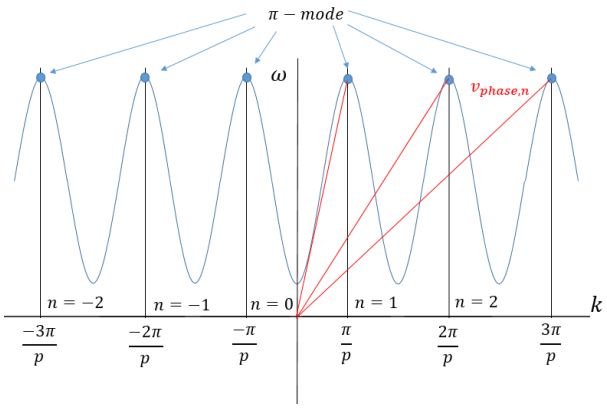


Figure 2.4: Dispersion Relation for an arbitrary cavity mode. n describes the spatial harmonic index. π-mode is shown. The phase velocity of each spatial harmonic is the gradient of the red line.

2.3.1 Equivalent Circuits

Equivalent circuits are often used to gain an understanding of the characteristics of periodic structures [35]. Considering an infinite number of coupled cells as shown in Fig. 2.5. The series impedance and shunt admittance are denoted Z and Y, respectively.

In order to calculate a useful result when considering equivalent circuits, a general circuit of n identical cells is considered. The n^{th} cell has voltage V_n and current I_n. From Kirchhoff's law,

$$I_{n-1} = I_n + YV_n, \tag{2.33}$$

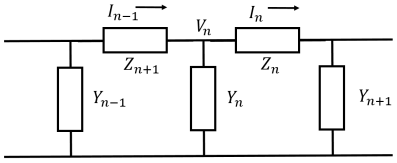


Figure 2.5: Schematic of periodic chain of circuit oscillators.

$$I_n = I_{n+1} + YV_{n+1}. \tag{2.34}$$

The voltage drop from one cell to the next is the product of the cell current and shunt impedance, from Ohm's law,

$$V_n - V_{n+1} = ZI_n. \tag{2.35}$$

Eliminating the voltage and using Floquet theorem to relate the currents in adjacent cells produces the following result [35],

$$\cos(\phi) = 1 + \frac{YZ}{2}. \tag{2.36}$$

Equation 2.36 describes the general result for finding the phase advance per cell for a given shunt admittance and series impedance. Two examples of equivalent circuits that describe coupled cavity accelerators are the electrically and magnetically coupled circuits.

Electrically Coupled Circuit

The electrically coupled circuit has a series impedance, L_s and series capacitance, C_s. As the circuit is electrically coupled the shunt mechanism (parallel) is a capacitor of capacitance C_p. The series impedance and shunt admittance of a cell are given;

$$Z = i\omega L_s + \frac{1}{i\omega C_s} = i\omega L_s \left(1 - \frac{\omega_0^2}{\omega^2}\right), \tag{2.37}$$

$$Y = i\omega C_p. \tag{2.38}$$

Where $\omega_0^2 = \frac{1}{L_s C_s}$ is a constant. Using the values of Z and Y with 2.36 and re-arranging for the resonant frequency;

$$\omega = \omega_0 \sqrt{\frac{2C_s}{C_p}(1 - \cos(\phi)) + 1}. \tag{2.39}$$

The above equations provides the dispersion relation for the electrically coupled circuit. The resonant frequency is clearly a function of phase advance per cell. The $0-$mode is the

normal mode when $\phi = 0$, and thus $\omega = \omega_0$. The $\pi-$mode frequency can be shown to be approximately $\omega = \omega_0(1 + \frac{2C_s}{C_p})$ for $C_s \ll C_p$. In this regime the frequency of the $\pi-$mode is higher in frequency than the $0-$mode. This circuit is equivalent to an iris loaded waveguide that couples electromagnetic energy from one cell to the next through the iris, or aperture. For normal modes that may be used for acceleration, such as the $\frac{2\pi}{3}$ mode, the dispersion curve has a positive gradient. Therefore $\frac{\partial\omega}{\partial\phi} > 0$ and the group velocity is positive.

Magnetically Coupled Circuit

The magnetically coupled circuit likewise has a series impedance and capacitance, however couples cells with an inductor of impedance L_p. Using 2.36 and solving for ω as in the previous case;

$$\omega = \frac{\omega_0}{\sqrt{\frac{2L_p}{L_s}(1 - \cos(\phi)) + 1}}. \tag{2.40}$$

Again, the 0-mode has frequency $\omega = \omega_0$. However, for $L_p \ll L_s$, the $\pi-$mode frequency goes as $\omega_0(1 - \frac{2L_p}{L_s})$, and has a frequency lower than the $0-$mode. Thus, a $\frac{2\pi}{3}$ normal mode has a negative group velocity and hence is called a backward-wave mode. Figure 2.6 displays the dispersion curve for both an electrically and magnetically coupled array of periodic oscillators. The group velocity is the gradient of the curve at a given phase advance.

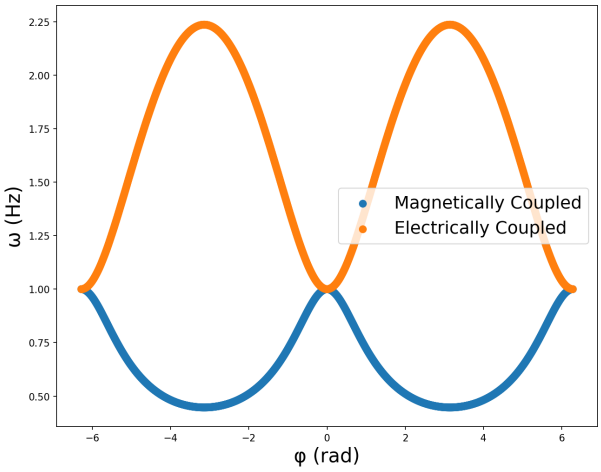


Figure 2.6: Dispersion relation of both electrically and magnetically coupled periodic circuit.

For both circuit examples, the idea of an inter-cell coupling value can be initiated. Calculating the frequency of the $\frac{\pi}{2}-$mode is often used to normalise the resonant frequency. The normalised frequency for each circuit are given;

$$\omega_e^2 = \omega_{\pi/2,e}^2(1 - k_e \cos(\phi)), \tag{2.41}$$

$$\omega_m^2 = \frac{\omega_{\pi/2,m}^2}{(1 - k_m \cos(\phi))}. \tag{2.42}$$

Where the inter-cell coupling constants are given;

$$k_e = \frac{2C_s}{C_p + 2C_s}, \tag{2.43}$$

$$k_m = \frac{2L_p}{L_s + 2L_p}. \tag{2.44}$$

The width of the pass band is an important characteristic of the dispersion relation. The width is given as the difference in frequency between the 0 and $\pi-$mode. For the electrically coupled circuit,

$$\omega_\pi - \omega_0 = \omega_0(1 + \frac{2C_s}{C_p} - 1) = \omega_0 \frac{2C_s}{C_p}. \tag{2.45}$$

Dividing the width with the frequency of the $\pi/2-$mode produces the fractional bandwidth

$$\frac{\omega_\pi - \omega_0}{\omega_{\pi/2}} = \frac{2C_s\omega_0}{C_p\omega_0\sqrt{1 + \frac{2C_s}{C_p}}} \approx \frac{2C_s\omega_0}{C_p\omega_0(1 + \frac{C_s}{C_p})} = \frac{2C_s}{C_p + C_s}. \tag{2.46}$$

Where it was used $C_s \ll C_p$. In this limit,

$$k_e \approx \frac{\omega_\pi - \omega_0}{\omega_{\pi/2}}, \tag{2.47}$$

and so the fractional bandwidth is the same as the coupling constant. A large coupling constant produces a larger fractional bandwidth and the normal modes on the dispersion curve span a large range of resonant frequencies. Small coupling constants suggest the coupling from one cell to the next is smaller and the allowed values of resonant frequencies spans a much smaller range. For the magnetically coupled circuit, the fractional bandwidth is

$$\frac{\omega_\pi - \omega_0}{\omega_{\pi/2}} = \frac{\omega_0(-\frac{2L_p}{L_s})}{\omega_0}\sqrt{1 + \frac{2L_p}{L_s}} \approx \frac{-2L_p}{L_s}\left(1 + \frac{L_p}{L_s}\right) = \frac{-2L_p}{L_s\left(1 - \frac{L_p}{L_s}\right)} = \frac{-2L_p}{L_s - L_p}, \tag{2.48}$$

and

$$k_m \approx \frac{\omega_\pi - \omega_0}{\omega_{\pi/2}}. \tag{2.49}$$

Thus, the inter-cell coupling constant can be described purely by the frequencies of the $0, \pi/2$ and π normal modes.

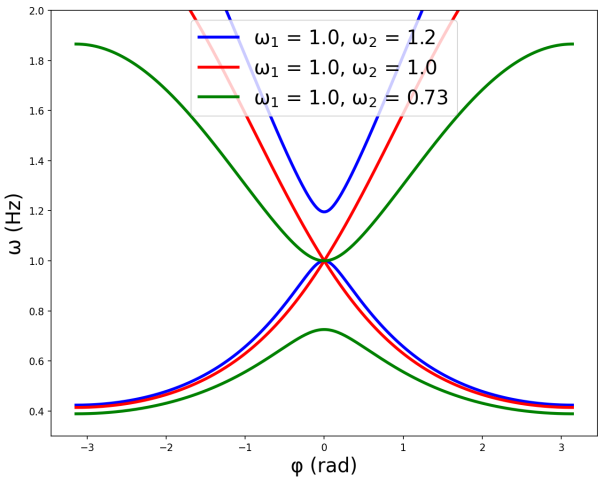


Figure 2.7: Dispersion curve of periodic array of oscillators with both inductive and capacitive admittance. Multiple solutions are shown for different values of ω_2.

Electrically and Magnetically Coupled Circuit

As a final example, the shunt mechanism is treated as both capacitive and inductive. There are both series components, L_s and C_s, and shunt components, L_p and C_p. Let $\omega_1^2 = 1/(L_s C_s)$ and $\omega_2^2 = 1/(L_p C_p)$. Using Eqn. 2.36, the resonant frequency is quadratic in nature and has two solutions, describing the upper and lower pass bands, as there exists two methods for coupling to take place between cells. At some special case (a certain set of values for C_p, L_p, C_s, L_s and $\omega_1 = \omega_2$), the two pass bands meet, removing the stop band - this is called confluence. At confluence the slope of one (previously separate) pass band continues into the other pass band. This means the group velocity slope does not tend to 0 as the turning point becomes a point of maximum gradient, leading to improved rf stability. Figure 2.7 displays the dispersion curve of a periodic array of oscillators with both inductive and capacitive shunt mechanisms. By changing the value of L_p or C_p, ω_2 can be tuned equal to ω_1, achieving confluence. For $\omega_2 > \omega_1$, the upper pass band moves too high in frequency for confluence to occur. As ω_2 is too large, $L_p C_p$ is too small. For $\omega_2 < \omega_1$, $L_p C_p$ is too large.

2.3.2 Coupled Oscillators

In order to gain a deeper understanding of the physics of coupled oscillators the case of three coupled oscillators is outlined [35]. The system is comprised of two end cell oscillators, with capacitance $2C_0$ and coupled to the central cell with inductance L_0, as shown in Fig. 2.8. The resonant frequencies of the end cells is therefore $\omega_0 = \frac{1}{\sqrt{2C_0 L_0}}$, this is the same resonant frequency as the central cell, which has capacitance C_0 however double the total inductance. Kirchhoff's Voltage Law says the sum of voltage in a closed loop is zero. Applying Kirchhoff's

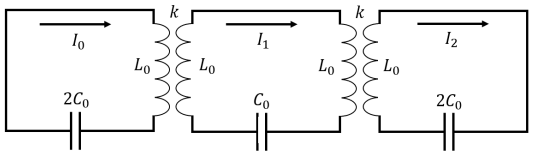


Figure 2.8: Schematic showing three coupled oscillators.

Law to each closed loop produces a set of three equations that can be solved as an eigenvalue problem (see [35]). Each normal mode is described by an eigenvector, I, representing the current in each cell. The eigenvalue represents the frequency of each normal mode.

$$\omega_{0\text{-mode}} = \frac{\omega_0}{\sqrt{1+k}},\ \vec{I} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{2.50}$$

$$\omega_{\pi/2\text{-mode}} = \omega_0,\ \vec{I} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{2.51}$$

$$\omega_{\pi\text{-mode}} = \frac{\omega_0}{\sqrt{1-k}},\ \vec{I} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \tag{2.52}$$

k represents the coupling constant, and is given as the ratio of the mutual inductance to the central inductance, L_0.

The general result for $N+1$ coupled oscillators have the following eigenvector elements [35];

$$v_{q,n} = \cos\left(\frac{\pi q n}{N}\right) e^{i\omega_q t}, \tag{2.53}$$

the indices q and n refer to the mode and cell respectively. The eigen-frequencies of mode q are

$$\omega_q = \frac{\omega_0}{\sqrt{1 + k\cos(\frac{\pi q}{N})}}. \tag{2.54}$$

Checking the above equations with $N = 2$ produces the correct current eigenvector and eigenvalues for a 3-cell system. The bandwidth of the dispersion relation for a general $N+1$ coupled oscillators is (for small k);

$$\omega_{\pi-\text{mode}} - \omega_{0-\text{mode}} = \omega_0\left(\frac{1}{\sqrt{1-k}} - \frac{1}{\sqrt{1+k}}\right) \approx \omega_0(1 + k/2 - 1 + k/2) = \omega_0 k, \tag{2.55}$$

and independant of N. This result is important; the number of normal modes increases linearly with N, however the bandwidth is fixed. Thus, coupled oscillator systems with more cells have smaller mode spacing. From Eqn. 2.54, the value of k can be calculated as the following;

$$k = \frac{\omega_\pi - \omega_0}{\omega_{\pi/2}}. \tag{2.56}$$

Effects of individual frequency errors

The above description provides the perfect case of $N+1$ oscillators. In reality, cells can not be manufactured ideally and cells will have frequency errors. The effect of frequency errors causes perturbations to the normal mode eigenvectors and eigenvalues. The analytical result is obtained using Perturbation Theory [35]. The first order approximation for the current eigenvector for mode q is given,

$$\vec{v}_{pert,q} = \vec{v}_q + \sum_{r \neq q} a_{qr} \vec{v}_r. \tag{2.57}$$

Where $\vec{v}_r$ is the unperturbed eigenvector for mode r and a_{qr} is some assumed small parameter given via the following

$$a_{qr} = \frac{[\vec{v}_q \mathbf{P} \vec{v}_r]}{\frac{1}{\omega_r^2} - \frac{1}{\omega_q^2}}, \tag{2.58}$$

and $\mathbf{P}$ is the Perturbation matrix.

Carrying out the above calculation produces the following result for the perturbed eigenfrequencies and eigenvectors. Denoting the unperturbed frequency as ω_0.

$$\omega_{q=0} = \frac{\omega_0}{\sqrt{1+k}\sqrt{1+\frac{\delta\omega_c}{\omega_0}}} \tag{2.59}$$

$$\vec{v}_{q=0} = \begin{pmatrix} 1 + \frac{1+k}{2k}\left(\frac{\delta\omega_c}{\omega_0} - 4\frac{\delta\omega_e}{\omega_0}\right) \\ 1 - \frac{1+k}{2k}\frac{\delta\omega_c}{\omega_0} \\ 1 + \frac{1+k}{2k}\left(\frac{\delta\omega_c}{\omega_0} + 4\frac{\delta\omega_e}{\omega_0}\right) \end{pmatrix} \tag{2.60}$$

$$\omega_{q=1} = \frac{\omega_0}{\sqrt{1 - 4(\delta\omega_e/\omega_0)^2}} \tag{2.61}$$

$$\vec{v}_{q=1} = \begin{pmatrix} 1 + 4/k^2 \frac{\delta\omega_c\delta\omega_e}{\omega_0^2} - 2/k^2\left(\frac{\delta\omega_e}{\omega_0}\right)^2 \\ \frac{-2}{k}\frac{\delta\omega_e}{\omega_0} \\ -1 - 4/k^2 \frac{\delta\omega_c\delta\omega_e}{\omega_0^2} + 2/k^2\left(\frac{\delta\omega_e}{\omega_0}\right)^2 \end{pmatrix} \tag{2.62}$$

$$\omega_{q=2} = \frac{\omega_0}{\sqrt{1-k}\sqrt{1+\frac{\delta\omega_c}{\omega_0}}} \tag{2.63}$$

$$\vec{v}_{q=2} = \begin{pmatrix} 1 - \frac{1-k}{2k}\left(\frac{\delta\omega_c}{\omega_0} - 4\frac{\delta\omega_e}{\omega_0}\right) \\ -1 - \frac{1-k}{2k}\frac{\delta\omega_c}{\omega_0} \\ 1 - \frac{1-k}{2k}\left(\frac{\delta\omega_c}{\omega_0} + 4\frac{\delta\omega_e}{\omega_0}\right) \end{pmatrix}. \tag{2.64}$$

The introduction of individual cell frequency errors produces some interesting insights to the three coupled oscillator system. Note the $\pi/2-$ mode results shows the second order perturbation theory correction. The two end cell magnitudes vary only with frequency errors squared. This result is not true of the 0 and $\pi-$ mode, which have first order frequency error contributions. This result shows a benefit of the $\pi/2-$mode.

Effects of power losses

For lossy systems, energy is lost in each oscillator from Ohmic (or power) dissipation. To best describe a coupled cavity system in steady-state a power generator is needed to account for losses. Power is coupled into a drive cell and distributes power along the chain. The voltage sum in the drive cell is the same as the non-lossy case with an additional drive term [35]. The eigenvector solutions for the three modes take the form

$$\vec{v}_{q=0} = \begin{pmatrix} 1 \\ e^{-i\frac{3\sqrt{1+k}}{kQ}} \\ e^{-i\frac{4\sqrt{1+k}}{kQ}} \end{pmatrix} \tag{2.65}$$

$$\vec{v}_{q=1} = \begin{pmatrix} 1 \\ \frac{1}{kQ}e^{i\frac{\pi}{2}} \\ -1 + \frac{2}{k^2Q^2} \end{pmatrix} \tag{2.66}$$

$$\vec{v}_{q=2} = \begin{pmatrix} 1 \\ -e^{i\frac{3\sqrt{1-k}}{kQ}} \\ e^{i\frac{4\sqrt{1-k}}{kQ}} \end{pmatrix}. \tag{2.67}$$

Thus the magnitude of the 0 and π mode fields remains 1 to first order, however have a shifted phase in the centre and end cell, called the power flow phase shift. The centre and end cell in the $\pi/2-$mode do not experience a power flow phase shift. The end cell (to second order) observes a power flow droop, a real effect that decreases the magnitude of the field in cells further from the drive cell. This phenomenon decreases with both the coupling constant and Q factor.

The equivalent circuit model has been useful for gaining understanding of the nature of coupled oscillators. The above discussion apply to coupled cavity accelerators. It has been

shown that individual cell frequency errors and Ohmic losses cause unattractive features in the 0 and $\pi-$ normal modes, in that the cell fields reduce in magnitude and experience power flow phase shifts. As a result, these modes are less tolerant to manufacturing errors. This effect can be visualised by picturing the shape of the dispersion relation. Modes that reside closer in frequency observe larger perturbative effects. The $\pi/2-$ mode lies on the dispersion relation where the gradient is largest and therefore mode spacing is maximised. Figure 2.9 displays the dispersion curve of a 9-cell coupled oscillator chain, the $\pi/2$ mode is shown with a larger point size. A given normal mode has a natural spacing $\Delta\omega = \frac{\omega}{Q}$ thus higher Q will reduce the likelihood of exciting multiple normal modes.

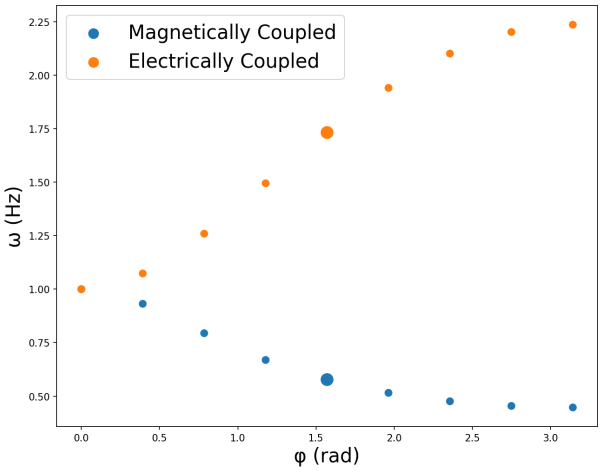


Figure 2.9: Dispersion curve of a 9-cell chain of coupled oscillators, coupled both electrically and magnetically.

The $\pi/2-$mode can be altered such that the nominally unexcited cells can be made physically short in the longitudinal dimension (the requirement for the correct frequency will change the radius of the cell). This is called a bi-periodic structure and provides the increased tolerance to frequency errors and power losses whilst the high shunt impedance of a π mode. The unexcited cells can also be moved physically off axis which are coupled to excited cells via magnetic coupling. This type of structure is called side-coupled, which is beneficial for the same reasons as the bi-periodic structure. Nevertheless, increasing the coupling constant of a chain of oscillators requires increased modulation of the magnetic field, which reduces the shunt impedance. As a result, accelerator design faces a conflict between increasing the coupling constant (to improve cavity tolerance to frequency errors and power losses) and maximising the shunt impedance for efficient transfer of energy to a particle beam. If a cavity is too long, the number of cells increases, increasing the required coupling and dropping the shunt impedance.

When designing a side-coupled $\pi/2$ structure, the non-excited cells are off-axis, and are geometrically different from the accelerating cells. In order to realise the structure, the $\pi/2$

resonant frequency of the accelerating and side-coupled cells must be equal. For non-zero differences in frequency, the power flow phase shift reappears, which was characteristically suppressed in the $\pi/2$ normal mode. The power flow droop still persists and therefore both effects require minimising. The field in an accelerating cell of a side-coupled structure is displayed in section 6.2.2.

2.4 Power Coupling

In the previous section, the lossy system required a generator attached to the drive cell. For a cavity system, power coupling can be achieved in a variety of ways, such as a magnetic-coupling loop on a coaxial transmission line, which induces a magnetic field when current flows. For higher power coupling, rf power is fed into the cavity via a waveguide.

A rf generator produces a current that flows through a waveguide to the cavity through a power coupler. Reflected power is dumped into a matched load via a circulator. The external Q factor is the Q of the external load, Q_{ex} and is given

$$Q_{ex} = \frac{\omega_0 U}{P_{ex}} \tag{2.68}$$

Where U is the total energy in the system. The total power loss is the sum of the external power loss from reflections and the power lost in the cavity, P_c, $P_t = P_{ex} + P_c$. The total power has an associated Q factor called the loaded Q,

$$Q_L = \frac{\omega_0 U}{P_t} = \frac{\omega_0 U}{P_{ex} + P_c}. \tag{2.69}$$

Thus,

$$\frac{1}{Q_L} = \frac{1}{Q_{ex}} + \frac{1}{Q_0}. \tag{2.70}$$

Defining $\frac{P_{ex}}{P_c}$ to be the waveguide to cavity coupling strength;

$$\frac{P_{ex}}{P_c} = \frac{Q_0}{Q_{ex}} = \beta. \tag{2.71}$$

It can be shown using an equivalent circuit that for a generator driven at the resonant frequency, the waveguide loaded impedance is related to the cavity load impedance by the factor β, $Z_L = \beta Z_c$. The reflection coefficient is given;

$$\Gamma = \frac{Z_L - Z_c}{Z_L + Z_c} = \frac{1 - \beta}{1 + \beta}. \tag{2.72}$$

$\Gamma = 0$ requires $\beta = 1$, this is called critically coupled and produces a matched state. Impedance matching also maximises the power flow into the cavity. The reflected power for a given input power is

$$P_{ref} = P_{in}\Gamma^2. \tag{2.73}$$

Energy conservation requires

$$E_{in} = E_{ref} + E_c \rightarrow P_{in} = P_{ref} + P_c, \tag{2.74}$$

therefore;

$$P_c = P_{in} - P_{ref} = P_{in}\left(1 - \Gamma^2\right) = P_{in}\left(1 - \frac{1-\beta}{1+\beta}\frac{1-\beta}{1+\beta}\right) = P_{in}\frac{4\beta}{(1+\beta)^2}, \tag{2.75}$$

which is maximum for $\beta = 1$. The above assumes a constant reflection coefficient during steady state. In reality, the power inside the cavity must fill up to the maximum value during the transient period. The cavity time constant is given as $\tau = \frac{2Q_L}{\omega_0}$. The transient reflection coefficient is;

$$\Gamma(t) = (1 - e^{(-t/\tau)})\frac{2\beta}{1+\beta} - 1, \tag{2.76}$$

the power in the cavity is therefore;

$$P_c(t) = P_{in}(1 - \Gamma(t)^2) = P_{in}\left(1 - \left((1 - e^{(-t/\tau)})\frac{2\beta}{1+\beta} - 1\right)^2\right). \tag{2.77}$$

For $\beta = 1$;

$$P_c(t) = P_{in}\left(1 - e^{(-2t/\tau)}\right). \tag{2.78}$$

For different values of β, different cavity reflections are observed with time. When $\beta << 1$,

$$\Gamma(t) = (1 - e^{(-t/\tau)})\frac{2\beta}{1+\beta} - 1 \approx (1 - e^{(-t/\tau)})2\beta - 1 \rightarrow |\Gamma(t \rightarrow \infty)| = |2\beta - 1|. \tag{2.79}$$

For $\beta >> 1$;

$$|\Gamma(t)| \approx |1 - 2e^{(-t/\tau)}|. \tag{2.80}$$

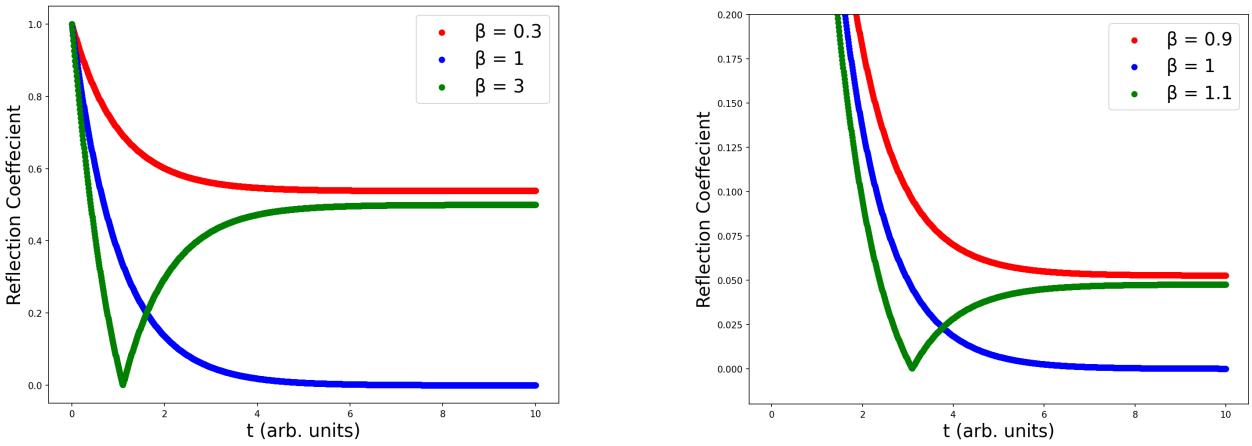
Indeed for $\beta = 1$;

$$|\Gamma(t)| = e^{(-t/\tau)} \rightarrow |\Gamma(t \rightarrow \infty)| = 0, \tag{2.81}$$

as expected. Figure 2.10 displays the reflection coefficient when power coupling to a Standing Wave cavity during the transient period for different coupling coefficient, β. (b) displays the response for β close to 1. For β errors of $\approx$ 10%, the cavity will still experience reflections $\approx$ 5%.

2.4.1 Traveling and Standing Wave Operation

Standing Wave (SW) cavities are resonant structures that are driven in either the π or $\pi/2$ normal modes. Power is coupled into the input coupler and the field builds up as the reflection coefficient tends to 0 over multiple time constants. This can be visualised



(a) (b)

Figure 2.10: Reflection coefficient during transient filling period of Standing Wave cavity for different coupling constant, β. Different values of β are shown in (a) and (b).

as multiple reflections within the cavity building the field evenly throughout the structure. Power emitted from the cavity destructively interferes with the reflected power until the stead state solution is reached. In Traveling Wave (TW) operation, there still exists a dispersion relation, and a particular mode is excited. However, power flows from the drive cell to the next cell, sequentially, depositing electromagnetic energy proportional to the group velocity of the wave. There are two main methods of particle acceleration using a traveling wave cavity, constant impedance and constant gradient. Constant impedance structures have identical cells at each period. Ohmic power loss per unit length is

$$\frac{dP_w}{dz} = -\frac{\omega U}{Q}. \tag{2.82}$$

The shunt impedance per unit length is

$$Z = \frac{E_l^2}{-\frac{dP_w}{dz}}, \tag{2.83}$$

where E_l is the magnitude of the longitudinal field. The wave energy travels at the group velocity, $v_g = P_w/U$. It is possible to show that;

$$\frac{1}{P_w}\frac{dP_w}{dz} = -\frac{\omega}{Qv_g}, \tag{2.84}$$

which has solutions;

$$P_w(z) = P_0 e^{-\frac{\omega}{Qv_g}z}, \tag{2.85}$$

and describes the wave power exponentially decreasing with z. Hence, α_0 is the attenuation per unit length and is defined

$$\alpha_0 = \frac{\omega}{2Qv_g}. \tag{2.86}$$

The longitudinal field amplitude also varies exponentially with z, indeed,

$$E_l^2 = \frac{\omega Z P_w}{Qv_g} \tag{2.87}$$

thus;

$$E_l = \sqrt{\frac{\omega Z P_0}{Qv_g}} e^{-\alpha_0 z} = E_0 e^{-\alpha_0 z}, \tag{2.88}$$

and the longitudinal field magnitude decreases exponentially with z similar to the wave power. The energy deposited in each cell available for particle acceleration is

$$U(z) = \frac{P_w(z)}{v_g}, \tag{2.89}$$

and decreases along the structure. The total attenuation over a cavity of length L_{cav}, $\tau_0 = \alpha_0 L_{cav}$. The time taken for the cavity to fill with energy is $t_{fill} = \frac{L_{cav}}{v_g} = \tau_0 \frac{2Q}{\omega}$. It is possible to create a system by which the wave power and group velocity decrease at the same rate, so as to keep $U(z)$ constant with z, and each cell provides the same gradient. The only way this is possible is to make v_g dependant on z;

$$P_w(z) = U v_g(z). \tag{2.90}$$

Writing $P_w(z)$ and $v_g(z)$ as power series in z and absorbing the constant U,

$$a_0 + a_1 z + a_2 z + ... + a_n z^n = b_0 + b_1 z + b_2 z^2 + ... + b_n z^n. \tag{2.91}$$

Therefore, the wave power and group velocity behave identically with z. The attenuation per unit length is therefore a function of z, $\alpha_0(z) = \frac{\omega}{2Qv_g(z)}$.

$$\frac{dP_w}{dz} = -2\alpha_0(z) P_w(z), \tag{2.92}$$

$$\frac{dP_w}{P_w(z)} = -2\frac{\omega}{2Qv_g(z)} dz = -2\frac{\omega U}{2QP_w(z)} dz. \tag{2.93}$$

Thus,

$$\frac{dP_w}{dz} = -\frac{\omega U}{Q} = \text{constant}, \tag{2.94}$$

as U is treated constant. The wave power and group velocity are therefore linearly dependant on z. This assumes the Q factor is constant with z. Thus;

$$P_w(z) = P_0 + (P_1 - P_0)\frac{z}{L_{cav}}, \tag{2.95}$$

$$v_g(z) = v_{g,0} + (v_{g,1} - v_{g,0})\frac{z}{L_{cav}}. \tag{2.96}$$

For final power P_1 and final group velocity, $v_{g,1}$. This requires the attenuation per unit length to also vary along the structure,

$$\alpha_0(z) = \frac{\omega}{2Qv_{g,0}(1 + (\frac{v_{g,1}}{v_{g,0}} - 1)\frac{z}{L_{cav}})} = \frac{\omega}{2Qv_{g,0}}\left[1 - (\frac{v_{g,1}}{v_{g,0}} - 1)\frac{z}{L_{cav}} + \left((\frac{v_{g,1}}{v_{g,0}} - 1)\frac{z}{L_{cav}}\right)^2\right], \tag{2.97}$$

where a binomial expansion was taken. As the wave power and group velocity reduce along the structure, the quantity $\frac{v_{g,1}}{v_{g,0}} - 1$ is negative, and each term in square brackets is positive. Therefore it is clear the attenuation increases with z. This is often referred to as tapering the group velocity. The above describes a constant gradient structure. Tapering is achieved by reducing the cell aperture or coupling slots along the structure. Reducing the real forward power flow reduces the group velocity, as the power flow is the integral of the Poynting vector perpendicular to the cross sectional area between cells. The fill time for a constant gradient cavity is

$$t_{fill} = \int_0^{L_{cav}} \frac{z}{v_{g,0} + (v_{g,1} - v_{g,0})\frac{z}{L_{cav}}} dz \approx \frac{2L_{cav}}{v_{g,0} + v_{g,1}}, \tag{2.98}$$

and therefore has longer fill times than a constant impedance cavity of the same length and starting group velocity. Typical group velocities are in the range of a few percent of the speed of light, however is an important quantity to be optimised for a given design.

SW cavities fill with a time dependant reflection coefficient, as discussed. The transient stage observes initially 100% reflection of the wave power which reduces to near 0 over multiple time constants, $\frac{2Q}{\omega}$ (a 95% filled cavity takes around three time constants). TW cavities have generally a shorter fill time. As a result, a TW cavity can accept a beam sooner than a SW cavity, resulting in more particles being accelerated per unit time.

The above description of cavity filling in both standing and TW cavities allows a simplified visualisation of how field builds up in both cases. Standing wave structures see field growing equally in all cell simultaneously. TW structures see field build in each cell subsequently. Hence, it is often referred that SW structures fill in time, whilst TW structures fill in space. As previously suggested, long standing wave structures are not viable, as the increased number of cells reduces the shunt impedance and therefore rf efficiency. This issue is not observed for TW structures. As power flows from the input guide to the output load, increasing the number of cells does not have any effect on the upstream cells, the downstream cells receive less power and therefore less stored energy - however - the wave power can be modulated

by changing the group velocity profile along the cavity. This freedom is not possible in a standing wave cavity, without suffering large power flow droop and power flow phase shift.

RF Network Quantities

Given the previous sections on basic rf acceleration, it is important to introduce various rf system parameters that define how often and the duration at which rf power is coupled into a cavity via an rf network. RF systems can operate with either continuous wave, (CW), or pulsed signals. Pulsed systems may use pulse compressors to increase peak power and shorten a pulse from a power source, such as a Klystron. The advantage of pulsed systems over continuous systems is the ability to run at a lower average power whilst producing large peak power signals. The pulse length, t_p, describes the duration of the rf pulse. The repetition rate, (RR), describes the frequency of pulses, and is defined as the number of pulses per second. Using the pulse length and repetition rate, the duty cycle (DC) is defined as;

$$DC = RR \times t_p \tag{2.99}$$

and defines the fraction of time the rf signal is 'on'. For the proton radiotherapy linac discussed in this project, the value of the pulse length is 5 μs and the RR is 200 Hz (thus there is a pulse every $\frac{1}{200} = 0.005$ s). The DC is therefore $5 \times 10^{-6} \times 200 = 0.001$, or 0.1%.

Average and Pulsed Heating

The average power, P_{avg}, is defined as

$$P_{avg} = P_{peak} \times DC, \tag{2.100}$$

where P_{peak} describes the peak power of the rf pulse. By limiting the average power a cavity experiences, there is less pulsed heating due to Ohmic losses. If the cavity becomes too hot, surface cracking and deformation occur which will result in a resonant frequency shift, this is termed average heating. By using cooling channels, flowing with water, heat can be extracted from the system to avoid temperature rises in the conducting material. A common figure of merit is the average power deposited per unit length, $\frac{P_{avg}}{L_{cav}}$. Where L_{cav} is the length of the cavity in which the average power is being deposited. For normal conducting copper structures, a conservative value for the average power per unit length for an S-band cavity is 4.5 kW/m [51].

If the pulse has a peak power that is too high, pulsed heating becomes a limiting factor. Pulse heating is the instantaneous rise in the material surface temperature - that has not yet diffused into the bulk of the material - and is confined to a small volume within the skin depth of the material. The time scale of pulsed heating is of the order 10 μ s. An approximate limit on the surface temperature rise is $\sim$ 40 K for normal conducting cavities. The temperature

gain can be calculated using the following

$$\Delta T = \frac{2P_d\sqrt{t_p}}{\sqrt{\pi\rho\kappa c_e}}, \tag{2.101}$$

where ρ is the material density, κ is the thermal conductivity and c_e is the specific heat capacity of the material. P_d is the deposited power and is given

$$P_d = \frac{R_s H_{peak}^2}{2}. \tag{2.102}$$

For peak magnetic field, H_{peak} and surface resistance R_s. The surface resistance is a function of the frequency of the rf power, ω, and skin depth, σ;

$$R_s = \sqrt{\frac{\mu_0}{\omega}2\sigma}, \tag{2.103}$$

where μ_0 is the vacuum magnetic permeability. In fact, surface resistance is also a function of temperature change. For high power operation, the effect of increased temperature causes an increased rate of power loss, however is only of the order of several percent, [58].

Hence whilst it is important to limit the average power, to reduce average heating, one can not simply reduce the DC for the same peak power, as pulsed heating is a function of the peak power and must not exceed $\sim$ 40 K [59].

2.5 Introduction to Beam Dynamics

The content discussed in this section are very well documented in the following book, [29]. So far, core concepts regarding rf particle acceleration have been briefly discussed. A major component of particle acceleration considers the effect of the electromagnetic fields on the particle dynamics.

In this section, transverse and longitudinal beam dynamics will be briefly introduced.

2.5.1 Transverse Beam Dynamics

When an electrically charged particle traverses some element that produces an electromagnetic field (e.g. an rf single cell that is excited in the TM_{010} mode), or a dipole magnet, the electromagnetic fields will interact with the particle, from Lorentz equation;

$$\vec{F} = \mathrm{q}(\vec{E} + \vec{v} \times \vec{B}). \tag{2.104}$$

In order to keep a particle along a required path, magnets may be used to bend (e.g. dipole) or focus the particle trajectory (e.g. dipole and quadrupole) as needed. Consider a simple

accelerating system, with only dipole and quadrupole magnets, drift lengths, and rf cavities. The particle trajectory may make a circular shape, as in the case of storage rings or circular colliders. Alternatively, the trajectory may be in a linear line, and the system is refered to as a linear accelerator, or a linac (where dipoles are not required, as there is no beam bending).

In general, a particle trajectory is designed (called a design orbit) so that particles are focused when the beam size (horizontal displacement) is large, in order to keep particles confined. In general, the particle trajectory is defined by the beam line lattice, a periodic collection of elements, that define how a particle bunch is transported along an accelerator. As a result, particle trajectories are systems that following simple harmonic motion, with particles being focused when the horizontal displacement is too large. The equation of motion describing the transverse particle position, x, is called Hill's equation [29],

$$\frac{d^2x}{ds^2} = -k(s)x, \tag{2.105}$$

where s is the longitudinal variable and $k(s)$ is the focusing strength. The solution to Hill's equation is the following;

$$x(s) = \sqrt{2J_x\beta_x(s)}\cos(\phi_x(s)), \tag{2.106}$$

where $\beta_x(s)$ is called the beta function, and is a periodic function with the same periodicity as $k(s)$. The variables J_x and ϕ_x are known as the action-angle variables and are called the betatron amplitude and phase, respectively. The betatron amplitude is a constant of integration. Therefore, every particle has an invariant betatron amplitude at all positions along a lattice.

A particle also has a divergence, given $x' = dx/ds$. Together, x, x' make the phase space coordinates of a particle trajectory. The divergence can be calculated by differentiating Eqn. 2.106,

$$\frac{dx}{ds} = \sqrt{2J_x}\frac{d\beta_x(s)}{ds}\frac{1}{2}\beta_x(s)^{-1/2}\cos(\phi_x) - \sqrt{2J_x\beta_x(s)}\sin(\phi_x)\frac{d\phi_x(s)}{ds}. \tag{2.107}$$

The alpha function is defined as;

$$\alpha_x(s) = -\frac{1}{2}\frac{d\beta_x(s)}{ds}. \tag{2.108}$$

Therefore the divergence becomes

$$x' = -\sqrt{\frac{2J_x}{\beta_x(s)}}\alpha_x(s)\cos(\phi_x) - \sqrt{\frac{2J_x}{\beta_x(s)}}\sin(\phi_x) = -\sqrt{\frac{2J_x}{\beta_x(s)}}(\alpha_x(s)\cos(\phi_x) + \sin(\phi_x)). \tag{2.109}$$

The gamma function is defined as;

$$\gamma_x(s) = \frac{1+\alpha_x(s)^2}{\beta_x(s)}. \tag{2.110}$$

Together, $\beta_x(s), \alpha_x(s), \gamma_x(s)$ are the Twiss (or Courant-Snyder) parameters, and are strictly functions of the lattice.

So far, everything discussed has been relative to a single particle. It is required to extend to discussions to a beam of particles by using the average value of the square of the phase space coordinates, $\langle x^2 \rangle, \langle x'^2 \rangle, \langle xx' \rangle$. Using the fact $\langle \cos(x)^2 \rangle = \langle \sin(x)^2 \rangle = \frac{1}{2}$,

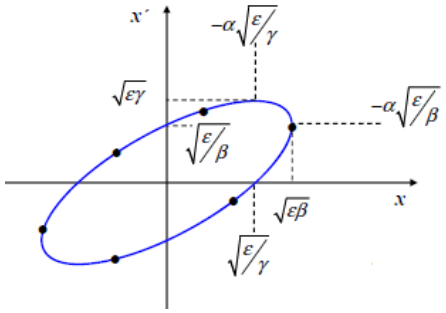
$$\langle x^2 \rangle = \beta_x \langle J_x \rangle, \quad \langle x'^2 \rangle = \gamma_x \langle J_x \rangle, \quad \langle xx' \rangle = -\alpha_x \langle J_x \rangle. \tag{2.111}$$

As each particle has an independant value of J_x, which is constant with position along the beam line (provided the lattice is comprised of symplectic transforms, see later), then the value $\langle J_x \rangle$ is also constant, and is called the geometric beam emittance, ε_g;

$$\varepsilon_g = \sqrt{\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2}. \tag{2.112}$$

The phase space coordinates can be combined to produce the equation for an ellipse, of constant area, $\pi\varepsilon_g$, as shown in Fig. 2.11,

$$\gamma_x x^2 + 2\alpha_x xx' + \beta_x x'^2 = \varepsilon_g. \tag{2.113}$$



(a)

Figure 2.11: Transverse beam ellipse described by Twiss parameters [60].

For an arbitrary point along the beam line lattice, the values of the Twiss parameters produces a rotated ellipse of area $\pi\varepsilon_g$.

Transfer Maps

Transfer maps are linear transformations that describe how the phase space evolves when transported through a given lattice, from point s to $s+L$. For a 1D phase space, the transfer map, R, acts on the initial phase space (x_0, x'_0) to produce the output phase space (x_1, x'_1).

The new coordinates can be written as a function of the old coordinates using a Taylor expansion,

$$x_1 = \frac{\partial x_1}{\partial x_0} x_0 + \frac{\partial x_1}{\partial x_0'} x_0', \tag{2.114}$$

$$x_1' = \frac{\partial x_1'}{\partial x_0} x_0 + \frac{\partial x_1'}{\partial x_0'} x_0'. \tag{2.115}$$

The above mappings can be written in matrix form;

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial x_0} & \frac{\partial x_1}{\partial x_0'} \\ \frac{\partial x_1'}{\partial x_0} & \frac{\partial x_1'}{\partial x_0'} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix}. \tag{2.116}$$

The mapping in Eqn. 2.116 is a linear transformation from one coordinate system, (x_0, x_0') to new coordinate system (x_1, x_1'). The matrix constructed from partial derivatives is the Jacobian matrix [61]. For a change of variables (mathematically equivalent to transforming from old to new coordinates), an area element is scaled by a factor given by the Jacobian determinant, J,

$$\delta x_0 \delta x_0' = J \delta x_1 \delta x_1'. \tag{2.117}$$

When $J = 1$, the area of a phase space element is conversed. Hence, a particle bunch occupies the same phase space area before and after the transformation, and the geometric emittance is conserved.

As the Jacobian matrix is comprised of first order partial derivatives only, it is the best linear approximation that can be made. A transformation is symplectic if it obeys the following criteria;

$$R^T \Omega R = \Omega, \tag{2.118}$$

where Ω is a non-singular, skew-symmetric matrix and is often taken to be the block matrix;

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \tag{2.119}$$

I_n is the $n \times n$ identity matrix. All Symplectic matrices have determinant $+1$ [62] and hence the geometric emittance is conserved under the transformation. This is called Liouville's theorem, the area of a phase space distribution is constant with respect to s.

In general, the geometric emittance is not Lorentz invariant, and so the normalised emittance, ε_n, is defined;

$$\varepsilon_n = \varepsilon_g \gamma_r \beta_r. \tag{2.120}$$

Where γ_r is the Lorentz factor and β_r is the normalised particle velocity. The normalised emittance is a Lorentz invariant quantity. For some transformation that does not have unit determinant, the geometric emittance will change. The Jacobian determinant is therefore;

$$J = \frac{\varepsilon_{g1}}{\varepsilon_{g0}} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}. \tag{2.121}$$

As an example, consider a transformation that increases the value of $\gamma_r\beta_r$. From Eqn. 2.121, the Jacobian determinant is less than one, and the phase space area reduces, this is adiabatic damping.

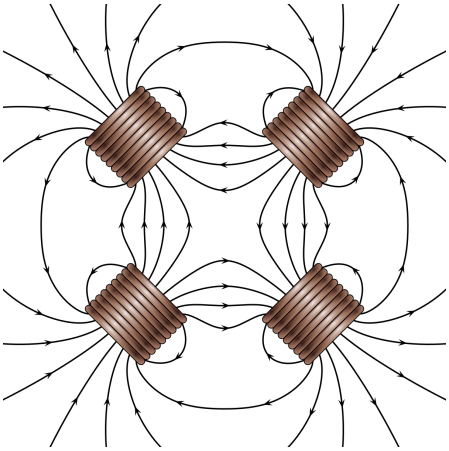
Equation 2.116 introduced the linear transfer map. Relevant common lattice element transfer maps are shown below [29].

$$\text{Drift Space} = \begin{pmatrix} 1 & L_{drift} \\ 0 & 1 \end{pmatrix} \tag{2.122}$$

$$\text{Focussing Quadrupole} = \begin{pmatrix} \cos(\sqrt{k}l_q) & \frac{1}{\sqrt{k}}\sin(\sqrt{k}l_q) \\ -\sqrt{k}\sin(\sqrt{k}l_q) & \cos(\sqrt{k}l_q) \end{pmatrix} \tag{2.123}$$

$$\text{Defocusing Quadrupole} = \begin{pmatrix} \cosh(\sqrt{k}l_q) & \frac{1}{\sqrt{k}}\sinh(\sqrt{k}l_q) \\ \sqrt{k}\sinh(\sqrt{k}l_q) & \cosh(\sqrt{k}l_q) \end{pmatrix} \tag{2.124}$$

Where l_q is the quadrupole length and k is the quadrupole k-strength, and is a normalised quantity with definition $k = \frac{q}{p_0}\frac{\partial B_y}{\partial x}$. Where the momentum p_0 and vertical magnetic field, B_y are used. The quadrupole field is shown in Fig. 2.12.



(a)

Figure 2.12: Quadrupole Magnetic field profile [63].

The quadrupole field strength increases linearly with transverse distance, and focuses in either x or y, not both. For a quadrupole that focuses in x, it will defocus in y. This can be seen in Eqn. 2.124, as the focussing quadrupole is comprised of trigonometric functions (the element has positive $k(s)$ in Eqn. 2.105) whilst the defocussing quadrupole is comprised of hyperbolic functions (the element has negative $k(s)$ in Eqn. 2.105). Beam line lattice are often compiled of alternating polarity quadrupoles to keep a beam focused in both transverse planes simultaneously. For a very short quadrupole length, the term $\sqrt{k}l_q << 1$ and the

trigonometric/hyperbolic terms may be approximated, this method is utilised is described in depth in Section 4.

Betatron Phase Advance

Consider a circular beam line lattice, comprised of dipoles, quadrupoles, drift spaces and rf cavities. As the Twiss parameters are functions of the lattice, they are constant with respect to integer revolutions on the lattice. Suppose the Twiss parameters are measured at point s_0. Plotting the (x, x') phase space coordinates of a particle at point s_0 for many turns will produce an ellipse of area $2\pi J_x$. Whilst J_x is conserved under transformation R, the betatron phase, ϕ_x, increases. The amount it advances is given by μ_x, the phase advance ($\mu_x(s_0, s_1) = \phi_x(s_1) - \phi_x(s_0)$), and is calculated as follows;

$$\mu_x(s_0, s_1) = \int_{s_0}^{s_1} \frac{ds}{\beta_x(s)}. \tag{2.125}$$

A transfer map can be equivalently described as a function of the Twiss parameters and phase advance,

$$R = \begin{pmatrix} \sqrt{\frac{\beta_{x1}}{\beta_{x0}}}(\cos(\mu_x) + \alpha_{x0}\sin(\mu_x)) & \sqrt{\beta_{x0}\beta_{x1}}\sin(\mu_x) \\ \frac{\alpha_{x0}-\alpha_{x1}}{\sqrt{\beta_{x0}\beta_{x1}}}\cos(\mu_x) - \frac{1+\alpha_{x0}\alpha_{x1}}{\sqrt{\beta_{x0}\beta_{x1}}}\sin(\mu_x) & \sqrt{\frac{\beta_{x0}}{\beta_{x1}}}(\cos(\mu_x) - \alpha_{x1}\sin(\mu_x)) \end{pmatrix}. \tag{2.126}$$

For one complete turn, the transfer map reduces to the following form

$$R_T = \begin{pmatrix} \cos(\mu_{x,T}) + \alpha_x \sin(\mu_{x,T}) & \beta_x \sin(\mu_{x,T}) \\ -\gamma_x \sin(\mu_{x,T}) & \cos(\mu_{x,T}) - \alpha_x \sin(\mu_{x,T}) \end{pmatrix}, \tag{2.127}$$

where $\gamma_x = \frac{1+\alpha_x^2}{\beta_x}$ and $\mu_{x,T}$ is the phase advance over one turn, and is related to the tune, $Q_x = \frac{\mu_{x,T}}{2\pi}$. The tune describes in number of betatron oscillations per turn. Let M denote the one turn map, $R_i R_{i-1} \cdots R_2 R_1 R_0 = M$ for i individual transfer maps. Over one period, the trace of Eqn. 2.126 is

$$\mathrm{Tr}(M) = 2\cos(\mu_{x,T}). \tag{2.128}$$

A stable system is one in which $\mathrm{Tr}(M) \leq 2$ and $-1 \leq \cos(\mu_{x,T}) \leq 1$. Visualisation of the phase advance can be done with the use of the normalising matrix,

$$N_x = \begin{pmatrix} \frac{1}{\sqrt{\beta_x}} & 0 \\ \frac{\alpha_x}{\sqrt{\beta_x}} & \sqrt{\beta_x} \end{pmatrix}. \tag{2.129}$$

Acting the normalising matrix on the phase space coordinates to produce the normalised phase space coordinates;

$$x_N = \frac{x}{\sqrt{\beta_x}}, \tag{2.130}$$

$$p_{x,N} = \frac{\alpha_x}{\sqrt{\beta_x}} x + \sqrt{\beta_x} x'. \tag{2.131}$$

Combining x_N and $p_{x,N}$ in the following way,

$$x_N^2 + p_{x,N}^2 = \gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2 = \varepsilon_{g,x} \tag{2.132}$$

recovers Eqn. 2.113. Thus, the normalised phase space $(x_N, p_{x,n})$ produces a circle with area $\pi\varepsilon_g$. Consider the transformation from point $s_0 \to s_1$,

$$R\vec{x_0} = \vec{x_1}. \tag{2.133}$$

The normalising matrices transform as the following,

$$N_{x,0}\vec{x_0} = \vec{x_{N,0}}, \quad N_{x,1}\vec{x_1} = \vec{x_{N,1}}. \tag{2.134}$$

Therefore,

$$\vec{x_0} = N_{x,0}^{-1}\vec{x_{N,0}}, \quad \vec{x_1} = N_{x,1}^{-1}\vec{x_{N,1}}. \tag{2.135}$$

Substituting Eqn. 2.135 into Eqn. 2.133,

$$RN_{x,0}^{-1}\vec{x_{N,0}} = N_{x,1}^{-1}\vec{x_{N,1}}. \tag{2.136}$$

Acting on the left with $N_x(s_1)$;

$$N_{x,1}RN_{x,0}^{-1}\vec{x_{N,0}} = \vec{x_{N,1}}. \tag{2.137}$$

Where $N_{x,1}RN_{x,0}^{-1}$ is calculated by combining Eqns. 2.126 and 2.129,

$$N_{x,1}RN_{x,0}^{-1} = \begin{pmatrix} \cos(\mu_x) & \sin(\mu_x) \\ -\sin(\mu_x) & \cos(\mu_x) \end{pmatrix}. \tag{2.138}$$

Thus, the phase advance is the rotation angle due to the transfer map, R, in normalised phase space. Whilst much of what has been described was motivated by considering a circular machine, everything discussed is valid for a linear lattice, provided there is periodicity.

2.5.2 Longitudinal Beam Dynamics

RF particle acceleration relies on a particle continuously observing an accelerating field. Types of accelerating structure were briefly discussed in Chapter 1.

The longitudinal beam dynamics (LBD) equivalence of Betatron motion in transverse beam dynamics (TBD) is called Synchrotron motion, and particles perform synchrotron oscillations. The motion considers how the longitudinal variables, (z, p_z) vary as a function of position along a lattice. Similarly to TBD, there is a longitudinal tune, which is the number of synchrotron oscillations per lattice period. Synchrotron oscillations arise when considering

circular machines, and depict how particles move ahead of, and behind the synchronous particle, due to factors like the circumference of the orbit, and particle energy.

When considering motion in a linear system, there are no synchrontron oscillations due to changes in path length. However, particles will still oscillate about the synchronous particle, as described below.

The energy gain of a particle when traversing a single cell of length, L, is calculated as follows;

$$\Delta E = qV_0 = q\int_{-L/2}^{L/2} E_z(z,t)\cos\left(\frac{\omega z}{\beta_z(z)c} + \phi\right)dz, \tag{2.139}$$

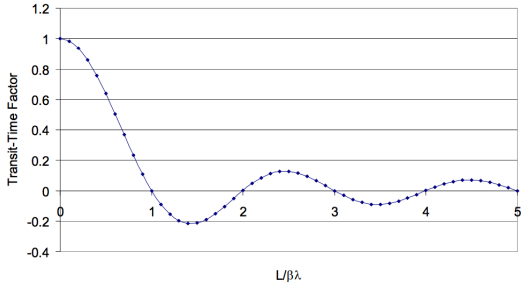
where $E_z(z,t)$ is the longitudinal electric field, ω is the frequency of the field, $\beta_z c$ is the velocity in the longitudinal direction and ϕ is the phase of the field. If the cavity is operating in a π-mode, the particle must cross an rf cell in half an rf period, $P/2$. This is called the synchronicity condition. For a general phase advance of ψ, the synchronicity condition requires the cell length to be

$$L = \frac{\psi\beta c}{\omega}. \tag{2.140}$$

The Transit Time Factor, T, is the factor that accounts for a time varying electric field as the particle traverses the gap;

$$T = \frac{\int_{-L/2}^{L/2} E_z(z,t)\cos(\frac{\omega z}{\beta_z(z)c} + \phi)dz}{\int_{-L/2}^{L/2} E_z(z,t)dz}. \tag{2.141}$$

The transit time factor is shown as a function of cell length in Fig. 2.13. When the cell length is equal to an integer value of $L/\beta\gamma$, the particle observes all phases. As a result, total energy gain is zero, $T = 0$. The general result for the energy gain is therefore



(a)

Figure 2.13: Transit time factor as a function of cell length.

$$\Delta E = qV_0 T\cos(\phi). \tag{2.142}$$

Electrons in accelerators reach $\beta \approx 1$ after a few cells. The initial cell lengths change from one cell to the next, to ensure synchronicity. However, after $\beta \approx 1$, the cell length is constant, and all particles travel at roughly the same velocity. In this regime, particles do not oscillate about a synchronous particle. The beam dynamics are different to proton accelerators, due to the difference in particle mass (protons are $\sim 2000\times$ more massive than electrons). As a result, the change in velocity for a proton over an rf cell is small. Longitudinal beam dynamics for small amounts acceleration is described below.

By considering a series of accelerating cells, coupled differential equations (with respect to longitudinal distance, s) can be constructed describing both the phase and energy gain deviation between an arbitrary particle and the synchronous particle [35];

$$\gamma_s^3 \beta_s^3 \frac{d(\phi - \phi_s)}{ds} = -2\pi \frac{(E - E_s)}{mc^2 \lambda}, \tag{2.143}$$

$$\frac{d(E - E_s)}{ds} = qE_0 T(\cos(\phi) - \cos(\phi_s)). \tag{2.144}$$

Where γ_s and β_s are the Lorentz factor and normalised velocity of the synchronous particle. m is the mass of the particle and λ is the wavelength. As the acceleration rate of protons is small, the values of E_0T, ϕ_s and $\gamma_s\beta_s$ are assumed constant.

By denoting the following;

$$w = \frac{E - E_s}{mc^2}, \;\; A = \frac{2\pi}{\gamma_s^3 \beta_s^3 \lambda}, \;\; B = \frac{qE_oT}{mc^2}, \tag{2.145}$$

eqns. 2.143 and 2.144 become;

$$\phi' = \frac{d\phi}{ds} = -Aw \tag{2.146}$$

and,

$$w' = \frac{dw}{ds} = B(\cos(\phi) - \cos(\phi_s)) \tag{2.147}$$

respectively. Where $\frac{d\phi}{ds} = \frac{\phi - \phi_s}{ds}$, and $\frac{dw}{ds} = \frac{(E - E_s)}{ds}$. Differentiating Eqn. 2.146,

$$\frac{d\phi'}{ds} = -Aw' = -AB(\cos(\phi) - \cos(\phi_s)). \tag{2.148}$$

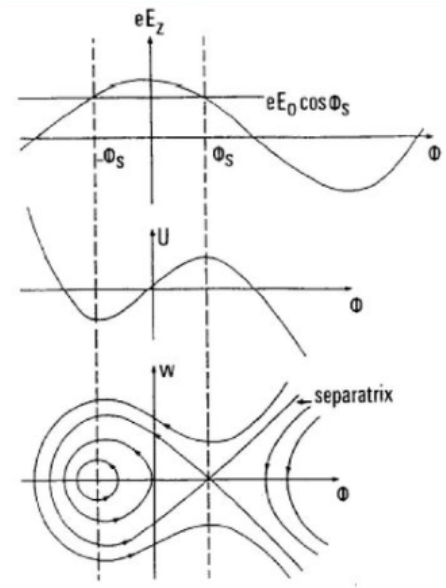
Integrating with the substitution $ds = d\phi/\phi'$;

$$d\phi' \phi' = -AB(\cos(\phi) - \cos(\phi_s))d\phi, \tag{2.149}$$

which can be written,

$$\frac{Aw^2}{2} + B(\sin(\phi) - \phi\cos(\phi_s)) = H_\phi. \tag{2.150}$$

H_ϕ is the constant of integration and the Hamiltonian of the system. The first term of Eqn. 2.150 is the kinetic energy term, and the second is the potential energy term. The longitudinal canonical coordinates w, ϕ describe the evolution of the physical system with respect to s. The observed electric field, potential energy, and phase space are shown in Fig. 2.14.



(a)

Figure 2.14: The observed longitudinal electric field, potential energy, and phase space separatrix of a system describing the longitudinal phase space (separatrix) of particles in an accelerating scheme [35].

The separatrix describes the area of phase space in which particle longitudinal motion is stable, about the synchronous particle. The area of stability is often refered to as the 'bucket'. From this treatment, particles within the 'bucket' will oscillate about the synchronous particle, as the particle energy shifts from being dominated by the kinetic or potential term. The phase range of ϕ defining the region of stable trajectories is a function of ϕ_s. The width is a maximum for $\phi_s = -\pi/2$, as this produces no net acceleration.

When acceleration is taken into account, and the Lorentz factor is not constant, the separatrix changes from a 'fish' to a 'golf club'. The production of the 'golf club' shape is discussed in Chapter 5.

Chapter 3

3 GHz RF Cavity Cell Design

3.1 RF Breakdown and Peak Surface Fields

One of the most important aspects of rf cavity design is the optimisation of the single rf cell. This concerns the geometry of a single cell, such that multiple objectives are simultaneously maximised and minimised. Section 2.2 describes multiple figures of merit (or objectives) for an rf cell, such as the shunt impedance, Z, which measures the rf efficiency of a single cell. Section 2.2 also describes the effect of pulsed heating, the instantaneous heating that occurs in the skin-depth of the conducting wall. The maximum pulsed heating is a constraint that must not exceed ~ 40 K [59], and thus is a function of the peak surface magnetic field, H_{peak}. The maximum surface magnetic flux density is related via $B_{peak} = \mu_0 H_{peak}$, where μ_0 is the vacuum permeability. The value of B_{peak} is calculated using computational microwave solvers, and is a function of the single cell geometry and accelerating gradient. The normalised peak magnetic flux density is used, $\frac{B_{peak}}{E_{acc}}$, as the microwave solver normalises 1 J of energy for a given simulation. Hence, the peak magnetic flux density is a cell objective that requires minimising. Another objective on normal conducting cavities is the peak surface electric field, $\frac{E_{peak}}{E_{acc}}$. This objective relates to the risk of rf breakdown [64].

RF breakdown is a phenomenon in which transmitted rf power drops to low levels over a time frame of the order 100 ns, due to the production of plasma spots on the surface of the cavity [65]. The rf breakdown phenomenon behaves differently in both TW structures to SW structures. In SW structures, incident power may be reflected towards the source which adds risk to the rf circuit. In TW cavities, input power may be absorbed by the breakdown. Multiple events can lead to surface deformation and subsequent detuning of individual cells [51], leading to shorter life times of structures. The breakdown rate quantifies the rate of breakdown events during operation and is given;

$$BDR = \frac{N_{events}}{N_p L_{cav}}. \tag{3.1}$$

Where N_p is the number of rf pulses and L_{cav} is the cavity length. The unit of BDR often

used is breakdowns per pulse per meter (bpp/m). Whilst the cause of rf breakdown is not well understood, surface imperfections have a direct impact on the rate of events. Surface imperfections cause the local electric field to enhance by a factor called the field enhancement factor. Large local electric field can cause electrons in the surface to be emitted from the bulk, called field emission [66]. Electrons can be subsequently accelerated by electromagnetic fields, causing dark current. The Defect model [67] is a theoretical model that suggests breakdown effects are caused by electric fields inducing charge on the structure surface, leading to tensile stress. The BDR is shown to go as;

$$BDR \propto e^{\varepsilon_0 E^2 \Delta V / k_B T}, \tag{3.2}$$

where ε_0 is the dielectric constant of a vacuum, E is the electric field, ΔV is the relaxation volume of the defect, k_B is the Boltzmann constant and T is the temperature. Another study found a power law relationship between the BDR, accelerating gradient, E_{acc}, and pulse length, t_p [68], [69],

$$BDR \propto E_{acc}^{30} t_p^5. \tag{3.3}$$

However, the power law relationships are a function of frequency. Another study calculated for an S-band cavity ($\sim$3 GHz), that BDR goes as $E_{acc}^{16} t_p^3$ [70]. For a pulse length of 5 μs, the S-band power law implies an accelerating gradient limit of $\sim$ 42 MeV/m for a BDR of 1×10^{-6} bpp/m, which is an acceptable limit for medical linacs [71]. Both Eqns. 3.2 and 3.3 imply the peak surface electric field is an objective the requires minimising. Standard peak surface electric field are in the range of 160 - 240 MV/m [51], however the limit of 240 MV/m is highly aggressive. Other studies suggest a more conservative value of 170 MV/m should not be exceeded [70]. In addition, the peak magnetic flux density must also be minimised in order to suppress the probability of breakdown events. The first quantitative limit of the peak surface electric field was proposed by W.D. Kilpatrick and is called the Kilpatrick Criterion [35], however is not often used as an indication of how resistant a cavity is to breakdown in the present day.

Another objective that relates to the BDR is the Modified Poynting Vector (S_c) (first proposed in 2009,[68]). S_c describes the power flow density over the structure walls and is given;

$$S_c = \mathrm{Re}[\vec{S}] + g_c \mathrm{Im}[\vec{S}], \tag{3.4}$$

where $\vec{S}$ is the Poynting vector, and is equal to $E \times H$. The real and imaginary components of S describe the active and reactive power flow along a structure. g_s is a factor that relates the magnitude of the contribution of the reactive power flow to S_c, and is a function (however fairly insensitive) of the local electric field value. The reason for the introduction of S_c as an objective is due to the field emission power flow, a quantity that describes the power that feeds the process of field emission. As the active power flow is in phase with the field emission power flow, the coefficient to $\mathrm{Re}[\vec{S}]$ is 1. As the reactive power flow is $\pi/2$ out of phase (and thus is zero when power flow field emission is maximum), the coefficient g_c is smaller. For

this study, $g_c = 1/6$. As the Modified Poynting vector (MPV) is the cross product of both the electric and magnetic field, the normalised quantity is given $\frac{\sqrt{S_c}}{E_{acc}}$. Experimental work completed suggests there is also a scaling law for the BDR as a function of S_c,

$$BDR \propto S_c^{15} t_p^5. \tag{3.5}$$

A value of S_c = 4-5 W/(μm)^2 is an often used [68] as a constraint for a BDR of 1×10^{-6} bpp/m for a pulse length of 200 ns. Thus for a pulse length of 5 μs, S_c should not exceed the limits $\sim$ 1.4-1.7 W/(μm)^2. Similarly to the MPV, the ratio of power flow to iris circumference, P/C, which has been shown to better predict breakdown rates relative to the peak surface electric field [72]. Whilst there can be limits placed on the peak surface electric, magnetic and MPV fields, breakdown events still occur, due to the statistical nature of rf breakdown.

3.2 Multi-Objective Problems

In this work, there are four objectives used to determine the calibre of a single cell geometry;

$$\text{Objectives} = \left[Z, \frac{B_{peak}}{E_{acc}}, \frac{E_{peak}}{E_{acc}}, \frac{\sqrt{S_c}}{E_{acc}} \right]. \tag{3.6}$$

The shunt impedance requires maximising and the peak surface fields require minimising, and the objectives are conflicting. As there are multiple objectives, multi-objective genetic algorithms were used.

Design problems often have the requirement of optimising multiple conflicting objectives [73]. Multi-objective (MO) problems can be tackled with the use of a weight function, F, which can be solved as a single objective problem,

$$F = \sum_{n=1}^{N} w_n O_n. \tag{3.7}$$

Where w_n are the individual objective weightings, that are often pre-defined scalars, and are often affected by personal influence [74]. Due to compressing multiple objectives into one single objective, information on individual objective values are not known for a given value of F.

Another method to solve MO problems is the use of genetic algorithms [75]. Genetic algorithms solve MO problems by simulating natural selection, and is therefore based on survival of the fittest [76]. Multi-objective genetic algorithms (MOGAs) produce a set of non-dominated, or Pareto optimal, solutions. However, MOGA still have the option to define objective weightings, which can effect the outcome of the converged Pareto front.

A non-dominated solution is a given point in decision space in which one objective can not be improved without another objective becoming less optimal. If a point in

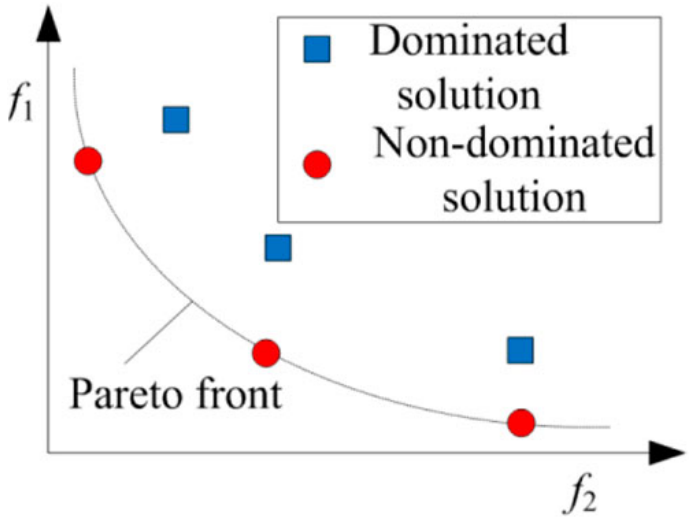


Figure 3.1: Schematic showing two objectives that require maximising, and the region of non-dominated and dominated points [77].

objective space can be improved in one objective and all other objectives remain the same or improve, the point is dominated, see Fig. 3.1. The set of non-dominated solutions produce the Pareto optimal front. The set of inputs is the decision/input space, which maps to the output/objective space. The input space has dimensionality equal to the number of input parameters and the objective space is an N-dimensional space (for N objectives). A strength of MOGAs is that they maintain diversity in Pareto optimal solutions. Figure 3.2 demonstrates the mapping from the decision space to the objective space, and the construction of the Pareto front. MOGAs have been used previously in the accelerating field to aid with multi-objective problems, such as in dc photoinjectors [78] and beam dynamics [79].

3.2.1 Introduction to MOGAs

Multi-objective genetic algorithms take an initial population from the decision space, once the output space is received, the best solutions are selected. Crossover concerns the mixing of individual solution traits, after which, mutation occurs. This set of events creates the next generation offspring, and the process is repeated, a work flow is shown in Fig. 3.3. Choices affecting the selection process, cross-over rate and mutation rate vastly effect the performance of MOGAs. Multiple MOGAs have been developed over the years [81]. In this work, two algorithms are used. Firstly, the fast non-dominated sorting genetic algorithm (NSGA-II). This MOGA has low computational requirements [82] relative to other elitist algorithms and thus allows for faster convergence to the Pareto front. The second MOGA used is the archive-based micro genetic algorithm (AMGA) [83]. The AMGA uses a small population size and

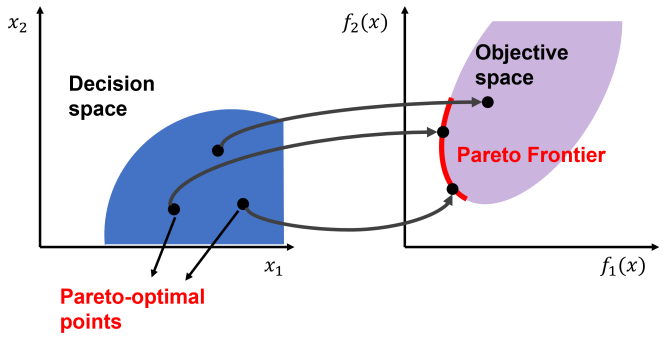


Figure 3.2: Pareto front [80].

keeps an archive of diverse Pareto optimal solutions to maintain its search history.

3.2.2 Applying MOGAs to Normal Conducting Single Cells

In this work, the input space is the rf single cell geometry components. The output space is created from the objectives shown in 3.6, and is therefore four dimensional. Producing the output space from a set of input parameters requires the use of a Microwave solver code, such as CST [56].

Conventionally, single cell geometry is constructed using circles and straight lines, as shown in Fig. 3.4. Often, single cells are optimised by sweeping parameters (e.g. Outer Nose Radius) whilst keeping other parameters constant to produce parameter curves. Multiple parameter curves are then compared to arrive at a final solution. Alternatively, single weight functions can be constructed from multiple objectives and subsequently optimised, [84]. Recently, work was completed that defined the single cell geometry using non-uniform rational basis B-spline (NURBS) [85], allowing for much more complex geometry shapes. In the same work, a single objective function was defined,

$$\text{Optimisation Function} = \frac{Z}{\text{Max}(\frac{E_{peak}}{E_{acc}}, 2)}, \tag{3.8}$$

that was optimised using a differential-evolution algorithm. The results suggested an achievable shunt impedance that was 10% better than the shunt impedance obtained using conventional circles.

The input space in this work is a set NURBS points, shown in Fig. 3.5, for a half cell. The position of each spline point is defined by an x and y coordinate. Spline points are bound by the position of surrounding spline points, so that unrealistic cavity shapes are not created. For example, the coordinates of spline points describing the geometry of the nose-cone are

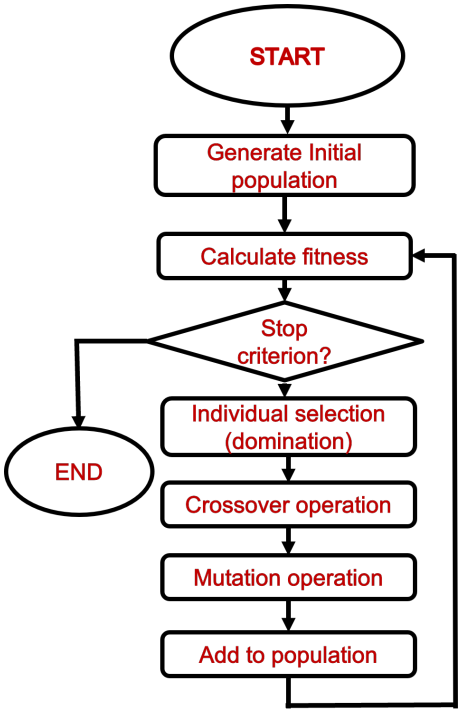


Figure 3.3: MOGA work flow [80].

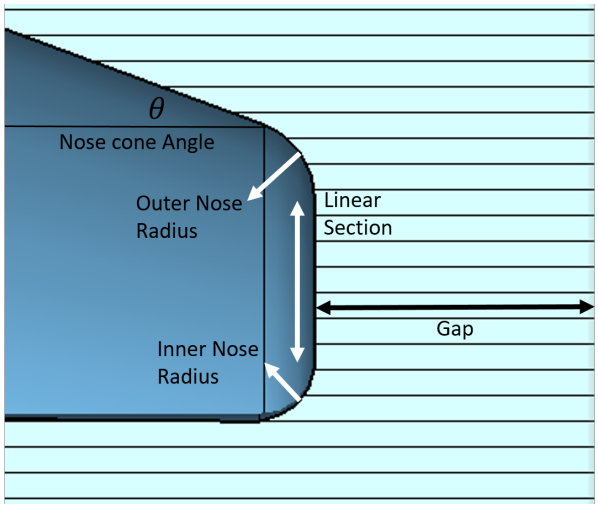


Figure 3.4: Conventional nose cone geometry

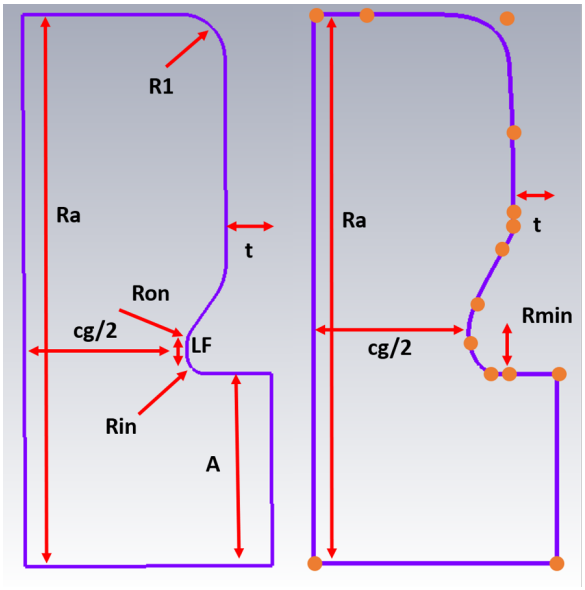


Figure 3.5: Conventional and NURBS single cell geometry.

bound such that the nose-cone is a convex shape. Single cell parameters are still defined at the initialisation of the single cell, such as t, the septum thickness, and $cg/2$, the minimum gap. Within the defined region, the spline can create complex shapes that create the decision space. A realistic spline geometry is shown in Fig. 3.6.

Extensive work was completed by another Lancaster student, Sam Smith, who compared both MOGAs, in addition to constructing a method to handle frequency errors [80, 86]. With the conclusions of this work, an S-band single cell was optimised using NCGA-II, the simulation parameters are shown in Fig. 3.8. The choice of objective weight factors changes the region of convergence for the MOGA, good results were obtained using the weights displayed. A larger weighting for Z was used in order to pull the Pareto front towards solutions with better rf efficiency. $\frac{B_{peak}}{E_{acc}}$ was given a slightly smaller weighting than the other peak surface fields. The magnetic field is lower in the region occupied by the nose cones, and therefore $\frac{B_{peak}}{E_{acc}}$ remains relatively small and is less sensitive to changes to the single cell geometry. $\frac{B_{peak}}{E_{acc}}$ becomes a larger issue when designing side-coupled cells and coupling slots. Both the surface electric and MPV fields are maximum on the nose cone of a single cell, and thus have equal weighting. Figure 3.7 displays the peak surface fields on a single cell.

The electromagnetic simulations were completed in CST Studio Suite [56], and the design process was completed with the Isight software [87]. In order to ensure reliable results, a mesh convergence study was undertaken. This study swept the number of mesh-cells per unit wavelength in CST. The calculated frequency changed as a function of number of mesh-cells until convergence was achieved. The number of mesh-cells used is a compromise between computation time and numerical accuracy. A high mesh density was used in order to ensure

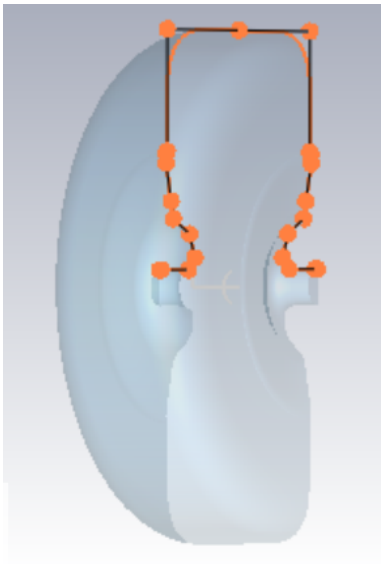
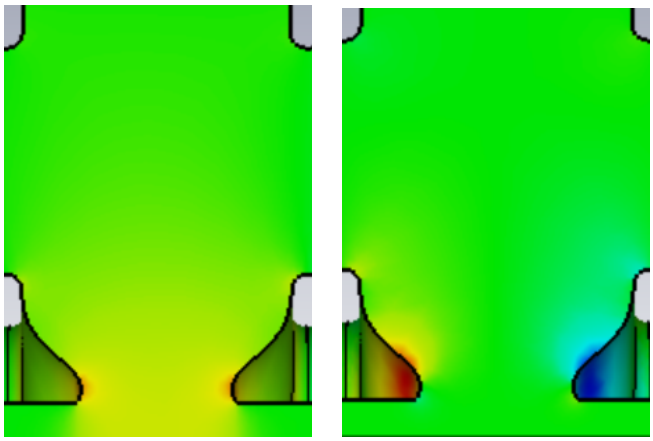
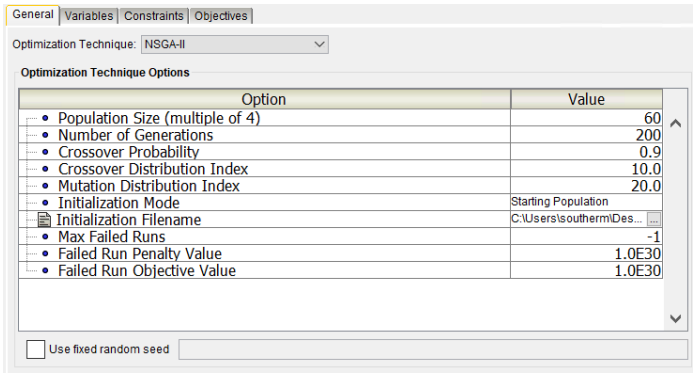


Figure 3.6: Example of a single cell geometry constructed with a NURBS.

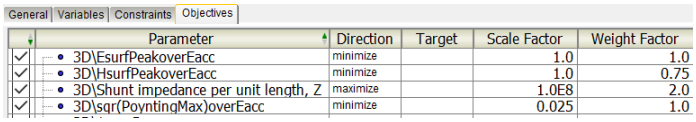


(a) (b)

Figure 3.7: Peak Surface Electric and MPV fields.



(a)



(b)

Figure 3.8: Isight Software showing (a) NCGA-II Initialising parameters, (b) Objective parameters

reliable results.

Once an initial run has been completed, the Pareto optimal points in the results can be used to seed a second simulation, increasing the probability of convergence.

Table 3.1 displays the single cell constant parameters. AVO placed a lower limit on the cavity aperture of 2.5 mm. In order to maximise Z, this minimum aperture value was chosen for the design. The septum thickness was chosen to be 2.5 mm. Septum thickness' of 2 mm have been realised previously for S-band single cells, [88]. Work completed in [51] demonstrates the maximum accelerating gradient is achieved for minimum septum thickness. However, if the septum becomes too thin, thermal effects can cause issues as there is less copper bulk to extract heat. As a result, 2.5 mm was initially chosen as a compromise between both effects.

Table 3.1: 3 GHz Single Cell π-mode constant parameters. Distances in (mm).

frequency, f (GHz)	L_{cell}	septum, t	minimum gap, cg	aperture, a
2.9985	25.7	2.5	$\frac{L_{cell}}{4}$	2.5

3.2.3 Visualisation of Single Cell Objective Space

Once the MOGA has run, a set of Pareto optimal solutions are obtained. In order to select a suitable single cell candidate, solutions must be graphically represented such that the

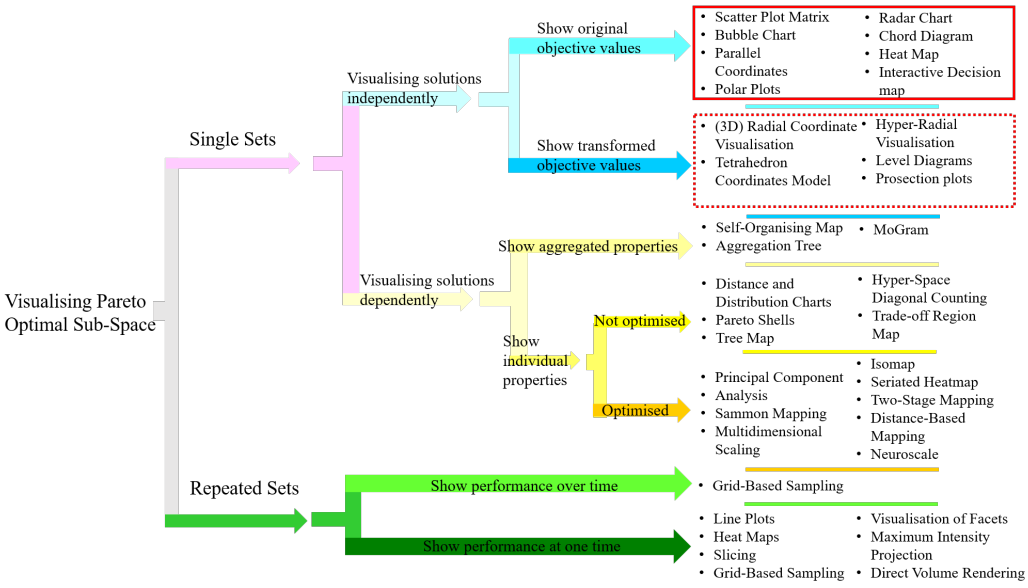


Figure 3.9: Methods to visualise multi-dimensional solutions

relationship between objectives can be understood. For two solutions, a simple scatter plot shows the trade-off between objectives. However as the number of objectives increases, visualising the Pareto front becomes more difficult.

Visualising a multi-dimensional set of solutions can be done in many ways, as described by Fig. 3.9. When designing a single cell, the Pareto space is a single set of points, and each solution must be visualised independently, to keep solutions separate. Each solution can be shown with original objective values (red square in Fig. 3.9) or with transformed objective values (red dashed square in Fig. 3.9). Transformed objective values is not an ideal method, as each objective of a single cell must be determinable from a given visualisation method. This is illustrated with Fig. 3.10, which shows the 3D-RadVis method [89] for the Pareto-optimal solution space. The 3D-RadVis is an example of a transforming objective values. The method places objective nodes on the circumference of a unit circle, at even intervals (e.g. for four objectives, nodes are placed every 90°). Each solution is placed within the unit circle, with a 'spring' connecting the solution to each objective. The spring constant strength is proportional to the objective value. Thus, solutions with a high objective value will be dragged closer to that objective node. Figure 3.10 shows three objective nodes, for each peak surface field, and a colour dimension showing Z.

The visualisation method shows that solutions with low $\frac{B_{pk}}{E_{acc}}$ have high Z. However, the individual values of the peak surface fields are not determinable. The information provided simply shows how 'good' or 'bad' a solution is in a particular objective. As a result, this visualisation method is not appropriate, whilst it does provide insight into the

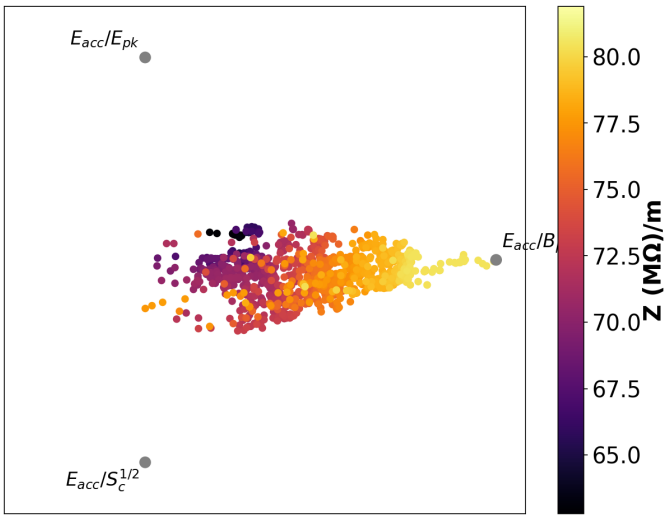
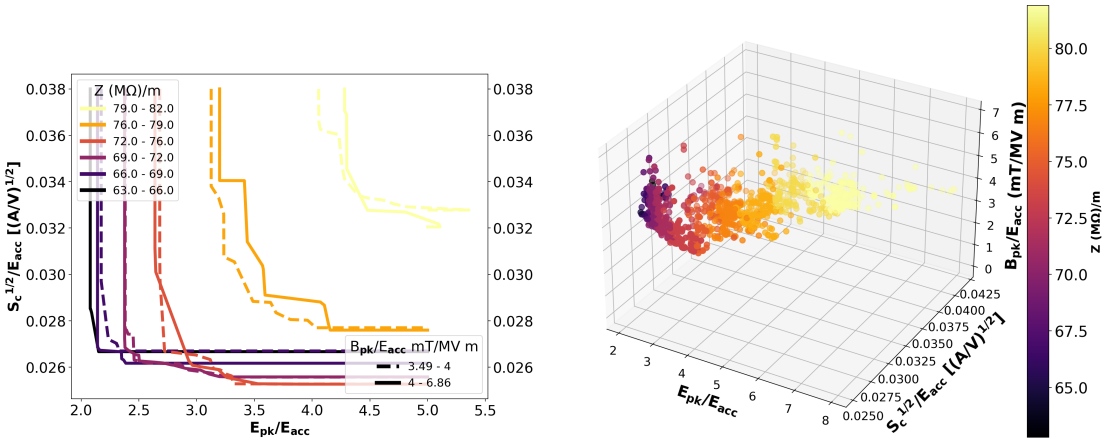


Figure 3.10: 3D-RadVis method for Pareto-optimal solutions.

relationship between objectives. As a result, further visualisation methods must show original objective values in order to aid the selection process. Figure 3.11 (a) shows a decision map of Pareto-optimal solutions. Decision maps bin solutions into discrete sets and displays the information as 2D scatter plots with a color and line style dimension. In Fig. 3.11 (a) the binned objectives are $\frac{B_{pk}}{E_{acc}}$ and Z. This method allows the comparison between multiple objectives. For example, solutions with maximum Z also are maximum in $\frac{E_{pk}}{E_{acc}}$ and $\frac{S_c^{1/2}}{E_{acc}}$. However, decision maps do not allow for comparisons between objective values within a given Pareto front, as only the bin range is shown. As a result, solutions in the same bin can have differences in Z up to 3 MΩ/m and $\frac{B_{pk}}{E_{acc}} \sim 1.5$ mT/MV/m.

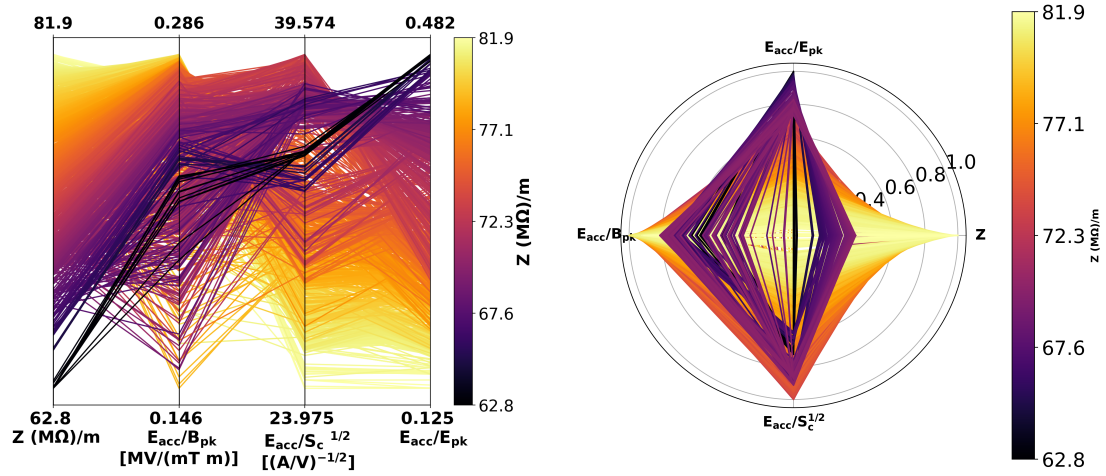
Figure 3.11 (b) displays a 3D bubble plot, with three spatial dimensions for each peak surface field and a color dimension showing Z. The bubble plot shows all solutions as a singular point on the graph, which is a benefit over previously discussed visualisation methods. However, the plot is portrayed on a 2D plane, and therefore the actual objective value is difficult to ascertain due to parallax error.

Figure 3.12 displays the (a) Parallel Coordinate Plot (PCP) and the (b) Radar Chart (RC). Note, the objective nodes show the reciprocal of peak surface field, so that all objectives require maximising. Both method plot each solution as a line that crosses an objective axis at the objective value. Both methods are essentially the same, however the PCP separates objective nodes horizontally whilst the RC separates objective nodes with different orientation. The PCP has benefits over the RC, as the physical space on which solutions are displayed is larger than the RC plot, for the same number of pixels. This means there is less white-space for a PCP plot, relative to the RC. For many Pareto optimal solutions, PCP and RC become very populated, and it is difficult to accurately compare many solutions.



(a) (b)

Figure 3.11: Original objective preserving visualisation methods. (a) displays the decision map and (b) displays a bubble plot.

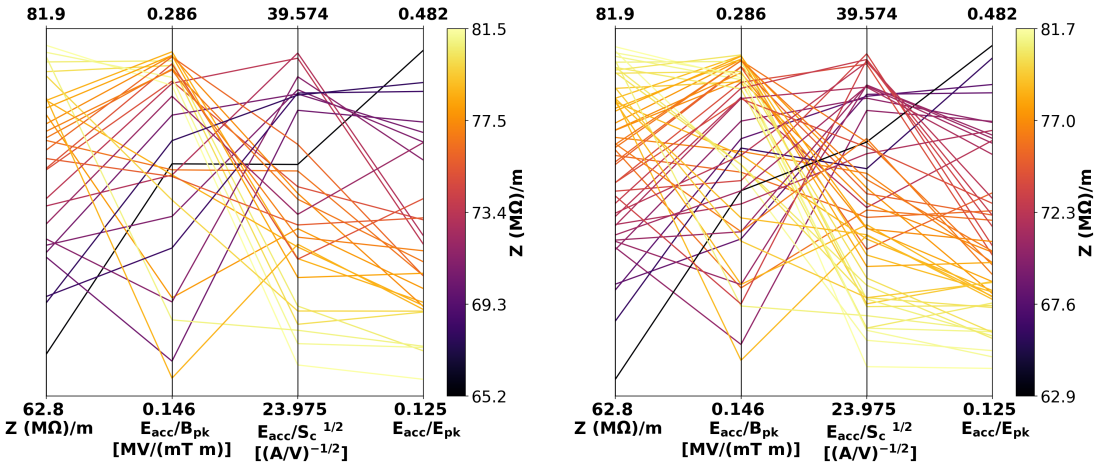


(a) (b)

Figure 3.12: Original objective preserving visualisation methods. (a) displays the Parallel Coordinate Plot and (b) displays a Radar Plot.

In addition, the order in which solutions are plotted affects how a PCP/RC appear. For example, Fig. 3.12 plots solutions in an order that decreases Z. Thus, high Z solutions become hidden. RC plots have benefits when considering the general trade-off between objectives. From Fig. 3.12(b), solutions from $\frac{E_{acc}}{E_{pk}}$ to Z create a concave shape, implying solutions can not be high in both objectives simultaneously. Solutions from $\frac{E_{acc}}{S_c^{1/2}}$ to $\frac{E_{acc}}{B_{pk}}$ create a straight line, suggesting solutions can be maximum in both objectives simultaneously.

In order to reduce the noise from the PCP plot, solutions were clustered into groups of similar objective values. Clusters are then plotted as a PCP with each objective value being the mean value of the objective within the cluster. Figure. 3.13 displays PCP for 30 and 50 clusters. The plots become easier to interpret, and clusters can be compared with each other. As the number of clusters decreases, the relative error of the cluster increases, as individual solutions must fit into fewer clusters.



(a) (b)

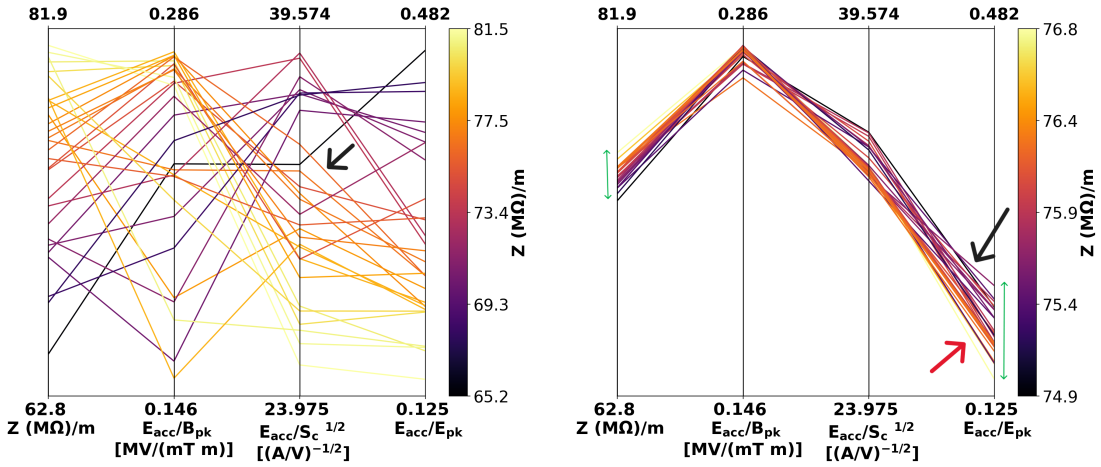
Figure 3.13: PCP showing 30 (a) and 50 (b) clustered solutions.

Clustering of solutions is similar to creating bins, as was shown in the decision map. However, the decision map displayed previously had six bins for the entire range of Z. For 30 clusters, the range of objective values within a cluster is far less. In reality, for an entire Z range of 62.8 - 81.9 MΩ/m, 30 clusters results in $\sim$ 0.6 MΩ/m maximum deviation in Z. Clustering solutions has provided a method to produce a more readable PCP, in which a solution can be selected whilst observing the entire Pareto-optimal space. A selection process is described in the following section.

3.2.4 Single Cell Candidate Selection

The previous section described many visualisation methods to observe the entire Pareto optimal space. Visualisation methods must keep original objective values, so that each solution can have determinable objectives. Clustering similar solutions together provides a method to reduce the noise of a PCP whilst preserving the core findings of the MOGA. The clustering was achieved using a K-means clustering algorithm [90].

The next step requires selecting a candidate single cell geometry. From Fig. 3.13 (a), solutions that perform maximally in one objective tend to perform as a minimum in other objectives. When selecting a final solution, the choice is influenced by the type of cavity being designed, e.g. collider accelerator or a medical linac. In addition, human bias will inevitably play a roll in the selection. Figure 3.14(a) shows the selected cluster to be expanded. The reason for this cluster selection are as follows. Firstly, the cluster has high Z whilst simultaneously having high $\frac{E_{acc}}{S_c^{1/2}}$. All solutions with higher values of $\frac{E_{acc}}{S_c^{1/2}}$ have lower Z and $\frac{E_{acc}}{B_{pk}}$. Whilst the solution is just under the average value for $\frac{E_{acc}}{E_{pk}}$, all other objectives are comfortably over the respective average objective value. Figure 3.14(b) shows the chosen cluster expanded into individual solutions. The solution depicted with the black arrow was selected. This choice was made as the relative increase in $\frac{E_{acc}}{E_{pk}}$ was larger than the relative increase in Z, and Z is already greater than the mean value, for all solutions.



(a) (b)

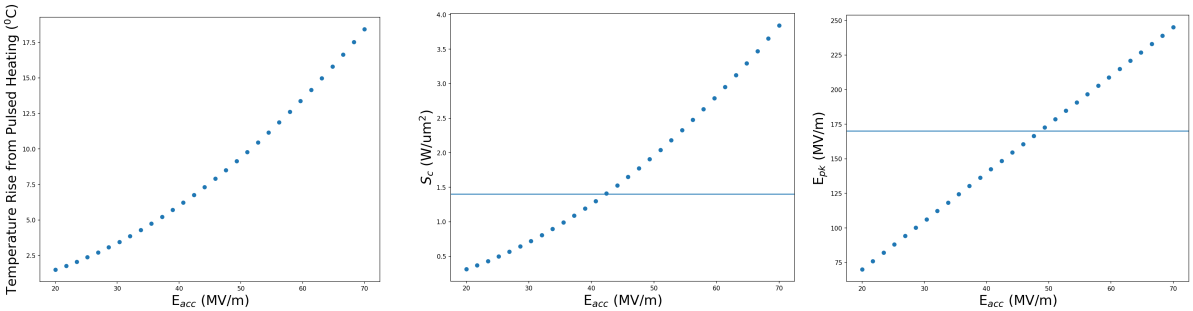
Figure 3.14: PCP showing selected cluster (a) and expanded cluster (b) into individual solutions.

Table 3.2 displays the objective values of the chosen single cell. Figure 3.15 displays the pulsed heating temperature rise and peak surface field limit for the chosen single cell design. The conservative limits (for a pulse length of 5 μs and RR of 200 Hz) of $S_c \sim$

1.4 W/(μm)^2 and $E_{pk} \sim 170$ MV/m are shown. The gradient is limited by the MPV and is approximately 40 MeV/m.

Table 3.2: Chosen single cell parameters for $\pi/2$ mode structure.

Z (MΩ/m)	75.6
B_{pk}/E_{acc} (mT/MV/m)	3.5
$\sqrt{S_c}/E_{acc}$ (A/V)$^{1/2}$	0.028
E_{pk}/E_{acc}	3.5
Septum thickness (mm)	2.5
Cell length (mm)	25.7
Aperture radius (mm)	2.5
Frequency (GHz)	2.9985



(a) (b) (c)

Figure 3.15: Pulsed heating temperature rise and peak field limits as a function of accelerating gradient for chosen single cell design.

3.3 Side-Coupled Cell Design

Recall from Section 2.3.1, a coupled oscillator system can have both electric and magnetic shunt admittance elements, producing two separated pass bands, that generally have a stop band between them. The term confluence describes the condition when the two pass bands meet, creating a single continuous pass-band. Section 2.3.2 described, using Perturbation theory, the effect of individual oscillator frequency errors on the normal modes of oscillation. Frequency errors effect the $\pi/2$ normal mode magnitudes only in second order, showing the stability of the mode. Operating a resonant structure in a $\pi/2$ mode and moving the unexcited cells off-axis, creating a side-coupled structure, allows the high shunt impedance of a π mode with the stability of the $\pi/2$ mode. The side-coupled cells must oscillate at the

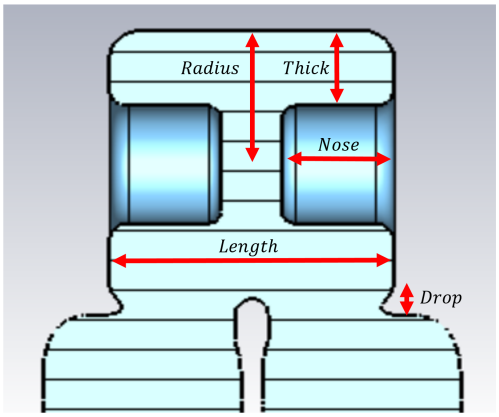


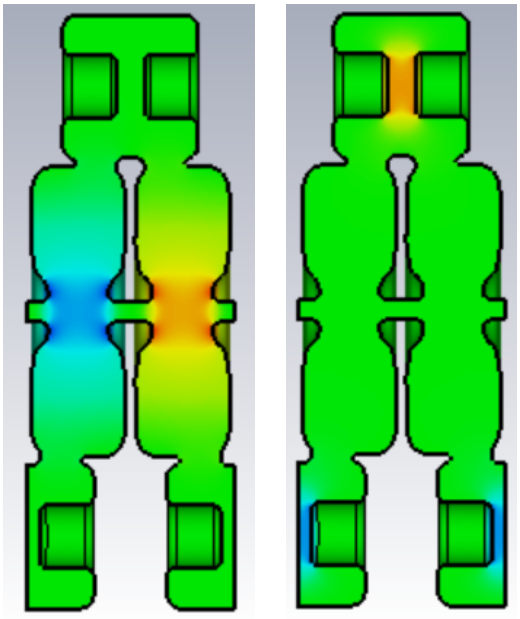
Figure 3.16: Geometry of side-coupled cell.

correct frequency such that confluence is reached, whilst maintaining the required inter-cell coupling constant, k. Side-coupled cells must also ensure low surface magnetic fields, and maximise the shunt impedance.

The first step required designing of the coupled cell to oscillate at the correct frequency. Figure 3.16 shows the side-coupled cell geometry. As the side-coupled cell must be the same frequency of the accelerating cell, one would expect a side-coupled cell to be a similar radius to the accelerating cell. This is not ideal, as it increases the transverse size of the cavity, requiring more space and copper material. In order to reduce the size of the side-coupled cell, the capacitance can be increased (recall $\omega_0 = \frac{1}{\sqrt{LC}}$). The nose parameter is an important parameter as it strongly effects the capacitance of the side-coupled cell. However, as the nose gap decreases, the capacitance increases, and the electric field increases between the noses. Hence, small machining errors in the side-coupled cell nose may produce large changes to the resonant frequency, from Slater's perturbation theory (discussed in Chapter 2). The side-coupled cell must be optimised such that the two $\pi/2$ modes are excited at the same frequency, and at the design frequency of 2.9985 GHz. The band gap is given by the percentage difference between the two mode frequencies.

The optimisation was completed in CST microwave solver, using the Trust region framework optimiser. The simulated model consists of two accelerating cells, and two side-coupled cells. However one side-coupled cell is split into two half cells. This allows a boundary condition to force a phase advance of 2π, ensuring periodicity. As a result, both $\pi/2$ modes can be calculated per simulation. The coupling constant is a quantity that changes as a function of the expected cavity length (described in a latter Chapter 6). The optimised side-coupled cell geometry is shown in Fig. 3.17, in addition to the longitudinal electric fields of both $\pi/2$ modes.

The next stage of the design considered the geometry of the coupling slot between the side-



(a) (b)

Figure 3.17: $\pi/2$ mode excited in both accelerating and side-coupled cell.

coupled cell and accelerating cell, as the dimensions effect the inter-cell coupling constant, shunt impedance and B_{pk}/E_{acc}. The slot width refers to the shorter edge of the slot, whilst the slot length refers to the longer edge of the slot, see Fig. 3.18. As the cavity length was not known, the value of the required k is not a singular value to optimise towards. However, different slot geometries provide different relationships between the shunt impedance, k, and B_{pk}/E_{acc}. The slot design therefore considers the relationships between the objectives. Each slot must produce confluence for both $\pi/2$ modes. The slot geometries are shown in Fig. 3.18. All slot geometries have two degrees of freedom, the slot width and the the slot length. The slot width is a bounded parameter, as the slot can not encroaches the edge of the accelerating cell. The slot length is limited to the diameter of the coupling cell, see Fig. 3.18. Recall the required coupling is calculated in Section 2.3.1. The slot widths and lengths are independently swept for each geometry, the relationships between k, Z and $\frac{B_{pk}}{E_{acc}}$ are shown in Fig. 3.19. The results suggest that all slot geometries behave similarly, and that larger k results in lower shunt impedance. For larger coupling, the peak surface magnetic field also increases. The shunt impedance and coupling constant are expected to behave in this way. As the slot area increases, the larger the inductance due to the magnetic field, increasing the inter-cell coupling. As the magnetic field perturbation increases, the Ohmic power losses increase and the shunt impedance drops. In addition, highly perturbed magnetic field increases the peak surface magnetic field. From Fig. 3.19, the rectangle slot provides a larger shunt impedance for the same value of k. The elliptical and rectangular slots seem

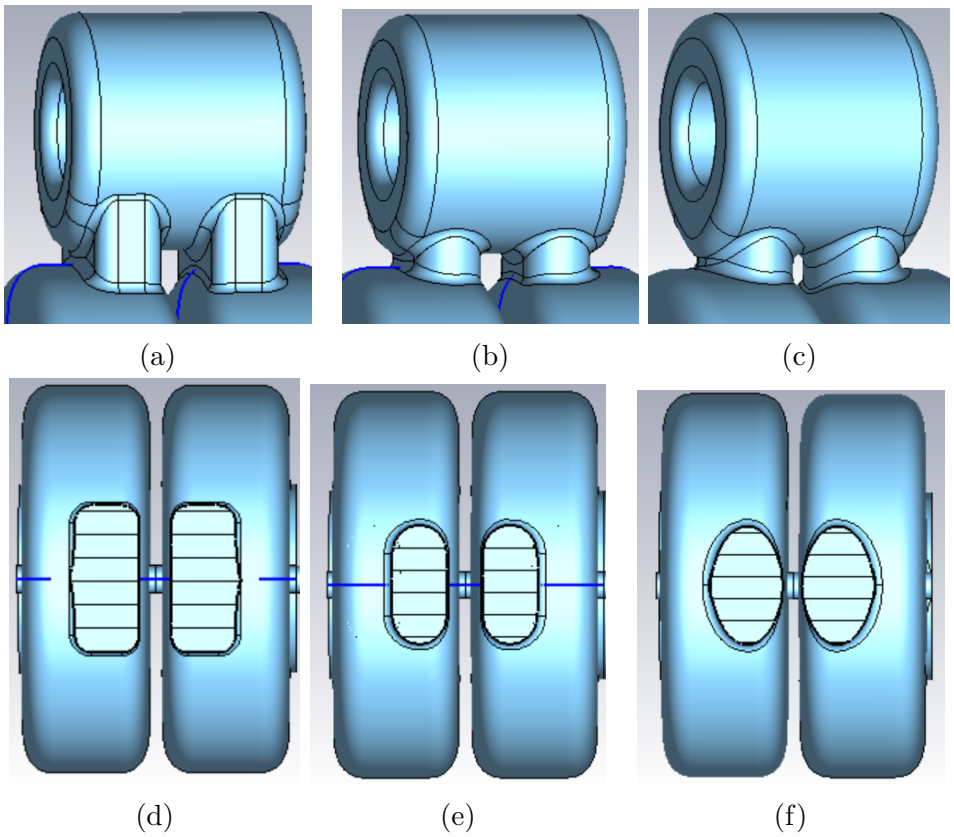
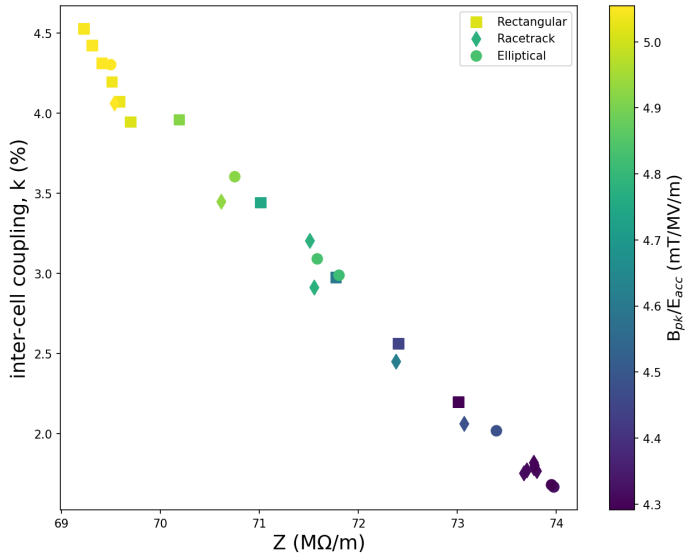


Figure 3.18: Side-Coupled Cell slot geometries, including the birds-eye view of the slot.

to perform similarly well, however the peak surface magnetic field is slightly larger for the elliptical slot. As the rectangular slot is easier to machine than an elliptical join it was chosen as the final slot geometry.

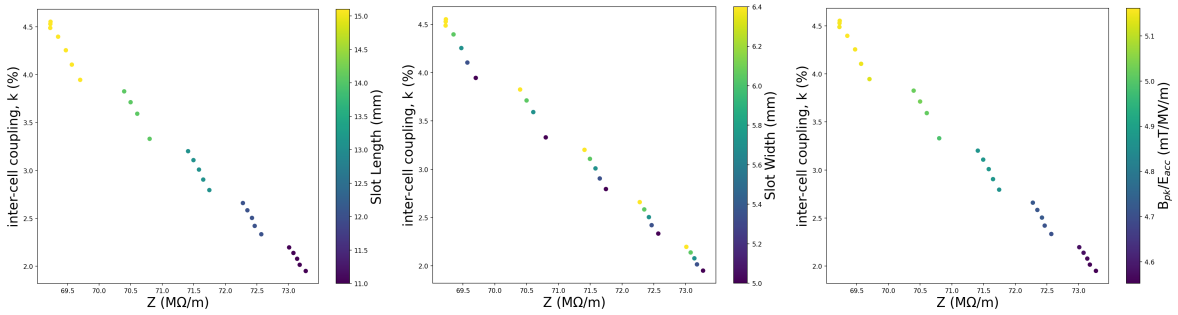


(a)

Figure 3.19: Inter-cell coupling as a function of shunt impedance and surface peak magnetic field for three different coupling slot geometries.

The rectangular slot is further explored by completing nested parameter sweeps of the slot width and length. the results are shown in Fig. 3.20. Figure 3.20 (a) displays the inter-cell coupling as a function of the shunt impedance and slot length whilst Fig. 3.20 (b) shows the slot width with the colour dimension. For each slot length, the slot width is swept between the achievable bounds. It can be seen that for a given slot width or length, sweeping the other parameter produces a straight line of relatively fixed gradient. The magnitude of the gradient when sweeping the slot width is larger than the gradient when sweeping the slot length. Thus, a change in the slot width produces a greater increase in k for a given reduction in Z than changing the slot length. In addition, from Fig. 3.20 (c), slots of the same length and different width have very similar values for $\frac{B_{pk}}{E_{acc}}$. When the slot length is increased, the value of $\frac{B_{pk}}{E_{acc}}$ jumps to an increased value. This is an intuitive result, as the magnetic field peaks on the coupling slot, increasing the slot width will not change the value of the peak magnetic field further. However, increasing the slot length will change the location of the slot edge with respect to the magnetic field, producing a different value of $\frac{B_{pk}}{E_{acc}}$. The above discussion implies the optimum slot geometry occurs when the width is at a maximum achievable value, with any additional coupling being achieved by increasing the slot length.

The final step of the slot design requires ensuring confluence is achieved. When the



(a) (b) (c)

Figure 3.20: Inter-cell coupling as a function of shunt impedance, slot dimension and surface peak magnetic field for the rectangular slot geometry.

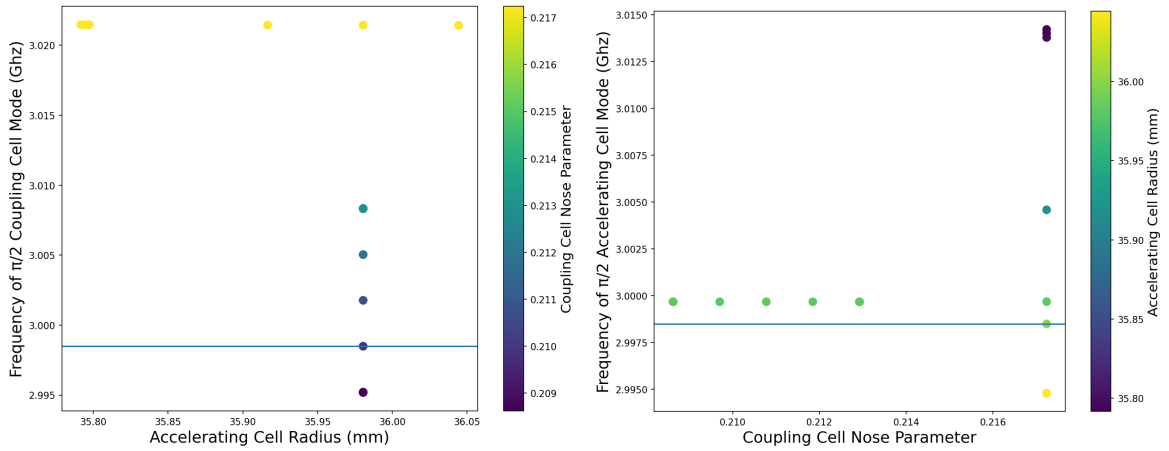
nested parameter sweep was completed, the frequency of both $\pi/2$ modes shift. In order to reach confluence for a given choice of the slot width and length, changes are made to the accelerating cell radius, and the coupling cell nose. Scaling the accelerating cell simply moves the coupling cell radially outward or inward, leaving the fields within unchanged. This results in no change to resonant frequency, from Slater's Perturbation theory. Similarly, changing the nose parameter of coupling cell will not effect the electric or magnetic field lines in the $\pi/2$ accelerating cell mode, allowing for tuning of the the coupling cell $\pi/2$ mode whilst keeping the $\pi/2$ accelerating cell mode frequency constant. This allows each mode to be brought back onto the design frequency independently. Figure 3.21 shows the effect of changing the nose parameter and accelerating cell parameter on the respective $\pi/2$ resonant frequencies. Further, the value of k, Z and $\frac{B_{pk}}{E_{acc}}$ does not change when confluence is again obtained. Thus, sweeping the slot parameters to explore the effect on k, Z and $\frac{B_{pk}}{E_{acc}}$, whilst ignoring the frequency values is a valid approach.

Once the side-coupled cell has been optimised for confluence, it is completed. If the design cavity is expected to be shorter or longer, the slot geometry can be altered as discussed previously to meet the required coupling constant, ensuring the width is maximised before increasing the length of the slot. Confluence can then be reached again efficiently by changing the coupling cell nose parameter and accelerating cell radius.

3.4 Traveling Wave Cell Coupling Slot Design

Results from transverse beam dynamics studies, which are presented in Chapter 4, show that a TW structure could also be a viable option for the cavity design, as the cavity lengths could reach lengths $\sim$ 1 m. In this region, TW structures generally dominate SW structures.

The first stage required designing TW single cells with phase advances of $2\pi/3$ and $4\pi/5$, respectively. As the TW single cells have different cell lengths, it was not initially known if



(a) (b)

Figure 3.21: Resonant frequency of both $\pi/2$ modes as a function of cell nose parameter and accelerating cell radius. Design frequency of 2.9985 GHz is shown.

single cell optimisation was required. As the NURBS is a function of cell length (via cg), the solution chosen for the SW cell is also a strong candidate for both TW single cells. In order to ensure a non-dominated solution, the NCGA-II was seeded with the SW solution and ran on the $4\pi/5$ single cell. A TW single cell is simulated differently to a SW cell. The cell is split into two half cells, with the aperture in the centre, as shown in Fig. 3.22. This is required so that periodic boundary conditions can enforce the correct phase advance per cell. Table 3.3 shows the chosen single cell objectives for the TW single cell candidates.

Table 3.3: Objective values of chosen candidate solutions for $4\pi/5$ and $2\pi/3$ mode TW single cells.

Phase Advance	Z (MΩ/m)	B_{pk}/E_{acc} (mT/MV/m)	$\sqrt{S_c}/E_{acc}$ $(\mathrm{A/V})^{1/2}$	E_{pk}/E_{acc}
$2\pi/3$	68	2	0.025	3.4
$4\pi/5$	78	2	0.028	3.9

The coupling slot design is an important aspect of TW cell design. The objectives are the achievable group velocity, Z, and B_{pk}/E_{acc}, as the magnetic field peaks on the coupling slots. The shape of the coupling slot can have an effect on the objectives described. Previous work [51] has compared multiple slot designs, with different slot shapes being slightly better for different phase advances per cell. The circular slot was found to work well for many cell types, and has benefits over other shapes due to the ease of machining. For this reason, the circular slot was used as the slot shape for the TW single cell. The remaining design choices are the number of slots, and the slot position.

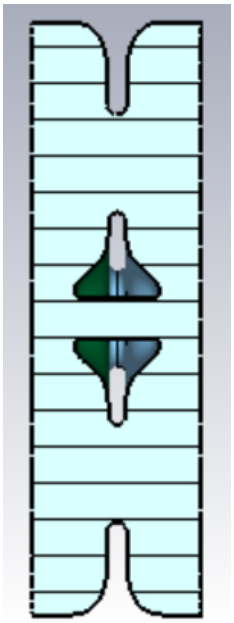
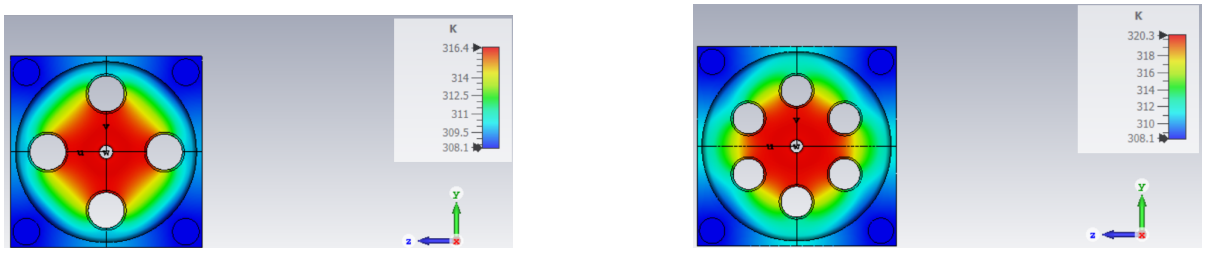


Figure 3.22: TW cell design with coupling slots.

Number of Circular Coupling Slot Study

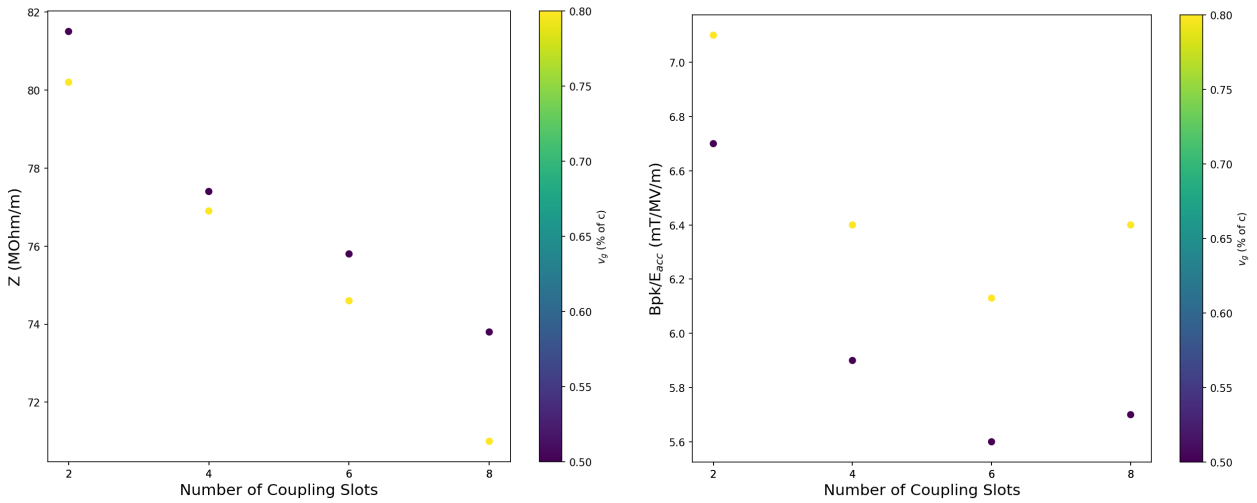
The number of coupling slots is a relatively straightforward design parameter. In this work, the number of slots was largely governed by the highest achievable shunt impedance for a given group velocity, and by the temperature rise in the cell wall due to Ohmic heating. As the number of coupling slots increases, the volume of bulk copper in the cell wall is reduced, and the temperature rise due to Ohmic heating is larger. This is because heat must diffuse through the copper wall outwards, towards cooling channels that are placed off-axis.

The group velocity has a power law relationship with respect to coupling slot radius, as



(a) (b)

Figure 3.23: Steady state temperature rise of TW single cell with four cooling channels for both four and six coupling slots. Both cells have $v_g = 1\%$ of c.



(a) (b)

Figure 3.24: Z (a) and $\frac{B_{peak}}{E_{acc}}$ (b) as a function of number of slots for two different group velocities.

shown in Fig. 3.25. One might expect the temperature rise in the copper wall will be related to the volume of wall, which will decrease with increasing number of slots and increasing slot radius. Figure 3.23 shows the steady state temperature rise from for a single cell with four cooling channels, each cooling channel was operated at a flow rate of 2 litres/min at a temperature of 308 K. Both single cells have been optimised for a group velocity of 1% of c. The temperature rise from four slots is $\sim$ 8 K, whilst the rise is 12 K for six slots. One reason for this may be the increased spacing in the copper bulk for heat to diffuse in the four slot single cell. The shunt impedance as a function of number of slots is shown in Fig. 3.24(a). The shunt impedance values are shown for two different values of group velocity, 0.5 and 0.8 % of c, values that are likely to be used for the conceptual linac design. Low values of the group velocity were simulated because a main design criteria for the design is high rf power efficiency, which is achieved by running at lower group velocities. Results show that fewer slots increase the shunt impedance. This may be because increasing the number of slots increases power losses, as the magnetic field is perturbed in multiple locations. $\frac{B_{peak}}{E_{acc}}$ as a function of number of slots is shown in Fig. 3.24(b). The general trend is that B_{peak}/E_{acc} increases for fewer slots, with the exception of 8 slots. One reason for the increased peak magnetic field for 8 slots may be due to the increased number of perturbations to the field, causing localised concentration of the field.

Four coupling slots was chosen for the TW cell design. This choice was motivated as four slots maintains higher shunt impedance than six slots, and the value of $\frac{B_{peak}}{E_{acc}}$ was not much greater than for six slots. In addition, a six slot single cell produced higher average

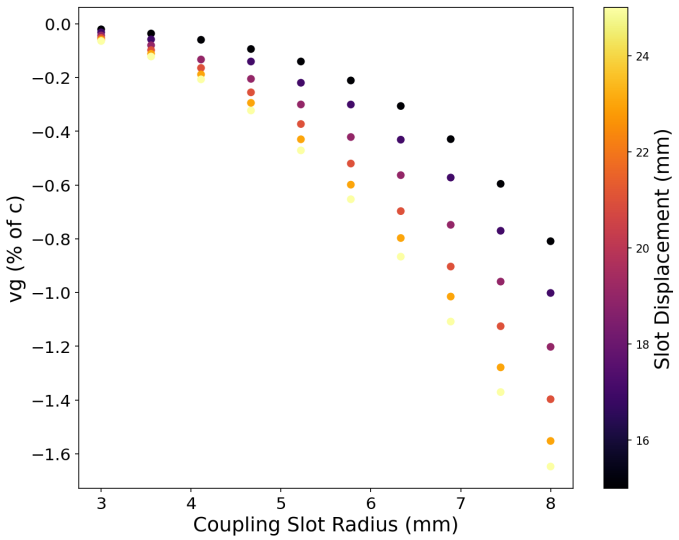


Figure 3.25: Group velocity of a TW cell as a function of slot displacement and radius.

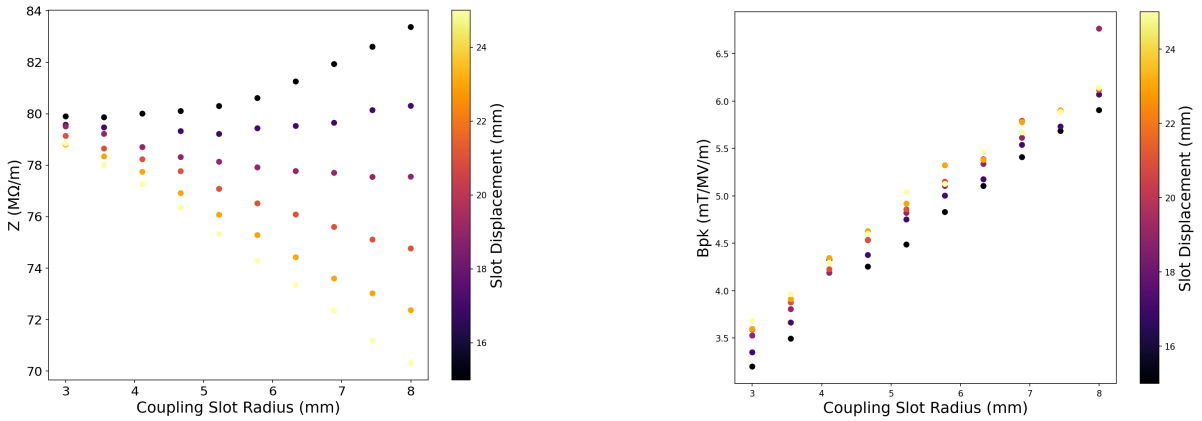
temperatures than four slots. Two slots was omitted due to the large $\frac{B_{peak}}{E_{acc}}$, which was around 7 mT/MV/m. As previously discussed, increased peak surface fields contribute to rf breakdown, even when below a supposed safe limit. Whilst four slots still provided a good shunt impedance, the increased shunt impedance achieved with two slots was not worth the vast increase in B_{peak}/E_{acc}.

Circular Coupling Slot Position Study

The position of the coupling slot is an important design choice as it effect the shunt impedance for a given group velocity, in addition to B_{peak}/E_{acc}. Limiting the peak magnetic field whilst maximising the shunt impedance for a given group velocity are the main objectives considered in the study.

The position of the slot relative to the centre of the cell is called the displacement, and is denoted R. The radius of a given coupling slot is r. The displacement was swept from 15 mm to 25 mm, and for each displacement the slot radius was swept from 3 mm to 8 mm. These limits were bound, as the minimum displacement of 15 mm could not realise a slot or radius larger than 8 mm as it would start to intrude on the nose cone.

Figure 3.25 displays the group velocity as a function of slot displacement and radius. Results imply that for the same slot radius, larger displacements produce higher group velocity. For a slot displacement of 15 mm, the maximum achievable group velocity is 0.8 % of c. This value roughly doubles for a slot displacement of 25 mm. As a principal design criteria is high rf efficiency, coupling slots that allow group velocities over $\sim$ 1.5 % of c were not ideal, due to the reduction in rf efficiency (increased power for a given accelerating



(a) (b)

Figure 3.26: Z and $\frac{B_{peak}}{E_{acc}}$ as a function of slot displacement and radius for a TW cell.

gradient). Figure 3.26 displays Z and $\frac{B_{peak}}{E_{acc}}$ as a function of slot displacement and radius. The relationship between shunt impedance and slot displacement and radius is interesting. For displacements less than ~ 19, the shunt impedance reduces with decreasing slot radius. This result is not expected. Generally larger slots have increased losses and therefore smaller Z, as the accelerating voltage is unaffected. One possible reason for this result is the following. For small slot displacement, as the slot radius increases, the more the nose-cone is encroached upon, and therefore losses due to the magnetic field on the nose cone are reduced, as less nose cone exists. One the displacement is greater than the nose cone height, this effect is no longer possible, and the shunt impedance decreases with increasing radius, as expected. For increasing displacement, the relationship between Z and slot radius becomes linear, and the slope decreases (as the slope is negative) as a faster rate with increasing slot displacement. As a result, larger slot displacement become much less efficient than slots with smaller displacement. The reason for this is because the magnetic field increases with radius, thus higher displacement slots occupy higher magnetic field regions, and losses are increased. As for the peak magnetic field, increasing slot radius increases the peak field value fairly linearly. TW cells with smaller slot displacement generally had smaller $\frac{B_{peak}}{E_{acc}}$, as the magnetic field is smaller in these regions. The coupling slot radius has a larger effect on the peak surface magnetic field, and is relatively linear.

Figure 3.27 shows the achievable group velocity as a function of shunt impedance and slot displacement for a TW cell. The figure suggests, for a given group velocity, smaller slot displacement provides the largest shunt impedance. From the previous discussions, all slot displacements have benefits. Whilst larger slot displacements have lower shunt impedance, the group velocity is largest for a given coupling slot radius. As a result, should $B_{peak}E_{acc}$ be the principal objective, the value can be minimised by having smaller slot radii, and higher group velocities can be achieved by having larger slot displacement. If the shunt impedance is

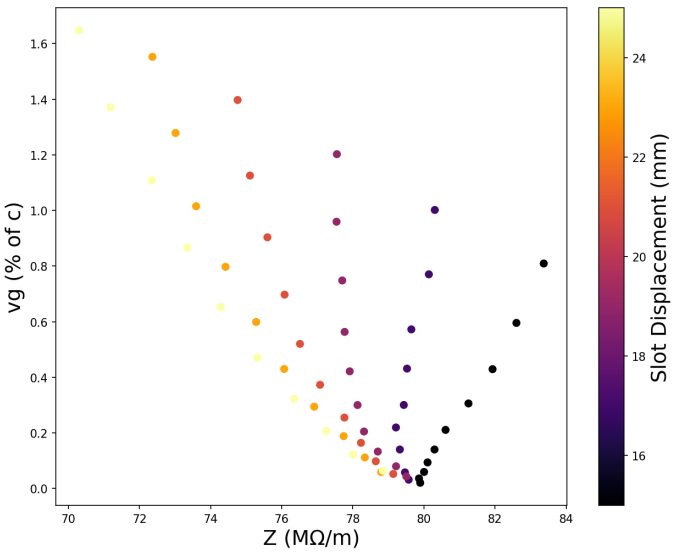


Figure 3.27: Group velocity of a TW cell as a function of shunt impedance and slot displacement.

more important in a given design, smaller slot displacements provide larger shunt impedance, with the trade-off being larger B_{peak}/E_{acc} for a given group velocity. In order to maintain higher shunt impedance and have the ability to realise larger group velocity, if required, a displacement of 19 mm was chosen.

Single Cells of Different Length

Sometimes, the exact length of an rf cell may not be know, as the number of cells has not been established, therefore the synchronous beta may be slightly different to the simulated value. The amount by which objective values change due to a small change in cell length are shown in Table 3.4. The synchronous beta was swept from 0.51 to 0.52 in multiple steps, the objectives were calculated as each step and the maximum deviation from the initial objective value was taken. A change in β_s of 0.01 is equivalent to a change in cell length of ~ 0.5 mm. This relates to a change in structure output energy (150 MeV input energy) of ~ 155 MeV to ~ 170 MeV. For a gradient of 35 MeV/m (with transit time factor of 0.5, this is a valid approximation, see Chapter 5), this is equivalent to a change in cavity length of ~ 0.3 m to 1.2 m, or a change in number of cells per structure of 11 to 47. The results suggest that small perturbations in cell length change the objectives by small amounts, specifically in the peak field values. The largest change is the change in Z, however a change of 1.7 (%) is still a small deviation. As a result, single cell solutions for a particular β_s are also good solutions for single cells of slightly different β_s. This allows a large time saving, as single cells do not need to be re-optimised for small changes in β_s. However, for protons at lower energy (or

indeed electrons), changes in cavity length impact changes in β_s by larger amounts than for protons at 150 MeV.

Table 3.4: The change in single cell objectives due to a change in cell length.

$\Delta\ \beta_s$	— $\Delta\ B_{pk}/E_{acc}$ (%)	$\Delta\ E_{pk}/E_{acc}$ (%)	$\Delta\ \sqrt{S_c}/E_{acc}$ (%)	Δ Z (%)
0.01	0.25	0.9	1	1.7

3.5 Conclusion of Single Cell Design

This chapter discussed the design process of a single cell, and the side-coupled cells and coupling slots for SW and TW cells, respectively. The single cell design is optimised for maximum shunt impedance whilst minimising the peak surface fields. The optimisation was completed using MOGA with the cell geometry being described with NURBS to increase the complexity of possible geometries and push the objectives to limits that are not achievable with conventional nose cone design techniques. The side-coupled cells are designed to achieve maximum inter-cell coupling for a given shunt impedance whilst keeping the peak surface magnetic field small. In order to keep the coupling constant a fluid variable (due to the unknown length of the cavity) the relationship between coupling and shunt impedance was ascertained. The TW coupling slot design was motivated with similar objectives - maximising the group velocity whilst keeping the shunt impedance high and peak surface fields low. The next step will discuss the beam dynamics of the linac system, and explore any constraint that may be placed upon cavity length and aid the choice of a SW or TW cavity.

Chapter 4

Transverse Beam Dynamics Studies

4.1 Twiss Parameters with Acceleration

When designing an rf linear accelerator, the transverse and longitudinal beam dynamics of the beam are important aspects. Transversely, if the beam size is too large at any point within the cavity, particles will collide with the cavity walls and less current is accelerated to the end target. For proton radiotherapy, the beam current is small ($\sim$1 nA) and ideally 100% transmission is realised. As a result, the beam size must constrained at particular points along the linac in order reduce the risk of beam loss. In this section, two potential focusing schemes are discussed, the FODO-like scheme, and the minimum aperture scheme (MAS). The FODO-like scheme places quadrupoles of alternating polarity between every rf cavity. The MAS matches the acceptance ellipse of a cavity to the beam ellipse at the cavity entrance, in order to minimise the cavity aperture for a given cavity length. In order to construct a more accurate description of the transverse beam dynamics, longitudinal acceleration is incorporated by producing a transfer map for an rf cavity. Within the new framework, important transverse beam dynamic quantities, such as the Twiss parameters and betatron phase advance, must be generalised to incorporate acceleration.

Recall the geometric emittance of a particle beam describes the area of the particle distribution in 2D phase space (coordinates, x, x' or y, y'), and is an invariant of motion if only symplectic transformations are considered, as shown by Liouville's Theorem [29]. Recall (from Eqn. 2.113) the phase space coordinates are related to the geometric emittance, $\varepsilon_{g,x}$, and Twiss parameters as follows.

$$\varepsilon_{g,x} = \beta_x x'^2 + 2\alpha_x x x' + \gamma_x x^2$$

Where β_x is related to the physical transverse beam size of the beam, α_x is the related to the rate of change of β_x with longitudinal displacement, and γ_x is related to the physical transverse velocity of the beam in the x plane. Together the parameters $\beta_x, \alpha_x, \gamma_x$ are called the Twiss (or Courant-Snyder) parameters [29]. Recall β is related to the physical beam size,

$\sigma = \sqrt{\varepsilon\beta}$, α is proportional to the beam divergence, $\alpha = -\frac{1}{2}\frac{d\beta}{ds}$, and $\gamma = \frac{1+\alpha^2}{\beta}$. Equation 2.113 also has a form for the y, y' coordinate phase space, with a separate set of Twiss parameters and geometric emittance (that need not be equal to the emittance in x).

It is required to use the normalised emittance defined as $\varepsilon_n = \varepsilon_g \gamma_r \beta_r$ where β_r is the particle beta and γ_r is the Lorentz factor. Suppose a particle in phase space has coordinates (x_0, x_0'), and traverses an element that can be described with transfer matrix, R. The new coordinates are (x_1, x_1'). As the Twiss parameters are a function of a lattice, they also transform $\beta_{x0} \rightarrow \beta_{x1}$. In general, the Lorentz factor will also change $\gamma_{r0} \rightarrow \gamma_{r1}$ and $\beta_{r0} \rightarrow \beta_{r1}$,

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix}. \tag{4.1}$$

The normalised emittance is invariant with respect to energy gain, therefore from Eqn. 2.113,

$$\begin{aligned} \gamma_{r1}\beta_{r1}(\beta_{x1}x_1'^2 + 2\alpha_{x1}x_1x_1' + \gamma_{x1}x_1^2) = \\ \gamma_{r0}\beta_{r0}(\beta_{x0}x_0'^2 + 2\alpha_{x0}x_0x_0' + \gamma_{x0}x_0^2). \end{aligned} \tag{4.2}$$

Where the invariance of the normalised emittance with respect to changes in energy has been used. Inverting the transformation in Eqn. 4.1,

$$\begin{pmatrix} x_0 \\ x_0' \end{pmatrix} = \frac{1}{\det(R)} \begin{pmatrix} R_{22} & -R_{12} \\ -R_{21} & R_{11} \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \end{pmatrix}. \tag{4.3}$$

Substituting x_0 and x_0' from Eqn. 4.3 into Eqn. 4.2 and comparing coefficients produces the Twiss transfer matrix. The Twiss transfer matrix relates the Twiss parameters at one point on the lattice to another point, and is a function of the phase space transfer map elements, R,

$$\begin{pmatrix} \beta_{x1} \\ \alpha_{x1} \\ \gamma_{x1} \end{pmatrix} = \frac{1}{\det(R)^2} \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} \begin{pmatrix} R_{11}^2 & -2R_{11}R_{12} & R_{12}^2 \\ -R_{11}R_{21} & R_{11}R_{22} + R_{12}R_{21} & -R_{12}R_{22} \\ R_{21}^2 & -2R_{21}R_{22} & R_{22}^2 \end{pmatrix} \begin{pmatrix} \beta_{x0} \\ \alpha_{x0} \\ \gamma_{x0} \end{pmatrix}. \tag{4.4}$$

4.2 RF Cavity Transfer Map

In order to more effectively track a particle distribution though a cavity, a transfer map must be defined. Analytical transfer maps for rf cavities have been produced before, in the relativistic case [91]. A less rigorous method is outlined below, assuming a π-mode rf cavity (however the results hold true for all cavity types), with cell length L_{cell} and particle synchronous phase ϕ_0. The derivation assumes the on-axis electric field profile goes as $\sin(x)$. In general, the field profile can be described with a Fourier series, however the important

result can be achieved using the first harmonic only. The field oscillates at the resonant angular frequency, ω,

$$E_z = \sin\left(\frac{\pi z}{L_{cell}}\right)\sin\left(\omega t + \phi_0\right). \tag{4.5}$$

The first order expansion of the non-zero transverse field components are given below (see Chapter 5 for the derivation),

$$E_r = -\frac{r}{2}\frac{dE_z}{dz}, \quad B_\theta = \frac{\omega r}{2c^2}E_z. \tag{4.6}$$

From the Lorentz force, the momentum gain in the x plane, for constant particle velocity is as follows;

$$\Delta p_x = q\cos\left(\theta\right)\left(\frac{1}{\beta_z c}\int E_r dz + \int B_\theta dz\right), \tag{4.7}$$

where θ is the azimuthal angle of a particle in the rf cavity. Substituting the non-zero fields from Eqns. 4.5 and 4.6 into Eqn. 4.7 produces an integral equation in terms of $\frac{dE_z}{dz}$ which is known from Eqn. 4.5. In order to simplify the integral whilst still gaining the relevant result it is assumed the particle beta, β_z, is equal to the synchronous particle beta, β_s, and it is constant.

$$\Delta p_x = q\cos(\theta)\frac{r}{2c}\int_0^{L_{cav}} \frac{\pi}{\beta_s L_{cell}}\cos\left(\frac{\pi z}{L_{cell}}\right)\sin\left(\omega t + \phi_0\right) + \frac{\omega\beta_s}{c}\sin\left(\frac{\pi z}{L_{cell}}\right)\cos\left(\omega t + \phi_0\right)dz. \tag{4.8}$$

The temporal term for B_θ becomes $\cos(\omega t + \phi_0)$ as the magnetic field lags the electric field by $\pi/2$. Performing a change of variable $t = \frac{z}{\beta_s c}$ and assuming a π-mode structure, the synchronicity condition is used $\frac{\pi}{L_{cell}} = \frac{\omega}{\beta_s c}$.

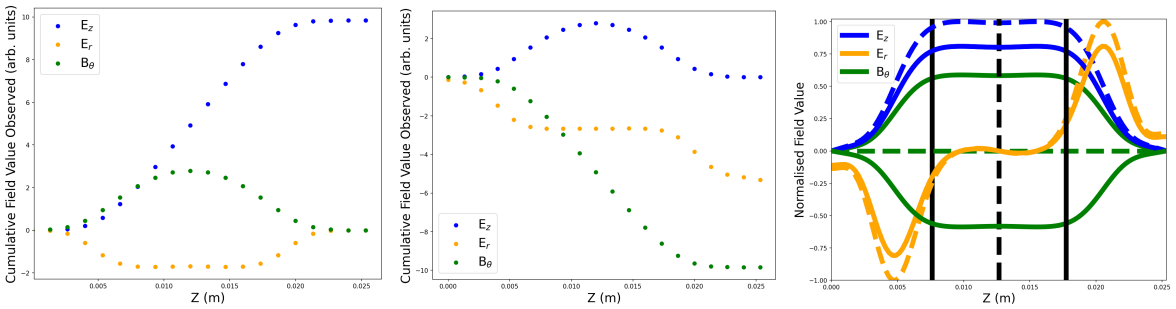
$$\Delta p_x = q\cos(\theta)\frac{r}{2c}\left[-\frac{1}{\beta_s}\frac{N\pi\sin(\phi_0)}{2} + \frac{\omega\beta_s}{c}\frac{NL_{cell}\sin(\phi_0)}{2}\right] \tag{4.9}$$

Where $L_{cav} = NL_{cell}$ and N is the number of cells. Simplifying the above, noting that $r\cos(\theta) = x_0$.

$$\Delta p_x = \frac{qx_0}{4c}N\pi\left(\beta_s - \frac{1}{\beta_s}\right)\sin(\phi_0) \tag{4.10}$$

The above describes a focusing/defocusing term, depending on ϕ_0. The longitudinal momentum gain, Δp_z can be computed with the same method as described above, the result is shown below;

$$\Delta p_z = \frac{q}{\beta_s c}\int_0^{L_{cav}} E_z dz = \int_0^{L_{cav}} \sin\left(\frac{\omega z}{\beta_s c}\right)\sin\left(\frac{\omega z}{\beta_s c} + \phi_0\right)dz = \frac{q}{\beta_s c}\frac{NL_{cell}}{2}\cos(\phi_0). \tag{4.11}$$



(a) Minimal Defocusing (b) Defocusing (c) Fields of (a).

Figure 4.1: Plots (a) and (b) show the cumulative field observed by a particle through an rf cavity due to different input rf phases. (c) Field components inside a single cell at three separate Z positions (vertical black lines) as observed by a particle with velocity equal to the design beta at a minimal defocusing phase.

The momentum gain in the transverse and longitudinal dimensions, as described by Eqns. 4.10 and 4.11, depend on $\sin(\phi_0)$ and $\cos(\phi_0)$, respectively. Hence maximally longitudinally accelerated beams are minimally transversely defocused. Figure 4.1 displays the cumulative field observed by a particle traveling at a minimal defocusing (a) and defocusing (b) phase. The field components observed by a synchronous particle traversing a single cell at a minimal defocusing phase are shown in Fig. 4.1 (c), for three separate longitudinal positions.

From Eqn. 4.10 the cavity map elements can be derived. Using s to denote longitudinal displacement and that $x' = \frac{p_x}{p_z} = \frac{dx}{ds}$.

$$\Delta p_x = p_{x1} - p_{x0} = x_1' p_{z1} - x_0' p_{z0} \tag{4.12}$$

$$x_1' = \frac{\Delta p_x + x_0' p_{z0}}{p_{z1}} \tag{4.13}$$

Inserting the form for Δp_x from Eqn. 4.10 into Eqn. 4.13 and simplifying using natural units gives the following

$$x_1' = \frac{N\pi}{4p_{z1}} \left(\beta_{z0} - \frac{1}{\beta_{z0}} \right) \sin(\phi_0) x_0 + \frac{p_{z0}}{p_{z1}} x_0' = R_{21} x_0 + R_{22} x_0'. \tag{4.14}$$

In order to find R_{11} and R_{12} Eqn. 4.14 is integrated with respect to s, and with $p_{z1} \to p_z(s)$ to introduce an explicit s dependence on the final longitudinal momentum,

$$\int_0^{L_{cav}} x'(s) = \int_0^{Lcav} dx(s) = \frac{N\pi}{4} \left(\beta_{z0} - \frac{1}{\beta_{z0}} \right) \sin(\phi_0) x_0 \int_0^{L_{cav}} \frac{1}{p_z(s)} ds + p_{z0} x_0' \int_0^{L_{cav}} \frac{1}{p_z(s)} ds. \tag{4.15}$$

$p_z(s)$ can be written in a form that is convenient for integration;

$$p_z(s) = \gamma_{rs}\beta_{rs}mc = \gamma_{rs}\sqrt{1 - \frac{1}{\gamma_{rs}^2}}mc = \sqrt{\gamma_{rs}^2 - 1}mc. \tag{4.16}$$

Thus,

$$\int_0^{L_{cav}} \frac{1}{p_z(s)}ds = \frac{1}{mc}\int_0^{L_{cav}} \frac{1}{\sqrt{\gamma_{rs}^2 - 1}}ds = L'. \tag{4.17}$$

A functional form for γ_{rs} must account for the rf phase and is assumed to increase/decrease linearly to the final value. In reality, the Lorentz factor will behave similar to a staircase function, due to the gaps between accelerating cells,

$$\gamma_{rs} = \gamma_{r0} + \cos(\phi_0)\Delta\gamma\frac{s}{L_{cav}}. \tag{4.18}$$

Where $\Delta\gamma$ is the maximum achievable gain, $\gamma_{r1} - \gamma_{r0}$. The rf phase dependence goes as $\cos(\phi_0)$, due to the Δp_z dependence. In order to proceed, a common hyperbolic identity is used;

$$\frac{d}{da}\cosh^{-1}(a) = \frac{1}{\sqrt{a^2 - 1}} \qquad a > 1. \tag{4.19}$$

As γ_{rs} is greater than 1 by definition, there is no issue with the requirement of $a > 1$. In order to use Eqn. 4.19 a substitution must be implemented;

$$a = \gamma_{rs} = \gamma_{r0} + \Delta\gamma\cos(\phi_0)\frac{s}{L_{cav}}, \tag{4.20}$$

$$ds = \frac{L_{cav}}{\Delta\gamma\cos(\phi_0)}da. \tag{4.21}$$

The integral in Eqn. 4.17 becomes

$$\frac{L_{cav}}{\Delta\gamma\cos(\phi_0)}\int_{\gamma_{r0}}^{\gamma_{r0}+\Delta\gamma\cos(\phi_0)} \frac{da}{\sqrt{a^2-1}} = \frac{L_{cav}}{\Delta\gamma\cos(\phi_0)}\left(\cosh^{-1}(\gamma_{r0} + \Delta\gamma\cos(\phi_0)) - \cosh^{-1}(\gamma_{r0})\right), \tag{4.22}$$

therefore,

$$L' = \frac{L_{cav}}{\Delta\gamma\cos(\phi_0)mc}\left(\cosh^{-1}(\gamma_{r0} + \Delta\gamma\cos(\phi_0)) - \cosh^{-1}(\gamma_{r0})\right). \tag{4.23}$$

Substituting L' into Eqn.4.15 and using $\int_0^{L_{cav}} x'(s) = x_1 - x_0$;

$$x_1 = x_0 + \frac{N\pi}{4}\left(\beta_{z0} - \frac{1}{\beta_{z0}}\right)\sin(\phi_0)L'x_0 + p_{z0}L'x_0' = R_{11}x_0 + R_{12}x_0'. \tag{4.24}$$

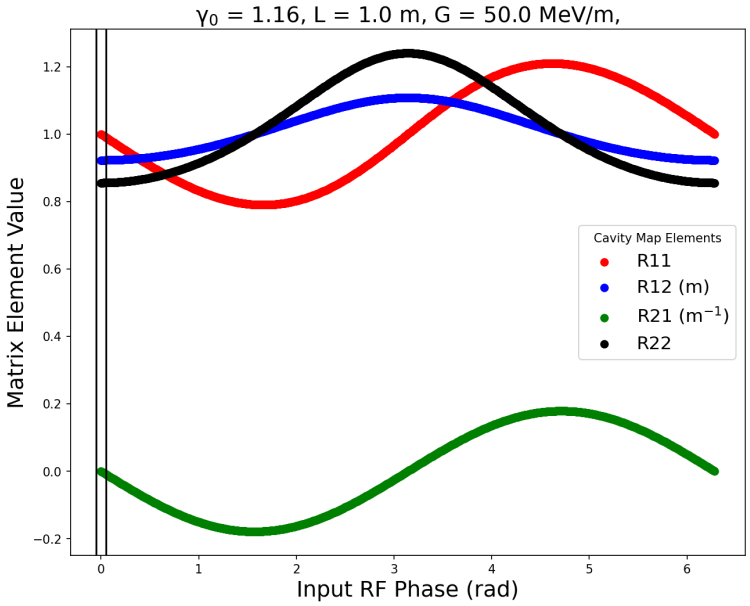


Figure 4.2: Cavity Map element values as a function of rf phase.

Providing the the remaining map elements, R_{11} and R_{12}. The full result for the phase dependent rf cavity map is shown below.

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \begin{pmatrix} 1 + \frac{N\pi}{4}\left(\beta_{z0} - \frac{1}{\beta_{z0}}\right)\sin(\phi_0)L' & \gamma_{r0}\beta_{z0}mcL' \\ \frac{N\pi}{4\gamma_{rs}\beta_{zs}mc}\left(\beta_{z0} - \frac{1}{\beta_{z0}}\right)\sin(\phi_0) & \frac{\gamma_{r0}\beta_{z0}}{\gamma_{rs}\beta_{zs}} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \tag{4.25}$$

The determinant of Eqn. 4.25 is calculated to be

$$\det(R) = \frac{\gamma_{r0}\beta_{z0}}{\gamma_{rs}\beta_{zs}} \tag{4.26}$$

as expected for a system undergoing adiabatic damping. Figure 4.2 shows the values of the cavity map elements as a function of rf phase. The synchronous phase, ϕ_0, is often selected such that the total energy gain is dp_z is maximum over the cavity, however is also selected to balance transverse and longitudinal focusing. From Eqn. 4.11, $\phi_0 = 0$ (shown by the two vertical lines in Fig. 4.2) minimises the defocusing term and simultaneously maximises longitudinal acceleration. In reality, the phase is often shifted to produce a larger region of longitudinal and transverse stability.

For $\phi_0 = 0$, the cavity map in Eqn. 4.25 reduces to a simpler form, with $R_{11} = 1$ and $R_{21} = 0$. Using the following identity,

$$\cosh^{-1}(x) = \ln(x + \sqrt{x^2 - 1}), \tag{4.27}$$

a simplified form for L' in R_{12} is determined;

$$L'(\phi_0 = 0) = \frac{L_{cav}}{\Delta\gamma mc}\left(\ln(\gamma_{r0} + \Delta\gamma + \sqrt{(\gamma_{r0} + \Delta\gamma)^2 - 1}) - \ln(\gamma_{r0} + \sqrt{\gamma_{r0}^2 - 1})\right). \quad (4.28)$$

The resulting transfer matrix is as follows;

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \begin{pmatrix} 1 & L_{cav}\frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{z1}+\gamma_{r1}}{\gamma_{r0}\beta_{z0}+\gamma_{r0}}\right) \\ 0 & \frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}\beta_{z1}} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix}. \quad (4.29)$$

Similar to the phase dependant cavity map, Eqn. 4.29 has determinant $\frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}\beta_{z1}}$. Thus, a particle bunch will occupy a smaller area in phase space after acceleration. Whilst the cavity map shown in Eqn. 4.29 is highly simplified, it approximates a cavity map better than a drift space, and will be used for cavity elements moving forward.

4.2.1 Betatron Phase Advance With Acceleration

The betatron phase advance (not to be confused with the rf phase) of a beam element, μ, represents the increase in the action angle variable of a particle [29] over that element. It was shown in Chapter 2 that the phase advance describes the rotation angle of normalised phase space (a circle as opposed to a rotated ellipse) due to some transformation. The natural next step is to determine how linear acceleration alters the phase advance, which can be determined using $\mu = \int_{s0}^{s1} \frac{1}{\beta(s)} ds$. The determinant of the rf independant cavity matrix, shown above, is $\frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}\beta_{z1}}$. The coefficient of the matrix that transforms Twiss parameters (shown in Eqn. 4.4) is $\frac{1}{\det(R)^2}\frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}\beta_{z1}}$, which becomes $\frac{\gamma_{r1}\beta_{z1}}{\gamma_{r0}\beta_{z0}}$ for an rf cavity with $\phi_0 = 0$. For non-zero acceleration, the coefficient of the Twiss parameter transformation matrix is therefore greater than one and the β function will grow relative to zero acceleration. Thus, the integral $\int \frac{1}{\beta}$ will decrease. The phase advance will therefore evolve more slowly in an accelerating system.

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} \begin{pmatrix} \sqrt{\frac{\beta_{x1}}{\beta_{x0}}}\left(\cos(\mu_x) + \alpha_{x0}\sin(\mu_x)\right) & \sqrt{\beta_{x0}\beta_{x1}} \\ \frac{(\alpha_{x0}-\alpha_{x1})\cos(\mu_x)-(1+\alpha_{x0}\alpha_{x1})}{\sqrt{\beta_{x0}\beta_{x1}}} & \sqrt{\frac{\beta_{x0}}{\beta_{x1}}}\left(\cos(\mu_x) - \alpha_{x1}\sin(\mu_x)\right) \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \quad (4.30)$$

4.3 The FODO-like Focusing scheme

Recall the FODO scheme is a periodic focusing scheme comprised of drift spaces separated by quadrupoles of alternating polarity. When designing a linac beam-line, the drift spaces may be replaced with rf cavities to accelerate the beam. In order to gain insight into this scheme analytically, the periodicity of the scheme is exploited, and the focusing quadrupole is split

in half. In order to gain an analytical understanding of the scheme, the cell is split into two half-FODO cells (see Fig. 4.3). Short drift lengths are incorporated between lattice elements to account for real world applications. The Twiss parameters tracked through a half-FODO cell will not return to the original values, as only half of a period has been traversed. The beta function is no longer perfectly periodic (as it increases with $\gamma_r\beta_r$), but it is seen later that the beam size is kept periodic. As a result, the scheme is called the FODO-like scheme.

The aim of solving the FODO-like scheme analytically, is to provide the quadrupole magnet k-strength, k, for a given quadrupole of length, l_q, such that the maximum beam size, σ, is minimised at the rf cavity entrance and exit; providing the smallest beam aperture possible for a given cavity length within the FODO-like scheme. Fundamentally, this requires minimising the Twiss β function at the cavity entrance and exit. Recall the Twiss transformation matrix from Eqn. 4.4 describes how the Twiss parameters transform due to a transfer map, R, which in general can be a product of multiple transfer maps, $R = ABC$. It can be shown that $\det R = \det A \det B \det C$. As quadrupoles and drift length transfer maps have unit determinant, only the cavity map determinant contributes to $\det R$. As a result, for a lattice comprised of quadrupoles, drift spaces and rf cavities, the determinant is $\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}$. The coefficient of the Twiss transformation matrix is therefore;

$$\frac{1}{det(R)^2}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}, \tag{4.31}$$

and the beta function transforms as follows;

$$\beta_1 = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}(R_{11}^2\beta_0 - 2R_{11}R_{12}\alpha_0 + R_{12}^2\gamma_0). \tag{4.32}$$

The input Twiss parameters (with subscript 0) are the values at some point in the lattice, which are free to choose. In this work, the input Twiss parameters are defined at some point within a quadrupole, such that α_0, the beam divergence, is zero. For zero acceleration, the start point is exactly half way through the quadrupole, however this is not true for non-zero acceleration. As $\alpha = -\frac{1}{2}\frac{d\beta}{ds}$, the input Twiss parameters are also defined at the point where β_0 is at an extreme. Recall, there are associated Twiss parameters for both the x and y transverse planes. In this work, β_x is at maximum when β_y is at a minimum, however these boundary conditions are not always true. In the following derivation, β_x is at a maximum when the input Twiss parameters are defined, however the choice is arbitrary,

$$\beta_{x0} = \mathbf{Max}, \beta_{y0} = \mathbf{Min}, \alpha_{x0} = \alpha_{y0} = \mathbf{0}. \tag{4.33}$$

To solve the problem analytically, transfer maps for quadrupoles, drift lengths and rf cavities must be known. Recall the transfer maps for quadrupoles and drift lengths;

$$R_{\mathrm{fq}} = \begin{pmatrix} \cos(\sqrt{k}l_q) & \frac{1}{k}\sin(\sqrt{k}l_q) \\ -k\sin(\sqrt{k}l_q) & \cos(\sqrt{k}l_q) \end{pmatrix} \tag{4.34}$$

$$R_{\mathrm{dq}} = \begin{pmatrix} \cosh(\sqrt{k}l_q) & \frac{1}{k}\sinh(\sqrt{k}l_q) \\ k\sinh(\sqrt{k}l_q) & \cosh(\sqrt{k}l_q) \end{pmatrix} \tag{4.35}$$

$$L_{\mathrm{gap}} = \begin{pmatrix} 1 & l_d \\ 0 & 1 \end{pmatrix}. \tag{4.36}$$

Quadrupole elements R_{12}, R_{21} have a dependence on the the k-strength (the quantity to be solved for) both inside and outside of trigonometric/hyperbolic functions, and thus becomes transcendental and impossible to solve analytically. Therefore, the trigonometric/hyperbolic functions are expanded to second order, henceforth called the semi-thin lens approximation;

$$\cos(\sqrt{k_1}l_{q1}) \approx 1 - k_1 l_{q1}^2/2$$
$$\sin(\sqrt{k_1}l_{q1}) \approx \sqrt{k_1}l_{q1}$$
$$\cosh(\sqrt{k_1}l_{q1}) \approx 1 + k_1 l_{q1}^2/2$$
$$\sinh(\sqrt{k_1}l_{q1}) \approx \sqrt{k_1}l_{q1}.$$

The semi-thin lens approximation retains all terms up to and including the order $k^n l_q^{n+2}$ (for arbitrarily high n). Second order expansions were chosen as it will provide greater accuracy and is still soluble. Recall, trigonometric/hyperbolic functions expanded to first order is the well-known thin lens approximation.

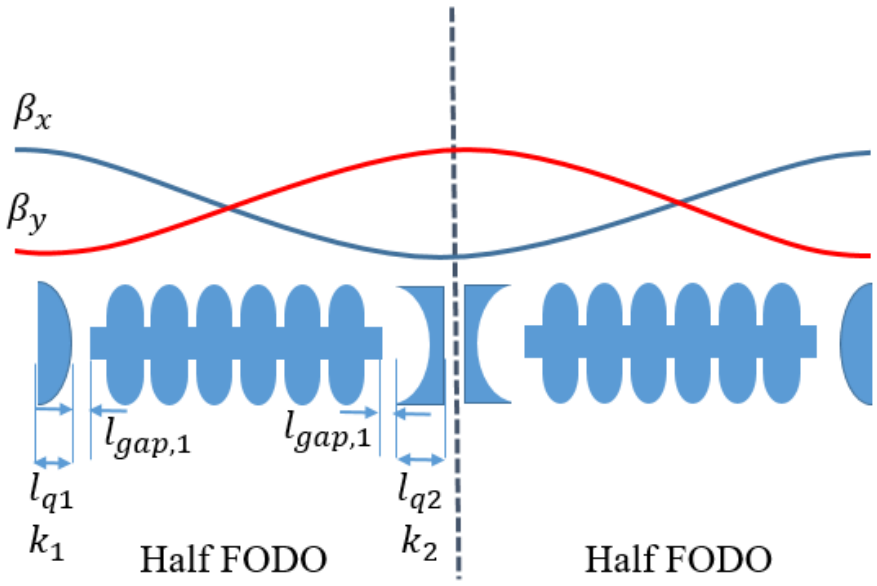


Figure 4.3: Schematic of the FODO-like focusing scheme with two RF cavities and quadrupoles showing how the β function changes.

The method requires transforming the Twiss parameters through a half-FODO cell (see Fig. 4.3); from some point within a quadrupole, to some point in the next quadrupole, at

which point,

$$\beta_{x0} = \mathbf{Min}, \beta_{y0} = \mathbf{Max}, \alpha_{x0} = \alpha_{y0} = \mathbf{0}. \tag{4.37}$$

The 2×2 transfer map for this transformation (in x) is denoted $R_{\text{half},x}$. In general, the transverse transfer map is 4×4 as it transforms both the x and y phase space, with the x transfer map being top left 2×2 section. The transfer map in x for a half-FODO cell is,

$$\begin{pmatrix} R_{\text{half},11} & R_{\text{half},12} \\ R_{\text{half},21} & R_{\text{half},22} \end{pmatrix} = \begin{pmatrix} 1 + \frac{k_2 l_{q2}^2}{2} & l_{q2} \\ k_2 l_{q2} & 1 + \frac{k_2 l_{q2}^2}{2} \end{pmatrix} \begin{pmatrix} 1 & L_{eff} \\ 0 & \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} \end{pmatrix} \begin{pmatrix} 1 - \frac{k_1 l_{q1}^2}{2} & l_{q1} \\ k_1 l_{q1} & 1 - \frac{k_1 l_{q1}^2}{2} \end{pmatrix}, \tag{4.38}$$

where k_1 and k_2 define the k-strengths of the focusing and defocusing quadrupoles, respectively (recall x is initially focused). l_{q1} and l_{q2} define some section length of the focusing and defocusing quadrupoles, respectively, and are defined such that α (in both x and y) is 0 and the end of the half-FODO cell. Going forward, the quadrupole section lengths, l_{q1}, l_{q2}, are simply refered to as quadrupole lengths. The center matrix of Eqn. 4.38 is the multiplication [drift][cavity][drift] and has been computed already. The cavity matrix used is shown in Eqn. 4.29. L_{eff} is the R_{12} element of the [drift][cavity][drift] transfer map and is given by;

$$L_{eff} = l_{gap}\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} + 1\right) + L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1} - \gamma_{r0}} \ln\left(\frac{\gamma_{r1}\beta_{r1} + \gamma_{r1}}{\gamma_{r0}\beta_{r0} + \gamma_{r0}}\right) \tag{4.39}$$

and is identical to R_{34} (the equivalent R_{12} element for the y transverse plane) as acceleration and drift lengths are azimuthally symmetric. The Twiss parameters transformed to the end point of a half-FODO are denoted with subscript 1. The Twiss α function in both the x and y planes are, from Eqn. 4.32;

$$\alpha_{x1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(-R_{\text{half},11}R_{\text{half},21}\beta_{x0} - \frac{R_{\text{half},12}R_{\text{half},22}}{\beta_{x0}}\right) \tag{4.40}$$

$$\alpha_{y1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(-R_{\text{half},33}R_{\text{half},43}\beta_{y0} - \frac{R_{\text{half},34}R_{\text{half},44}}{\beta_{y0}}\right). \tag{4.41}$$

Recall $\alpha_{x0} = \alpha_{y0} = 0$ and therefore $\gamma_{(x/y)0} = \frac{1}{\beta_{(x/y)0}}$. In addition, $\det(R_{\text{half},x}) = \det(R_{\text{half},y}) = \det(R_{\text{half}})$. The matrix R_{half} provides the necessary information to transform the Twiss parameters to a point in the lattice where the beam divergence returns to 0 for some set of quadrupole, cavity, and drift length parameters. Solving Eqns. 4.40 and 4.41 for both β_{x0} and β_{y0} before taking the positive definite solution;

$$\beta_{x0} = \sqrt{\frac{-R_{\text{half},12}R_{\text{half},22}}{R_{\text{half},11}R_{\text{half},21}}}, \quad \beta_{y0} = \sqrt{\frac{-R_{\text{half},34}R_{\text{half},44}}{R_{\text{half},33}R_{\text{half},43}}}. \tag{4.42}$$

Enforcing the condition of the FODO system that the beam size in x at the start of the half-FODO is equal to the beam size in y at the end of the half-FODO;

$$\sigma_{x0} = \sigma_{y1}, \quad \sigma_{x1} = \sigma_{y0}. \tag{4.43}$$

The beam size in x at any point along the lattice is given

$$\sigma_x = \sqrt{\frac{\varepsilon_n \beta_x}{\beta_r \gamma_r}}. \tag{4.44}$$

Using Eqns 4.43 and 4.44 the relation between the Twiss β, separated by a half-FODO cell, are as follows;

$$\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\beta_{x0} = \beta_{y1}, \qquad \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\beta_{y0} = \beta_{x1} \tag{4.45}$$

$$\beta_{x0}\beta_{x1} = \beta_{y0}\beta_{y1}. \tag{4.46}$$

It is useful to define the aspect ratio of the transverse beam sizes at the each end of a half-FODO, denoted r;

$$\frac{\beta_{x0}}{\beta_{y0}} = \frac{\beta_{y1}}{\beta_{x1}} = r. \tag{4.47}$$

Translating the β functions to the end of a half-FODO cell gives the following;

$$\beta_{x1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},11}\beta_{x0} + \frac{R^2_{\text{half},12}}{\beta_{x0}}\right) \tag{4.48}$$

$$\beta_{y1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},33}\beta_{y0} + \frac{R^2_{\text{half},34}}{\beta_{y0}}\right). \tag{4.49}$$

Translating the Twiss γ provides another set of equations;

$$\gamma_{x1} = \frac{1}{\beta_{x1}} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},21}\beta_{x0} + \frac{R^2_{\text{half},22}}{\beta_{x0}}\right) \tag{4.50}$$

$$\gamma_{y1} = \frac{1}{\beta_{y1}} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},43}\beta_{y0} + \frac{R^2_{\text{half},44}}{\beta_{y0}}\right). \tag{4.51}$$

Using Eqns. 4.49 and 4.51 the product of the β function before and after the half-FODO are given;

$$\beta_{x0}\beta_{x1} = \frac{\beta_{x0}}{\gamma_{x1}} = \frac{\beta_{x0}}{\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},21}\beta_{x0} + \frac{R^2_{\text{half},22}}{\beta_{x0}}\right)}, \tag{4.52}$$

$$\beta_{y0}\beta_{y1} = \frac{\beta_{y0}}{\gamma_{y1}} = \frac{\beta_{y0}}{\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R^2_{\text{half},43}\beta_{y0} + \frac{R^2_{\text{half},44}}{\beta_{y0}}\right)}. \tag{4.53}$$

From Eqn. 4.46, the above products are equal, providing the following requirement;

$$R^2_{\text{half},21} + \frac{R^2_{\text{half},22}}{\beta^2_{x0}} = R^2_{\text{half},43} + \frac{R^2_{\text{half},44}}{\beta^2_{y0}}. \tag{4.54}$$

Inserting β_{x0} and β_{y0} from Eqn. 4.42

$$R_{\text{half},21}^2 - \frac{R_{\text{half},22}^2}{\frac{R_{\text{half},12}R_{\text{half},22}}{R_{\text{half},11}R_{\text{half},21}}} = R_{\text{half},43}^2 - \frac{R_{\text{half},44}^2}{\frac{R_{\text{half},34}R_{\text{half},44}}{R_{\text{half},33}R_{\text{half},43}}}. \tag{4.55}$$

Simplifying the above produces the following;

$$R_{\text{half},21}\left(R_{\text{half},21} - \frac{R_{\text{half},22}R_{\text{half},11}}{R_{\text{half},12}}\right) = R_{\text{half},43}\left(R_{\text{half},43} - \frac{R_{\text{half},44}R_{\text{half},33}}{R_{\text{half},34}}\right). \tag{4.56}$$

Multiplying Eqn. 4.56 with both $R_{\text{half},12}$ and $R_{\text{half},34}$ separately to gain two separate equations;

$$R_{\text{half},21}\det(R_{\text{half}}) = R_{\text{half},43}\left(R_{\text{half},43}R_{\text{half},12} - R_{\text{half},44}R_{\text{half},33}\frac{R_{\text{half},12}}{R_{\text{half},34}}\right) \tag{4.57}$$

$$R_{\text{half},21}\left(R_{\text{half},21}R_{\text{half},34} - R_{\text{half},22}R_{\text{half},11}\frac{R_{\text{half},34}}{R_{\text{half},12}}\right) = R_{\text{half},43}\det(R_{\text{half}}). \tag{4.58}$$

Solving Eqns. 4.57 and 4.58 for $\det(R_{\text{half}})$ and equating;

$$\frac{R_{\text{half},43}}{R_{\text{half},21}}\left(R_{\text{half},43}R_{\text{half},12} - R_{\text{half},44}R_{\text{half},33}\frac{R_{\text{half},12}}{R_{\text{half},34}}\right) = \frac{R_{\text{half},21}}{R_{\text{half},43}}\left(R_{\text{half},21}R_{\text{half},34} - R_{\text{half},22}R_{\text{half},11}\frac{R_{\text{half},34}}{R_{\text{half},12}}\right) \tag{4.59}$$

$$\frac{R_{\text{half},43}^2}{R_{\text{half},21}^2} = \frac{R_{\text{half},21}R_{\text{half},34} - R_{\text{half},22}R_{\text{half},11}\frac{R_{\text{half},34}}{R_{\text{half},12}}}{R_{\text{half},43}R_{\text{half},12} - R_{\text{half},44}R_{\text{half},33}\frac{R_{\text{half},12}}{R_{\text{half},34}}}. \tag{4.60}$$

The above requirement must be met to ensure constant maximum beam size. The requirement used to arrive at the above constraint, $\beta_{x0}\beta_{x1} = \beta_{y0}\beta_{y1}$ can be solved using information unused thus far;

$$\beta_{x0}\beta_{x1} = \frac{1}{\det(R_{\text{half}})^2}\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R_{\text{half},11}^2\beta_{x0}^2 + R_{\text{half},12}^2\right) \tag{4.61}$$

$$\beta_{y0}\beta_{y1} = \frac{1}{\det(R_{\text{half}})^2}\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(R_{\text{half},33}^2\beta_{y0}^2 + R_{\text{half},34}^2\right). \tag{4.62}$$

Similarly, the two expressions are equal. Substituting the forms for β_{x0}, β_{y0} as functions of matrix elements (Eqn. 4.42);

$$-\frac{R_{\text{half},11}R_{\text{half},12}R_{\text{half},22}}{R_{\text{half},21}} + R_{\text{half},12}^2 = -\frac{R_{\text{half},33}R_{\text{half},34}R_{\text{half},44}}{R_{\text{half},43}} + R_{\text{half},34}^2. \tag{4.63}$$

Dividing by $R_{\text{half},12}R_{\text{half},34}$

$$-\frac{R_{\text{half},11}R_{\text{half},22}}{R_{\text{half},21}R_{\text{half},34}}+\frac{R_{\text{half},12}}{R_{\text{half},34}}=-\frac{R_{\text{half},33}R_{\text{half},44}}{R_{\text{half},43}R_{\text{half},12}}+\frac{R_{\text{half},34}}{R_{\text{half},12}}. \tag{4.64}$$

Introducing terms to have like-denominators

$$\frac{-R_{\text{half},11}R_{\text{half},22}+R_{\text{half},12}R_{\text{half},21}}{R_{\text{half},21}R_{\text{half},34}}=\frac{-R_{\text{half},33}R_{\text{half},44}+R_{\text{half},34}R_{\text{half},43}}{R_{\text{half},43}R_{\text{half},12}} \tag{4.65}$$

which, using $\det R_{\text{half},x}=\det R_{\text{half},y}$, provides the result following result;

$$\frac{R_{\text{half},12}}{R_{\text{half},34}}=\frac{R_{\text{half},21}}{R_{\text{half},43}}. \tag{4.66}$$

Substituting the constraint shown in Eqn. 4.66 into the constraint from Eqn. 4.60,

$$\begin{aligned}\frac{R_{\text{half},43}^2}{R_{\text{half},21}^2}=\frac{R_{\text{half},21}R_{\text{half},34}-R_{\text{half},22}R_{\text{half},11}\frac{R_{\text{half},34}}{R_{\text{half},12}}}{R_{\text{half},43}R_{\text{half},12}-R_{\text{half},44}R_{\text{half},33}\frac{R_{\text{half},12}}{R_{\text{half},34}}}=\\ \frac{R_{\text{half},21}R_{\text{half},12}\frac{R_{\text{half},43}}{R_{\text{half},21}}-R_{\text{half},22}R_{\text{half},11}\frac{R_{\text{half},34}}{R_{\text{half},12}}}{R_{\text{half},43}R_{\text{half},34}\frac{R_{\text{half},21}}{R_{\text{half},43}}-R_{\text{half},44}R_{\text{half},33}\frac{R_{\text{half},12}}{R_{\text{half},34}}}=\frac{\det(R_{\text{half}})}{\det(R_{\text{half}})},\end{aligned} \tag{4.67}$$

and therefore;

$$\frac{R_{\text{half},43}^2}{R_{\text{half},21}^2}=1. \tag{4.68}$$

The final constraint for the individual transfer map elements are therefore;

$$R_{\text{half},21}=R_{\text{half},43},\quad R_{\text{half},12}=R_{\text{half},34} \tag{4.69}$$

$$R_{\text{half},11}R_{\text{half},22}=R_{\text{half},33}R_{\text{half},44}. \tag{4.70}$$

Where the positive solution is taken when solving Eqn. 4.68, as the negative solution is not wanted. Relationships between quadrupole parameter values, k_1, l_{q1}, k_2, l_{q1} can be calculated using the constraints in Eqn. 4.69. Semi-thin lens approximations of $R_{\text{half},12}$ and $R_{\text{half},34}$, are found by expanding Eqn. 4.38. Equating the map elements produces the following

$$\begin{aligned}\left(\frac{\sqrt{k_1}l_{q1}}{\sqrt{k_1}}+L_{eff}\left(1-\frac{k_1l_{q1}^2}{2}\right)\right)\left(1+\frac{k_2l_{q2}^2}{2}\right)+\frac{\sqrt{k_2}l_{q2}}{\sqrt{k_2}}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\left(1-\frac{k_1l_{q1}^2}{2}\right)=\\ \left(\frac{\sqrt{k_1}l_{q1}}{\sqrt{k_1}}+L_{eff}\left(1+\frac{k_1l_{q1}^2}{2}\right)\right)\left(1-\frac{k_2l_{q2}^2}{2}\right)+\frac{\sqrt{k_2}l_{q2}}{\sqrt{k_2}}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\left(1+\frac{k_1l_{q1}^2}{2}\right)\end{aligned} \tag{4.71}$$

expanding and simplifying, ignoring all higher order terms that are not included in the semi-thin lens approximation;

$$L_{eff}\left(k_2 l_{q2}^2 - k_1 l_{q1}^2\right) = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} k_1 l_{q1}^2 l_{q2} - k_2 l_{q2}^2 l_{q1}. \tag{4.72}$$

The remaining constraint,

$$R_{\text{half},21} = R_{\text{half},43}, \tag{4.73}$$

yields the following result

$$k_2 l_{q2} = k_1 l_{q1} \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}. \tag{4.74}$$

Using the constraint from Eqn. 4.74 and inserting into Eqn. 4.72

$$L_{eff}\left(k_2 l_{q2}^2 - k_1 l_{q1}^2\right) = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} k_1 l_{q1}^2 l_{q2} - \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} k_1 l_{q1}^2 l_{q2} = 0. \tag{4.75}$$

Taking the non-trivial solution (for $L_{eff} \neq 0$);

$$k_1 l_{q1}^2 = k_2 l_{q2}^2. \tag{4.76}$$

Using Eqns. 4.74 and 4.76 produces relationships between consecutive quadrupole k-strengths and lengths;

$$l_{q1} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} l_{q2}, \tag{4.77}$$

$$k_1 = \frac{k_2}{\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\right)^2}. \tag{4.78}$$

It can be shown Eqns. 4.74 and 4.78 are always solutions by inserting the results into Eqn. 4.69 for the thick lens case.

Recall, the quadrupole lengths discussed above, l_{q1} and l_{q2}, are the lengths of the quadrupole sections, which are some proportion of the entire quadrupole. The length of the other section of the quadrupole is explored later in this section. For zero acceleration, the quadrupole is split exactly in half and the real quadrupole would have length $2l_{q1}$. It is required to produce function forms of β_{x0}, β_{y0} and the aspect ratio, r, as functions of quadrupole, drift length and cavity parameters (lattice parameters).

Firstly, dividing β_{y1} by β_{x1};

$$\frac{\beta_{y1}}{\beta_{x1}} = \frac{R_{\text{half},33}^2 \beta_{y0} + \frac{R_{\text{half},34}^2}{\beta_{y0}}}{R_{\text{half},11}^2 \beta_{x0} + \frac{R_{\text{half},12}^2}{\beta_{x0}}} = r, \tag{4.79}$$

before multiplying by $\frac{\beta_{y0}}{\beta_{x0}}$ (equivalent to dividing by r);

$$\frac{\beta_{y1}\beta_{y0}}{\beta_{x1}\beta_{x0}} = \frac{R^2_{\text{half},33}\beta^2_{y0} + R^2_{\text{half},34}}{R^2_{\text{half},11}\beta^2_{x0} + R^2_{\text{half},12}} = 1, \tag{4.80}$$

thus,

$$R^2_{\text{half},33}\beta^2_{y0} + R^2_{\text{half},34} = R^2_{\text{half},11}\beta^2_{x0} + R^2_{\text{half},12} \rightarrow \frac{R^2_{\text{half},33}}{R^2_{\text{half},11}} = r^2. \tag{4.81}$$

Where the fact $R_{\text{half},12} = R_{\text{half},34}$ has been used. As r is positive definite, the positive solution is taken when the final equality above is square rooted. Producing semi-thin lens approximations for $R_{\text{half},11}$ and $R_{\text{half},33}$ produces an approximation for r;

$$r = \frac{R_{\text{half},33}}{R_{\text{half},11}} \approx \frac{1 + L_{eff}k_1 l_{q1} + k_1 l^2_{q1} - \frac{L_{eff}k^2_1 l^3_{q1}}{2}}{1 - L_{eff}k_1 l_{q1} - k_1 l^2_{q1} - \frac{L_{eff}k^2_1 l^3_{q1}}{2}}. \tag{4.82}$$

Another form that gives the ratio of beam sizes, r, is the following;

$$\frac{\gamma_{x1}}{\gamma_{y1}} = \frac{\beta_{y1}}{\beta_{x1}} = r. \tag{4.83}$$

Inserting approximations for γ_{x1}, γ_{y1} from Eqn. 4.51 into the above equation yields;

$$\frac{R_{\text{half},22}}{R_{\text{half},44}} = r = \frac{R_{\text{half},33}}{R_{\text{half},11}}. \tag{4.84}$$

The input Twiss parameters, β_{x0} and β_{y0} can be determined as functions of lattice parameters and r, by inserting the semi-thin lens approximations for elements of R_{half} into Eqn. 4.42;

$$\beta_{x0} = \frac{\sqrt{r}}{k_1 l_{q1}}\sqrt{1 + \frac{l_{q1}}{L_{eff}}}, \tag{4.85}$$

$$\beta_{y0} = \frac{1}{\sqrt{r}k_1 l_{q1}}\sqrt{1 + \frac{l_{q1}}{L_{eff}}}. \tag{4.86}$$

Recall the aim of the method is to produce the lattice elements such that the maximum beam size is minimised at the cavity entrance and exit. The Twiss β (in x) at the cavity entrance, β_{xc0}, is determined using Eqn. 4.4 where the transfer map is comprised of the focusing quadrupole (of strength k_1 and length l_{q1}) and a drift space of length l_{g1}. The transfer matrix will be referred to as R_{entrance},

$$R_{\text{entrance}} = \begin{pmatrix} 1 & l_{g1} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 - \frac{k_1 l^2_{q1}}{2} & l_{q1} \\ k_1 l_{q1} & 1 - \frac{k_1 l^2_{q1}}{2} \end{pmatrix} = \begin{pmatrix} 1 - \frac{k_1 l^2_{q1}}{2} + k_1 l_{q1} l_{g1} & l_{q1} + l_{g1}\left(1 - \frac{k_1 l^2_{q1}}{2}\right) \\ k_1 l_{q1} & 1 - \frac{k_1 l^2_{q1}}{2} \end{pmatrix}. \tag{4.87}$$

Using the Twiss parameter mapping matrix, β_{xc0} is defined as follows;

$$\beta_{xc0} = R^2_{\text{entrance},11}\beta_{x0} + \frac{R^2_{\text{entrance},12}}{\beta_{x0}}, \tag{4.88}$$

and is minimised by differentiating Eqn. 4.88 with respect to the quadrupole k-strength, k_1 (although it would be just as valid to minimise with respect to the l_{q1}) and setting to zero. As the beam size in x at the start of the half-FODO cell is equal to the beam size in y at the end of the cell, minimising the beam at the cavity entrance will also minimise the beam size in y at the cavity exit. The total derivative is given;

$$\frac{d\beta_{xc0}}{dk_1} = 2R_{11}\frac{dR_{11}}{dk_1}\beta_{x0} + R^2_{11}\frac{d\beta_{x0}}{dk_1} + 2R_{12}\frac{dR_{12}}{dk_1}\frac{1}{\beta_{x0}} - R^2_{12}\frac{1}{\beta^2_{x0}}\frac{d\beta_0}{dk_1} = 0. \tag{4.89}$$

Terms like $\frac{dR_{\text{entrance},11}}{dk_1}$ can be calculated straightforwardly, using Eqn. 4.87,

$$\frac{dR_{\text{entrance},11}}{dk_1} = -\frac{l^2_{q1}}{2} + l_{q1}l_{g1} \tag{4.90}$$

$$\frac{dR_{\text{entrance},12}}{dk_1} = -\frac{l^3_{q1}}{2}. \tag{4.91}$$

It is convenient to expand the derivative of β_{x0} as a function of r. Differentiating β_{x0} with respect to k_1 is completed by using the semi-thin lens approximation for β_{x0} from Eqn. 4.85;

$$\frac{d\beta_{x0}}{dk_1} = \sqrt{\frac{1}{l^2_{q1}} + \frac{1}{l_{q1}L_{eff}}}\frac{d}{dk_1}\left(\frac{\sqrt{r}}{k_1}\right) = \sqrt{\frac{1}{l^2_{q1}} + \frac{1}{l_{q1}L_{eff}}}\left(\frac{1}{2r^{\frac{1}{2}}k_1}\frac{dr}{dk_1} - \frac{r^{\frac{1}{2}}}{k^2_1}\right). \tag{4.92}$$

This can be written in a more useful form by taking out a factor of β_{x0} as defined in Eqn. 4.85;

$$\frac{d\beta_{x0}}{dk_1} = \beta_{x0}\left(\frac{1}{2r}\frac{dr}{dk_1} - \frac{1}{k_1}\right). \tag{4.93}$$

Rearranging Eqn. 4.89 for $\frac{\beta_{x0}}{dk_1}$, equating to the above equality and dividing by β_{x0}, with $R_{\text{entrance}} = R$;

$$\frac{1}{2r}\frac{dr}{dk_1} - \frac{1}{k_1} = \frac{-2R_{11}\frac{dR_{11}}{dk_1} - 2R_{12}\beta^{-2}_{x0}\frac{dR_{12}}{dk_1}}{R^2_{11} - R^2_{12}\beta^{-2}_{x0}} \tag{4.94}$$

$$\left(\frac{1}{2r}\frac{dr}{dk_1} - \frac{1}{k_1}\right)\left(R^2_{11} - R^2_{12}\beta^{-2}_{x0}\right) = -2R_{11}\frac{dR_{11}}{dk_1} - \frac{2R_{12}}{\beta^2_{x0}}\frac{dR_{12}}{dk_1}. \tag{4.95}$$

Multiplying by r produces;

$$\frac{R^2_{11}}{2}\frac{dr}{dk_1} - \frac{R^2_{12}}{2\beta^2_{x0}}\frac{dr}{dk_1} - \frac{rR^2_{11}}{k_1} + \frac{rR^2_{12}}{k_1\beta^2_{x0}} = -2rR_{11}\frac{dR_{11}}{dk_1} - \frac{2rR_{12}}{\beta^2_{x0}}\frac{dR_{12}}{dk_1}. \tag{4.96}$$

The derivative of r with respect to k_1 must now be computed;

$$\frac{dr}{dk_1} = \frac{l_{q1}^4 L_{eff}^2 (L_{eff} + l_{q1}) k_1^2 + 2 l_{q1} (L_{eff} + l_{q1})}{\left(1 - L_{eff} k_1 l_{q1} - k_1 l_{q1}^2 - \frac{L_{eff} k_1^2 l_{q1}^3}{2}\right)^2}. \tag{4.97}$$

However, the semi-thin lens approximation ignores terms of order $k^n l_q^{n+2}$, for any positive integer n. As L_{eff} is the same order of magnitude as the physical cavity length, it will be of the order 1. As a result, Eqn. 4.97 reduces to the following;

$$\frac{dr}{dk_1} = \frac{2 l_{q1} L_{eff}}{\left(1 - L_{eff} k_1 l_{q1} - k_1 l_{q1}^2 - \frac{L_{eff} k_1^2 l_{q1}^3}{2}\right)^2}. \tag{4.98}$$

The final step requires inserting the analytical forms for β_{x0} (Eqn. 4.85), r (Eqn. 4.82), R_{11}, R_{12} (Eqn. 4.87), $\frac{dR_{11}}{dk_1}$, $\frac{dR_{12}}{dk_1}$ (Eqn. 4.91) into Eqn. 4.96 and simplifying omitting and terms that are beyond second order. A cubic in k_1 is produced.

$$-\frac{l_{q1}^4 L_{eff}^2}{2} k_1^3 + l_{q1}^2 (2 l_g L_{eff} - 2 l_{q1} L_{eff} - L_{eff}^2) k_1^2 - l_{q1} (L_{eff} + l_{q1}) k_1 + 1 = 0 \tag{4.99}$$

Many methods exist to solve a cubic equation, however here the cubic is written as a depressed cubic [92] and a formula demonstrated by Francois Viete is implemented. A general cubic has the form;

$$ax^3 + bx^2 + cx + d = 0. \tag{4.100}$$

Solving the cubic can be done by writing the cubic as a depressed cubic. The first step is to change variables;

$$x = t - \frac{b}{3a} \tag{4.101}$$

dividing the remaining cubic with a leaves the depressed cubic;

$$t^3 + pt + \mathrm{q} = 0 \tag{4.102}$$

where $p(a, b, c)$ and $\mathrm{q}(a, b, c, d)$ are functions of the original cubic coefficients,

$$p = \frac{3ac - b^2}{3a^2} \tag{4.103}$$

$$q = \frac{2b^3 - 9abc + 27a^2 d}{27a^3}. \tag{4.104}$$

The roots are calculated using the following;

$$t_m = 2\sqrt{\frac{-p}{3}} \cos\left[\frac{1}{3} \cos^{-1}\left(\frac{3q}{2p}\sqrt{\frac{-3}{p}}\right) - \frac{2\pi\theta_m}{3}\right]. \tag{4.105}$$

The term inside the square root(s) of Eqn. 4.105 may be negative, as p may be less than 0. As a result, t_k is treated as complex. It is useful to display the complex number in the form $Re^{i\theta}$ and take the cube root. This gives one solution to the cubic with the largest real component. The remaining two solutions can be found by rotating the complex numbers by one third of a full period;

$$\theta_m = \frac{2\pi m}{3}, \qquad m = (0, 1, 2) \tag{4.106}$$

therefore,

$$k_1 = 2\sqrt{\frac{-p}{3}}\cos\left(\frac{1}{3}\cos^{-1}\left(\frac{3q}{2p}\sqrt{\frac{-3}{p}}\right) - \theta_m\right) - \frac{b}{3a} \tag{4.107}$$

where

$$p = \frac{3ac - b^2}{3a^2},\ q = \frac{2b^3 - 9abc + 27a^2d}{27a^3} \tag{4.108}$$

and

$$a = \frac{-L_{eff}^2 l_{q1}^4}{2} \tag{4.109}$$

$$b = l_{q1}^2(-2L_{eff}l_{q1} - L_{eff}^2 + 2l_g L_{eff}) \tag{4.110}$$

$$c = -l_{q1}(L_{eff} + l_{q1}) \tag{4.111}$$

$$d = 1. \tag{4.112}$$

The cubic can be simplified to first order by keeping terms of the order $k_1 l_{q1}$ only. Producing an equation for k_1 that is more straightforward to solve;

$$k_1 l_{q1} = \frac{\sqrt{5} - 1}{2L_{eff}}. \tag{4.113}$$

The above equation is the thin lens approximation for k_1, and whilst much easier to solve, is highly inferior to the semi-thin lens approximation for the optimised value of k_1 (see later).

Equation 4.107 provides an analytical form for calculating first quadrupole k-strength in order to minimise the maximum beam size at the cavity entrance. As a result, the optimal beam size at the cavity entrance can be calculated as a function of lattice parameters, as the equation for β at the cavity entrance, β_{xc0}, was derived and shown in Eqn. 4.88. This method therefore allows a self-contained analytical method to determine the minimum beam size at the cavity entrance as function of cavity length. The performance of the method will be analysed in Section 4.6.

4.4 Concatenating Multiple half-FODO Cells

As there is nothing special about the first half-FODO cell that allowed the derivation of the half-FODO constraints, the constraints must hold true for subsequent half-FODO cells. The next step is to derive a set of telescoping equations. Firstly, quadrupole lengths must be re-defined to incorporate a second index that describes if it is the first or second section of a complete quadrupole, that is;

$$l_{q1} = l_{q1,2}, \;\; l_{q2} = l_{q2,1}, \tag{4.114}$$

where the second index describes which section of the quadrupole the length describes (first or second), as shown in Fig. 4.4. The parameter L_{eff} also adopts a subscript, indicating the half-FODO cell index, $L_{eff,1}$ is the first half-FODO cell value of L_{eff}. Recall $L_{eff,1}$ is a function of the cavity gradient, cavity length, and drift length;

$$L_{eff,1} = l_{g1}(1 + \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}) + L_{cav1}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1} - \gamma_{r0}} \ln\left(\frac{\gamma_{r1}\beta_{r1} + \gamma_{r1}}{\gamma_{r0}\beta_{r0} + \gamma_{r0}}\right). \tag{4.115}$$

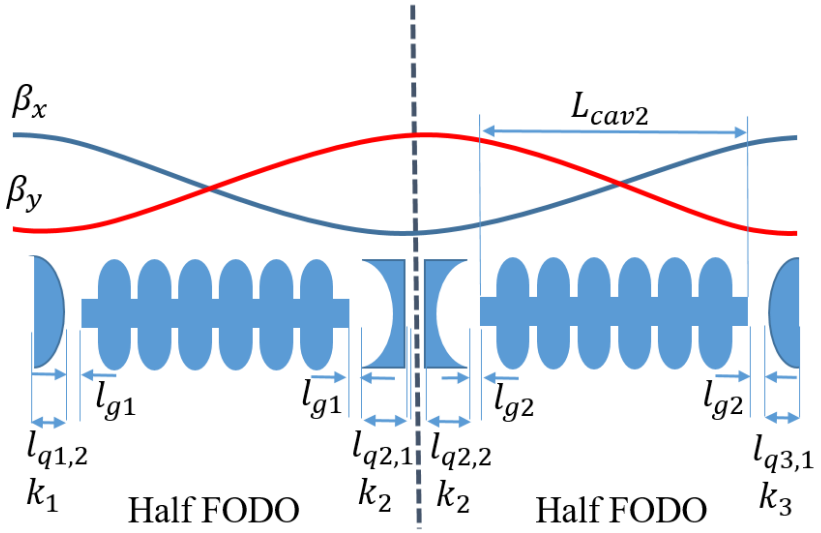


Figure 4.4: Schematic of the FODO-like focusing scheme with two RF cavities. Lattice parameters are labelled to aid with nomenclature.

For a set of N half-FODO cells, there exists N focusing quadrupole sections and N defocusing quadrupole sections, a total of 2N quadrupole sections (N complete quadrupoles). The k-strength are related via the following relationship, as described by Eqn. 4.78;

$$k_1 = \frac{k_2}{\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\right)^2} = \frac{k_3}{\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\right)^2} = \frac{k_4}{\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\right)^2\left(\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r2}\beta_{r2}}\right)^2} = \cdots = \frac{k_{2N}}{\prod_{i=0}^{N-1}\left(\frac{\gamma_{ri}\beta_{ri}}{\gamma_{r(i+1)}\beta_{r(i+1)}}\right)^2}. \tag{4.116}$$

Where $k_2 = k_3$ as they are sections of the same quadrupole, separated into two separate half-FODO cells (see Fig. 4.4). The relationship between quadrupole lengths in the n^{th} and $(n+1)^{th}$ half-FODO cells are determined using Eqn. 4.77,

$$l_{qn,2} = \frac{\gamma_{r(n-1)}\beta_{r(n-1)}}{\gamma_{rn}\beta_{rn}} l_{\mathrm{q}(n+1),1}. \tag{4.117}$$

Instead of assuming a value for $l_{qn,2}$ (as was required for $l_{q1,2}$), it is later shown, using Eqn. 4.122, how the quantity can be derived.

Equation 4.45 determines how the Twiss β must evolve along a lattice of half-FODO cells for a given acceleration. Equations 4.85 and 4.86 determine the Twiss β as functions of the aspect ratio, r, quadrupole k-strength, quadrupole length, and L_{eff}. In order that Twiss β transform correctly from one half-FODO cell to the next, Eqns. 4.45 and 4.85, are used. The result places a constraint on consecutive values of L_{eff};

$$L_{eff,2} = \frac{L_{eff,1}}{\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}}. \tag{4.118}$$

Generally, the iterative form for $L_{eff,(n+1)}$ is given;

$$L_{eff,(n+1)} = \frac{L_{eff,1}}{\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r(n)}\beta_{r(n)}}}. \tag{4.119}$$

The requirement of Eqn. 4.119 demonstrates that downstream values of L_{eff} are constrained by the values of $L_{eff,1}$ and the change in Lorentz factor. Cavity lengths have discrete lengths as they are comprised of an integer number of single cells (which must be a length that allows the synchronicity condition to bet met). The calculation of L_{eff} is also a function that depends on the acceleration from an rf cavity (change in Lorentz factor). In order to satisfy Eqn. 4.119, either the gap length, $l_{g,n}$, or the cavity length, $L_{cav,n}$, can be altered (indeed both can be altered simultaneously). Below, the case is shown for changing $l_{g,n}$, which solves the case for when a cavity is already designed, and alteration to the length is not possible. If a value of the first drift length, $l_{g,1}$, is defined, subsequent values can be calculated such that the constraint on consecutive L_{eff} is fulfilled for any cavity length and change in Lorentz factor. In order to produce the set of telescoping equations for $l_{g,n}$, firstly $l_{g,2}$ is determined. This is completed by rearranging $L_{eff,2}$ (from the equivalent form in Eqn. 4.115) for $l_{g,2}$ and enforcing Eqn. 4.118. $l_{g,2}$, is given by

$$l_{g,2} = \frac{\frac{L_{eff,1}}{\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}} - l_{cav,2}\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r2}-\gamma_{r1}} \ln\left(\frac{\gamma_{r2}\beta_{r2}+\gamma_{r2}}{\gamma_{r1}\beta_{r1}+\gamma_{r1}}\right)}{\frac{\gamma_{r1}\beta_{r1}}{\gamma_{2}\beta_{r2}}+1}. \tag{4.120}$$

The general form for the n_{th} half-FODO cell is

$$l_{g,n} = \frac{L_{eff,n} - l_{cav,n}\frac{\gamma_{r(n-1)}\beta_{r(n-1)}}{\gamma_{rn}-\gamma_{r(n-1)}} \ln\left(\frac{\gamma_{rn}\beta_{rn}+\gamma_{rn}}{\gamma_{r(n-1)}\beta_{r(n-1)}+\gamma_{r(n-1)}}\right)}{\frac{\gamma_{r(n-1)}\beta_{r(n-1)}}{\gamma_{rn}\beta_{rn}}+1}. \tag{4.121}$$

This result shows that provided the cavity lengths and Lorentz factors are known along the lattice, the subsequent gap lengths can be determined such that the constraint on consecutive values of L_{eff} is met (Eqn. 4.119).

Indeed, the cavity length in consecutive half-FODO cells can be altered in order to satisfy Eqn. 4.119, or a combination of both l_g and L_{cav}. It is shown later, that changing l_g only is not an ideal solution, as the real estate gradient drops.

The last constraint to be satisfied requires a constant aspect ratio, r. From Eqn. 4.82;

$$r_1 = \frac{1 + L_{eff,1}k_1 l_{q1,2} + k_1 l_{q1,2}^2 - \frac{L_{eff,1}k_1^2 l_{q1,2}^3}{2}}{1 - L_{eff,1}k_1 l_{q1,2} - k_1 l_{q1,2}^2 - \frac{L_{eff,1}k_1^2 l_{q1,2}^3}{2}} = r_2 = \frac{1 + L_{eff,2}k_2 l_{q2,2} + k_2 l_{q2,2}^2 - \frac{L_{eff,2}k_2^2 l_{q2,2}^3}{2}}{1 - L_{eff,2}k_2 l_{q2,2} - k_2 l_{q2,2}^2 - \frac{L_{eff,2}k_2^2 l_{q2,2}^3}{2}}. \tag{4.122}$$

One method to satisfy the constraints in Eqn. 4.122 is the following;

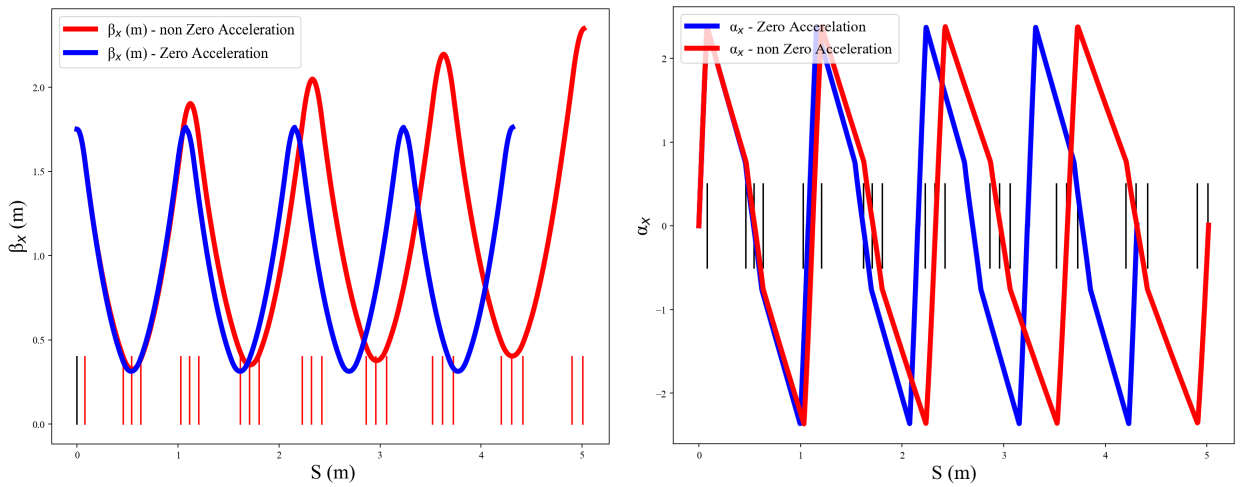
$$l_{q2,2} = \frac{l_{q1,2}}{\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}}. \tag{4.123}$$

Which provides a solution for Eqn. 4.117, the quadrupole length in the next half-FODO cell. By satisfying the constraints outlined in this section, the Twiss β functions transforms as described by Eqn. 4.45, and increase linearly with $\gamma_r\beta_r$. In addition, the beam size and aspect ratio return to the same constant value, at the boundary of each half-FODO cell.

In this section, the FODO-like scheme was analytically solved whilst incorporating longitudinal acceleration from RF cavities. Provided with the first quadrupole length, l_{q1}, first drift length, l_{g1}, cavity lengths, $l_{cav,n}$ and change in particle Lorentz factors, γ_{rn}, optimised FODO cell lattice elements can be determined along the entire lattice. The values calculated are such that the beam size is minimised at the cavity entrance/exit in order to maximise transverse beam capture for a given cavity length. A script was written in Python [93] that determined the complete set of optimised quadrupole parameters and drift lengths (k, l_q, l_g), as a function of cavity lengths and the Lorentz factors at each cavity. The Twiss parameters and phase advance are also calculated along the lattice.

4.5 FODO-like Lattice Analytical Results

The Twiss β_x and α_x functions are shown in Fig. 4.5. β_x alternates between a maximum and minimum at the end of each subsequent half-FODO cell. For zero acceleration, β_x returns to the same maximum and minimum. When acceleration is considered, the maximum β_x increases along the lattice, as the Lorentz factor increases after each rf cavity. Another point of interest is that the FODO cell period is not constant along the lattice. Horizontal lines show the placement of lattice elements (for the accelerating case). Figure 4.5 (b) displays



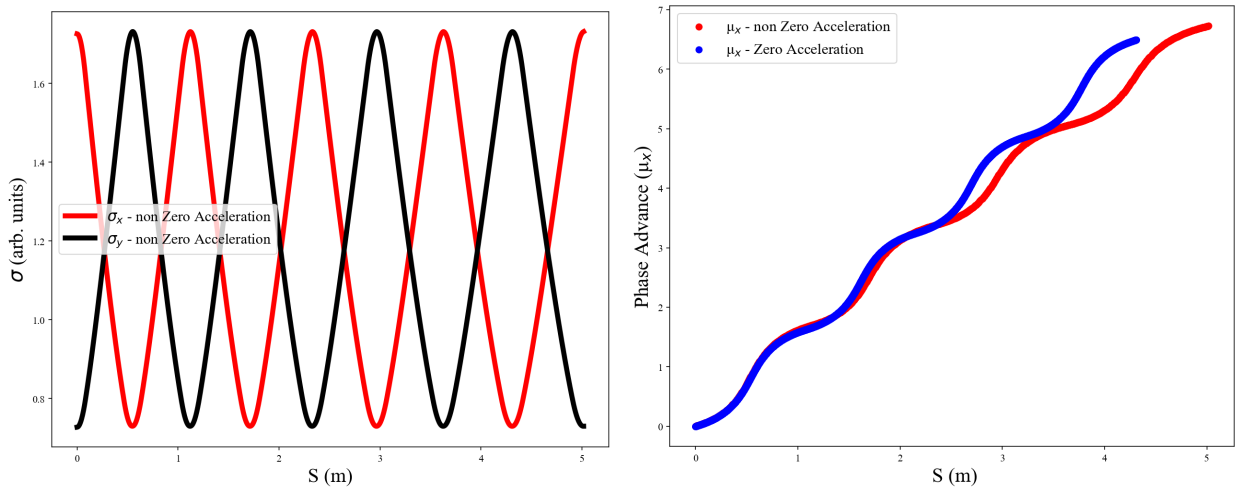
(a) (b)

Figure 4.5: Plot displaying calculated values of β_x (a) and α_x (b) along a FODO-like lattice comprised of 4 FODO cells in both an accelerating and non accelerating scheme.

the Twiss α_x function over the lattice, however is only calculated at a boundary between lattice elements, and extrapolated between with a straight line. α_x tends to zero at the end of each half-FODO cell, at a distance $l_{qn,1}$ into each quadrupole, as expected. The physical beam size in both transverse planes is shown in Fig. 4.6 (a). The maximum beam size is constant along the lattice, as constrained. At the first cavity entrance, the beam size in x is bigger than y. At the end of the cavity, the y beam size is bigger. The maximum beam size (in x or y) is the same at both cavity ends. Figure 4.6 (b) shows the betatron phase advance (in x) along the lattice. At a given point along the lattice, the phase advance in the non-accelerating case is larger than the accelerating case. Hence, the phase advance per unit length is smaller when acceleration is non-zero.

Figure 4.7 shows the value of β as calculated using two methods. The first methods calculates the expected β by the Lorentz propagation Eqns. 4.45. The second method propagates β along the lattice using the Twiss transformation matrix (Eqn. 4.49) with lattice parameters calculated using the iterating Eqns. 4.116, 4.117, 4.121and 4.123. Both methods calculate the same value for β along the lattice, showing the self-consistent nature of the method.

Figure 4.8(a) shows the beam size at the end of each half-FODO cell, for two different lattice schemes. The two schemes displayed are the theoretical scheme/FODO-like, and the constant lattice parameter/standard FODO scheme. The FODO-like scheme provides a constant beam size at the end of each half-FODO, as defined in the derivation of lattice elements. The standard FODO scheme uses the value of k_1 (calculated using Eqn. 4.107), l_{q1} and l_{g1} for all half-FODO cells. From Fig. 4.8(a), the beam size in the standard FODO



(a) (b)

Figure 4.6: Plot displaying calculated values of σ_x (a) and μ_x (b) along a FODO-like lattice comprised of 4 FODO cells in an accelerating scheme.

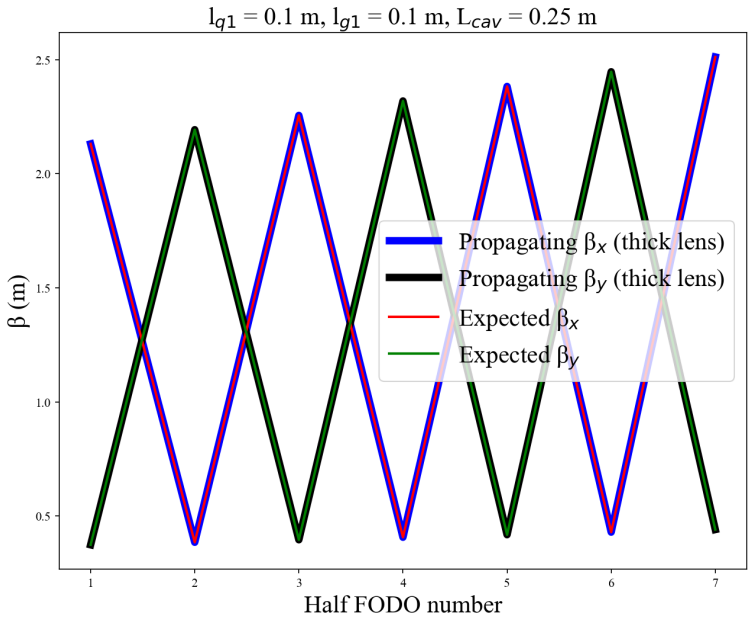
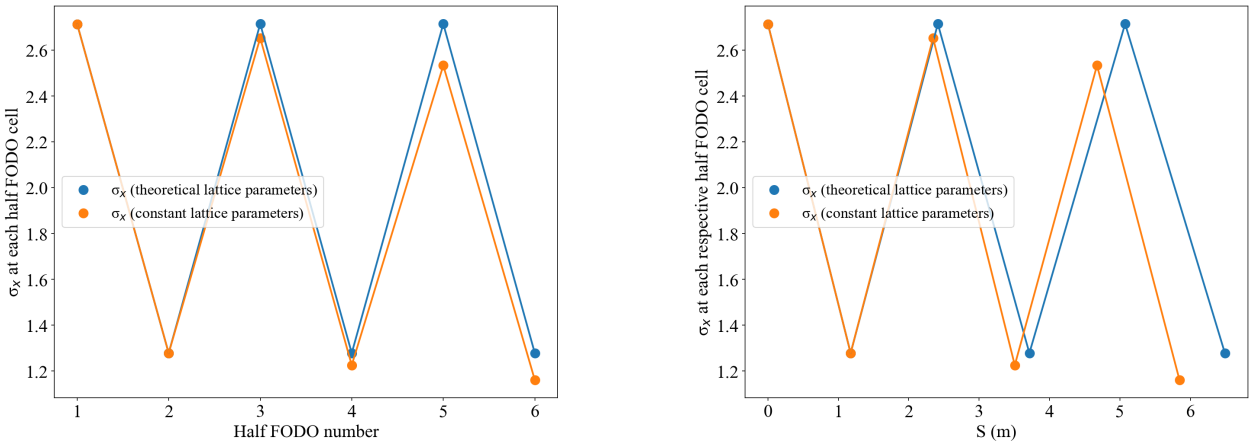


Figure 4.7: Twiss β function calculated with two methods (Eqns. 4.45 and 4.49). The starting lattice parameters are shown.

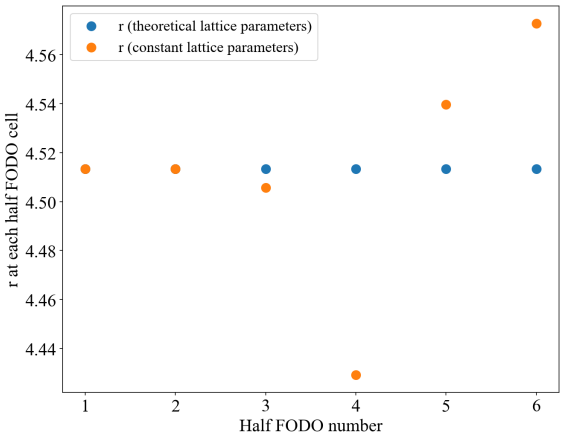


(a) (b)

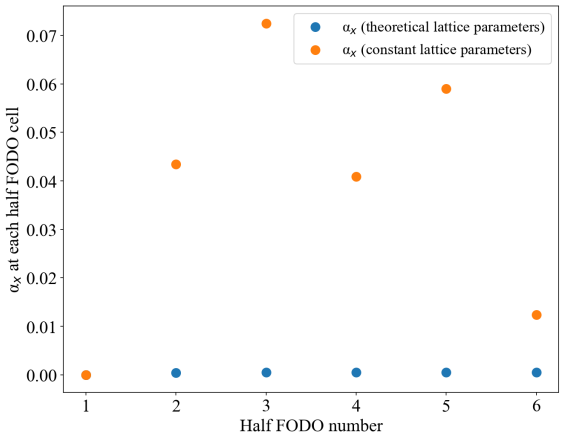
Figure 4.8: Beam size as a function of (a) half-FODO number and displacement (b) for constrained/theoretical lattice and constant lattice parameters.

scheme is not constant at the boundary between half-FODO cells. The increasing Lorentz factor is not counteracted by an increasing β_x and thus the beam size shrinks along the lattice. Figure 4.8(b) displays the beam size for both schemes as a function of longitudinal distance, s. The beam size in the standard FODO scheme does not return to the same constant value at the end of every half-FODO cell, but decreases along the lattice as the Lorentz factor increases. From Eqns. 4.77 and 4.123, consecutive values of quadrupole length and L_{eff} increase with Lorentz factor in the FODO-like scheme. As a result, consecutive half-FODO cell lengths increase gradually. This does not occur in the standard FODO scheme. The increase in period length is a function of the $\gamma_r\beta_r$, and as acceleration tends to 0, the lattice parameters tend to the standard FODO scheme values, as expected. This implies the role of acceleration in a FODO scheme is to stretch the lattice longitudinally.

Figure 4.9 (a) displays the ratio of beam sizes (r) at the end of each half-FODO cell (bigger transverse beam size divided by smaller transverse beam size). The FODO-like lattice produces the constant beam ratio, as expected. For the standard FODO scheme, the ratio is not constant, showing the beam is either over or under focused due to a miss-match in lattice parameters. This is because oscillations in the Twiss parameters are not coinciding with the end of consecutive half-FODO cells; the phase space ellipse does not return to the initial distribution after one period, it will be slightly misaligned. The misalign continues to grow with more periods (half-FODO cells). The FODO-like scheme is designed around maintaining this exact requirement. This result is again observed in Fig. 4.9 (b), which shows α_x at the end of every half-FODO cell. α_x in the FODO-like lattice returns to 0 at each half-FODO cell boundary, as required. The standard FODO scheme does constrain α_x to be 0 at each half-FODO boundary. As the value is consistently greater than 0, this suggests β_x



(a) β ratio at each half-FODO cell for constrained/theoretical lattice and constant lattice parameters



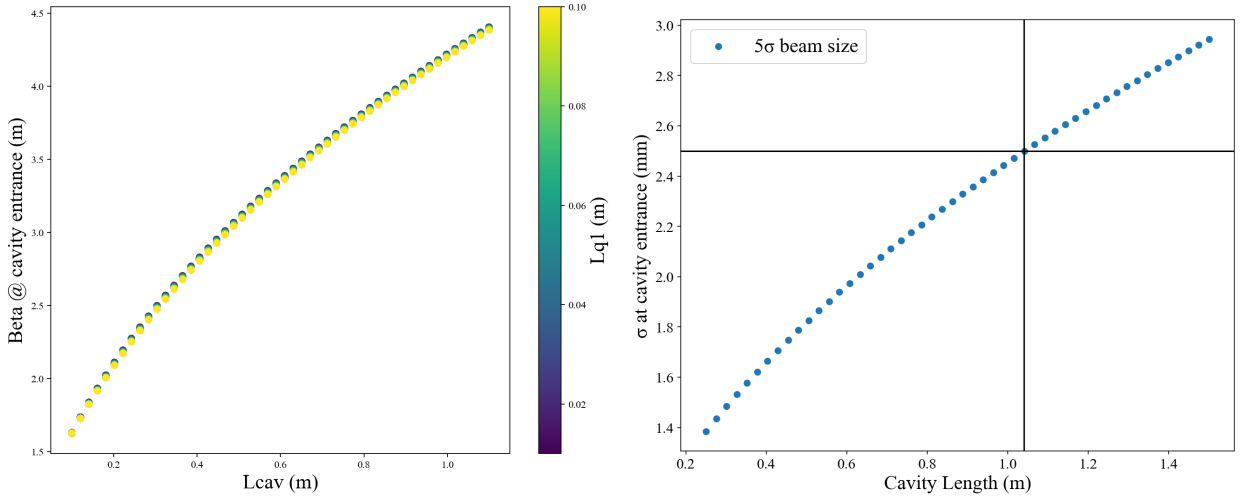
(b) α_x at each half-FODO cell for constrained/theoretical lattice and constant lattice parameters

Figure 4.9: The constrained FODO parameters at each half-FODO cell, r and α_x as calculated by the theoretical/constrained lattice and constant parameter lattice.

is decreasing to a minimum is has not yet reached, or decreasing after a maximum it has already passed. Whilst the scheme is relatively stable over the lattice, it is not periodic, as in the constrained system.

It was shown previously the transverse beam size was minimised with respect to the quadruple k-strength, as opposed to the quadrupole length. The reason for this is the insensitivity of β_{xc0} with respect to changes in l_{q1}. Figure 4.10 (a) shows the optimal minimised β_{xc0} as a function of cavity length, L_{cav}, and quadrupole length, l_{q1}. Each point has been optimised with respect to k_1. For large changes in l_{q1}, the optimum value of β_{xc0} is essentially unchanged. Thus, optimising β_{xc0} with respect the l_{q1} is not efficient. In fact, the optimal value of l_{q1} to minimise β_{xc0} for a given k_1 was $\sim$1 m, with a very broad minimum.

The main motivation behind the analytical FODO-like method was to determine the lattice parameters such that the maximum beam size at the cavity entrance/exit was minimised for a given cavity length. Figure 4.10 (b) shows the minimum beam size at the rf cavity entrance for different cavity lengths in the FODO-like focusing scheme. The limit on the rf single cell aperture is 2.5 mm. The normalised transverse rms beam emittance is 0.032 π mm mrad [94]. Once cavity lengths reach $\sim$1 m long, the 5σ transverse maximum beam size becomes greater than the 2.5 mm aperture. The study concludes that cavity lengths up to 1 m in length can be used until transverse losses become too great. This allows the option of exploring longer TW structures in addition to the previously assumed SW structure. The reason for this is the optimal performance of TW structures over SW structures as the cavity length increases. However, for lengths up to $\sim$1 m, either type could perform better. The



(a) (b)

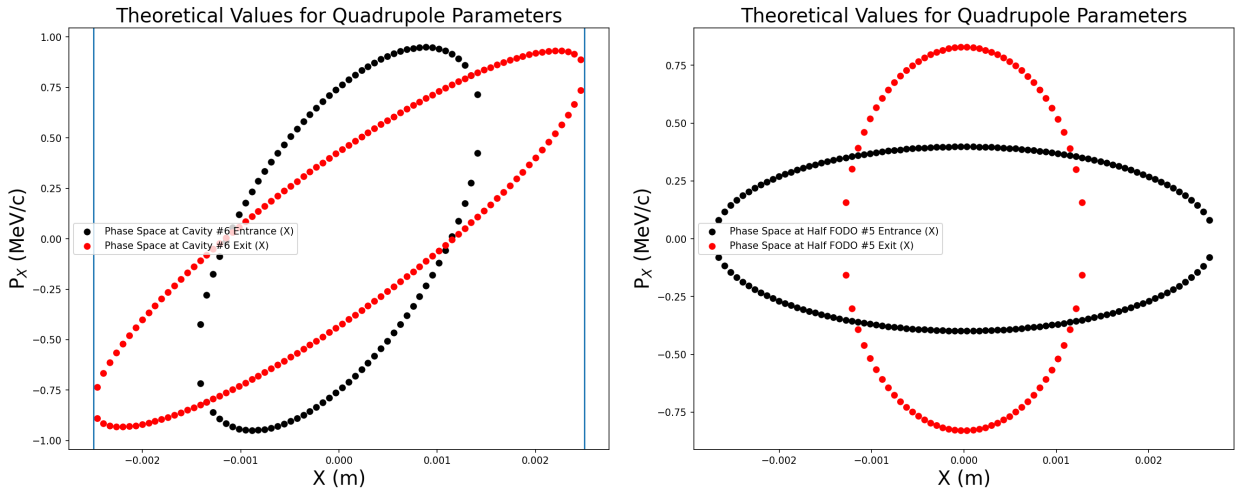
Figure 4.10: (a) Beta function at the cavity entrance as a function of cavity length and l_{q1} with constant k-strength. (b) Maximum beam size as a function of cavity length using the FODO-like focusing scheme.

next Chapter 6, describes the performance comparison between both accelerator types using a detailed analytical method.

As the analytical FODO-like method also calculates the Twiss parameters along the lattice, the phase space beam ellipse is known at any point. Consider a lattice comprised of three FODO cells, with each cell comprised of two rf cavities. The quadrupole and drift space parameters are determined using the FODO-like scheme, and the quadrupole k-strengths has been calculated such that the maximum beam size is minimised at the entrance of the cavity. Each cavity length is $\sim$1 m. Figure 4.11 (a) displays the 5σ phase space (x) beam ellipse at the entrance and exit of the sixth rf cavity. Figure 4.10 (b) demonstrated that cavity lengths $\sim$1 m will produce a 5σ beam size at the cavity entrance/exit, this is confirmed as the output phase space ellipse (red ellipse) fits within the beam aperture at the exit. The rf cavity map acts on the input ellipse to produce the output ellipse. If the cavity length had been longer, the beam ellipse will have defocused beyond the aperture limit and particles will have been lost on the cavity walls.

Figure 4.11 (b) shows the phase space (x) of a particle beam at the start and end of the fifth half-FODO cell. At the entrance of the half-FODO cell (black ellipse) β_x is maximum and $\alpha_x = 0$. The red ellipse describe the phase space ellipse at the end of the half-FODO system, at this point in the lattice β_x is minimum, and again $\alpha_x = 0$. As the constraints are enforced the lattice parameters evolve with the Lorentz factor, and at the boundary between half-FODO cells the aspect ratio and beam sizes return to the constant value.

Figure 4.12 shows the input and output beam ellipse over a half-FODO cell in standard



(a) (b)

Figure 4.11: 5σ phase space of beam ellipse at defined position in beam lattice. The system is described using the FODO-like scheme.

FODO scheme (constant lattice parameters). Whilst the phase space ellipse at the start exit of the half-FODO cell does not exactly coincide with minimum β_x and $\alpha_x = 0$, the error is small. Whilst the exact constraints to ensure periodicity between half-FODO cells is not met, the solution is stable over many cavities. The benefit of the standard FODO scheme is clear when considering the longitudinal acceptance. As the lattice parameters do not grow in length with the Lorentz factor, the longitudinal beam spread is smaller than the FODO-like system.

For linac systems with short cavities, the FODO-like system does not increase in length considerably, whilst retaining constant aspect ratio between FODO cells. This is because the increase in FODO cell length is most influenced by the increase in $l_{g,n}$. In the FODO-like treatment, the gap length absorbs any difference in required cavity length such that Eqn. 4.119 holds true. As the cavity length increases, the correction to consecutive $l_{g,n}$ increases. For cavity lengths $\sim$1 m, (and gradient 50 MeV/m) the gap length increases by $\sim$ 30% per FODO cell. For cavity lengths $\sim$ 0.3 m the gap length increases by $\sim$ 5% per cell. For drift spaces $\sim$ 0.1 m, a 5% increase in drift length is small. As a result, the FODO-like system (with constant cavity length) is more appropriate for linacs with shorter structures, as the classic FODO results are retained, and the real estate gradient does not decrease. For longer cavity lengths, the decrease in real estate gradient, due to the increasing gap length, does not justify the retainment of classic FODO requirements, such as constant aspect ratio. As discussed previously, other methods can be used to satisfy Eqn. 4.119, such as altering the cavity length. Figure 4.13 (a) displays the beam size as the boundary between half-FODO cells for the standard FODO and FODO-like schemes. In order to satisfy

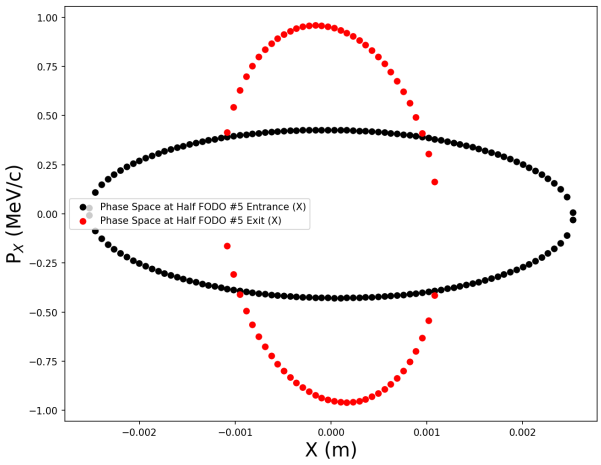
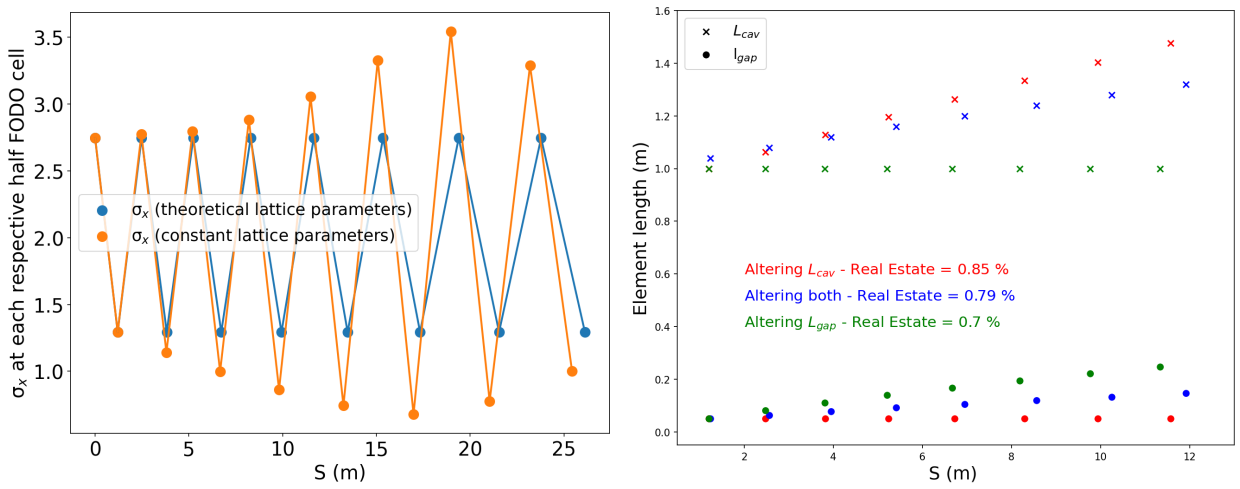


Figure 4.12: 5 σ Phase Space of beam ellipse at entrance and exit of fifth half-FODO cell using the semi-thin lens approximation. Lattice parameters are kept constant for each FODO cell.

Eqn. 4.119, the cavity length is increased, as opposed to the drift length. The FODO-like scheme produces constant beam size at each boundary, as required. The FODO scheme becomes unstable after multiple FODO cells. As all lattice parameters are constant, there is no correction for the increasing cavity length, and the standard FODO scheme breaks down. This is unlike the case of constant cavity length, where the standard FODO schemes performs well and remains stable. By definition, the FODO-like scheme is stable for all methods of satisfying Eqn. 4.119. Figure 4.13 (b) displays the length of l_g and L_{cav} along a lattice for three different methods for satisfying Eqn. 4.119. As expected, increasing the cavity length with constant drift length provides the maximum real estate gradient (the real estate here is defined as the ratio of accelerating length to total length). It can be shown that the case of constant drift length has a higher real estate than the standard FODO scheme, as consecutive cavity lengths increase faster than consecutive quadrupole lengths. Keeping the cavity length constant produces the lowest real estate gradient as the drift length increases most rapidly. The solution in blue shows a case where the constraint on $L_{eff,n}$ is met by changing $l_{g,n}$ and $L_{cav,n}$ simultaneously.

In reality, cavity lengths may be difficult to design to a particular length. In this case, a simultaneous correction to l_g and L_{cav} can be employed.



(a) (b)

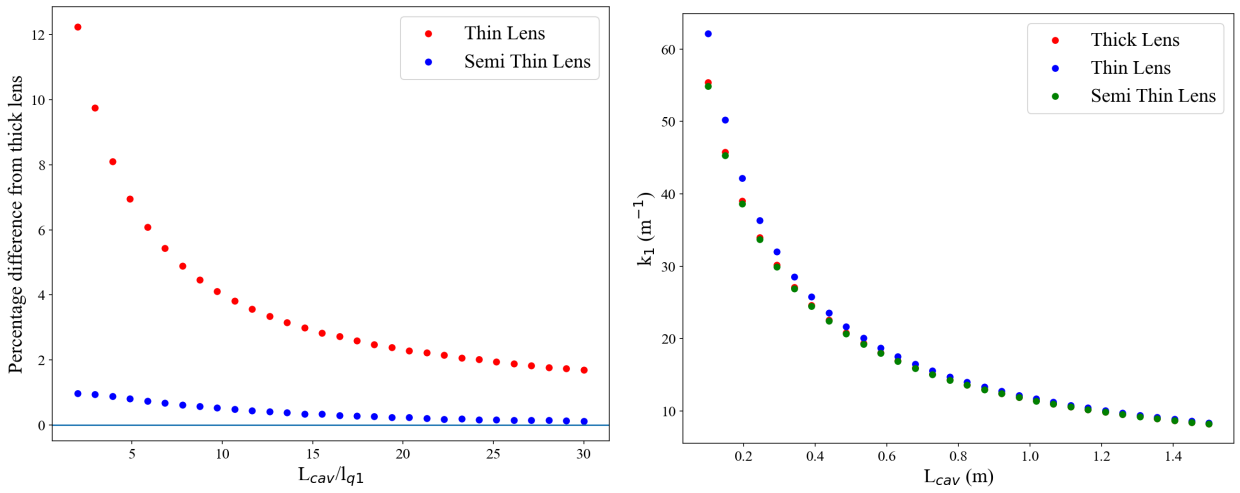
Figure 4.13: 5σ phase space of beam ellipse at defined position in beam lattice. The system is described using the FODO-like scheme.

4.6 Comparison between the Semi-thin and Thick Lens FODO-like Scheme

In order to determine the accuracy of the semi-thin lens (analytical) approximation, it was compared to the thick lens regime. The thick lens regime was solved using computational methods for the optimum k-strength such that the beam size at the cavity entrance was minimised.

Figure 4.14 (a) displays the percentage difference between the k-strength that minimises the maximum beam size at the cavity entrance as calculated by the thin, semi-thin, and thick lens regime. The percentage difference is relative to the thick lens regime. The semi-thin lens regime calculates a k-strength value that is very close to the thick lens regime, with errors under 1% for all considered cavity lengths. As the cavity length increases, the semi-thin lens approximation tends to a vanishing error, demonstrating the accuracy of the method. The thin lens regime shows high inaccuracy at short cavity lengths, and performs better with increasing cavity length, as expected. For the same value of L_{cav}/l_{q1}, the error of the thin and semi-thin lens regimes increase as l_{q1} increases.

Figure 4.14 (b) shows the three calculated k-strength values that minimise the beam size at the cavity entrance in a FODO-like scheme, as a function of cavity length, for a constant quadrupole length. As the cavity length increases, the optimised value of k_1 decreases. At very short cavity lengths, both the semi-thin and thick lens regimes calculate very similar vales for the k-strength. As the cavity length increases, the thin lens approximation tends toward the thick lens calculation. The semi-thin lens approximation for k_1 is almost exact,



(a) (b)

Figure 4.14: (a) The percentage difference between optimal value of k_1 as calculated by thick lens and the thin and semi-thin regimes, as a function of L_{cav}/l_{q1}, for constant $l_{q1} = 0.05$ m, $l_g = 0.05$ m. (b) Optimal value of k_1 as calculated by the thick, semi-thin, and thin lens regime, as a function of cavity length.

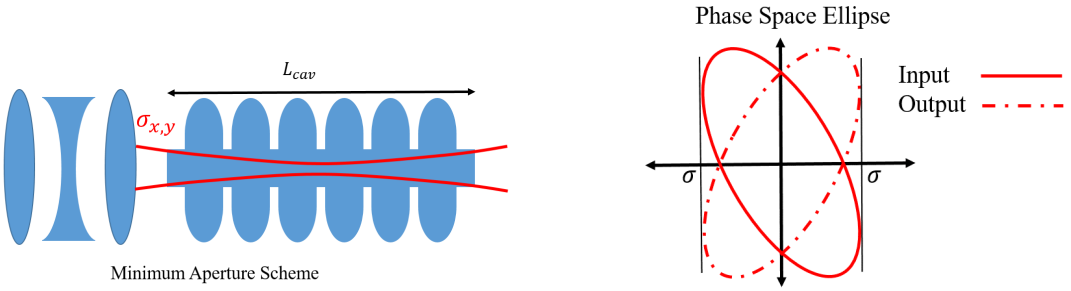
even for cavity lengths around 20 cm.

4.7 Minimum Aperture Scheme with Acceleration

In the previous section, a FODO-like scheme was discussed within a framework that allowed for acceleration. Another potential focusing scheme for the linac is the Minimum Aperture Scheme (MAS), which is also considered. The MAS uses short matching sections between cavities such that the beam ellipse at the cavity entrance is identical to the acceptance ellipse of the cavity. This results in an input beam that is simultaneously focusing in both the x and y planes. For zero acceleration, the beam size is minimised (beam waist) at the centre of the cavity. Figure 4.15 shows a schematic of the MAS focusing scheme alongside the expected input/output transverse phase space ellipse. The maximum beam size at the cavity entrance and exit is equal to the aperture of the cavity, thus motivating the name 'minimum aperture'. For a given beam emittance, the MAS allows the longest achievable cavity length.

When acceleration is considered, the location of the beam waist moves into the first half of the cavity. This is because as the beam is accelerated (as defined by the simple cavity map, with no defocusing terms), adiabatic damping takes place and the beam divergence reduces along the cavity. Thus, for a given step in longitudinal displacement, s, the maximum beam divergence, is reduced (recall $x' = \frac{p_x}{p_z}$). In order to reach the maximum beam size at the

cavity exit, the beam must have minimised before the mid-point, as additional longitudinal distance is required for a given amount of divergence.



(a) (b)

Figure 4.15: (a) Schematic of the MAS. The maximum beam size in both transverse planes occurs at the start and end of the cavity. The waist occurs at some point within the cavity. (b) The phase space ellipse at the entrance (input) and exit (out) of the cavity.

Recall the FODO-like scheme, which produced the smallest beam size possible after traversing a single quadrupole and drift space. In the MAS, the matching section is not constrained in this way, and can therefore produce a beam ellipse that matches the acceptance ellipse of a cavity. The MAS does, however, place a limit on the cavity length, due to finite value of the geometric emittance, and therefore finite value of beam divergence.

Recall the simple cavity transfer map, R (from Eqn. 4.29),

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \begin{pmatrix} 1 & L_{cav}\frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{z1}+\gamma_{r1}}{\gamma_{r0}\beta_{z0}+\gamma_{r0}}\right) \\ 0 & \frac{\gamma_{r0}\beta_{z0}}{\gamma_{r1}\beta_{z1}} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix}.$$

The following discussion calculates the maximum cavity length in the MAS. It is assumed the beam has already traversed the matching section and is at the cavity entrance. The MAS requires the beam size at the cavity entrance and exit to be equal. This is equivalent to

$$\sqrt{\frac{\beta_{xc0}\varepsilon_n}{\gamma_{r0}\beta_{r0}}} = \sqrt{\frac{\beta_{xc1}\varepsilon_n}{\gamma_{r1}\beta_{r1}}}. \tag{4.124}$$

For a given normalised emittance, ε_n, input/output Lorentz factor γ_{r0}/γ_{r1}, input/output relativistic beta, β_{r0}/β_{r1}, and input/output Twiss β, β_{xc0}/β_{xc1}. This evaluates to the following;

$$\beta_{xc0} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\beta_{xc1}. \tag{4.125}$$

Recall that the Twiss parameters can be translated from one point in the lattice to another using the Twiss transfer matrix (Eqn. 4.4) provided the mapping elements are known. The Twiss β will transform as the following;

$$\beta_{xc1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}(R_{11}^2\beta_{xc0} - 2R_{11}R_{12}\alpha_{xc0} + R_{12}^2\gamma_{xc0}). \tag{4.126}$$

Where R is the simple cavity transfer map. Substituting elements;

$$\beta_{xc0}\frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\Bigg(\beta_{xc0} - 2L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)\alpha_{xc0} + \left(L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)\right)^2\gamma_{xc0}\Bigg). \tag{4.127}$$

Using that $\gamma_x = \frac{1+\alpha_x^2}{\beta_x}$ and solving for a quadratic in α_{xc0},

$$0 = -2\alpha_{xc0}\beta_{xc0} + L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)(1+\alpha_{xc0}^2). \tag{4.128}$$

However, β_{xc0} is known, as the geometric emittance and cavity aperture (beam size) are defined variables. Solving the quadratic in Eqn. 4.128 provides a solution for α_{xc0};

$$\alpha_{xc0} = \frac{\beta_{xc0} \pm \sqrt{\beta_{xc0}^2 - \left(L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)\right)^2}}{L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)}. \tag{4.129}$$

There exists only one solution such that the beam is focusing at the cavity entrance and defocusing at the exit, with a minimum at some point along the cavity. In order to have only one solution, the radical term in Eqn. 4.129 must be 0, thus;

$$\beta_{xc0} = L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right). \tag{4.130}$$

The value for α_{xc0} is given;

$$\alpha_{xc0} = 1, \tag{4.131}$$

and

$$\gamma_{xc0} = \frac{2}{L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)}. \tag{4.132}$$

The value of α_{xc0} is expected to be positive, recall $\alpha_x = -\frac{1}{2}\frac{d\beta_x}{ds}$ thus a focusing beam will have $\frac{d\beta_x}{ds} < 0$ and $\alpha_x > 0$. The values of α_{xc1} and γ_{xc1} are found using the Twiss transfer matrix,

$$\alpha_{xc1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\Big(-R_{11}R_{21}\beta_{xc0} + (R_{11}R_{22}+R_{12}R_{21})\alpha_{xc0} - R_{12}R_{22}\gamma_{xc0}\Big). \tag{4.133}$$

Substituting the simple cavity map elements;

$$\alpha_{xc1} = \frac{\gamma_{r1}\beta_{r1}}{\gamma_{r0}\beta_{r0}}\left(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\alpha_{xc0} - L_{cav}\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}-\gamma_{r0}}\ln\left(\frac{\gamma_{r1}\beta_{r1}+\gamma_{r1}}{\gamma_{r0}\beta_{r0}+\gamma_{r0}}\right)\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\gamma_{xc0}\right) \tag{4.134}$$

before simplifying,

$$\alpha_{xc1} = \alpha_{xc0} - 2 = -1. \tag{4.135}$$

Lastly, for γ_{xc1},

$$\gamma_{xc1} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}}\gamma_{xc0}. \tag{4.136}$$

The MAS provides the required values of the Twiss parameters at the cavity entrance such that the cavity can be as long as possible, for a given cavity aperture. A matching section is required in order to match the correct values of the Twiss parameters in both x and y. As the matching sections may be comprised of multiple quadrupoles and drift spaces, they will be longer than the focusing section in a FODO cell focusing. The number of quadrupoles increases directly proportionally to the number of cavities in the FODO-like scheme.

Consider a linac comprised of 10 cavities. Suppose the focusing scheme used is the FODO-like scheme. There would be a total of 10 quadrupoles and 20 drift lengths. Suppose the beam emittance allowed a 10 m long cavity in the MAS. In this case, the MAS matching section may be comprised of fewer than 10 quadrupoles and 20 drift lengths. In this case, the MAS may provide a higher real estate gradient. Thus, each linac system must accurately compare both schemes. In practise, the power requirements of very long rf structures becomes a limiting factor, and must be taken into account. Chapter 6 discusses the power requirements of different types of rf accelerating structure in detail.

It was shown previously that the FODO-like scheme could produce a stable solution for cavity lengths $\sim$1 m (for a given beam emittance and cavity aperture, and 5σ beam size). The maximum cavity lengths using MAS are calculated by re-arranging for L_{cav} in Eqn. 4.130, for a defined value of β_{xc0};

$$L_{cav} = \frac{\beta_{xc0}}{\ln\left(\frac{\sqrt{\gamma_{r1}^2-1}+\gamma_{r1}}{\sqrt{\gamma_{r0}^2-1}+\gamma_{r0}}\right)}\left(\frac{\gamma_{r1}-\gamma_{r0}}{\sqrt{\gamma_{r0}^2-1}}\right). \tag{4.137}$$

Where $\gamma_r\beta_r = \sqrt{\gamma_r^2-1}$. Substituting $\gamma_{r1} = \gamma_{r0} + \delta\gamma_r$ and assuming $\delta\gamma_r << \gamma_{r0}$ allows simplifications to be made.

$$L_{cav} = \frac{\beta_{xc0}}{\ln\left(1+\frac{\delta\gamma_r}{\sqrt{\gamma_{r0}^2-1}}\right)}\left(\frac{\delta\gamma_r}{\sqrt{\gamma_{r0}^2-1}}\right), \tag{4.138}$$

for small $x, \ln(1+x) = x$ and the final solution is

$$L_{cav} = \beta_{xc0}. \tag{4.139}$$

The above result is a good approximation for $\delta\gamma^2 << \gamma_{r0}$, which is likely for proton accelerators, as $\delta\gamma_r$ would have to be of the order 1 GeV. In low energy electron linacs this assumption may not be accurate, as $\delta\gamma_r$ may be of a similar order to γ_{r0}.

Recall, the normalised emittance is taken to be $\approx 0.032 \times 10^{-6}$ πm rad, [94]. For protons at 150 MeV the value of $\beta_r\gamma_r$ is ~ 0.58. The designed cavity has an aperture of 2.5 mm, which is treated as the 5σ beam size. The maximum β_{xc0} is therefore

$$\beta_{xc0}|_{5\sigma} = \sqrt{\frac{\sigma^2\beta_r\gamma_r}{\varepsilon_n}} \approx \sqrt{\frac{(2.5\times10^{-3})^2\times 0.58}{0.032\times10^{-6}}} = 10.6 \text{ m} \sim L_{cav}|_{max}. \tag{4.140}$$

Hence, the MAS can realise cavity lengths up to ~ 10 m. This is around 10 $\times$ longer than cavities in the FODO-like scheme. As both scheme can allow long structures, TW cavities were subsequently explored. As TW cavities can be highly efficient at lengths of around 1 m, there may be little to gain in producing longer structures. In addition, longer structures require higher input power, which can raise issues at the input port. The cost of higher peak power increases exponentially and thus longer structures that produce high real estate gradients (due to the short matching sections relative to active length) may come with a very high cost.

4.8 Conclusion of Transverse Beam Dynamics

In this chapter, the transverse beam dynamics with non-zero acceleration was explored with respect to two focusing schemes, the FODO-like scheme and the MAS. An rf cavity transfer map was constructed that increased the longitudinal momentum of a particle beam. This was incorporated into an analytical description of the two focusing schemes, in which various important results were found. In the FODO-like scheme, the results provided the lattice parameters (quadrupole k-strength, quadrupole length, drift length) such that the maximum beam size was minimised at the cavity entrance. For both schemes, the maximum cavity length was established, such that the beam was transmitted at near 100% for a 5σ beam size. Given the normalised transverse emittance of the beam ($\sim$0.032 $\times$ 10^{-6} πm rad), the FODO-like scheme could realise cavity lengths around 1 m, whilst the MAS could realise lengths around 10 m. These values assume a cavity aperture of 2.5 mm. Cavity lengths over 1 m long are generally more difficult to construct, due to the higher input power and physical manufacturing of structures with many cells. In addition, the FODO-like scheme is a simpler scheme, as it does not require complicated matching sections. The quadrupole k-strengths and lengths have also been determined analytically for the FODO-like scheme, with high accuracy, due to the semi-thin lens approximation.

The next step in the design will establish the optimum, if any, cavity length for the final cavity design, as a function of input power. This essentially compares the optimum SW and TW cavity designs. Generally, shorter structures are more efficient as SW, and longer

structures are more efficient as TW. There are many factors that influence the choice of both cavity types, such as the fill time, total required power and real estate gradient. In order to complete this step, an analytical model was created that approximates the energy gain of a structure as function of input power, for both Standing and TW structures.

Chapter 5

Fast Cell to Cell Tracking Function

5.1 1D Single Cell Energy gain Function

A particle traversing an rf single cell excited in a TM_{010} cavity mode will gain or lose energy depending on the phase of the rf field as the particle enters the cell. For one cell, half of all available phases with produce net acceleration, and the other half will decelerate a particle. The synchronicity condition provides the cell length such that a particle traverses a cell in the time it takes for the field to advance by ψ, where ψ is the phase advance per cell. When the synchronicity condition is met, a particle traveling at the synchronous beta, β_s, will continue to see accelerating field in each subsequent cell. In reality, a particle's velocity increases as the energy increases, and the particle will move further from the synchronous beta. rf acceleration is appropriate for protons with $\beta \sim 0.3$ [95] or greater (the coupled cavity structure will start accelerating protons at 37.5 MeV in the medical linac) due to the effective shunt impedance per unit length and achievable gradient being superior to other acceleration methods, such as the RFQ and drift tube linac. Common practise for determining particle energy and position through rf structures is the use of time-stepping particle tracking codes, such as ASTRA [96], OPAL [97] and RF-Track [98]. In order to accurately simulate particle coordinates, the time steps are often short and tracking large numbers of particles through many rf cells becomes computationally expensive. As rf single cells have an energy gain that is of the order 1 MeV, protons at $\sim$ 40 MeV will experience slow changes in relativistic beta. This can not be said for electrons, which become relativistic after a few rf cells, changing velocity rapidly, due to electrons being $\sim$ 2000 times less massive. The small gain in proton relativistic beta over an rf cell allows them to be approximated at constant velocity over a given rf cell, which allows an analytical calculation for the energy gain of a proton over a single cell in one computation, allowing for fast tracking simulations.

Consider a single cell on-axis longitudinal electric field, $E_z(s)$, and s is the longitudinal displacement through the rf cell. A SW rf field will vary as $\cos(\phi_0 + \omega t)$. Assuming the proton velocity is constant over a single cell allows a change of variable from t to s using

$t = \frac{s}{\beta c}$. Therefore, the energy gain (technically, the voltage gain) of a particle as it travels over a single cell is given by;

$$E_{\text{out}} = E_{\text{in}} + \int_0^{L_{cell}} E_z(s)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right) ds. \tag{5.1}$$

Where E_{in} is the input energy, and the particle phase, ϕ_0, is defined as the phase when the particle is at the entrance of the cell. In general, the on-axis E_z can be written as a Fourier series, as the field in a coupled cavity linac is considered periodic for this work.

$$E_z(s) = E_0\left(\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2n\pi s}{P}\right) + b_n \sin\left(\frac{2n\pi s}{P}\right)\right) \tag{5.2}$$

Where E_0 is the maximum on-axis field gradient. For a $\pi-$mode structure, the field periodicity is two cells. Using a parity argument, it can be understood the on-axis E_z component over one wavelength does not differentiate between $+z$ and $-z$ directions, thus $a_0 = 0$. As the field is odd about $s = 0$, all a_n terms are zero, and only b_n terms survive when constructing the Fourier series coefficients. The integration region is [0, L_{cell}] (however the bounds can be any distance) and therefore only the first half of a Fourier series period ($L_{cell} = \frac{\lambda}{2}$) for SW is ever used. The reason for constructing a two-cell Fourier series is because a single cell Fourier series describing $E_z(s)$ would repeat every cell, and would require a_0, a_n, b_n terms. For structures with different phase advance per cell, the Fourier series will require both a_n and b_n terms, and is shown later in the case of TW accelerators. Figure 5.1 shows the first half period of a Fourier series description of the on-axis $E_z(s)$ component, and the corresponding Fourier terms. Keeping only b_n terms for the Fourier series description of $E_z(s)$.

$$E_z(s) = E_0 \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi n s}{2L_{cell}}\right). \tag{5.3}$$

Inserting Eqn. 5.3 into Eqn. 5.1 and using the synchronicity condition for a π phase advance requires $\pi = \frac{\omega L_{cell}}{\beta_s c}$.

$$E_{\text{out}} = E_{\text{in}} + \int_0^{L_{cell}} E_0 \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\omega s}{\beta_s c}\right) \cos\left(\phi_0 + \frac{\omega s}{\beta c}\right) ds \tag{5.4}$$

As the integral is over independant variable, s, the summation commutes with the integral and can be moved outside. The integral can be computed by using the trigonometric formula;

$$\sin(a)\cos(b) = \frac{1}{2}(\sin(a-b) + \sin(a+b)). \tag{5.5}$$

The integral now takes the form;

$$E_{\text{out}} = E_{\text{in}} + \frac{E_0}{2} \sum_{n=1}^{\infty} \int_0^{L_{cell}} b_n \left(\sin\left(\frac{n\omega s}{\beta_s c} - \phi_0 - \frac{\omega s}{\beta c}\right) + \sin\left(\frac{n\omega s}{\beta_s c} + \phi_0 + \frac{\omega s}{\beta c}\right)\right) ds \tag{5.6}$$

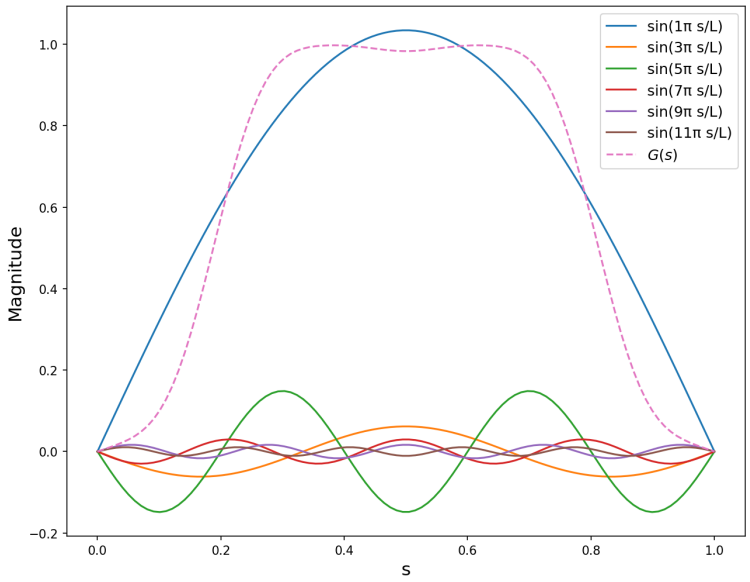


Figure 5.1: First half period of on-axis E_z profile, G(s), for a single cell with an aperture. The individual Fourier terms are also shown.

integrating the above produces;

$$E_{\text{out}} = E_{\text{in}} + \frac{E_0}{2}\sum_{n=1}^{\infty} b_n \left[\frac{\cos(s(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c})} - \frac{\cos(s(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c})} \right]_0^{L_{cell}}. \tag{5.7}$$

Taking out a factor of L_{cell} and using the synchronicity condition, $\frac{\pi\beta_s c}{\omega} = L_{cell}$ to simplify produces the following

$$E_{\text{out}} = E_{\text{in}} + \frac{E_0 L_{cell}}{2\pi}\sum_{n=1}^{\infty} b_n \left[\frac{\cos(\frac{\beta_s}{\beta}\pi - n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} - n)} - \frac{\cos(\frac{\beta_s}{\beta}\pi + n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} + n)} \right]. \tag{5.8}$$

The energy gain is thus a function of $\frac{\beta_s}{\beta}$, the phase when the particle enters the cell, ϕ_0, and the Fourier series coefficients, b_n. The effective transit time factor (different to the classic transit time factor in literature) is defined here as the term that multiplies $E_0 L_{cell}$, and describes the phase dependence of the energy gain over one cell, in addition to the time variation of the field as the particle traverses the cell. It can be seen the phase can not be separated from the effective transit time factor. Figure 5.2 shows the effective transit time factor for protons near 40 MeV accelerated over one cell, as a function of rf phase. Changes in proton energy (and therefore proton beta) do not have a huge effect on the shape

of the effective transit time factor, other than a slight horizontal and magnitude shift, as different proton beta will achieve maximum energy gain at a slightly different input phase, ϕ_0. The largest value of the effective transit time factor from the plot is approximately 0.5. This suggests the effective gradient of an rf cell for protons at this energy is roughly half that of the maximum gradient. Indeed it is possible to define a different integration region

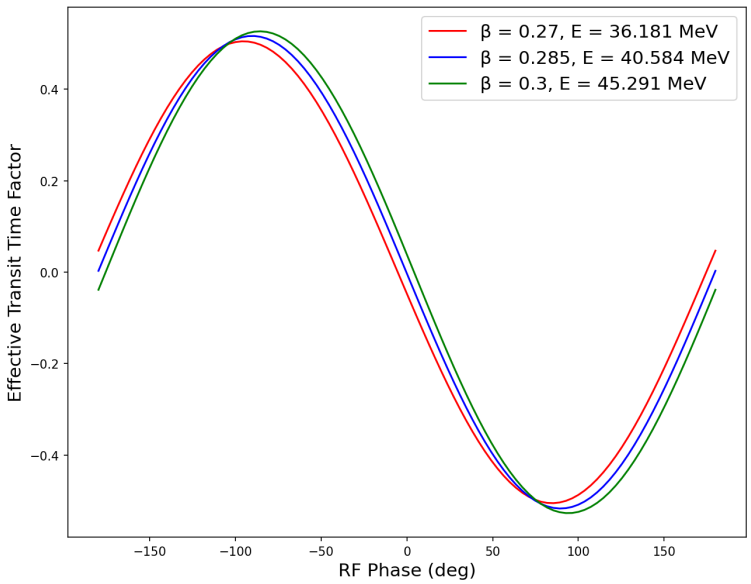


Figure 5.2: The Effective Transit Time Factor for protons around 40 MeV as a function of phase.

and determine energy gain over different length scales. However, as the relativistic beta is assumed constant over the integration region the accuracy of the analytical model will decrease with increasing integration region, as a particle will have gained/lost more energy. It is just as feasible to integrate less than a single cell length in one iteration, however this would increase the number of computations relative to the single cell iteration.

A particle traversing a cell at a different beta to the synchronous beta will not arrive at the following cell at the same phase, due to phase slippage. Phase slippage is the term used to describe the mismatch between relativistic beta and synchronous beta culminating in a different phase at the subsequent cell. This can be visualised as follows: consider a particle traversing a cell at a relativistic beta less than the synchronous beta. The phase advances such that the synchronous particle arrives at the next cell at the exact same phase, hence ϕ_0 = constant. The slower particle will arrive late relative to the phase, and the phase will have evolved an additional amount, due to the delay in the slow particle traversing the cell more slowly. As a result, the value of a particle's phase, ϕ_0, will need to be updated after each iteration, to account for any phase slippage. In order to calculate the phase at the start

of the subsequent cell;

$$\phi_{0,n+1} = \phi_{0,n} + \Delta\phi_n, \tag{5.9}$$

where $\phi_{0,n+1}$ and $\phi_{0,n}$ are the particle phase at the start of the $(n+1)^{th}$ and n^{th} cell. The phase slippage is defined as;

$$\Delta\phi_n = \omega\Delta t_n = \omega\left(\frac{L_{cell}}{\beta_n c} - \frac{L_{cell}}{\beta_s c}\right). \tag{5.10}$$

Δt_n is the time difference between a synchronous particle and a particle of velocity $\beta_n c$ traversing an rf cell of length L_{cell}. Therefore the phase slip of a particle over an rf cell is given by;

$$\Delta\phi_n = \frac{\omega L_{cell}}{c\beta_s}\left(\frac{\beta_s}{\beta_n} - 1\right). \tag{5.11}$$

The term outside the bracket is the phase advance per cell, ψ,

$$\Delta\phi = -\psi\left(1 - \frac{\beta_s}{\beta}\right). \tag{5.12}$$

Thus,

$$\phi_{0,n+1} = \phi_{0,n} - \psi\left(1 - \frac{\beta_s}{\beta}\right). \tag{5.13}$$

A particle with $\beta_n < \beta_s$ translates to $-\psi\left(1 - \frac{\beta_s}{\beta}\right) > 0$ and the phase at the next cell has evolved, $\phi_{0,n+1} > \phi_{0,n}$. There is an implicit approximation when calculating the phase slippage from one cell to the next. It is assumed each cell is oscillating out of phase by the phase advance per cell exactly. For cavity models in EM solver software, the phase advance is often exactly correct due to enforcing symmetry conditions. Generally structures may have slightly different phase advances from one cell to the next (even after tuning the phase advance in the latter stages of the design). This problem can be combated by calculating the phase advance from one cell to the next using EM solvers and using that value when updating the phase at each cell.

Figure 5.3 shows how the phase of a particle slips over the course of a cavity for protons in the 40 MeV range. Maximum acceleration occurs when the phase of a particle returns to the original value upon cavity exit. As a particle enters at a velocity less than the synchronous beta, the observed phase starts to grow as it evolves faster than a particle traverses an rf gap. As the particle increases in velocity, the rate at which the phase evolves reduces until the particle is faster than the synchronous particle and the observed phase begins to reduce until the original phase is observed, at the cavity exit. The off phase particle has a starting phase slightly too high, as it is slower than β_s, the phase grows as the particle slips until $\beta = \beta_s$. At this point, the particle travels more quickly than β_s and the particle phase starts

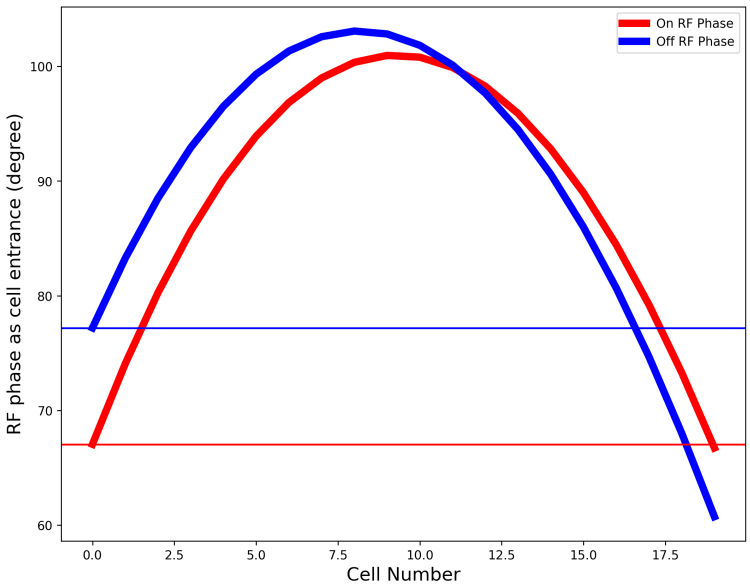


Figure 5.3: phase of an ideal (on phase) and not ideal (off phase) particle traversing an rf cavity.

to reduce. As the initial gain is too great, the particle slips too far before the end of the rf cavity.

The equations outlined in this section calculate the energy gain of a particle of velocity, βc, traversing a single cell of length L_{cell}. The particle enters the cell at an phase of ϕ_0. The cell length and synchronous beta are related by the previously described relationship $\psi = \frac{\omega L_{cell}}{\beta_s c}$. This derivation was calculated in 1D, where the value of E_0 is the maximum of the on-axis field $E_z(s)$, with units of MeV/m. The 1D equation shows very strong agreement with the well bench-marked tracking code ASTRA [96]. Figure 5.4 displays the longitudinal phase space as calculated by both ASTRA and the single cell function. There is a high level of agreement between the two methods and confirms the approximation of constant relativistic beta over an rf cell is valid. The method described can be expanded to a 3D model, tracking the 6D phase space of a particle, x, p_x, y, p_y, z, p_z. Moving forward, the single cell function will be refered to as FC2CT (Fast Cell-to-Cell Tracking).

5.2 Non-zero Field Components in a TM_{010} Mode

The previous section described the method used to obtain an analytical function that calculates the energy gain of a particle over a single rf cell for a given relativistic beta, phase, and cell length. The on-axis longitudinal electric field was described with a Fourier series.

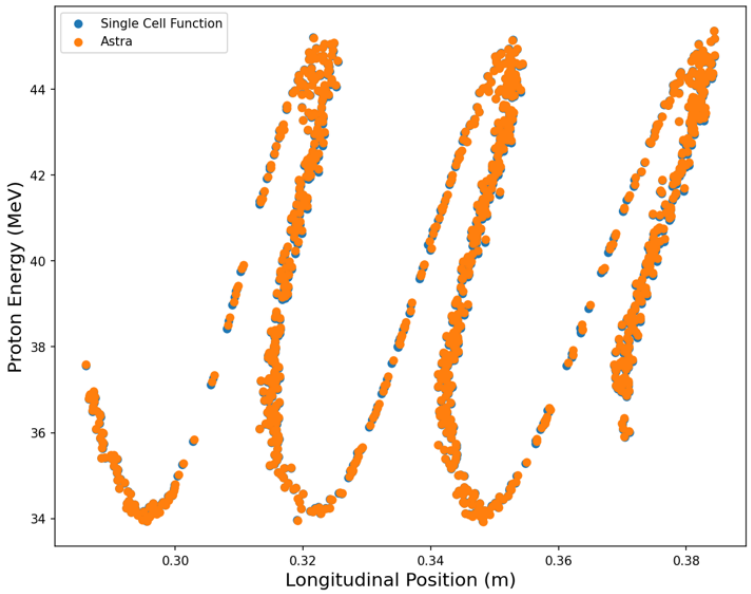


Figure 5.4: 1D phase space plot showing energy of proton beam tracked through 30 cell structure as calculated by ASTRA and single cell function. On-axis field profile was taken from π-mode single cell model in CST.

In order to extend FC2CT to calculate the 6D phase space of a particle, all non-zero field components must be taken into account, $E_r(s)$ and $B_\theta(s)$. For all TM monopole modes, B_z is zero, using $\nabla.\vec{B} = 0$ and as B_θ has azimuthal symmetry, $B_r = 0$. Recall E_r is related to $\frac{dE_z}{dz}$, and is zero in a pillbox cavity. When an aperture is introduced, $E_z(s)$ becomes a function of s and $E_r(s)$ is non-zero. The only zero electric field component is E_θ, from $\nabla \times \vec{B} = \frac{1}{c^2}\frac{d\vec{E}}{dt}$ in a vacuum. The fields $E_r(r, \theta, s)$ and $B_\theta(r, \theta, s)$ are first expanded as Taylor series about $r = 0$, in the small r limit. From the results obtained previously, the field values take on the forms of the Bessel functions of the first kind. The radial profile for E_z is a 0^{th} order Bessel function, whereas E_r and B_θ are 1^{st} order Bessel functions. As E_z is an even function about $r = 0$, only even powers of r will exist in the expansion. E_r and B_θ are odd in r, and only odd powers exist. For E_z it is convenient to expand in Cartesian coordinates, x and y, in order to make the derivation simpler.

$$E_z(x, y, z) = E_z(0, 0, z) + \frac{1}{2!}\frac{\partial^2 E_z(0, 0, z)}{\partial x^2}x^2 + \frac{1}{2!}\frac{\partial^2 E_z(0, 0, z)}{\partial y^2}y^2 \tag{5.14}$$

$$E_r(r, z) = \frac{\partial E_r(0, z)}{\partial r}r + \frac{1}{3!}\frac{\partial^3 E_r(0, z)}{\partial r^3}r^3 \tag{5.15}$$

$$B_\theta(r, z) = \frac{\partial B_\theta(0, z)}{\partial r}r + \frac{1}{3!}\frac{\partial^3 B_\theta(0, z)}{\partial r^3}r^3. \tag{5.16}$$

Using the wave equation;

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{E} = 0. \tag{5.17}$$

Separating the time variable from the spatial variables in Eqn. 5.17 and taking the z component;

$$\frac{\partial^2 E_z(0,0,z)}{\partial x^2} + \frac{\partial^2 E_z(0,0,z)}{\partial y^2} + \frac{\partial^2 E_z(0,0,z)}{\partial z^2} = -\frac{\omega^2}{c^2}E_z(0,0,z). \tag{5.18}$$

As the rf single cell has azimuthal symmetry (all field derivatives with respect to θ are 0);

$$\frac{\partial^2 E_z(0,0,z)}{\partial x^2} = \frac{\partial^2 E_z(0,0,z)}{\partial y^2}, \tag{5.19}$$

therefore,

$$\frac{\partial^2 E_z(0,0,z)}{\partial x^2} = -\frac{1}{2}\left(\frac{\omega^2}{c^2}E_z(0,0,z) - \frac{\partial^2 E_z(0,0,z)}{\partial z^2}\right). \tag{5.20}$$

From Eqn. 5.14,

$$E_z(x,y,z) = E_z(0,0,z) + \frac{1}{2!}\frac{\partial^2 E_z(0,0,z)}{\partial x^2}(x^2+y^2). \tag{5.21}$$

Combining Eqns. 5.20 and 5.21 before simplifying;

$$E_z(x,y,z) = E_z(0,0,z) - \frac{x^2+y^2}{4}\left(\frac{\omega^2}{c^2}E_z(0,0,z) - \frac{\partial^2 E_z(0,0,z)}{\partial z^2}\right). \tag{5.22}$$

Converting to cylindrical coordinates and introducing back the temporal term the E_z expansion is given;

$$E_z(r,z,t) = \left[E_z(0,z) - \frac{r^2}{4}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right)\right]\cos\left(\phi_0 + \omega t\right). \tag{5.23}$$

For expansions of E_r and B_θ the following Maxwell's equation is used;

$$\vec{\nabla}.\vec{E} = 0. \tag{5.24}$$

Expanded in cylindrical coordinates;

$$\frac{1}{r}\frac{\partial(rE_r(r,z))}{\partial r} + \frac{\partial E_z(r,z)}{\partial z} = 0, \tag{5.25}$$

where the term $\frac{1}{r}\frac{\partial E_\theta(r,z)}{\partial \theta}$ is zero from azimuthal symmetry. Substituting Eqn. 5.15 and 5.23 into Eqn. 5.25;

$$\frac{1}{r}\frac{\partial(r\left(\frac{\partial E_r(0,z)}{\partial r}r + \frac{1}{3!}\frac{\partial^3 E_r(0,z)}{\partial r^3}r^3\right)}{\partial r} + \frac{\partial\left(\left[E_z(0,z) - \frac{r^2}{4}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right)\right]\right)}{\partial z} = 0. \tag{5.26}$$

Omitting all terms of the form $\frac{\partial^n E_r(0,z)}{\partial r^n}$ for even n when differentiating.

$$\frac{\partial E_r(0,z)}{\partial r} + \frac{1}{3!}\frac{\partial^3 E_r(0,z)}{\partial r^3}r^2 + \frac{\partial E_r(0,z)}{\partial r} + \frac{3}{3!}\frac{\partial^3 E_r(0,z)}{\partial r^3}r^2 = -\frac{\partial E_z(0,z)}{\partial z} + \frac{r^2}{4}\left(\frac{\partial^3 E_z}{\partial z^3} + \frac{\omega^2}{c^2}\frac{\partial E_z(0,z)}{\partial z}\right) \quad (5.27)$$

Comparing coefficients of r;

$$\frac{\partial E_r(0,z)}{\partial r} = -\frac{1}{2}\frac{\partial E_z(0,z)}{\partial z} \quad (5.28)$$

$$\frac{\partial^3 E_r(0,z)}{\partial r^3} = \frac{6}{16}\left(\frac{\partial^3 E_z}{\partial z^3} + \frac{\omega^2}{c^2}\frac{\partial E_z(0,z)}{\partial z}\right). \quad (5.29)$$

Substituting the terms in Eqn. 5.28 and Eqn. 5.29 into Eqn. 5.15 and simplifying

$$E_r(r,z) = \left[-\frac{r}{2}\frac{\partial E_z(0,z)}{\partial z} + \frac{r^3}{16}\left(\frac{\partial^3 E_z}{\partial z^3} + \frac{\omega^2}{c^2}\frac{\partial E_z(0,z)}{\partial z}\right)\right]\cos(\phi_0 + \omega t). \quad (5.30)$$

Finally, the following Maxwell's equation is used for the radial expansion of B_θ

$$\vec{\nabla} \times \vec{B} = \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t}. \quad (5.31)$$

As B_θ is the only non-zero field component for the magnetic field in an rf cavity Eqn. 5.31 simplifies,

$$\frac{1}{r}\left(\frac{\partial\left(rB_\theta(r,z)\right)}{\partial r}\right) = \frac{1}{c^2}\frac{\partial E_z(r,z,t)}{\partial t}. \quad (5.32)$$

However, the magnetic field and the electric field differ by $\pi/2$, and hence a temporal term for the magnetic field will be of the form $\sin(\phi_0 + \omega t)$ (due to the initialisation of E_z having $\cos(\phi_0 + \omega t)$). As a result, differentiating Eqn. 5.23 with respect to time provides a sin term, and both temporal terms will cancel. Substituting the Taylor expansion from Eqn. 5.16 and using Eqn. 5.23 the derivation continues.

$$\frac{1}{r}\left(\frac{\partial\left(\frac{\partial B_\theta(0,z)}{\partial r}r^2 + \frac{1}{3!}\frac{\partial^3 B_\theta(0,z)}{\partial r^3}r^4\right)}{\partial r}\right) = \frac{\omega}{c^2}\left[E_z(0,z) - \frac{r^2}{4}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right)\right] \quad (5.33)$$

Again, all terms of the form $\frac{\partial^n B_\theta(0,z)}{\partial r^n}$ for even n are zero as the function is odd. Hence, computing the partial derivatives and comparing coefficients of r;

$$\frac{\partial B_\theta(0,z)}{\partial r} = \frac{\omega}{2c^2}E_z(0,z) \quad (5.34)$$

$$\frac{\partial^3 B_\theta(0,z)}{\partial r^3} = -\frac{6\omega}{16c^2}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right). \tag{5.35}$$

Substituting the terms in Eqns. 5.34 and 5.35 into Eqn. 5.16 and simplifying;

$$B_\theta(r,t) = \frac{\omega}{c^2}\left[\frac{r}{2}E_z(0,z) - \frac{r^3}{16}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right)\right]\sin\left(\phi_0 + \omega t\right). \tag{5.36}$$

5.3 Derivation of FC2CT in 3D

Recall that E_z can be written as a Fourier series,

$$E_z(s) = E_0\sum_{n=0}^{\infty} b_n \sin\left(\frac{n\pi s}{L}\right). \tag{5.37}$$

The field must be differentiated up to three times, as higher order derivatives appear in Eqns. 5.23, 5.30 and 5.36.

$$\frac{\partial E_z(s)}{\partial z} = \frac{E_0\pi}{L}\sum_{n=0}^{\infty} b_n n\,\cos\left(\frac{n\pi s}{L}\right) = E_z'(s) \tag{5.38}$$

$$E_z''(s) = -\frac{E_0\pi^2}{L^2}\sum_{n=0}^{\infty} b_n n^2\,\sin\left(\frac{n\pi s}{L}\right) \tag{5.39}$$

$$E_z'''(s) = -\frac{E_0\pi^3}{L^3}\sum_{n=0}^{\infty} b_n n^3\,\cos\left(\frac{n\pi s}{L}\right) \tag{5.40}$$

Derivatives of E_z are simply Fourier series with different Fourier coefficients. Equation 5.8 is the integrated version of Eqn. 5.4 and can be used to create a standard result for the integral of a Fourier series multiplied by a trigonometric function over the bounds [0, L]. Recall the result;

$$\int_0^L E_0\sum_{n=1}^{\infty} b_n\,\sin\left(\frac{n\omega s}{\beta_s c}\right)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds =$$
$$\frac{E_0 L}{2\pi}\sum_{n=1}^{\infty} b_n\left[\frac{\cos(\frac{\beta_s}{\beta}\pi - n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} - n)} - \frac{\cos(\frac{\beta_s}{\beta}\pi + n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} + n)}\right] \tag{5.41}$$
$$= \frac{E_0 L}{2\pi}F(b_n, \beta_s/\beta, \phi_0).$$

Where $F(b_n, \beta_s/\beta, \phi_0)$ is a function containing the summation. The momentum gain in each of the three planes is given by the Lorentz force.

$$d\vec{p} = \mathrm{q}\vec{E}dt + \mathrm{q}\vec{v}\times\vec{B}dt \tag{5.42}$$

Changing variables and expanding;

$$dp_x = \frac{\mathrm{q}}{v_z}E_x(s)ds + \frac{\mathrm{q}}{v_z}(v_yB_z - v_zB_y(s))ds \approx \frac{\mathrm{q}}{v_z}E_x(s)ds - \frac{\mathrm{q}}{v_z}(v_zB_y(s))ds \tag{5.43}$$

$$dp_y = \frac{\mathrm{q}}{v_z}E_y(s)ds - \frac{\mathrm{q}}{v_z}(v_xB_z - v_zB_x(s))ds \approx \frac{\mathrm{q}}{v_z}E_y(s)ds + \frac{\mathrm{q}}{v_z}(v_zB_x(s))ds \tag{5.44}$$

$$dp_z = \frac{\mathrm{q}}{v_z}E_z(s)ds - \frac{\mathrm{q}}{v_z}(v_xB_y(s) - v_yB_x)ds \approx \frac{\mathrm{q}}{v_z}E_z(s)ds. \tag{5.45}$$

It is convenient to project all particles into the first quadrant of the imaginary plane by taking

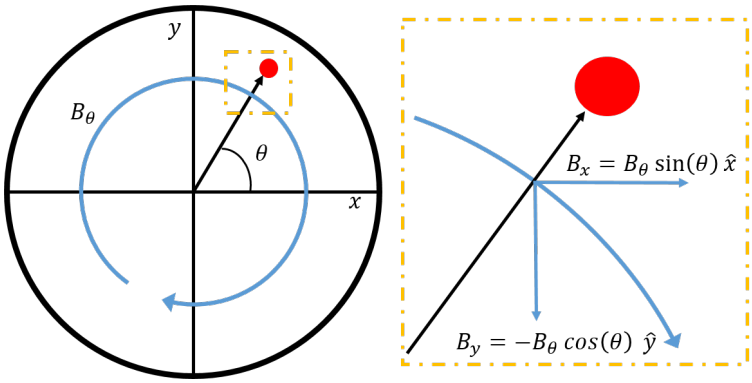


Figure 5.5: Particle as observed looking down the beam line of an rf single cell.

the magnitude of the x and y position. In this quadrant, for an accelerating field (where B_θ is clockwise, in this case, see Fig. 5.5), the component B_y is negative for all particles in the quadrant, which allows a simplification in the derivation. Note the same argument can be made for anti-clockwise B_θ, however the assumption must be carried through the entire derivation. The only transverse component of the electric field is E_r, and from Eqn. 5.5, $E_x = E_r\cos(\theta)$ and $E_y = E_r\sin(\theta)$. Inserting a periodic time dependence and making the position the independant variable;

$$dp_x = \mathrm{q}\ \mathrm{sign}(x)\ \cos(\theta)\int_0^L \left(\frac{E_r(s)}{v_z}\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right) + B_\theta(s)\cos\left(\phi_{B,0} + \frac{\omega s}{\beta c}\right)\right)ds \tag{5.46}$$

$$dp_y = \mathrm{q}\ \mathrm{sign}(y)\ \sin(\theta)\int_0^L \left(\frac{E_r(s)}{v_z}\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right) + B_\theta(s)\cos\left(\phi_{B,0} + \frac{\omega s}{\beta c}\right)\right)ds \tag{5.47}$$

where $\phi_{B,0} = \phi_0 - \frac{\pi}{2}$ (due to the phase difference between magnetic and electric field components) and the simplification;

$$-B_y = B_\theta\cos(\theta) \tag{5.48}$$

has been used as a particle position is defined in the first quadrant of the imaginary plane, with a clockwise B_θ. This does not reduce the number of computations, however means the functional form of the momentum gain is constant for any particle position. sign(x) signifies the numerical sign of the particle coordinate.

$$\theta = \tan^{-1}\left(|\frac{y}{x}|\right) \tag{5.49}$$

Using the functional form for E_z that is consistent with off-axis particles, from Eqn. 5.23 and the Lorentz force for dp_z, the integral for longitudinal momentum gain becomes

$$dp_z = \frac{\mathrm{q}}{v_z}\int_0^L \left[E_z(0,z) - \frac{r^2}{4}\left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2}E_z(0,z)\right)\right]\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds. \tag{5.50}$$

Inserting E_z and the derivatives of E_z as Fourier series;

$$\begin{aligned} dp_z = &\frac{\mathrm{q}}{v_z}\int_0^L E_0\sum_{n=0}^{\infty} b_n \sin\left(\frac{n\pi s}{l_{cell}}\right)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds \\ &- \frac{\mathrm{q}}{v_z}\frac{r^2}{4}\int_0^L \left(-\frac{E_0\pi^2}{l_{cell}^2}\sum_{n=0}^{\infty} b_n n^2 \sin\left(\frac{n\pi s}{l_{cell}}\right) + \frac{\omega^2}{c^2}E_0\sum_{n=0}^{\infty} b_n \sin\left(\frac{n\pi s}{l_{cell}}\right)\right)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds. \end{aligned} \tag{5.51}$$

Collecting like Fourier series;

$$\begin{aligned} dp_z = &\frac{qE_0}{v_z}\left(1 - \frac{\omega^2}{c^2}\frac{r^2}{4}\right)\int_0^L \sum_{n=0}^{\infty} b_n \sin\left(\frac{n\pi s}{l_{cell}}\right)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds \\ &+ \frac{qE_0}{v_z}\frac{\pi^2}{l_{cell}^2}\frac{r^2}{4}\int_0^L \sum_{n=0}^{\infty} b_n n^2 \sin\left(\frac{n\pi s}{l_{cell}}\right)\cos\left(\phi_0 + \frac{\omega s}{\beta c}\right)ds. \end{aligned} \tag{5.52}$$

From Eqn. 5.41 the above is integral can be evaluated. For ease, the notation $b_n n^i = b_{n,i}$ is introduced,

$$dp_z = \frac{qE_0}{v_z}\left(1 - \frac{\omega^2}{c^2}\frac{r^2}{4}\right)\frac{L}{2\pi}F(b_{n,0},\beta_s,\beta,\phi_0) + \frac{qE_0}{v_z}\frac{\pi^2}{l_{cell}^2}\frac{r^2}{4}\frac{L}{2\pi}F(b_{n,2},\beta_s,\beta,\phi_0). \tag{5.53}$$

Where again $F(b_{n,i},\beta_s,\beta,\phi_0)$ is given by

$$F(b_{n,i},\beta_s,\beta,\phi_0) = \sum_{n=1}^{\infty} b_n n^i \left[\frac{\cos(\frac{\beta_s}{\beta}\pi - n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} - n)} - \frac{\cos(\frac{\beta_s}{\beta}\pi + n\pi + \phi_0) - \cos(\phi_0)}{(\frac{\beta_s}{\beta} + n)}\right]. \tag{5.54}$$

The same process is completed for dp_x and dp_y. A complexity arises when integrating E_r. From Eqn. 5.30, the radial electric field has odd order derivatives of E_z. This means the Fourier series of E_r has only a_n coefficients (differentiating $\sin(x)$ and odd number of times leaves $\cos(x)$).

$$E_r(s,r) = -\frac{r}{2}\frac{E_0\pi}{L}\sum_{n=0}^{\infty} b_n n\,\cos\left(\frac{n\pi s}{L}\right) + \frac{r^3}{16}\left(-\frac{E_0\pi^3}{L^3}\sum_{n=0}^{\infty} b_n n^3\,\cos\left(\frac{n\pi s}{L}\right) + \frac{\omega^2}{c^2}\frac{E_0\pi}{L}\sum_{n=0}^{\infty} b_n n\,\cos\left(\frac{n\pi s}{L}\right)\right) \quad (5.55)$$

Integrating E_r with the separable time dependence, $\cos\left(\frac{\omega s}{\beta c}+\phi_0\right)$ over the cell length is different to the previous case, as the Fourier series for E_r is comprised of $\cos(\frac{n\pi s}{L})$, and so the result in Eqn. 5.41 no longer holds. The integral is now comprised of two $\cos(x)$ terms;

$$\int_0^L E_0\sum_{n=1}^{\infty} b_n\,\cos\left(\frac{n\omega s}{\beta_s c}\right)\cos\left(\phi_0+\frac{\omega s}{\beta c}\right)ds. \quad (5.56)$$

The result is similar to the previous case with $\sin(x)\rightarrow\cos(x)$ and the second term in the square brackets being negated.

$$\int_0^L E_0\sum_{n=1}^{\infty} b_n\,\cos\left(\frac{n\omega s}{\beta_s c}\right)\cos\left(\phi_0+\frac{\omega s}{\beta c}\right)ds = \frac{E_0 L}{2\pi}\sum_{n=1}^{\infty} b_n\left[\frac{\sin(\frac{\beta_s}{\beta}\pi - n\pi+\phi_0)-\sin(\phi_0)}{(\frac{\beta_s}{\beta}-n)} + \frac{\sin(\frac{\beta_s}{\beta}\pi + n\pi+\phi_0)-\sin(\phi_0)}{(\frac{\beta_s}{\beta}+n)}\right] = \frac{E_0 L}{2\pi}H(b_n,\beta_s/\beta,\phi_0). \quad (5.57)$$

Integrating the contribution from $E_r(s)$ using the result from Eqn. 5.57.

$$\frac{1}{v_z}\int_0^L E_r(s)\cos\left(\phi_0+\frac{\omega s}{\beta c}\right)ds = \frac{-r}{2}\frac{E_0\pi}{L}\frac{L}{2\pi}H(b_{n,1},\beta_s/\beta,\phi_0)+ \frac{r^3}{16}\left(\frac{-E_0\pi^3}{L^3}H(b_{n,3},\beta_s/\beta,\phi_0)+\frac{\omega^2}{c^2}\frac{E_0\pi}{L}H(b_{n,1},\beta_s/\beta,\phi_0)\right). \quad (5.58)$$

Lastly, integrating the contribution from B_θ;

$$B_\theta(s) = \frac{\omega E_0}{c^2}\left[\frac{r}{2}\sum_{n=0}^{\infty} b_n\sin\left(\frac{n\pi s}{L}\right) - \frac{r^3}{16}\left(-\frac{\pi^2}{L^2}\sum_{n=0}^{\infty} b_n n^2\,\sin\left(\frac{n\pi s}{L}\right)+\frac{\omega^2}{c^2}\sum_{n=0}^{\infty} b_n\sin\left(\frac{n\pi s}{L}\right)\right)\right] \quad (5.59)$$

which can be completed using the result in Eqn. 5.41, produces the following;

$$\int_0^L B_\theta(s)\cos\left(\phi_{B,0}+\frac{\omega s}{\beta c}\right)ds = \frac{\omega E_0}{c^2}\frac{r}{2}F(b_n,\beta_s/\beta,\phi_{B,0}) \\ -\frac{\omega E_0}{c^2}\frac{r^3}{16}\left(\frac{-\pi^2}{L^2}F(b_{n,2},\beta_s/\beta,\phi_{B,0})+\frac{\omega^2}{c^2}F(b_n,\beta_s/\beta,\phi_{B,0})\right). \tag{5.60}$$

Recall, the magnetic field lags the electric field by $\frac{\pi}{2}$, and therefore the value of ϕ_0 is replaced with $\phi_{B,0}$ when integrating momentum contributions due to the magnetic field. One could simply keep one term defining the phase, ϕ_0, however will need to subtract $\frac{\pi}{2}$ when the magnetic field is concerned. Putting together Eqns. 5.46, 5.58 and 5.60 gives an expression for the momentum change in x,

$$dp_x = -\text{qsign}(x)\cos(\theta)\frac{rE_0}{4}\left(H(b_{n,1},\beta_s/\beta,\phi_0)-\frac{2\omega}{c^2}F(b_n,\beta_s/\beta,\phi_{B,0})\right) \\ +\text{qsign}(x)\cos(\theta)\frac{r^3E_0}{16}\left(\frac{\omega^2}{c^2}\frac{\pi}{L}H(b_{n,1},\beta_s,\beta,\phi_0)-\frac{\pi^3}{L^3}H(b_{n,3},\beta_s/\beta,\phi_0)\right) \\ +\frac{\omega^3}{c^4}F(b_n,\beta_s/\beta,\phi_{B,0})-\left(\frac{\omega}{c^2}\frac{\pi^2}{L^2}F(b_{n,2},\beta_s/\beta,\phi_{B,0})\right). \tag{5.61}$$

In y the result is as follows;

$$dp_y = -\text{qsign}(y)\sin(\theta)\frac{rE_0}{4}\left(H(b_{n,1},\beta_s/\beta,\phi_0)-\frac{2\omega}{c^2}F(b_n,\beta_s/\beta,\phi_{B,0})\right) \\ +\text{qsign}(y)\sin(\theta)\frac{r^3E_0}{16}\left(\frac{\omega^2}{c^2}\frac{\pi}{L}H(b_{n,1},\beta_s/\beta,\phi_0)-\frac{\pi^3}{L^3}H(b_{n,3},\beta_s/\beta,\phi_0)\right) \\ +\frac{\omega^3}{c^4}F(b_n,\beta_s/\beta,\phi_{B,0})-\left(\frac{\omega}{c^2}\frac{\pi^2}{L^2}F(b_{n,2},\beta_s/\beta,\phi_{B,0})\right). \tag{5.62}$$

Where, recall;

$$F(b_{n,i},\beta_s/\beta,\phi_0)=\sum_{n=1}^{\infty}b_n n^i\left[\frac{\cos(\frac{\beta_s}{\beta}\pi-n\pi+\phi_0)-\cos(\phi_0)}{(\frac{\beta_s}{\beta}-n)}-\frac{\cos(\frac{\beta_s}{\beta}\pi+n\pi+\phi_0)-\cos(\phi_0)}{(\frac{\beta_s}{\beta}+n)}\right], \tag{5.63}$$

and

$$H(b_{n,i},\beta_s/\beta,\phi_0)=\sum_{n=1}^{\infty}b_n n^i\left[\frac{\sin(\frac{\beta_s}{\beta}\pi-n\pi+\phi_0)-\sin(\phi_0)}{(\frac{\beta_s}{\beta}-n)}+\frac{\sin(\frac{\beta_s}{\beta}\pi+n\pi+\phi_0)-\sin(\phi_0)}{(\frac{\beta_s}{\beta}+n)}\right]. \tag{5.64}$$

For the special case of a single spatial harmonic ($b_n = b_1$) for a particle with $\beta_s = \beta$. The apparent momentum gain in all three axis is zero. In order to solve this special case, the trigonometric terms can be expanded simplified for the limit β_s/β approaching 1. As the result is not physical, the result is not shown here.

5.4 FC2CT Applied to Traveling Wave Structures

The function used to analytically calculate momentum gain in 3D relies on the particle seeing a SW structure, with the E_z component symmetric about the center of the cell. If a particle is accelerated in a TW structure, the on-axis electric field is not in general symmetric about the centre of the cell. Whilst the difference for a TW single cell is subtle, the momentum gain can be approximated analytically.

Similar to the SW case, the real component of the on-axis E_z component using a Fourier series must be used. Technically, the periodicity of the Fourier series is a function of the chosen phase advance per cell of the given structure; the E_z component in a TW structure with a phase advance of $\frac{2\pi}{3}$ is periodic every three cells, and therefore the Fourier series describing it will have a periodicity of three cells.

The method to find the real component is as follows. Generally, a TW field can be described by a complex field with the on-axis E_z given as;

$$E_z(z,t) = M(z)e^{i(\omega t - kz + \phi_0)}. \tag{5.65}$$

Where $M(z)$ is some magnitude that is a function of z. The real part of the complex field is thus

$$E_z(z,t)|_{observed} = Re\left[E_z(z,t)\right] \tag{5.66}$$

at $t = 0$,

$$E_z(z,0) = M(z)e^{i(-kz+\phi_0)} \tag{5.67}$$

hence,

$$E_z(z,t) = E_z(z,0)e^{i\omega t} = E_z(z,0)\cos(\omega t) + iE_z(z,0)\sin(\omega t). \tag{5.68}$$

As $E_z(z,0)$ is a complex field, it can be written as the sum of the real and imaginary components;

$$E_z(z,0) = A(z) + iB(z). \tag{5.69}$$

Inserting Eqn. 5.69 into Eqn. 5.68 and taking the real component.

$$Re\left[E_z(z,t)\right] = A(z)\cos(\omega t + \phi_0) - B(z)\sin(\omega t + \phi_0) \tag{5.70}$$

Where $A(z)$ and $B(z)$ (not to be confused with the azimuthal magnetic field component, B_θ) are the real and imaginary field components of $E_z(z,t)$ taken from CST at any time, t, and can be described as Fourier series.

$$A(z) = \sum_{n=1}^{N} a_n(real)\cos\left(\frac{2n\pi z}{P}\right) + b_n(real)\sin\left(\frac{2n\pi z}{P}\right) \tag{5.71}$$

$$B(z) = \sum_{n=1}^{N} a_n(imag)\cos\left(\frac{2n\pi z}{P}\right) + b_n(imag)\sin\left(\frac{2n\pi z}{P}\right) \tag{5.72}$$

Where P is the period of the real/imaginary field, and is a function of the phase advance per cell as previously discussed.

In order to find the longitudinal momentum gain over one rf cell $Re[E_z(z,t)]$ must be integrated. Again, a change of variables $t = \frac{z}{\beta c}$ is performed, assuming constant particle velocity over the integration region,

$$\int_0^{L_{cell}} Re[E_z(z,t)]dz =$$
$$\int_0^{L_{cell}} \left[\sum_{n=1}^{N} a_n(real) \cos\left(\frac{2n\pi z}{P}\right) \cos\left(\frac{\omega z}{\beta c} + \phi_0\right) + b_n(real) \sin\left(\frac{2n\pi z}{P}\right) \cos\left(\frac{\omega z}{\beta c} + \phi_0\right) - \right.$$
$$\left. \sum_{n=1}^{N} a_n(imag) \cos\left(\frac{2n\pi z}{P}\right) \sin\left(\frac{\omega z}{\beta c} + \phi_0\right) + b_n(imag) \sin\left(\frac{2n\pi z}{P}\right) \sin\left(\frac{\omega z}{\beta c} + \phi_0\right) \right] dz. \tag{5.73}$$

In order to make the analytical computing of this integral more straight forward, a standard calculated integral form is determined, similar to the SW case, and is described in detail below.

5.4.1 Standard Integrals for TW FC2CT method

The four terms in Eqn. 5.73 to be integrated are simply the integral of two trigonometric functions multiplied together. As a result, the standard result shown in the SW derivation is possible. However, this method is still fairly long, and so a different method is shown for the TW case.

Firstly, it can be seen from Eqn. 5.70 that the real component of E_z is the sum of two terms, where both terms are Fourier series multiplied by a time varying trigonometric function. It has been previously shown how components from E_r and B_θ can be described by the on-axis Re$[E_z]$ field and the derivatives. As a result, any calculation will fundamentally require integrating Re$[E_z]$ or a derivative. The i^{th} derivative of Re$[E_z]$ is as follows;

$$\frac{d^i}{dz^i}\text{Re}[E_z] = \frac{d^i A(z)}{dz^i} \cos(\omega t + \phi_0) - \frac{d^i B(z)}{dz^i} \sin(\omega t + \phi_0). \tag{5.74}$$

Where $\frac{d^i A(z)}{dz^i}$ and $\frac{d^i B(z)}{dz^i}$ are simply Fourier series. As the real component of E_z is comprised of Fourier series multiplied by trigonometric functions, it will be prudent to find the standard integral of a Fourier series (FS) multiplied by a $\cos(x)$ function, as this will be some term, X, that is required to be calculated;

$$\text{X} = \int_{z_0}^{z_1} FS(s, re/im, i) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right) ds. \tag{5.75}$$

Where $FS(s, re/im, i)$ is a Fourier series representing the real/imag E_z component ($A(z)$ or $B(z)$) differentiated i times with respect to z. U is some factor such that the trigonometric function is $\cos(x)$ and not $\sin(x)$, as $\sin(x) = \cos(x - \frac{\pi}{2})$. Each Fourier series is written as the following:

$$FS(z, re/im, i) = \left(\frac{2\pi}{P}\right)^i \sum_{n=1}^{N} n^i \left[s_a a_n(im) \cos\left(\frac{2n\pi z}{P} + M\right) + s_b b_n(im) \cos\left(\frac{2n\pi z}{P} + N\right)\right]. \tag{5.76}$$

Where, again, all $\sin(x)$ terms have been converted to $\cos(x)$ with the use of variables, M and N (that is, M and N are chosen such that the trigonometric functions are $\cos(x)$). Table 5.1 displays the values of s_a, s_b, M and N for different i. The maximum number of derivatives required is three. This results in all integrals being of the following form

Table 5.1: Variable values for different differentiation index, i

i	s_a	s_b	M	N
0	1	1	0	-$\frac{\pi}{2}$
1	-1	1	-$\frac{\pi}{2}$	0
2	-1	-1	0	-$\frac{\pi}{2}$
3	1	-1	-$\frac{\pi}{2}$	0

$$K \int_{z_0}^{z_1} \cos(as + b)\cos(cs + d)ds = \frac{K}{2}\frac{1}{a^2 - c^2}\Big((a + c)[\sin((a - c)z_1 + b - d) - \sin((a - c)z_0 + b - d)] + (a - c)[\sin((a + c)z_1 + b + d) - \sin((a + c)z_0 + b + d)]\Big) = \frac{K}{2} I(a, b, c, d). \tag{5.77}$$

Where the values of a, b, c, d will be known. Substituting Eqn. 5.76 into Eqn. 5.75 and expanding;

$$\mathrm{X} = \int_{z_0}^{z_1} \left(\frac{2\pi}{P}\right)^i \sum_{n=1}^{N} n^i \Big[s_a a_n(im) \cos\left(\frac{2n\pi s}{P} + M\right) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right) + s_b b_n(im) \cos\left(\frac{2n\pi s}{P} + N\right) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right)\Big] ds. \tag{5.78}$$

Using the commutativity of the integral and summation,

$$\mathrm{X} = \left(\frac{2\pi}{P}\right)^i \sum_{n=1}^{N} n^i \Bigg[s_a a_n(im) \int_{z_0}^{z_1} \cos\left(\frac{2n\pi s}{P} + M\right) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right) ds + \\ s_b b_n(im) \int_{z_0}^{z_1} \cos\left(\frac{2n\pi s}{P} + N\right) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right) ds \Bigg]. \quad (5.79)$$

Inserting the standard integral result from Eqn. 5.77;

$$\left(\frac{2\pi}{P}\right)^i \sum_{n=1}^{N} \frac{n^i}{2} \left[s_a a_n(im) I\left(\frac{2n\pi z}{P}, M, \frac{\omega s}{\beta c}, \phi_0 + U\right) + s_b b_n(im) I\left(\frac{2n\pi s}{P}, N, \frac{\omega s}{\beta c}, \phi_0 + U\right) \right] \\ = Q(im/re, i, \phi_0 + U) = \mathrm{X}. \quad (5.80)$$

Where the integral is standardised as a function of i, Fourier coefficients describing the real or imaginary E_z component, and U. This is possible as inputs a and c from Eqn. 5.77 will be constant for any integral problem, and all s_a, s_b, M, N are defined for a given i.

Finally, the standard integral of a Fourier series multiplied by a cosine term is written as the result $Q(im/re, i, \phi 0 + U)$.

$$\int_{z_0}^{z_1} FS(s, re/im, i) \cos\left(\frac{\omega s}{\beta c} + \phi_0 + U\right) ds = Q(im/re, i, \phi_0 + U) \quad (5.81)$$

The method is reiterated as follows: from Lorentz force, momentum gain is calculated by integrating a given field component over the cell length. As each field component can be written as the real component of the on-axis E_z component and its derivatives, any integral will integrate $\mathrm{Re}[E_z]$ or a derivative. $\mathrm{Re}[E_z]$ (and derivatives) can be written as a sum of independant terms, where each term is a Fourier series multiplied by a time-varying trigonometric function. Any trigonometric term can be converted from $\sin(x) \rightarrow \cos(x)$ and thus any integral will be of the form $\int \cos(xs)\cos(ys)ds$. Each Fourier series term is comprised of a sum of two terms each of the form $\cos(xs)\cos(ys)$. Integrating a Fourier series multiplied by a trigonometric term can thus be written as a standard result, as shown in Eqn. 5.81. The standard result reached here is clearer by seeing an example worked through in the following section, where the momentum gain is calculated along each axis.

5.5 Derivation of FC2CT for TW Cavities

It was shown previously how $E_z(r, z)$ can be described as a function of the on axis field $E_z(0, z)$ at some radial displacement, r, see Eqn. 5.23. The result is provided again for ease;

$$E_z(r, z) = \left[E_z(0, z) - \frac{r^2}{4} \left(\frac{\partial^2 E_z}{\partial z^2} + \frac{\omega^2}{c^2} E_z(0, z) \right) \right]. \quad (5.82)$$

It was shown previously the real component of the field in a TW cavity is a combination of both the real and imaginary E_z components that independently oscillate out of phase with a fixed phase relation. Substituting the real component of the complex E_z component, Eqn. 5.70 into the above expansion for $E_z(r,z)$ to produce the observed E_z component off axis;

$$E_z(r,z,t) = \Bigg[A(z)\cos(\omega t+\phi_0) - B(z)\sin(\omega t+\phi_0) \\ - \frac{r^2}{4}\Bigg(\frac{\partial^2}{\partial z^2}\left[A(z)\cos(\omega t+\phi_0) - B(z)\sin(\omega t+\phi_0)\right] + \\ \frac{\omega^2}{c^2}\big(A(z)\cos(\omega t+\phi_0) - B(z)\sin(\omega t+\phi_0)\big)\Bigg)\Bigg]. \quad (5.83)$$

Recall from the Lorentz force equation

$$dp_z \approx \frac{\mathrm{q}}{\beta_z c}E_z(s)ds. \quad (5.84)$$

Substituting the form for $E_z(r,z,t)$ from Eqn. 5.83 with a change of variables, $t \to s$;

$$dp_z = \int_{z_0}^{z_1} A(s)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0\right)ds - \int_{z_0}^{z_1} B(s)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0-\pi/2\right)ds \\ - \int_{z_0}^{z_1}\frac{r^2}{4}A''(s)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0\right)ds + \int_{z_0}^{z_1}\frac{r^2}{4}B(s)''\cos\left(\frac{\omega s}{\beta_z c}+\phi_0-\pi/2\right)ds \\ - \int_{z_0}^{z_1}\frac{\omega^2 r^2}{4c^2}A(s)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0\right)ds + \int_{z_0}^{z_1}\frac{\omega^2 r^2}{4c^2}B(s)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0-\pi/2\right)ds. \quad (5.85)$$

Where $A(z)''$ denotes the double derivative of Fourier series $A(z)$, and $B(z)''$ denotes the double derivative of Fourier series $B(z)$ with respect to z. These terms can also be written in the defined form;

$$A(z)'' = FS(z,re,2), \;\; B(z)'' = FS(z,im,2). \quad (5.86)$$

As $A(z)$ represents the real field component of E_z and has been differentiated twice, $i=2$ and $B(z)$ refers to the imaginary field component. Inserting all Fourier series in Eqn. 5.85 as the defined notation and simplifying;

$$dp_z = \left(1-\frac{\omega^2 r^2}{4c^2}\right) \\ \times\left[\int_{z_0}^{z_1} FS(s,re,0)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0\right)ds - \int_{z_0}^{z_1} FS(s,im,0)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0-\pi/2\right)ds\right] \\ -\frac{r^2}{4}\left[\int_{z_0}^{z_1} FS(s,re,2)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0\right)ds - \int_{z_0}^{z_1} FS(s,im,2)\cos\left(\frac{\omega s}{\beta_z c}+\phi_0-\pi/2\right)ds\right]. \quad (5.87)$$

Inserting the standard integrals as defined in Eqn. 5.81;

$$dp_z = \left(1 - \frac{\omega^2 r^2}{4c^2}\right) \times \left[Q(re, 0, \phi_0) - Q(im, 0, \phi_0 - \pi/2)\right] - \frac{r^2}{4}\left[Q(re, 2, \phi_0) - Q(im, 2, \phi_0 - \pi/2)\right]. \quad (5.88)$$

Where $Q(im/re, i, \phi_0 + U)$ is defined in Eqn. 5.80. Recall the transverse momentum is given as follows;

$$dp_x = \text{sign}(x)\cos(\theta)\left(\frac{E_r(s)}{v_z}ds + B_\theta(s)ds\right) \quad (5.89)$$

$$dp_y = \text{sign}(y)\sin(\theta)\left(\frac{E_r(s)}{v_z}ds + B_\theta(s)ds\right). \quad (5.90)$$

The same process is thus completed for p_x, p_y which will be outlined here. Only one calculation is required for both the transverse planes, as the term in brackets is identical for the calculation of both dp_x and dp_y. The forms of $E_r(r, z)$ and $B_\theta(r, z)$ displayed as functions of $E_z(r = 0, z)$ and the derivatives are shown in Eqns. 5.30 and 5.36 respectively. Integrating E_r with respect to s,

$$\int_{z_0}^{z_1} E_r(r, s, t)ds = \int_{z_0}^{z_1}\left[-\frac{r}{2}\frac{\partial E_z}{\partial z} + \frac{r^3}{16}\left(\frac{\partial^3 E_z}{\partial z^3} + \frac{\omega^2}{c^2}\frac{\partial E_z}{\partial z}\right)\right]ds. \quad (5.91)$$

Where $\frac{\partial^i E_z}{\partial z^i}$ are implicitly functions of s and t and are substituted using Eqn. 5.74,

$$\int_{z_0}^{z_1} E_r(r, s, t)ds = \int_{z_0}^{z_1}\left(-\frac{r}{2} + \frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[A(s)'\cos\left(\frac{\omega s}{\beta_z c} + \phi_0\right) - B(s)'\cos\left(\frac{\omega s}{\beta_z c} + \phi_0 - \frac{\pi}{2}\right)\right]ds + \int_{z_0}^{z_1}\frac{r^3}{16}\left[A(s)'''\cos\left(\frac{\omega s}{\beta_z c} + \phi_0\right) - B(s)'''\cos\left(\frac{\omega s}{\beta_z c} + \phi_0 - \frac{\pi}{2}\right)\right]ds.$$

Converting the Fourier series for the real and imaginary fields, A and B, into the defined form (this is a purely aesthetic change in order to aid the production of $Q(re/im, i, \phi_0 + U)$)

$$\int_{z_0}^{z_1} E_r(r, s, t)ds = \int_{z_0}^{z_1}\left(-\frac{r}{2} + \frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[FS(s, re, 1)\cos\left(\frac{\omega s}{\beta_z c} + \phi_0\right) - FS(s, im, 1)\cos\left(\frac{\omega z}{\beta_z c} + \phi_0 - \frac{\pi}{2}\right)\right]ds + \int_{z_0}^{z_1}\frac{r^3}{16}\left[FS(s, re, 3)\cos\left(\frac{\omega z}{\beta_z c} + \phi_0\right) - FS(s, im, 3)\cos\left(\frac{\omega s}{\beta_z c} + \phi_0 - \frac{\pi}{2}\right)\right]ds.$$

Using the standard integral solution in Eqn. 5.81;

$$\int_{z_0}^{z_1} E_r(r,s,t)ds = \left(-\frac{r}{2}+\frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[Q(re,1,\phi_0)-Q(im,1,\phi_0-\frac{\pi}{2})\right]+ \frac{r^3}{16}\left[Q(re,3,\phi_0)-Q(im,3,\phi_0-\frac{\pi}{2})\right]. \quad (5.92)$$

It is important to note the above result is a purely solved equation. It provides an analytic approximation (up to some factors that provide the correct units) for the transverse momentum change due to the E_r field acting on a particle as it traverses a single cell. Similarly, the magnetic field effect on a particle is calculated by first inserting the field as E_z and the derivatives;

$$\int_{z_0}^{z_1} B_\theta(r,s,t)ds = \int_{z_0}^{z_1} \frac{\omega}{c^2}\left[\frac{r}{2}E_z(0,s,t)-\frac{r^3}{16}\left(\frac{\partial^2 E_z}{\partial z^2}+\frac{\omega^2}{c^2}E_z(0,s,t)\right)\right]ds. \quad (5.93)$$

For convenience, the result is stated;

$$\int_{z_0}^{z_1} B_\theta(r,s,t)ds = \frac{\omega}{c^2}\left(\frac{r}{2}-\frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[Q(re,0,\phi_{B,0})-Q(im,0,\phi_{B,0}-\frac{\pi}{2})\right]- \frac{\omega}{c^2}\frac{r^3}{16}\left[Q(re,2,\phi_{B,0})-Q(im,2,\phi_{B,0}-\frac{\pi}{2})\right]. \quad (5.94)$$

Where the following result was used,

$$\int_{z_0}^{z_1} \frac{\partial^n}{\partial s^n}E_z(r,s,t)ds = Q(re,n,\phi_0)-Q(im,n,\phi_0-\frac{\pi}{2}), \quad (5.95)$$

which is derived by combining Eqns. 5.74 and 5.81.

Using the results from Eqns. 5.92 and 5.94, the momentum change in the transverse planes are computed,

$$dp_{x/y} = \left[\frac{1}{\beta_z c}\left(-\frac{r}{2}+\frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[Q(re,1,\phi_0)-Q(im,1,\phi_0-\frac{\pi}{2})\right]+ \frac{1}{\beta_z c}\frac{r^3}{16}\left[Q(re,3,\phi_0)-Q(im,3,\phi_0-\frac{\pi}{2})\right]+ + \frac{\omega}{c^2}\left(\frac{r}{2}-\frac{r^3}{16}\frac{\omega^2}{c^2}\right)\left[Q(re,0,\phi_{B,0})-Q(im,0,\phi_{B,0}-\frac{\pi}{2})\right]- \frac{\omega}{c^2}\frac{r^3}{16}\left[Q(re,2,\phi_{B,0})-Q(im,2,\phi_{B,0}-\frac{\pi}{2})\right]\right] \times \cos(\theta)/\sin(\theta). \quad (5.96)$$

Similar to the SW case, ϕ_0 slips each cell due to the constant beta approximation and must be updated each cell via the following relation,

$$\phi_{n+1,0} = \phi_{n,0} - \frac{2\pi L_{cell}}{\lambda} + \frac{\omega L_{cell}}{\beta_{avg} c}, \tag{5.97}$$

where β_{avg} is the relativistic beta of the mean particle energy over the cell and λ is the design wavelength and is related to the phase advance per cell.

This method allows an accurate method of tracking a particle beam once the real and imaginary field maps are known. However, due to the large number of terms required per iteration, the application of the method are limited. A script was written in python designed to implement FC2CT for TW structures. When particle simulations were conducted the speed of the method was slower than the SW FC2CT. As the method requires both a_n and b_n Fourier coefficients in addition to both the real and imaginary E_z components. For calculations of the transverse momentum change in the SW FC2CT method, six summations are computed, however this value increases to 16 for the TW FC2CT method. A simplification of a true TW cavity can be approximated by treating the E_z component as symmetric about the centre of the rf cell. For a SW cavity there are no traveling components. As a result, the field builds evenly in each cell. In a TW cell, the field has a small traveling term which causes the field to build non-uniformly. Only when the cell is fully excited, does the field appear symmetric about the centre of the cell, see Fig. 5.6.

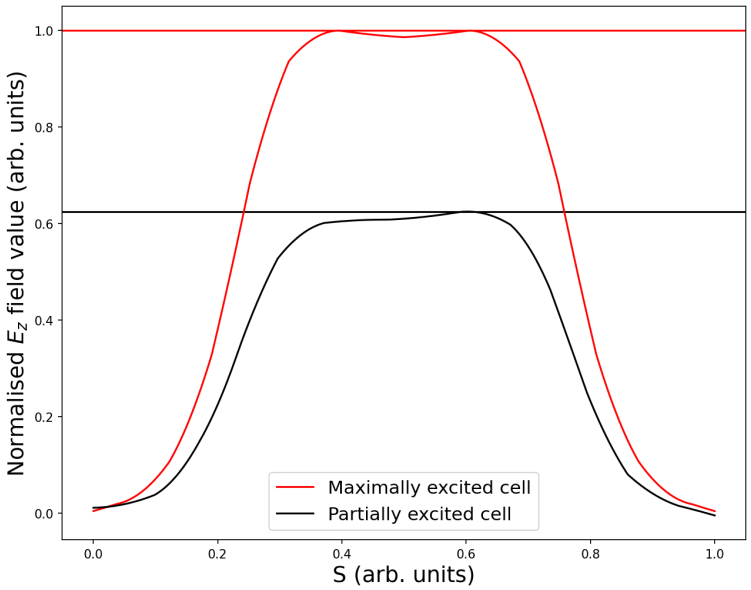


Figure 5.6: Real component of E_z component in a TW single cell at two different excitation levels.

By treating a TW cell as a SW cell with a different value for the phase advance per cell, the TW structure can be modelled as if it was a SW structure. This approximation thus treats each TW cell as identical, and excites uniformly within the cell. This approximation is only valid for small apertures, when E_z goes to zero between accelerating cells. For structures with a large aperture, E_z may not tend to zero between cells, and the field can not be accurately treated as a standing wave.

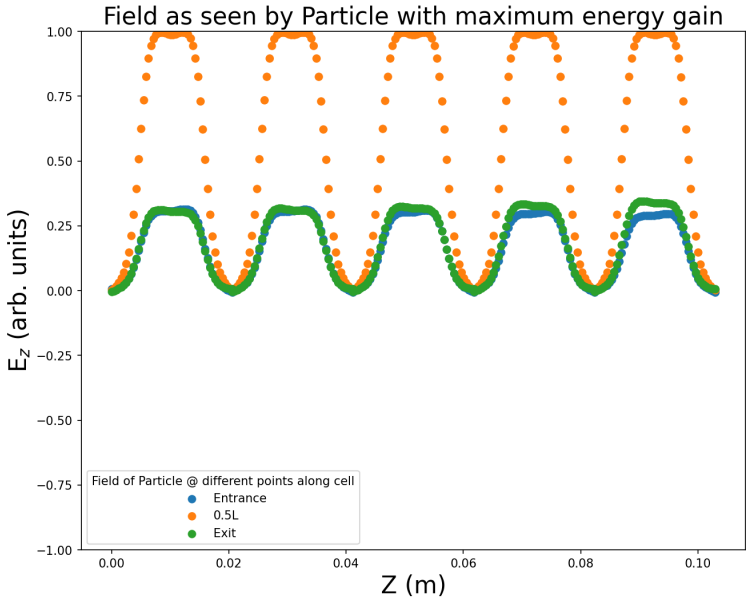


Figure 5.7: On-axis E_z component seen by a particle traversing a $\frac{4\pi}{5}$ phase advance TW cavity.

Figure 5.7 shows the on-axis field seen by a particle at a certain phase traversing a TW cavity such that the energy gain is maximised. However, as the particle is at the entrance/exit of the single cell at this point, the field is small in this region, and will be similar to the field as calculated by the SW approximation. For a π-mode structure, maximum energy gain occurs when the field is unexcited and increasing when the particle enters the cell. For structures with phase advance per cell less than π, the ideal energy gain occurs when the field is already excited and increasing. This is because the particle traverses the cell in less time than the field takes to advance by π.

5.6 Comparison of FC2CT to known Tracking Codes

The previous sections outline the calculation of the 6D phase space for both Standing and TW cavities.

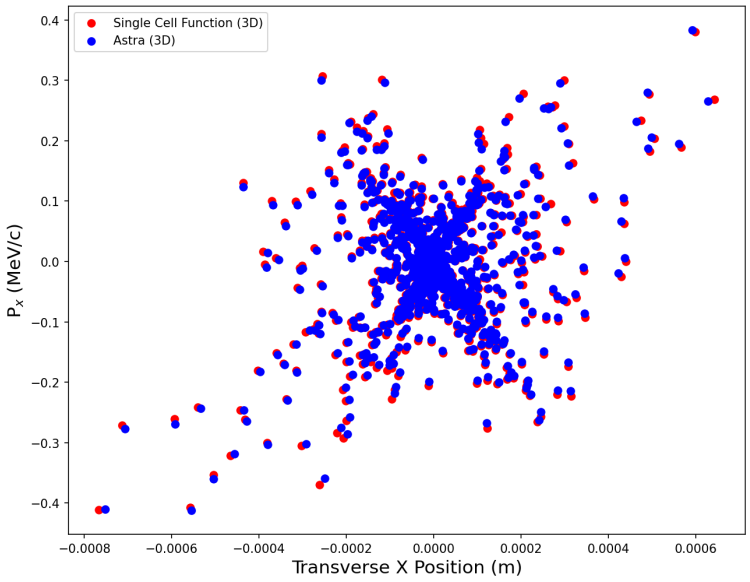


Figure 5.8: Phase Space plot showing transverse X phase space calculated in ASTRA and single cell method. Proton beam simulated at 150 MeV tracked through 50 π-mode cells with a gradient of 50 MeV/m.

5.6.1 Comparison of FC2CT and ASTRA for SW Cavities

The tracking code ASTRA [96] was used to benchmark the performance of the SW single cell functions. The 1D field map was extracted from CST microwave solver (for an S-band rf cell at 2.9985 GHz) and used for both the FC2CT and ASTRA simulations. The input particle beam was generated using the ASTRA generator application with a momentum spread in each dimension. The following phase space plots had different input beam phase spaces, so as to not compare the same input beam with every simulation.

Figure 5.8 shows the transverse phase space of a proton beam at an initial energy of 150 MeV after traversing a 50 cell SW structure as tracked by ASTRA and FC2CT. In order to compare accurately, the ASTRA simulation was initiated with the on-axis E_z component profile only; it was not run with 3D field maps. The two output distributions agree very strongly, with particle to particle deviation generally increasing with increasing x and P_x. The FC2CT method completed the simulation in a similar duration to ASTRA, in addition to being written in a non-optimised interpretative code, which are generally slower for simulation jobs.

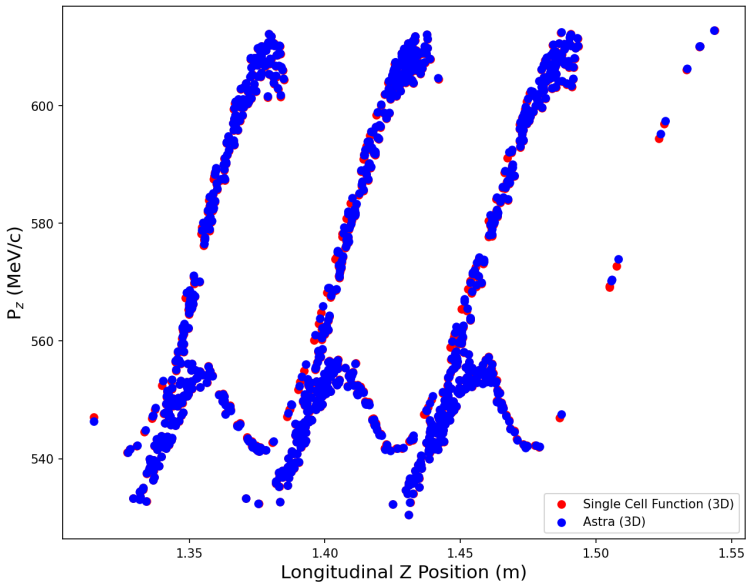


Figure 5.9: Phase Space plot showing longitudinal phase space calculated in ASTRA and single cell method. Proton beam simulated at 150 MeV tracked through 50 π-mode cells.

Figure 5.9 shows the longitudinal phase space of a proton beam starting at 150 MeV (P_z $\approx$ 550 MeV/c) through 50 SW cells at a gradient of 50 MeV/m. The initial distribution (which is the same as for Fig. 5.8) has a length of three rf periods and hence three periodic separatrix distributions are created. The distributions highly agree, notably at the 'head' of the separatrix, which is often the area of interest for particle tracking simulations. This makes intuitive sense, as the particles in the 'head' see an accelerating field for the majority of the single cell, and the constant beta approximation is accurate. Particles falling down the 'neck' have the highest discrepancy between the ASTRA and FC2CT calculations.

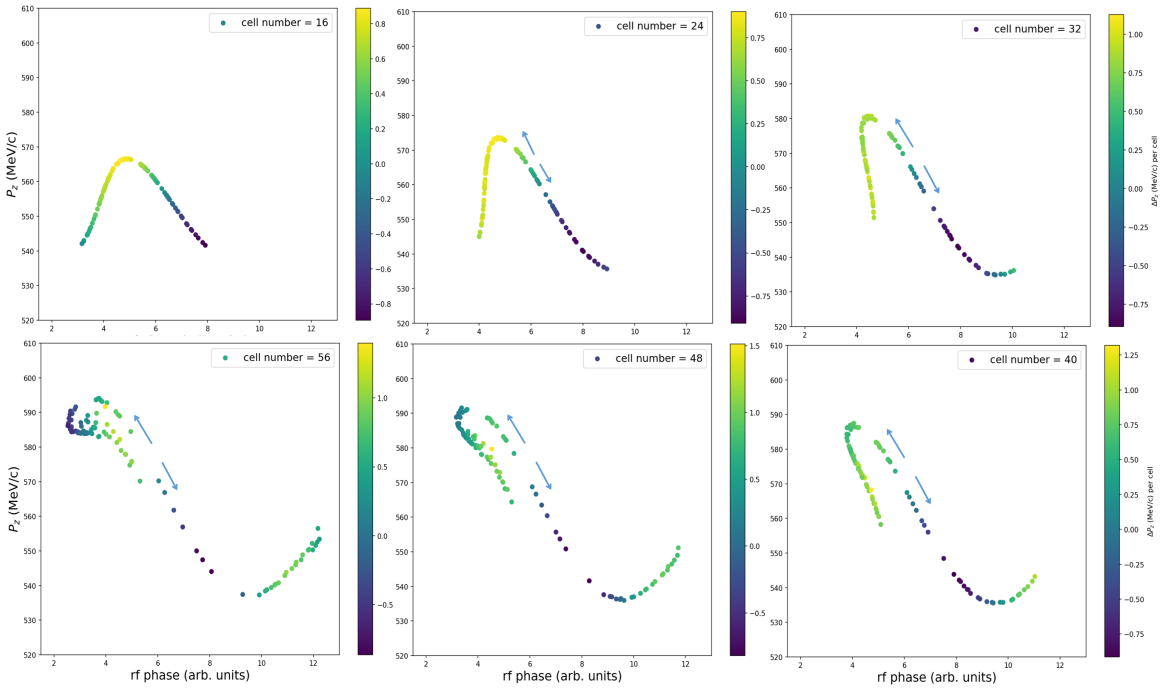


Figure 5.10: Phase space plot showing proton longitudinal momentum as a function of phase, at six different rf cells. The arrows show the discontinuity between particles moving up and down the 'neck'. Colour dimension shows ΔP_z of the previous rf cell.

The 'neck' particles start at an accelerating phase and gain momentum in the early cells, as they travel slower than $\beta_s c$ initially, the phase slippage is positive. At some point, the phase slips to a phase where the particle observes equal parts of a positive and negative E_z component whilst traversing the cell. Particles can be on either edge of an phase 'knife-edge', where the particle observes marginally more accelerating or decelerating phase over the next few rf cells. Particles on one side will slowly gain momentum and move towards the 'head' whilst the particles on the other side of the 'knife-edge' slowly lose momentum. This causes the 'neck' section to form, where particles on one side are pulled up the 'neck', towards the 'head', and the other particles are pulled further down the 'neck' leaving a sparsely occupied area of phase space where momentum gain cancels in each cell. As the momentum gain in this region is small, minuscule changes in relativistic beta may have large effects on whether a particle observes more acceleration or decelerating field and thus the constant relativistic beta approximation is less accurate in this region. This description is more easily visualised with the aid of a diagram, which is shown in Fig. 5.10. Particles in the 'neck' region have very low values of ΔP_z, with particles pulled up the 'neck' having slightly positive values

and particles pulled down, slightly negative values. In a time-stepping code, the particle velocity is updated multiple times per rf cell and the true value for ΔP_z is more accurately represented.

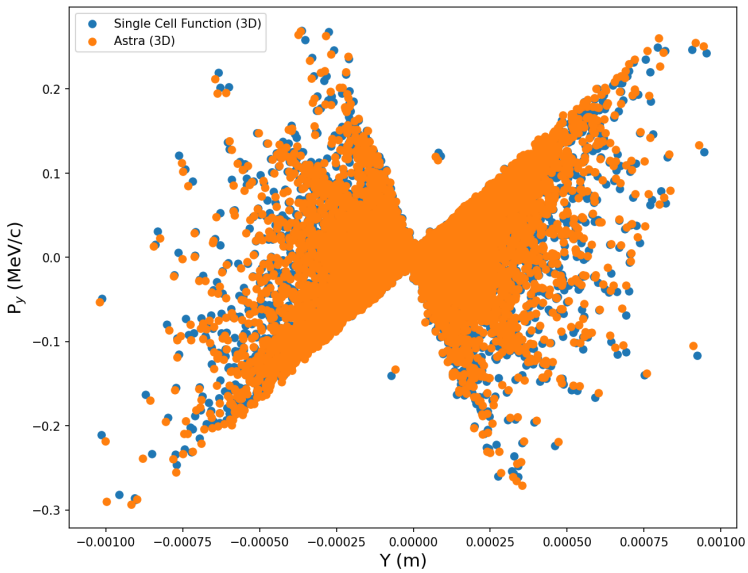


Figure 5.11: Phase Space plot showing transverse Y phase space calculated in ASTRA and single cell method. Proton beam simulated at 150 MeV tracked through 80 π-mode cells with a gradient of 50 MeV/m.

Figure 5.11 demonstrates the accuracy of FC2CT for longer rf structures. The result shows the y, P_y phase space distribution of two particle beams as tracked by ASTRA and FC2CT through 80 rf π-mode cells for protons starting at 150 MeV. Thus, the cavity is approximately 2m long. The figure shows the FC2CT method sustains the accuracy for longer structures, and any initial errors do not propagate to large errors by the end of a simulation. The initial distribution is different to that of Figures 5.8 and 5.9 in order to explore the agreement between ASTRA and FC2CT for a different set of input parameters.

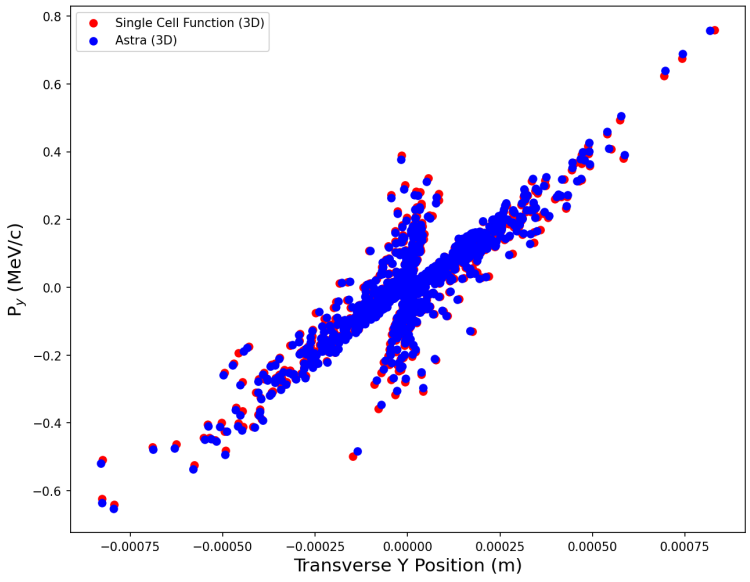


Figure 5.12: Phase Space plot showing transverse Y phase space calculated in ASTRA and single cell method. Proton beam simulated at 37.5 MeV tracked through 20 π-mode cells with a gradient of 50 MeV/m.

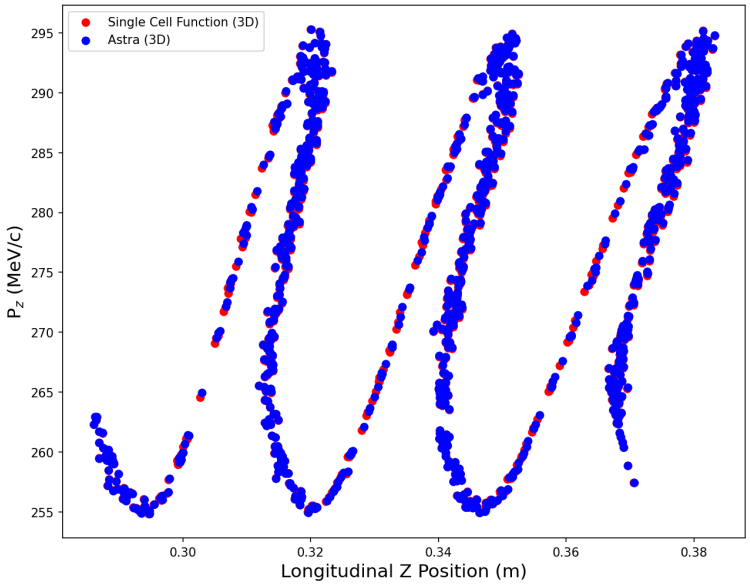
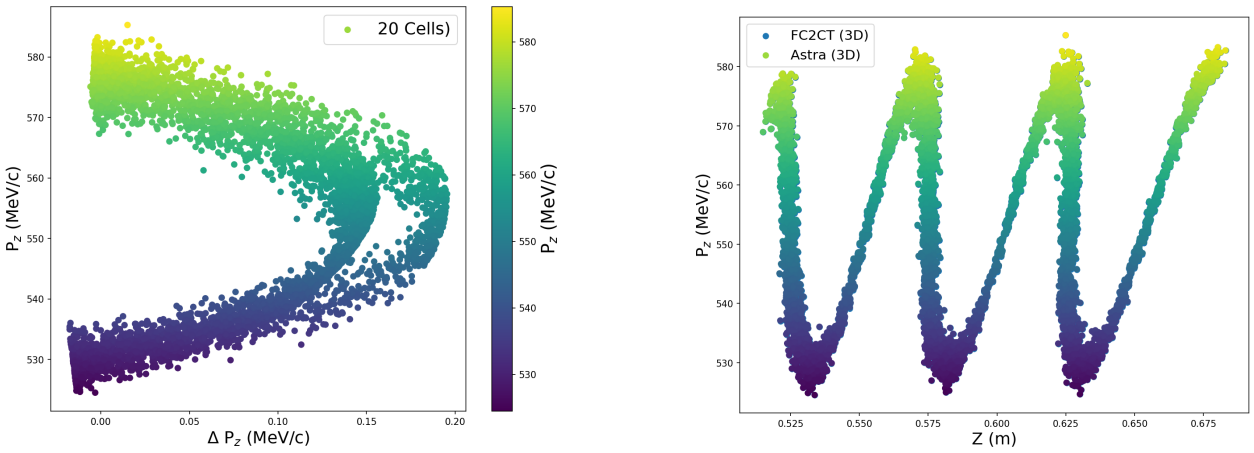


Figure 5.13: Phase Space plot showing longitudinal Z phase space calculated in ASTRA and single cell method. Proton beam simulated at 37.5 MeV tracked through 20 π-mode cells with a gradient of 50 MeV/m.

Protons at 150 MeV have relativistic beta ~ 0.5. The energy gain over an rf cell is of the order 1 MeV, the gain over one cell will not largely change beta. For lower energy protons, a gain of 1 MeV will have a larger impact on the fractional change in relativistic beta ($\frac{\delta\beta}{\beta}$) over an rf cell, and will reduce the accuracy of the constant beta approximation. To further qualitatively test the accuracy of FC2CT relative to ASTRA, a lower energy proton beam was simulated. Figures 5.12 and 5.13 show the transverse and longitudinal phase space plots respectively for a proton beam at 37.5 MeV tracked through 20 rf cells at a gradient of 50 MeV/m. The two distributions again show strong agreement at the lower proton energy. Again, the particles on the 'neck' in Fig. 5.13 have the most deviation from the distribution as simulated with ASTRA. Nevertheless, the two phase space plots show that even protons at 37.5 MeV can be modelled accurately using FC2CT. The FC2CT method iterates over each cell, thus for lower energy particles, the cell length is shorter and FC2CT iterates over a smaller longitudinal distance relative to particles of higher energy. This means FC2CT applied to a 37.5 MeV proton beam will require more iterations per unit length than protons at 150 MeV.

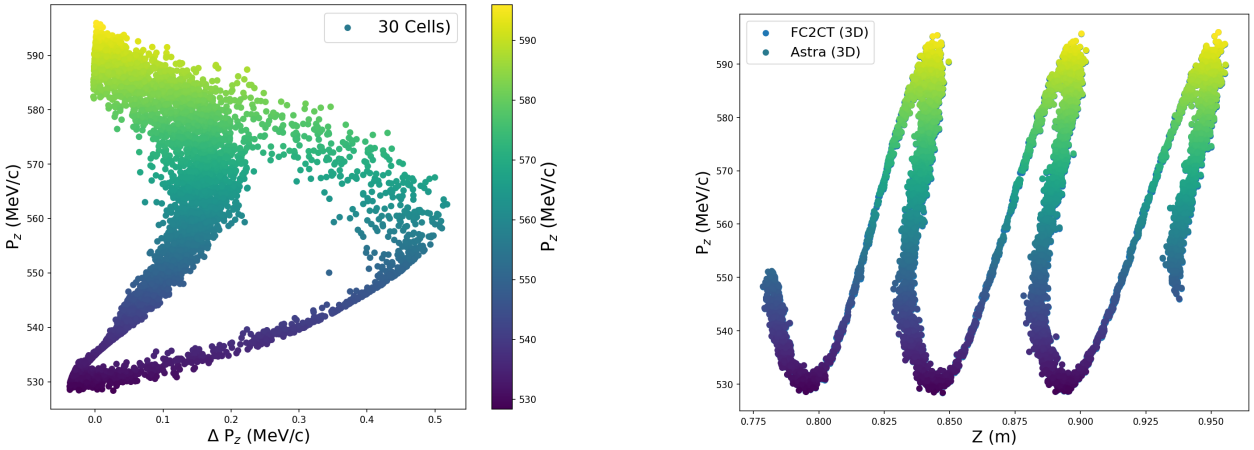


(a) (b)

Figure 5.14: (a) P_z as a function of difference in P_z as calculated by ASTRA and FC2CT (b) Longitudinal phase space of distribution after 20 rf cells.

In order to quantitatively access the accuracy of FC2CT to ASTRA, the difference between phase space coordinates is calculated and discussed below. Figure 5.14 shows the difference in calculated P_z from ASTRA and FC2CT. The input proton distribution had a longitudinal length of three rf cycles and had a nominal energy of 150 MeV, the gradient is 50 MeV/m. The phase space distribution is also shown for ease of understanding. Particles with high and low longitudinal momentum have small errors in calculated P_z. As the momentum value tends to the mean value, $\sim$ 555 MeV/c, the error in P_z increases. Particles not on the tail or head of a bunch appear on one of two necks. Particles on each neck have

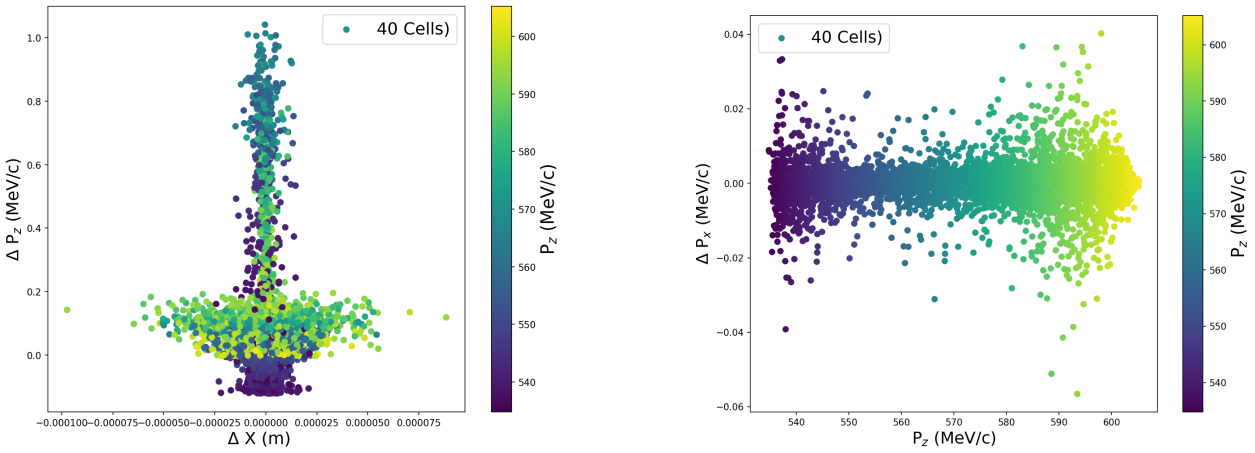
larger discrepancies in calculated P_z, which was discussed previously. The distribution in Fig. 5.14 (a) forms a closed loop, as particles in difference rf bunches appear at the same position in the error distribution plot. If the error distribution of the distribution in Fig. 5.14 (a) was visualised in 3D, it would appear similar to a saddle shape.



(a) (b)

Figure 5.15: (a) P_z as a function of difference in P_z as calculated by ASTRA and FC2CT (b) Longitudinal phase space of distribution after 30 rf cells.

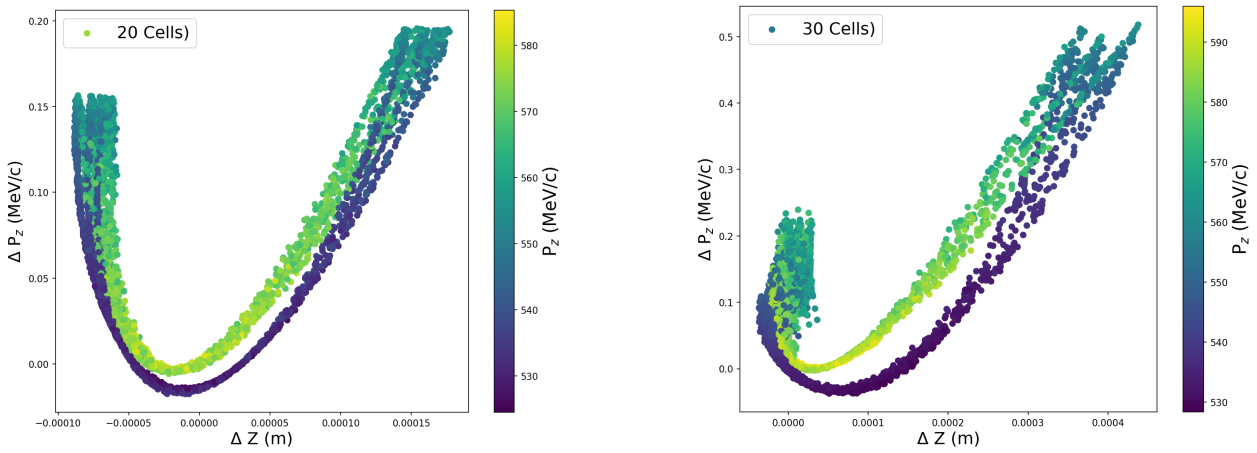
Figure 5.15 shows the discrepancy in calculated P_z after 30 SW cells, for the same initial distribution. Again, the longitudinal phase space is shown. In (a), the closed loop shape is again achieved. However, particles on the thinner neck with higher error in P_z from Fig. 5.14 continue to produce larger errors in P_z as more cells are traversed. The maximum P_z error increases from ∼0.2 MeV/c to ∼ 0.5 MeV/c after 10 more cells are traversed. The particles on the other, thicker, neck appear to remain in a region of P_z error ∼ 0.15 MeV/c. This causes the saddle shape to stretch as the error in P_z for particles in the neck continues to grow. Nevertheless, the maximum error in P_z (relative to ASTRA accuracy) is approximately 0.1%. As the number of cells traversed increases, particles at the tail and head of the bunch continue to have very small P_z discrepancy, demonstrating that errors do not propagate aggressively for particles in these regions. Additionally, the previous figures discussed suggest the error in P_z is always positive. As ΔP_z was calculated as $P_z(ASTRA) - P_z(FC2CT)$, The FC2CT method underestimates the longitudinal momentum, for (nearly) all particles in a bunch.



(a) (b)

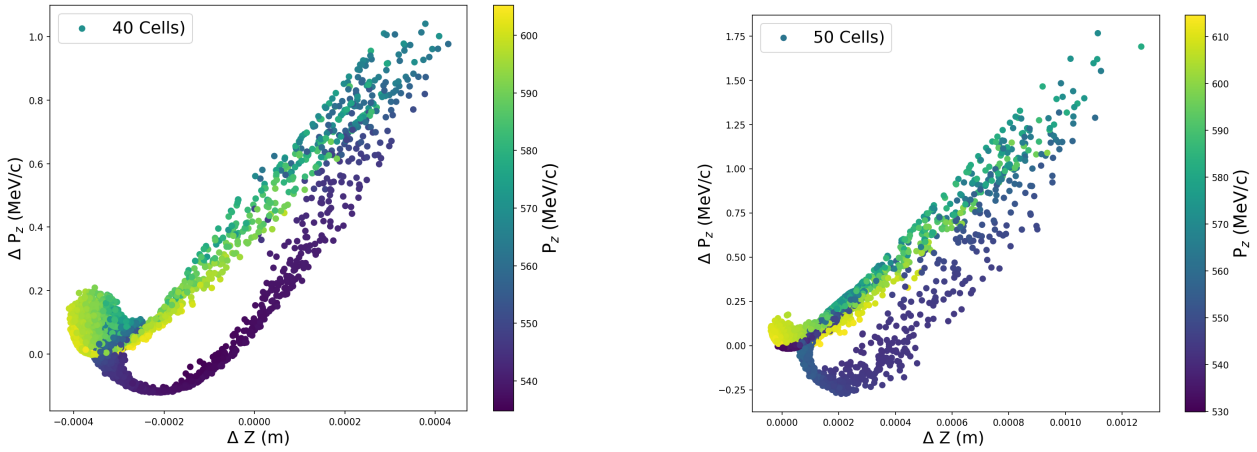
Figure 5.16: (a) Difference in P_z as calculated by ASTRA and FC2CT as a function of difference in X as calculated by ASTRA and FC2CT (b) Difference in P_x as calculated by ASTRA and FC2CT as a function of P_z. 40 rf cells.

Figure 5.16 (a) shows the error in P_z as a function of error in X, from tracking through 40 rf cells. Figure (a) shows the error in transverse position increases for a particular error in P_z, between 0 and 0.2 MeV/c. For particles with increased error in P_z, the error in X appears fairly uniform. The maximum position error in X is approximately 0.1 mm. As the iris aperture is 2.5 mm, this error is small, albeit larger than longitudinal errors. (b) shows the error in P_x as a function of P_z. This plot demonstrates a fairly uniform error in P_x as a function of P_z. This implies the FC2CT transverse momentum gain calculation introduces a relatively constant level of inaccuracy, independant of particle position within a bunch. There appears a slight increase in P_x error for higher P_z particles however is not substantial enough to draw any meaningful conclusions about. The maximum error in P_x is approximately 0.06 MeV/c. As the maximum value of P_x is 1 MeV/c, the error in P_x is approximately 6 %. This value increases to approximately 8% after 50 cells.



(a) (b)

Figure 5.17: (a) Difference in P_z as a function of difference in Z after 20 (a) and 30 (b) rf cells.



(a) (b)

Figure 5.18: (a) Difference in P_z as a function of difference in Z after 40 (a) and 50 (b) rf cells.

Figure 5.18 displays the error in Z as a function of the error in P_z for the same initial distribution tracked through 20, 30, 40, and 50 rf cells. The 3D distribution in Fig. 5.18 (a) shows the saddle shape that has been discussed previously, forming a closed loop. Particles at the tail and head of a bunch have low error in both P_z and Z. Particles with larger error in P_z, occurring on the neck of the longitudinal phase space plot, also experiencing the maximum error in longitudinal position. This is an expected result. If the momentum gain is incorrect, the phase slippage calculated after an iteration will be incorrect, and propagate the error. This leads to incorrect momentum gain calculation in subsequent cells, and a

feedback loop is created. From the figure, particles have both positive and negative values for ΔZ, thus FC2CT both over and underestimates average particle velocity over a structure, however the results imply the underestimation is greater in magnitude. a minimum in ΔP_z occurs near $\Delta Z = 0$.

As the number of rf cells increases, the error distribution tends to a linear plot, with a positive orientation starting at (0, 0). Generally, the error in both P_z and Z becomes a positive value, again implying FC2CT underestimates both P_z gain and average particle velocity. For particles with larger errors in P_z and Z, the distribution tends to an expanding region that locally blows up. These particles appear on the neck of the longitudinal phase space distribution and become difficult to accurately track in this region, as discussed previously. After 40 rf cells, the maximum error in P_z is roughly 0.2%, relative to ASTRA, and a position error or roughly 0.4 mm ($\sim 1/60$ of a cell length for protons at this energy in a SW structure). After 50 rf cells, the maximum error in P_z is approximately 0.3%. The error in position is approximately 1.2 mm, which is $\sim 1/20_{th}$ of a cell length. This error is small, and demonstrates the ability of FC2CT for protons accelerated through SW cavities.

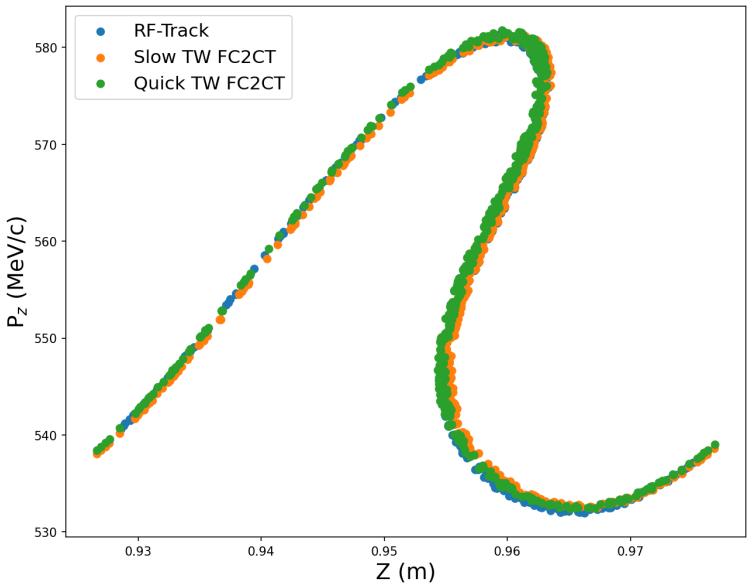


Figure 5.19: Plot displaying the longitudinal phase space calculated in both RF-Track and FC2CT. Proton beam simulated at 150 MeV tracked through 45 TW cells.

5.6.2 Comparison of FC2CT and RF-Track for TW Cavities

So far, it has been shown the FC2CT method is highly effective at approximation the phase space of a proton beam at varying energies. In order to test how accurate the TW version of FC2CT, another reliable tracking code was used, RF-Track [98]. The input distribution

for the comparison between RF-Track and FC2CT had a Gaussian momentum spread in x, P_x, y, P_y, however P_z was constant for all particles, at the design input momentum. This choice was made to infer the deviation between the output longitudinal phase space distributions, see later. The TW cavity was comprised of 45 cells with a phase advance per cell of $\frac{4\pi}{5}$. The gradient was 33.5 MeV/m and the cavity length was ≈ 1 m. Each simulation was ran using the same on-axis E_z component profile, taken from CST Microwave Solver.

Figure 5.19 displays the longitudinal phase space of both distributions as simulated by RF-Track and FC2CT. Two methods for the TW FC2CT simulation are shown, labelled slow and fast. The slow method uses both the real and imaginary E_z component defined over an rf period, as described previously. The fast method utilises the SW approximation for the TW cavity, by assuming a field profile that is symmetric about the centre of the rf cell at all times. The field can be created using only b_n terms in the Fourier series of the field profile and thus runs more quickly. It can be seen the distributions are very similar for all three methods with very slight deviation, particularly at the head of the separatrix. As the distribution was initialised with a fixed P_z, the shape of the output phase space distributions can be compared exactly. For particles with $P_z < 550$ MeV/c, the distribution shapes slightly differ. Thus, the rotation of the phase space as calculated by RF-Track is different to that of the FC2CT method.

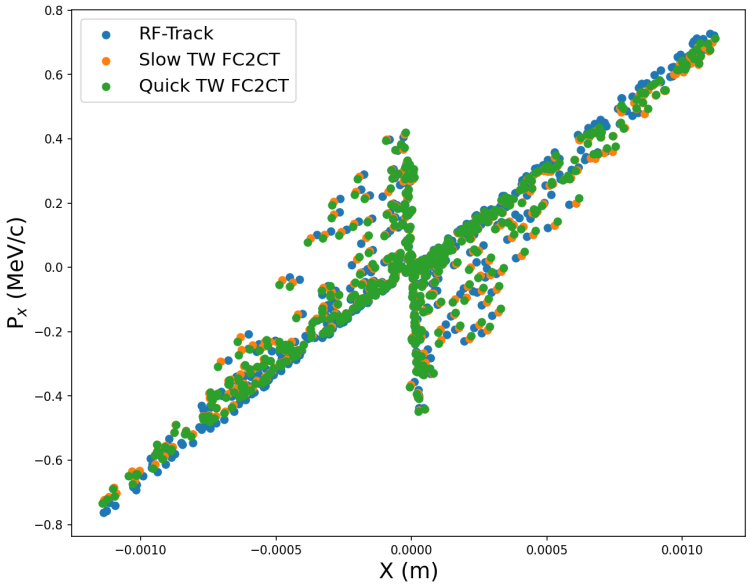


Figure 5.20: Phase Space plot showing transverse (X) phase space calculated in RF-Track and both FC2CT fast and slow methods.

Figure 5.20 displays a subset of the x, P_x phase space distribution as tracked by the three methods. The distributions highly agree for the RF-Track and FC2CT methods. The fast

FC2CT method appears to function as well as the slow method, however the slow method took $\sim 2-3$ times longer than the fast method. The increase in accuracy of the fast method is not clear, and does not justify the use of the method.

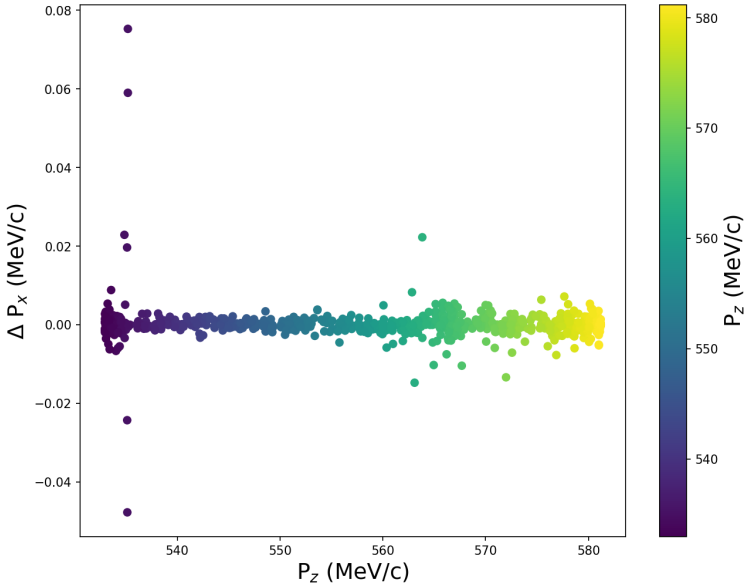


Figure 5.21: Difference in P_x as a function of P_z from calculations using FC2CT (fast) and RF-Track.

Figure 5.21 describes the error in P_x due to the FC2CT fast method. Similar to Fig. 5.16 (b), there is no clear structure, with particles of all momentum experiencing a relatively constant error in P_x. However, for low momentum particles, there are some fairly large discrepancies, with ΔP_x reaching 0.08 MeV/c.

The results suggest the TW FC2CT method can accurately track particle through TW structures. However, the slow method requires many more calculations per iteration, and is likely not viable over conventional tracking methods. Whilst the fast method removes this issue, a true TW is not being simulated, and the accuracy is inferior. Only one simulation was compared with RF-Track, due to time constraints. Future work comparing the effectiveness of FC2CT relative to RF-Track would be a beneficial area to explore.

5.7 The Constant Beta Approximation and Accuracy Improvements to FC2CT

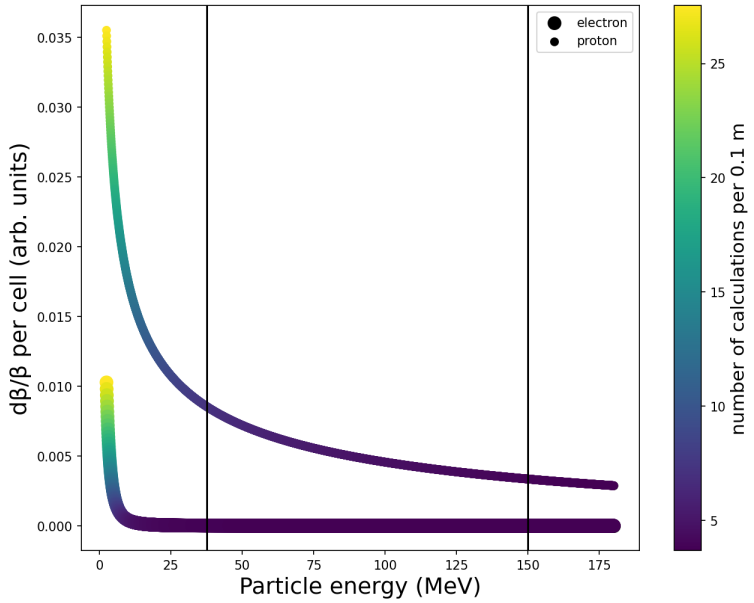


Figure 5.22: Plot showing change in relativistic beta divided by relativistic beta over a single cell for a gradient of 50 MeV/m for both electrons and protons at different energy. Color dimension shows how higher energy particles require fewer single cell calculations per unit longitudinal distance.

As shown with phase space plot comparisons, protons at 37.5 and 150 MeV can be accurately tracked using the FC2CT. Figure 5.22 shows the fractional change in relativistic beta per cell as a function of particle energy, for both electrons and protons. The gain per cell is calculated as $50L_{cell}$, and the starting particle energy is 2.5 MeV. The two vertical lines at 37.5 and 150 MeV show this value is much higher for protons than for electrons at the same energy. This is because at these energies, electrons are highly relativistic and changing the energy further will not have an effect on the relativistic beta. Protons can be simulated using FC2CT at a vast range of energies, however rf coupled cavity machines will not be used for low energy protons as the effective shunt impedance is greater with RFQs and drift tube linacs. $\frac{\delta\beta}{\beta}$ is used as an indication of the validity of the constant beta function. $\delta\beta$ alone can not be used as an indication. Protons initially at 1 MeV have $\beta \sim 0.046$. If the proton gains 1 MeV over an rf cell, the value of β becomes 0.065. Thus $\delta\beta \sim 0.02$. The value of $\frac{\delta\beta}{\beta}$ however is 0.4. Results shown in this section show strong agreement between FC2CT and ASTRA have $\frac{\delta\beta}{\beta} < 0.01$.

Accuracy Improvements to FC2CT

Up to now, FC2CT only considers one iteration per rf cell. However, the method can be altered to compute multiple iterations per cell. One potential method to improve FC2CT considers running an iteration over a single cell multiple times, updating the value of the constant beta used in subsequent iterations. As a result, the particle does not travel at the input beta through the cell, but a value that closer resembles the average particle beta over the cell.

Another improvement requires reducing the integration region, so that a single iteration integrates some section of a single cell. Recall the energy gain in 1D for a SW cavity from z_0 to z_1 for an arbitrary phase advance per cell is given,

$$E(z_1) = E(z_0) + \frac{E_0}{2}\sum_{n=1}^{\infty} b_n \left[\frac{\cos(s(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c})} - \frac{\cos(s(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c})} \right]_{z_0}^{z_1}. \quad (5.98)$$

Expanding the integral bounds;

$$E(z_1) = E(z_0) + \frac{E_0}{2}\sum_{n=1}^{\infty} b_n \left[\left(\frac{\cos(z_1(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c})} - \frac{\cos(z_1(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c})} \right) - \left(\frac{\cos(z_0(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} - \frac{n\omega}{\beta_s c})} - \frac{\cos(z_0(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c}) + \phi_0)}{(\frac{\omega}{\beta c} + \frac{n\omega}{\beta_s c})} \right) \right]. \quad (5.99)$$

Instead of integrating over the cell length, the bounds z_1 and z_0 are over some section. Defining the number of iterations per cell length as I, then $z_1 = (i+1)\frac{L_{cell}}{I}$, $z_0 = i\frac{L_{cell}}{I}$, where i is the section integer, $i = 0, 1, ..., I-1$. Each cell then requires I iterations to be completed. The phase, $\phi_{0,i}$ and relativistic beta, β_i, must be updated each section. In general, the 1D energy gain for a phase advance per cell, ψ, over one cell split into I calculations is

$$E(L_{cell}) = E_0 + \frac{G_0 L_{cell}}{2}\sum_{i=0}^{i=I-1}\sum_{n=1}^{\infty} b_n \left[\frac{1}{(\frac{\psi\beta_s}{\beta_i} - n\psi)} \left(\cos\left(\frac{i+1}{I}(\frac{\psi\beta_s}{\beta_i} - n\psi) + \phi_{0,i} \right) - \cos\left(\frac{i}{I}(\frac{\psi\beta_s}{\beta_i} - n\psi) + \phi_{0,i} \right) \right) - \frac{1}{(\frac{\psi\beta_s}{\beta_i} + n\psi)} \left(\cos\left(\frac{i+1}{I}(\frac{\psi\beta_s}{\beta_i} + n\psi) + \phi_{0,i} \right) - \cos\left(\frac{i}{I}(\frac{\psi\beta_s}{\beta_i} + n\psi) + \phi_{0,i} \right) \right) \right]. \quad (5.100)$$

Depending on the accuracy required and particle energy of a given simulation, the choice of I can be chosen to trade-off between accuracy and computing time. As this approach requires multiple calculations per cell, it loses the speed benefit of the FC2CT method, and essentially behave similarly to a conventional tracking code.

Figure 5.23 shows the percentage difference between initial FC2CT calculation (blue line) of particle energy and other calculations of varying computational duration, for a proton at 150 MeV over a single rf cell of gradient 50 MeV/m. The 'number of iterations' index refers to running the FC2CT method multiple times and updating the constant relativistic beta used in the subsequent calculations. For example, 'number of iterations' = 2, refers to running FC2CT twice, with the second calculation of the cell energy gain using the average value of β over the cell as calculated by the first iteration. 'Integration length' refers to the fraction of a single cell completed per iteration. For example 'integration length' = 0.5 refers to an iteration being half a cell length, thus requiring two iterations per cell length and $I = 2$. Hence, the orange and green lines in Fig. 5.23 require the same number of iterations (2) per cell. It can be seen that updating the relativistic beta by iterating over a single cell twice has a much smaller effect than integrating over half a cell twice. The figure also shows results from creating 4, 100 and 200 sections within a single cell. By 100 iterations per cell the value for particle energy converges, as computing 200 iterations per cell does not change the percentage difference from initial calculation from 100 iterations per cell. It is also interesting to note that the percentage difference due to using an average relativistic beta changes the value different to changing the iteration length.

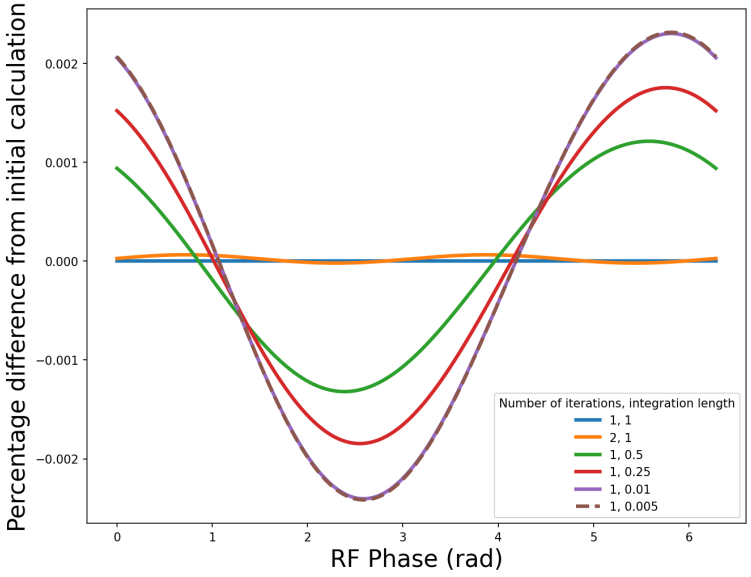


Figure 5.23: Percentage change from an initial FC2CT calculation of particle energy to calculations made with increasing computational time. Results are shown as a function of phase for protons at 150 MeV.

For certain rf phases, decreasing the integration length has larger effects on the calculated value of particle energy, relative to other phases. The maximum percentage difference occurs at different rf phases for different integration length, however the shift is minor. The

maximum percentage error from changing the number of iterations does not coincide with changing the integration length. The maximum percentage error is approximately 0.002 %, which is equivalent to the proton energy calculating being incorrect by ~ 3 keV. By reducing the integration length to half an rf cell, the energy gain calculation changes by approximately 1.5 keV. For large simulations with many particles, the small change in calculated phase space is not worth the investment in additional computation time, as a result developing FC2CT with reduced integration length is likely not viable over conventional tracking codes. The methods outlined to potentially improve the accuracy of FC2CT do not improve the calculated phase space by any meaningful amount, however increase computation time. In order to test the true accuracy of the methods, a higher order integrator would be required, such as an RK12 integrator.

5.8 Conclusion of FC2CT

Whilst it has not been completed, and would be an interesting next step for this work, it is likely a FC2CT script written in a compiled language code could produce highly accurate results that vastly reduce computation time relative to time stepping codes for approximate particle tracking simulations. In addition, further work comparing the validity of the FC2CT TW method relative to RF-Track would be an interesting development. Such comparisons include lower energy range protons, and cavities of varying length. Nevertheless, results show a strong agreement between RF-Track and FC2CT qualitatively, as shown by the similar distributions in phase space in Figs. 5.19 and 5.20. An important feature that FC2CT omits is space charge effects, and thus is most accurate when considering protons beams of low beam current. However, an interesting improvement could assess a method to introduce space charge into FC2CT.

Thus far, FC2CT determines the 6D phase space of particles traversing periodic field maps. In reality, rf cavities exert small transverse forces on particles at the structure entrance/exit. An additional future study could approximate the effect of these fringe fields on particle motion, allowing FC2CT to accurately track particles through more realistic structures.

In this chapter, an analytical method was created and implemented that allows for the calculations of the 6D phase space for a particle traversing an rf cavity comprised of multiple rf cells. The method, FC2CT, can be applied to both SW and TW cavities. Both methods were compared to reliable tracking codes, ASTRA and RF-Track, and shown strong agreement for both comparisons. The TW method losses the computational benefit achieved with the simpler SW method, and thus an approximation of the TW cavity was implemented, showing strong agreement with RF-Track. Due to the fast nature of the model, and the easy implementation of different rf cell gradients, the method will be utilised in the following section - concerning the comparison between SW and TW cavities for the final cavity design.

Chapter 6

RF Power Requirements

6.1 Power Distribution in Traveling and Standing Wave Structures

Chapter 4 explored two potential focusing schemes for the linac lattice, factoring in longitudinal acceleration. The results show a FODO focusing scheme can incorporate cavity lengths of the order 1 m whilst accepting a 5σ beam size at the cavity entrance/exit. This allowed the possibility of longer TW structures, in addition to SW structures. In order to fully compare SW and TW cavities (only constant gradient structures are considered), the electrical power required to obtain a 100 MeV gain (from 150 to 250 MeV) is an important figure merit in the linac design. The required power relates to both the initial cost and running cost of the machine over the course of its life. In recent years, there has been a global shift towards fostering environmental sustainability within the engineering industry. As a result, the environmental impact of a linac system should be one of the most important design objectives during the design. Other important objectives for the linac design include the total linac length (inc. focusing), and the filling time (ability to accelerate more protons per rf pulse), which both require minimising.

Table 6.1: Inputs and outputs for conceptual linac design. Objectives are shown with *. Constraints are shown with †.

Inputs	**Outputs**
Total Power*†	Total Linac Length (inc. focusing)*
Number of Structures†	Fill Time*
N_{cells} per Structure†	Average Power Loss†
Input Group Velocity (TW)†	Output Group Velocity (TW)†
Cell Phase Advance (TW)†	Gradient†

Table 6.1 displays the inputs and outputs for comparisons between linac candidates. Linac objectives are labelled with *, and linac constraints are labelled with †. Linac objectives describe outcomes that are to be maximised or minimised, whilst constrained values are limited to a range, or have an upper limit. For example, the gradient has an upper limit constraint of 40 MeV/m. For a given number of structures in a linac candidate, each structure must meet the required energy output, such that the total gain is achieved. The total linac length is calculated by summing all structures lengths within a linac candidate, with an additional length taking into account focusing length (see later). Results from the constrained comparison are shown in Section 6.3.2.

The energy gain calculation was performed using FC2CT, discussed in Chapter 5. FC2CT requires the on-axis E_z field (real and imaginary), and the magnitude of the on axis E_z field, E_0. Therefore to accurately compare SW and TW structures, the value of E_0 must be determinable for a given RF structure and input total power. SW and constant gradient TW structures distribute power differently, an accurate approximation of the power in a given cell is vital for the comparison. The electromagnetic solver CST, provides the value of E_0 for a simulated structure with 1 J total energy (1 W of input power for TW structures). In order to calculate the value of E_0 for a single cell with a different total energy, a scaling relation between the CST cell and real cell must be determined. Recall, from Chapter 3, the MPV places an upper limit on the gradient of $\sim$ 40 MeV/m. As the single cell $E_z(z)$ field profile is exported from CST, the Fourier coefficients are known, and the field profile can be reconstructed after small changes in cell length.

In order to accurately approximate the performance of a structure, optimised single cells must have been simulated and designed, for both SW and TW structures. This was discussed in Chapter 3. The following sections describe the Python script methods used to determine various candidate outputs, such as the cavity gradient, cavity energy gain, and group velocity (in the case of TW).

6.1.1 Gradient as a function of rf Power

As previously described, in order to determine the energy gain of a structure using FC2CT, the gradient (E_0) is required. The method used to calculate E_0 as a function of input power is now described.

The power in an rf cell can be calculated as follows,

$$P_{cell} = \frac{V_{cell}^2}{Z_{cell} L_{cell}}. \tag{6.1}$$

Where Z_{cell} is the shunt impedance per unit length. As the cell voltage is the integral of the electric field over the length, $V_{cell} = \int_0^{L_{cell}} E_z(z)dz = \eta E_0 L_{cell}$. Here, E_0 is the on-axis maximum gradient, and η is some scaling factor that accounts for the shape of $E_z(z)$, and is

constant. Inserting the cell voltage into Eqn. 6.1

$$P_{cell} = \frac{\eta^2 G_{cell}^2 L_{cell}}{Z_{cell}}. \tag{6.2}$$

The power in the cell is also related to the quality factor;

$$P_{cell} = \frac{\omega U_{cell}}{Q_{cell}}. \tag{6.3}$$

Where Q_{cell} is the intrinsic Q-factor. The power of the cell as calculated by CST is therefore, from Eqns. 6.2 and 6.3;

$$P_{CST} = \frac{\eta^2 G_{CST}^2 L_{cell,CST}}{Z_{CST}} \tag{6.4}$$

$$P_{CST} = \frac{\omega U_{CST}}{Q_{CST}}. \tag{6.5}$$

Where the value $L_{cell,CST}$ is not necessarily equal to L_{cell}, as the cell length varies as β_s changes. As η is constant, the shape of the on-axis E_z field is assumed to be the same for both the CST cell and real cell, just with different magnitudes. Rearranging the above expressions and solving for η yields;

$$\eta^2 = \frac{\omega Z_{CST} U_{CST}}{Q_{CST} G_{CST}^2 L_{cell,CST}}. \tag{6.6}$$

The maximum on-axis gradient for the real cell is thus

$$G_{cell}^2 = G_{CST}^2 \frac{L_{cell,CST}}{L_{cell}} \frac{Z_{cell}}{Z_{CST}} \frac{Q_{CST} P_{cell}}{\omega U_{CST}}. \tag{6.7}$$

As CST provides values for all values with subscript 'CST', the remaining unknowns are the cell power, P_{cell}, cell shunt impedance per unit length, Z_{cell}, and cell length, L_{cell}. Whilst Z_{cell} and L_{cell} will be similar to the CST values, Z_{cell} and L_{cell} will change with geometric parameters such as the coupling slot radius (for TW structures) and β_s, respectively. Hence must be treated as variables. However as the changes to the geometry are small, η is assumed constant. Figure 6.1 shows the E_z field profile as calculated by CST for $\beta_s = 0.52$. The Fourier series approximation of E_z was calculated for the same β_s, however the Fourier coefficients were determined from a field map with $\beta_s = 0.515$. The strong agreement between both profiles implies small changes to β_s are valid, and η is constant.

6.1.2 v_g Calculation in a Traveling Wave Structure

The power in a SW cell is calculated as $P_{cell} = \frac{P_{tot}}{N_{cells}}$, as each cell is treated identical, and only one power coupler is used. A constant gradient structure shares the power unevenly

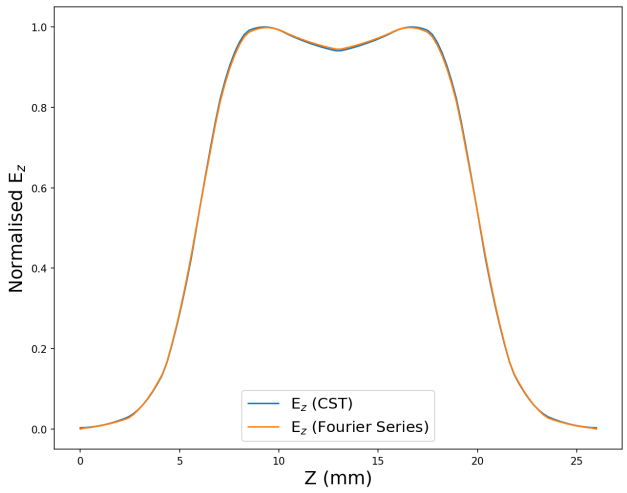


Figure 6.1: E_z field profile as calculated by CST and Fourier series for β_s of 0.52. Fourier series coefficients were calculated based of $\beta_s = 0.515$.

along the structure, to operate each cell at a constant gradient. This is done by tapering the group velocity along the cavity, as discussed previously. The attenuation per unit length at the n_{th} cell is denoted α_n. The group velocity in the n_{th} cell is given;

$$v_{g,n} = \frac{\omega}{2Q_n\alpha_n}. \tag{6.8}$$

The power in the first cell is simply the total power from the input coupler;

$$P_1 = P_{tot}. \tag{6.9}$$

The power in the second cell will be less than the first cell due to power attenuation,

$$P_n = P_{n-1}e^{-2L_{cell}\alpha_{n-1}}. \tag{6.10}$$

The group velocity is tapered from cell to cell such that the energy deposited in each cell (and thus the gradient) is constant,

$$U_{\text{constant}} = P_n v_{g,n}. \tag{6.11}$$

Therefore,

$$\frac{v_{g,n}}{v_{g,n-1}} = \frac{P_n}{P_{n-1}}. \tag{6.12}$$

The gradient in the n_{th} cell, from Eqn. 6.7, is

$$G_{cell} = G_{CST}\sqrt{\frac{L_{cell,CST}Q_{CST}}{Z_{CST}U_{CST}\omega}}\sqrt{\frac{P_{cell}Z_{cell}}{L_{cell}}}. \tag{6.13}$$

The group velocity must be known for a single cell, and is calculated in CST using two methods. The first method uses the fact;

$$v_g = \frac{\delta\omega}{\delta k}. \tag{6.14}$$

Introducing a small change to the phase advance per cell produces a (linear) change to ω. The wave-vector, k, is related to the phase advance per cell via $\frac{\phi}{L_{cell}} = k$. Thus,

$$v_g = \frac{\delta\omega}{\delta\phi} L_{cell}. \tag{6.15}$$

The second method integrates the Poynting vector over the cell aperture and coupling slots. Recall the design of the coupling slots was shown in Chapter 3,

$$P_z = \int_{area} \vec{E} \times \vec{H}|_z dA = \int_{area} [E_r H_\theta - E_\theta H_r] dA = v_g U. \tag{6.16}$$

Where $E_\theta H_r$ is zero for a TM_{010} mode. The group velocity is therefore;

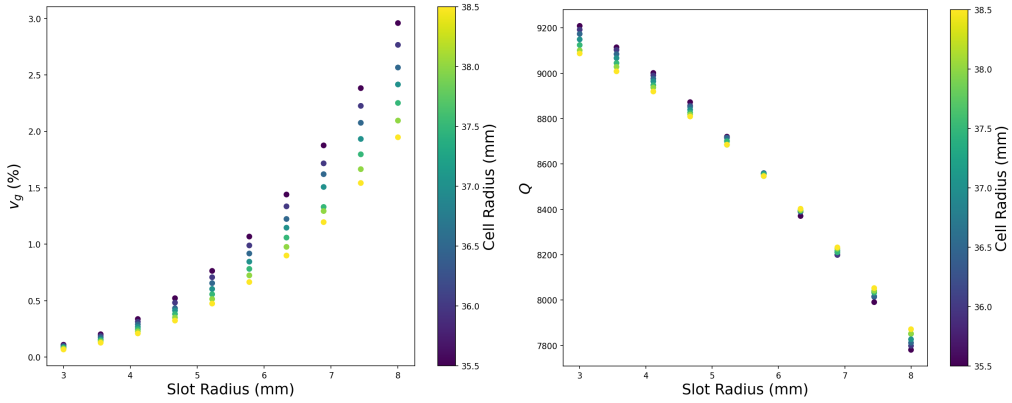
$$v_g = \int_{\text{aperture}} E_r H_\theta dA + \int_{\text{coupling slots}} E_r H_\theta dA. \tag{6.17}$$

In order to slow down the group velocity along the structure, the radius of the coupling slots was used. However, the coupling slot radius also effects important single cell parameters, namely Q, Z and frequency, f. As the cell must oscillate at the correct frequency, the cell radius is used to correct f after changes in the coupling slot radius. As Eqn. 6.7 requires Q and Z, these values must be known accurately, and thus must be determinable for different slot and cell radius.

The coupling slot and cell radius were swept and the cell parameters, including v_g are calculated, as shown in Figs. 6.2 and 6.3. By applying a least squares fit to the results, polynomial expressions can be obtained and the cell parameters are determinable for different cell geometries. The correct frequency and group velocity can be determined by creating a weight function that sums the frequency and group velocity errors (assuming a value for the wanted v_g). Using gradient descent, the weight function tends to zero, optimising the values of coupling slot radius and cell radius. Finally, the values of Q and Z are determined from the respective polynomials.

Effects of changing cell length on effective gradient

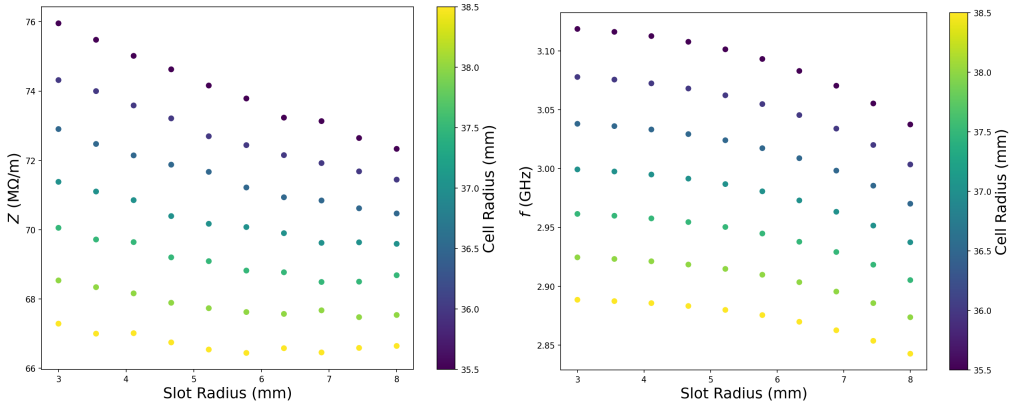
The cavity synchronous beta must be optimised to maximise the energy gain for a given number of cells. This was achieved by sweeping β_s and calculating the effective gradient as a function of number of cells using FC2CT.



(a) v_g as a function of cell coupling slot radius and cell radius.

(b) Q as a function of cell coupling slot radius and cell radius.

Figure 6.2: Cell v_g and Q as a function of single cell coupling slot radius and cell radius.



(a) Z as a function of cell coupling slot radius and cell radius.

(b) f as a function of cell coupling slot radius and cell radius.

Figure 6.3: Cell shunt impedance and frequency as a function of single cell coupling slot radius and cell radius.

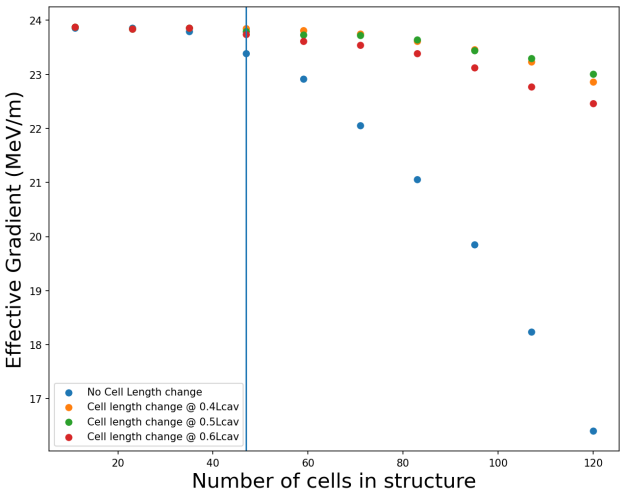


Figure 6.4: Effective gradient as a function of number of cells in a structure for changing cell length along structure. Cavity type is a $\frac{4\pi}{5}$ TW structure at a gradient of 50 MeV/m for protons at 150 MeV.

Figure 6.4 displays the effective gradient of a cavity ($4\pi/5$ TW structure) as a function of number of rf cells, for protons at 150 MeV. Four cavity types are displayed, three cavity solutions have a change in the cell length at some defined point within the cavity. The other cavity shown has constant cell length. The effective gradient is calculated using the FC2CT method, and the gradient is 50 MeV/m. The value of β_s is swept for each cavity (or twice if two cell lengths are considered within the same cavity) and the optimal effective gradient is taken. Cavities with a change in the cell length are expected to perform better (have higher effective gradient) as the total rf phase slippage is smaller. Thus allowing for the synchronicity condition to be continuously met for cavities with large energy gain (large change in relativistic β). The results suggest cavities with constant cell length perform as well as cavities with changing cavity length for structures with approximately 40 cells or fewer, this is equivalent to structure lengths $\sim$ 80 cm. The vertical mark is shown for a cavity of 48 cells, with length $\sim$ 1 m. Cavities of this length, with constant cell length have a marginally lower effective gradient, approximately 0.4 MeV/m, over cavities comprised of two cell lengths. Cavities of constant cell length over 1 m, see an aggressive drop in the effective gradient. Due to the increased cost of machining two different cell lengths for a marginal gain, cavities up to $\sim$ 1 m are comprised of one cell length. The cavity gradient in this project is aimed to be between 30 - 40 MeV/m, due to constraints on the MPV (see Chapter 3), thermal heat load (see later), and the aim to operate at high rf efficiency. As a result, a cavity comprised of one cell length will perform just as well as cavities of two cell lengths up to $\sim 40 \times \frac{50}{40} \approx 50$ rf cells (1 m long). This is because the rf phase slippage is effected by the energy gain, and thus the cavity length, not the number of cells.

At 60 cells, the effective gradient of cavities with constant cell length drop by approximately 4%, relative to the other cavity types. The plot also displays cases where the cell length changes at different point along the structure. When cavities are comprised of 40 - 70 cells, changing the cell length at 40% of the cavity length provides the optimal effective gradient. For cavities comprised of 70+ cells, changing the cell length at the centre becomes the optimum. For shorter structures, the relativistic beta changes more in the first few cells than the last few cells, due to relationship between relativistic beta and energy (a gain of 1 MeV for a proton at 150 MeV will change β by a greater amount than a 1 MeV gain at 160 MeV). Thus, changing the cell length before the mid-point will allow for a more optimal averaged cell length. However, this effect is very minor, and after ~ 70 cells changing the cell length at mid-point becomes optimal. For protons at higher energy, cavities of one cell length will perform as well as cavities with two cell lengths for longer structures (> 1 m), as the value of β changes less for a given energy gain. Conversely, for lower energy protons, cavities comprised of one cell length will become inferior at short cavity lengths < 1 m).

6.2 Structure Energy Gain Calculation

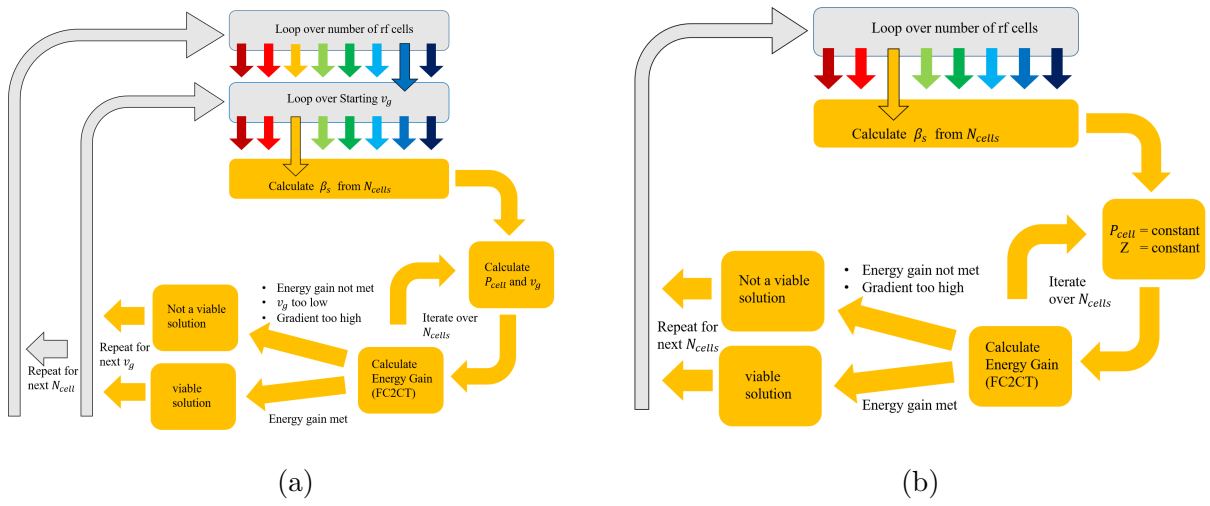


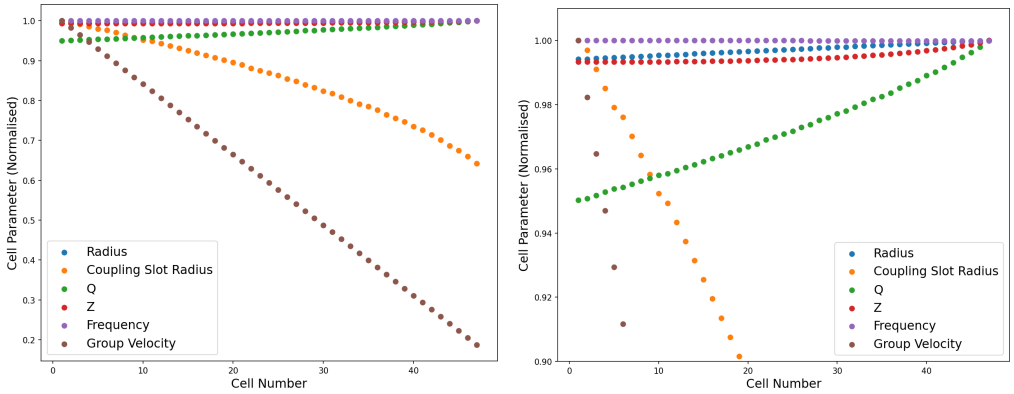
Figure 6.5: Energy gain method for TW (a) and SW(b) cavity types for a predefined input power.

Figure 6.5 displays the method of calculating the total energy gain for a given input power for both TW and SW cavity types. The method discussed was repeated for multiple values of input power.

6.2.1 Traveling Wave Structure Calculation

The TW method considers both the number of cells and the group velocity as the independant parameters that must be swept. For each value of N_{cells}, the optimal β_s is found, using FC2CT. For each value of N_{cells}, the starting v_g is swept. For a given v_g, the simulation iterates over consecutive cells until either the gradient exceeds the cavity limit (40 MeV/m), the required energy gain is met, or v_g becomes too low (< 0.1 % of c). For each cell, the coupling slot radius and cell radius are calculated such that the frequency and v_g are correct. The cell Q, Z are determined using the polynomial relationships and the attenuation, α, is determined using Eqn. 6.8. The FC2CT method is used to calculate the energy gain over each cell using the on-axis gradient, E_0, calculated using Eqn. 6.7. Once the simulation is completed, all solutions that provide the required energy gain are kept. For each solution, the cell parameters, input group velocity, β_s and input power are known.

Figure 6.6 displays the cell parameters as a function of cell number for a TW simulation. As the TW was simulated as constant gradient, the group velocity falls linearly. In order to taper the group velocity as described, the coupling slot decreases with cell number. The cell Q and Z both increase slowly to the final value, however the value do not change substantially. Figure 6.6 (b) shows that, in order to keep the cell frequency constant, the cell radius must increase slightly along the structure to balance the frequency shift due to the decreasing coupling slot radii.



(a) Original (b) Zoomed in

Figure 6.6: Cell parameters as a function of cell number along the cavity. Group velocity tapers from 0.8 % of c to 0.15 % of c.

6.2.2 Standing Wave Structure Calculation

The method for approximating the performance of a SW cavity is similar to the TW method. However for the SW method, one factor that must be considered is the effect of cavity length

on the shunt impedance per unit length.

From Fig. 6.6, a TW structure observes a slight increase in Z along the structure. However in a SW cavity, as the number of cells increases, the inter-cell coupling constant must increase, to ensure the same end cell phase shift and power flow droop, as discussed in Chapter 2. Increasing the inter-cell coupling can be achieved by increasing the aperture radius. Increasing the aperture radius moves the nose-cone into a region of higher magnetic field, which increasing magnetic losses. In addition, an increasing aperture will reduce the concentration of axial electric field lines. Both effects reduce the shunt impedance. Another other method to increase coupling is by introducing off-axis coupling slots, as previously discussed.

The effect of the coupling constant on the end cell field magnitude and phase can be understood from perturbation theory for a chain of coupled oscillators. Consider a side-coupled structure, comprised of accelerating cells and side-coupled coupling cells. Accelerating cells are conventionally labelled $2n$, and thus coupling cells are labelled $2n+1$. The Q factors for each type of cell are Q_a and Q_c respectively. The number of coupling cells is one fewer than the number of accelerating cells and so the total number of cells is odd. The stop-band is the region between the lower and upper pass-bands of the dispersion curve and has a magnitude equal to the difference in frequency between the accelerating and coupling cells, $\delta\omega = \omega_c - \omega_a$. From perturbation theory, the second order approximation of the field in an accelerating cell is given [35];

$$X_{2n} \approx (-1)^{n-m} X_{2m} \left[1 - \frac{2(m^2-n^2)}{k^2 Q_a Q_c}\right] e^{i\frac{4(m^2-n^2)}{k^2 Q_a}\frac{\delta\omega}{\omega_a}}. \tag{6.18}$$

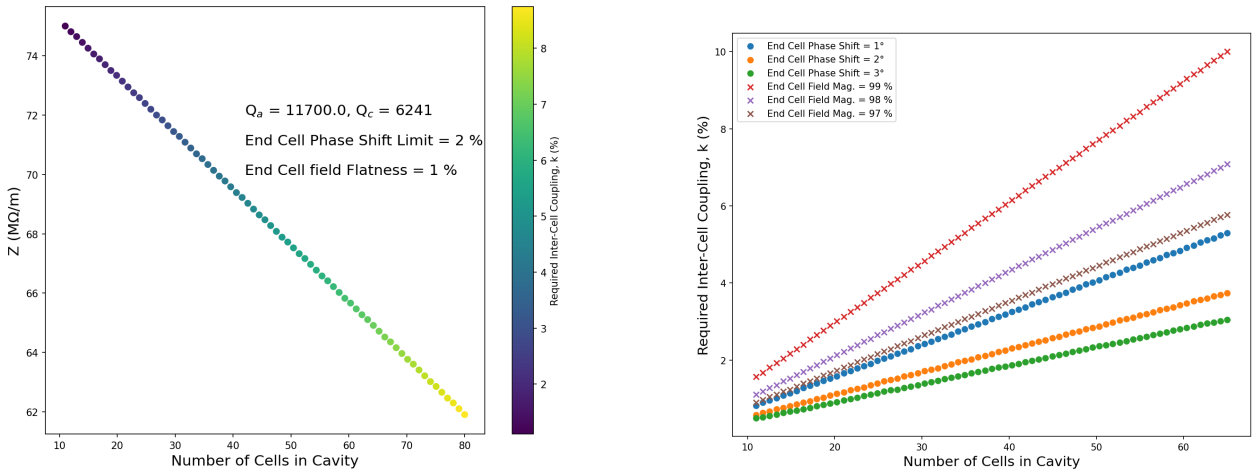
Where X_{2m} is the field in the drive cell, and k is the coupling constant between accelerating and coupling cells. The value of n determines the accelerating cell the field is calculated in, and can take values $n = 0, ..., m$. The value of m depends on the total number of cells as $m = \frac{N_{total}-1}{4}$. The second term in the square brackets of Eqn. 6.18 describes the power-flow droop. The complex term describes the power-flow phase shift. The power-flow phase shift vanishes when the resonant frequency of the accelerating and coupling cells are equal, at confluence.

The ratio of the end cell field magnitude to the drive cell magnitude and the end cell phase shift are two important values for cavity design that require maximising (to a magnitude of one) and minimising, respectively. Both objectives are functions of the coupling constant.

The ratio of the end cell field magnitude to the drive cell (maximum power flow droop) is given

$$\left|\frac{X_{end}}{X_{drive}}\right| = 1 - \frac{(N_{total}-1)^2}{8k^2 Q_a Q_c}, \tag{6.19}$$

hence for increasing k the closer the ratio of field magnitudes is to unity. As the number of cells increases, the gradient in the end cell reduces relative to the drive cell. The maximum



(a) (b)

Figure 6.7: (a) Shunt impedance as a function of number of cells and required coupling constant (for fixed end cell phase shift and end cell field flatness) for a side-coupled SW cavity. (b) Required coupling constant as a function of number of cells and end cell phase shift and field flatness, for a side-coupled SW cavity.

power flow phase shift at the end cell is given

$$\phi_{shift} = \frac{(N_{total}-1)^2}{4k^2 Q_a}\frac{\omega_c - \omega_a}{\omega_a}, \tag{6.20}$$

which increases with N_{total} and decreases with inter-cell coupling, k. The power flow phase shift is also a function of difference in frequency between accelerating and coupling cells. This shows the important of tuning a side-coupled structure to reach confluence. The field in the coupling cells is

$$X_{2n+1} = (-1)^{n-m} X_{2m}\left[\frac{2n+1}{kQ_a}\right] e^{i\frac{\pi}{2}}. \tag{6.21}$$

Note the magnitude of the field in a coupling cell is largest for cells closest to the drive cell ($n = m$) and smallest for $n = 0$.

Figure 6.7 (a) displays the shunt impedance as a function of number of cells in a side-coupled SW cavity. The SW cavity simulated is the one described in Section 3, and has been optimised. A cavity length of 1 m would be comprised of approximately 40 cells for a $\pi/2$ mode side-coupled structure. Thus, for a SW cavity of realistic cavity length range (0.1 m - 1 m) the shunt impedance range is approximately 69 - 75 MΩ/m. From Fig. 6.3 (a), the $4\pi/5$ TW cell shunt impedance is in the region of 66 - 76 MΩ/m. Thus, both optimised structures have similar ranges for the shunt impedance. The total power (for both TW and SW) requirement was modulated by a factor to account for losses in the rf network, 20% losses is often assumed, and is the value used here.

6.3 Simulation Results

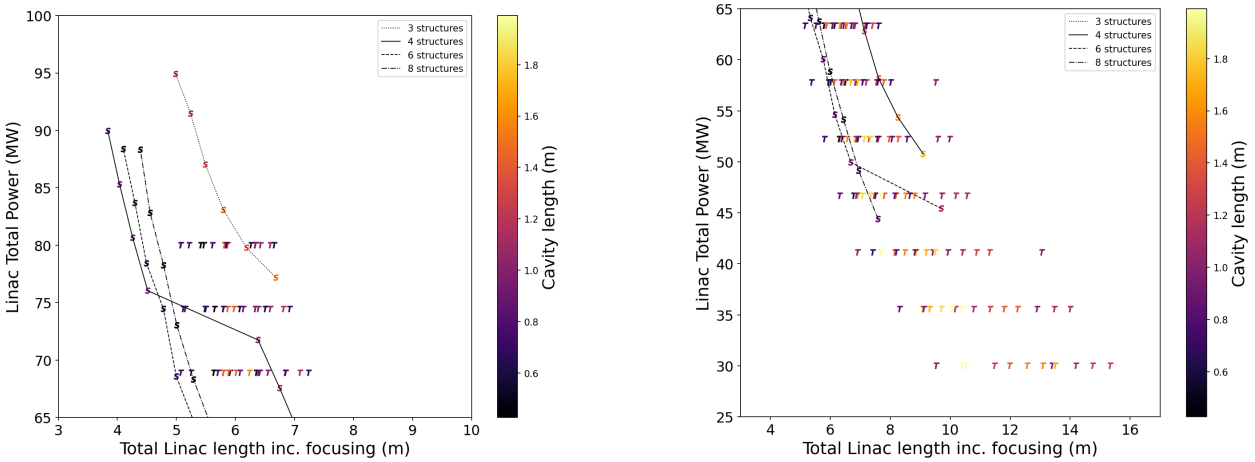
6.3.1 Unconstrained Scan Results

Firstly, a major deep scan of the parameter space was completed, ignoring any constraints as described in table 6.1. The only constraint was that the total energy gain must exceed 100 MeV. The motivation for this scan was to explore the entire parameter space of possible solutions for the cavity design. As cavity length was not constrained, both the FODO-like and MAS focusing schemes may be required (recall Chapter 4 discussed two focusing schemes which allowed for two values of maximum achievable cavity length), and therefore the total linac length included additional length due to focusing elements. Linac solutions using the MAS are assumed to have three quadruples and three drift lengths per cavity, whilst solutions that use the FODO-like scheme (cavity lengths under 1 m) have one quadrupole and two drift lengths per cavity. From Chapter 4, the quadrupole length and drift length are assumed to be 0.1 m.

Figure 6.8 displays the design space for an unconstrained linac system. The plot displays the total power required (power per cavity $\times$ number of cavities in linac) as a function of total linac length, including focusing (sum of cavity lengths plus additional length for focusing). Each point is a linac solution, not a single cavity. The minimum and maximum number of cavities per linac was two and eight, respectively.

The objective space creates a Pareto front with respect to the total power and total length objectives. For very high power systems, (shown in Fig 6.8(a)) SW solutions become optimal over TW solutions. This is because shorter cavity lengths (roughly, less than 0.5 m) are more efficient at distributing the rf power, due to reduced losses, and increased shunt impedance per unit length as a result of fewer cells. At this area in objective space, TW solutions can not compete with the SW design. The SW linac solution comprised of three structures is not optimal, as the individual cavity lengths become too long, and the shunt impedance drops. SW solutions with four, six and eight structures all compete fairly well at high power. Once the total power is reduced to $\sim$ 75 MW, four structures becomes a poor solution as the reduction in power reduces the operating gradient, which increases the length of the four SW structures (to $\sim$ 1 m), and forces the use of the MAS over the FODO-like scheme, which requires additional longitudinal length (relative to FODO) to focus, as discussed in Chapter 4. At a total power of $\sim$ 50 MW, the six cavities in the SW solution have lengths approaching 1 m and adopt the MAS, increasing the total linac length considerably.

At $\sim$ 60 MW total power, TW solutions dominate the Pareto front (shown in Fig 6.8(b)). At lower input power, the cavity gradients drop, and cavity lengths increase to reach the required energy gain. Longer cavity lengths are more efficient when operated as TW cavities over SW cavities, as the shunt impedance does not reduce with increasing number of cells. The different TW solutions observed have different cavity lengths and different numbers of cavities per linac solution, all solutions are operated in the $4\pi/5$ mode as constant gradient



(a) Standing Wave dominated solutions.

(b) Traveling wave dominated solutions.

Figure 6.8: Linac total power as a function of total linac length, including focusing from quadrupoles. Color dimension shows individual cavity length. Solutions are depicted as SW and TW by 'S' and 'T', respectively.

structures. The optimal TW solutions have cavity lengths ~ 1 m and thus use the FODO-like scheme. One TW solution that performs well is the linac comprised of four TW cavities with lengths ~ 1.8 m (yellow solutions). These solutions require the MAS for sufficient focusing, however due to the higher rf power efficiency obtained with longer TW structures, perform well with respect to other TW solutions. This because there are fewer cavities comprising the linac solution. Whilst these solutions show the potential strength of very long TW structures in the MAS, they are still not optimal to shorter TW structures. The focusing length of a two cavity linac in the MAS is assumed to be the same as a six cavity solution in the FODO-like scheme (same total number of quadrupoles (6)). When decreasing the total linac power, there is a point where the real estate gradient of the MAS becomes superior to the FODO-like scheme, as the ratio of accelerating to non-accelerating elements shifts towards MAS solutions. For example, suppose a MAS focused solution existed with one cavity of length 10 m, requiring only three quadrupoles. The real estate of this solution is very high. However, the is not viable for this linac design, as the required input power would be too large. As linac solutions with four or fewer cavities require the MAS scheme (lengths > 1 m), they are inferior to FODO-like linac solutions of six or more cavities.

The general result that SW cavities are superior at short cavity lengths and TW are superior at longer lengths is an expected result. The fact this result was retained implies the model is operating as intended. This result can be understood with the following discussion.

SW cavities comprised of fewer cells require smaller coupling for the same power flow droop and phase shift. Therefore have higher shunt impedance per unit length than longer SW cavities.

TW structures dump unused power into a load at the end of the cavity. If the cavity is run with a high group velocity, power flows through the structure quickly, little power is absorbed in each cell resulting in a low gradient, and most of the power is dumped into the load. If the group velocity is lower, each cell will absorb more rf power (due to higher α) and the (constant) gradient is higher. Hence, the group velocity of a TW structure is an important design choice.

For a given input power, the group velocity decreases linearly along a constant gradient TW structure, at a rate that is independant of the input power and the starting group velocity. The rate at which the power drops along the structure is also linear, however not independent of the starting group velocity. In order to limit the power dumped into the load, the rate at which power is used must increase with decreasing TW structure length (the gradient of $\frac{P_{cell}}{P_{input}}$ must become more negative). In order to accomplish this, the group velocity must be very low (more power is deposited per cell). There is, however, a limit to the group velocity in a TW structure (~ 0.1 % of c), it can not be made arbitrarily low. This is because low v_g structures have increased risk during operation, as the power may not reach the end of the structure. Lower group velocity can result in the gradient starting to decay along the cavity, and the structure is more sensitive to machining errors as v_g decreases. Hence, short TW structures must either accept high amounts of power into the load, or accept the risk of lower group velocities. Whilst one could use a lower power source, so that the unused power is less, the gradient will be low. Nevertheless, low group velocity structures have increased rf efficiency, as more energy is deposited per cell, for a given input power.

As the length of TW structures increases, the rate at which power is deposited can reduce (technically increase, as the rate is negative). This is because there are more cells to fill with electromagnetic energy (in addition, the shunt impedance does not decrease with cell number, unlike SW). Therefore, longer TW structures can operate at higher group velocity, with lower values of the ratio P_{cell}/P_{in}. Therefore there is a trade-off; reducing v_g to increase rf efficiency, or running with higher v_g, and dumping more rf power. In the limit the structure length becomes very long, the group velocity can be initially high whilst dumped power is low. In reality, structures can not be very long (for beam dynamics and machining reasons) and thus the decision between operating group velocity and dumped power is a compromise. The choice of group velocity also effects the fill time. In the context of medical accelerators, the lower the fill time, the more protons can be accelerated per rf pulse. The fill time is therefore an important objective.

6.3.2 Constrained Scan Results

From the results obtained in Fig. 6.8, the linac solution could proceed as either SW or TW design, as both solutions thrive in different areas of objective space.

As a result, a decision was made with the industrial partners, AVO, to constrain the design space and complete a less general comparison of all structure types. The input power

became a constrained value, with peak powers of 3.75 and 5.5 MW. Each klystron can power one or two cavities (by being split once). These values came from known klystrons that are commercially available [99]. The number of cavities was limited to six or eight. Any fewer than six requires TW structures that become very long and require the MAS, which from Fig. 6.8, do not outperform shorter TW structures, as previously discussed. Whilst eight or more cavities is possible, the total power required increases (dropping power efficiency), leading to higher initial and running costs. The reason five or seven cavities was not explored is to future proof the design. Should higher power klystrons be used (over 5.5 MW), they can be split between two cavities, which requires an even number of structures. Cavity lengths longer than 1 m are omitted, as the MAS increases the real estate gradient until very long cavities are considered. Thus all linac solutions are focused using the FODO-like scheme.

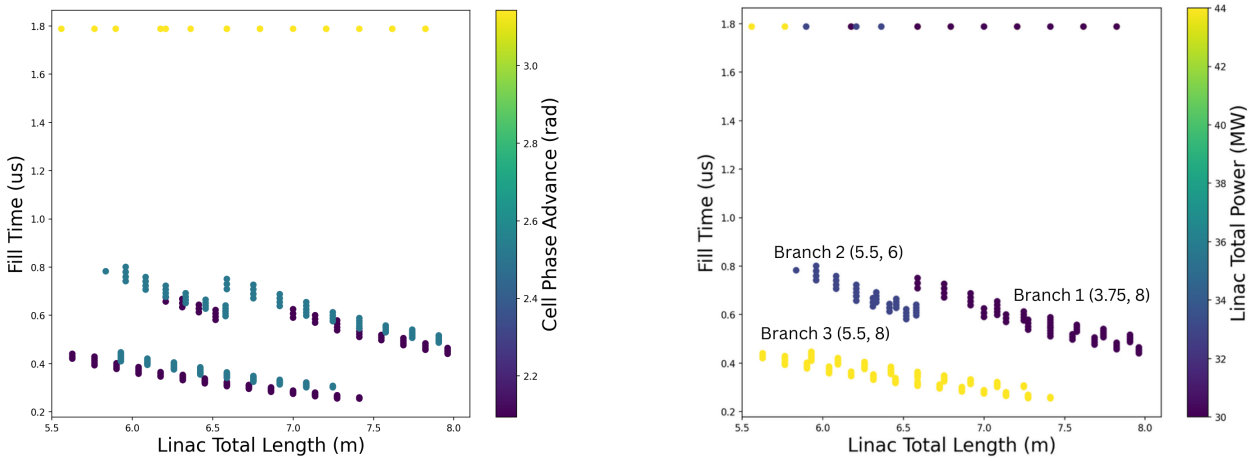
As described in Chapter 3, the peak field values are important figures of merit that require minimising. Another gradient limiting factor is the average power deposit per unit length, $\frac{P_{avg}}{L_{cav}}$. This figure of merit concerns the average heating due to the magnetic field within an rf cell. The average heating can cause deformation of the cell walls and may result in a shift in resonant frequency. Average heating is influenced by the repetition rate and pulse length of the power source. As the cavity is to be used for proton radiotherapy, a conservative choice for the average power per unit length of 4.5 kW/m was agreed, to heavily reduce the risk of operational detuning. S-band cavities designed for medical applications have previously been designed and tested with average power per unit length in the range 4.7 - 8 kW/m, [51], [88] and [100]. The calculation of average power was calculated assuming a repetition rate of 200 Hz and pulse length of 5 μs.

In order to produce a highly efficient structure, relatively aggressive limits on the input/output group velocity were chosen as 0.8/0.15% of c. Group velocities in this region have been achieved before for an S-band linac accelerating protons for medical applications, see [88], [101].

To produce a more complete design space for the constrained scan, two TW structures were simulated, with phase advances of $\frac{2\pi}{3}$ and $\frac{4\pi}{5}$. This required two optimised single cell models to be created in CST for each phase advance. Each model was comprised of four off-axis coupling slots as described in Chapter 3. In addition, scans had to be computed for each single cell model so that polynomial relationships could be determined between the cell radii, R, coupling slot radii, r, and the cell figures of merit, v_g, Q, Z and f.

Figure 6.9 displays the fill time of linac solutions as a function of total cavity length. Figure 6.9 (a) displays the cavity phase advance per cell with the colour dimension, whilst Fig. 6.9 (b) displays the total linac power required with the colour dimension. The SW cavity is shown with a phase advance of π (although it is technically the $\pi/2$ normal mode) and has a constant fill time as a function of cavity length. The SW solutions have a much larger filling time, however have similar values for linac total length. As a result, SW solutions were omitted from the solution space, as there was no discernible benefit over TW structures.

The TW solutions can be separated into three branches (shown in Fig. 6.9 (b)). Within



(a) (b)

Figure 6.9: Cavity fill time as a function of total linac length and (a) Phase advance per cell or (b) Linac total power.

a given branch of solutions, solutions are separated by phase advance per cell ($2\pi/3$ and $4\pi/5$), and different input v_g. Shorter linacs operate at higher gradients, and lower group velocity, thus have longer fill time. Branch 3 displays linac solutions with eight structures, each powered with 5.5 MW klystrons (44 MW total peak power). These solutions have the lowest fill time (< 0.5 μs). Due to the thermal load limit, additional power could not be used to improve the gradient (equivalent to decreasing the ratio P_{out}/P_{in}), therefore the higher total power allowed a higher group velocity for the same gradient, which minimises the fill time relative to the other branches.

Within branch 3, the $2\pi/3$ structure is superior to the $4\pi/5$ structure, as the fill time is marginally less for the same linac total length. In addition, only $2\pi/3$ structures could achieve total linac lengths under ~ 5.8 m with the given constraints. This implies, for higher average group velocity, the $2\pi/3$ structure performs better than the $4\pi/5$ structure. Branch 2 consists of solutions with 33 MW total power, with six structures of 5.5 MW peak power each. The branch has both $2\pi/3$ and $4\pi/5$ structures, however the $4\pi/5$ structures perform better, as $2\pi/3$ structures can not produce solutions with linac total length less than ~ 6.2 m. For these solutions, there is less available power to run at high group velocity for the same gradient. As a result, the fill time is longer (than branch 3), by a factor of ~ 2. Branch 1 solutions have 8 cavities each powered with 3.75 MW, for a total power of 30 MW. Branch 1 linacs operate at lower gradient, and have similar average v_g to branch 2 solutions, as the fill times are similar. Whilst branch 3 solutions require 3 MW less total power than branch 2 solutions, the additional total length is substantial. Similar to branch 2, $4\pi/5$ structures perform better for branch 1 linacs.

The comparison between $2\pi/3$ and $4\pi/5$ structures is considered using the result in

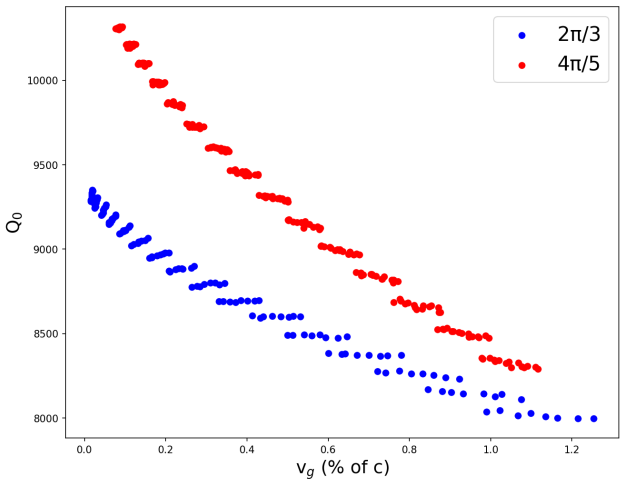


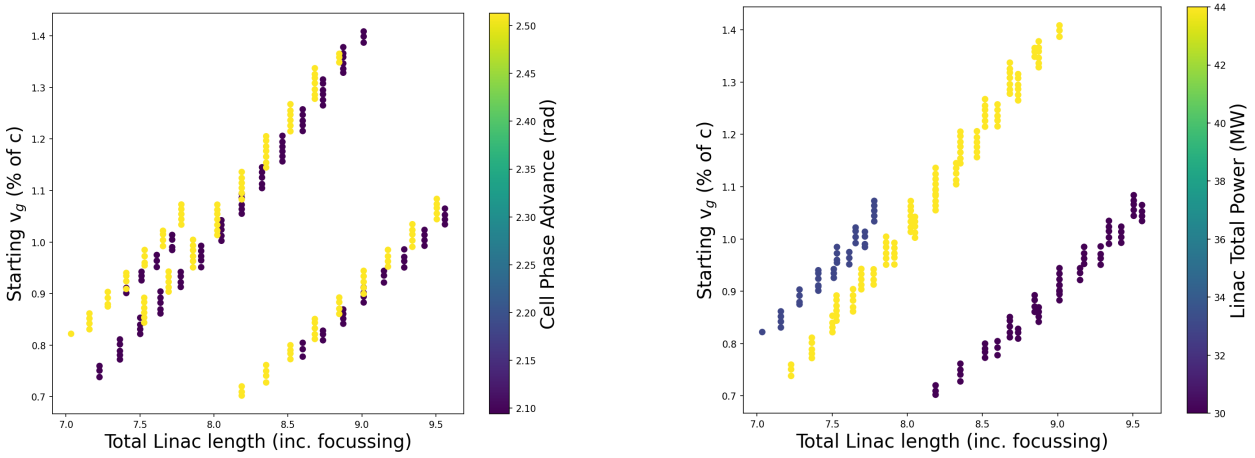
Figure 6.10

Figure 6.11: Q_0 as a function of cell group velocity for both the $2\pi/3$ and $4\pi/5$ TW structures.

Fig. 6.11. $2\pi/3$ structures have lower Q than $4\pi/5$ structures for the same v_g. Therefore, $2\pi/3$ structures have higher attenuation for a given v_g, from Eqn. 6.8. As described previously, shorter TW structures can not reduce v_g indefinitely in an aim to reduce the rate of P_{cell}/P_{in} (to decrease the dumped power). As $2\pi/3$ structures can attenuate more power per unit length for the same group velocity, there is less dumped power, relative to the $4\pi/5$ structure, as more power has been used to excite the individual cells. This is seen in branch 3 of Fig. 6.9, where $2\pi/3$ structures can operate with higher group velocity for the same gradient, and vice-versa.

When focusing elements are taken into account, the realistic total linac length is different for six structure linacs relative to eight structure linacs. For each cavity in a FODO-like scheme, there is a quadrupole and two drift lengths. The additional length to accommodate non-accelerating elements must be explored, to give the real estate gradient. Figure 6.12 displays the starting group velocity as a function of total linac length including focusing and structure type (a) or total power (b). Again, there are three branches of solutions each with different total power. Low total power solutions (branch 1) have $\sim$ 1 m of additional linac length, for starting group velocity > 0.8 % of c. Whilst these solutions require only 30 MW of total power, the real estate gradient is too low and therefore the solutions are omitted. For solutions with starting group velocity over 0.8 % of c, branch 2 solutions have higher real estate gradients relative to branch 3. The solution with the highest real estate gradient from branch 2 is a $4\pi/5$ structure, whilst the highest real estate for a branch 3 solution is a $2\pi/3$ structure.

When comparing the remaining branches (2 and 3) of solutions, there are distinct benefits of both designs. From each branch, one solution was put forward to a short-list of solutions,



(a) (b)

Figure 6.12: Starting group velocity as a function of total linac length including focusing and (a) Phase advance per cell or (b) Linac total power.

displayed in table 6.2.

Table 6.2: Potential solutions for conceptual linac design for the energy range 150 - 250 MeV.

Solution	1	2	3	4
Total Power (MW)	44	33	44	33
Total Linac Length inc. focusing (m)	8.05	7.78	7.65	7
Cavity Length (m)	0.8	1.1	0.75	0.96
Number of Structures	8	6	8	6
v_g (% of c) [start, end]	[1, 0.45]	[1, 0.31]	[0.9, 0.35]	[0.82, 0.18]
Phase Advance	$2\pi/3$	$4\pi/5$	$4\pi/5$	$4\pi/5$
Fill Time (μs)	0.3	0.6	0.36	0.75
Gradient (MeV/m)	30	29	32	33
P_{out}/P_{in}	0.45	0.31	0.38	0.22

Table 6.2 displays four conceptual design choices for the proposed linac, accelerating protons from 150 - 250 MeV. Solution 1 is the 'safest' linac, running at higher average group velocity, thus benefits from a shorter fill time. Solution 4 is the most 'aggressive' linac, requiring the least rf power, with the highest real estate gradient. Solutions 2 and 3 are in-between the safer and more aggressive solutions. Solutions 2 is comprised of 6 cavities of length 1.1 m. The issue of transverse focusing within the FODO-like scheme becomes a limiting factor, however the solution is attractive due to the low power requirement. Solution 3 is a more aggressive version of solution 1; an 8 structure linac running at lower group velocity, with slightly higher fill time and real estate gradient. Solution 4 is the most

aggressive linac, requiring the least power and achieving the highest real estate gradient, however has the longest fill time.

Fundamentally, the choice is whether to stick at 8 structures, using a less aggressive linac design. Or, to extract higher amounts of energy using lower group velocities and completing the required energy gain in 6 structures, saving on focusing length and total power. In the end, solution 4 was chosen, as it requires 25 % less total power, and has the highest real estate gradient. Solution 4 is still within the average deposited power limit of 4.5 kW/m, as well as the group velocity limits of [0.8, 0.1] (% of c). Solution 4 also has the lowest amount of dumped power, 22 %, and has the highest rf efficiency. The increased difficulty of machining the structures with lower v_g was an accepted trade-off, as structures with low v_g have been previously demonstrated.

The simulation results produces the number of cells, the group velocity of each cell, the cell radii, and coupling slot radii. As the simulation swept over synchronous beta, the cell length is also known, and kept constant. The first cavity (of six) is comprised of 47 cells, and is approximately 1 m. The input v_g is $\sim$ 0.82 % of c and the output value is $\sim$ 0.18 % of c. The gradient of the cavity is 33 MeV/m. From discussions of Fig. 6.4, the increase in real estate gradient from changing the cell length for a 1 m long cavity is around 0.5 MeV/m, for $E_0 = 50$ MeV/m. For a gradient of 33 MeV/m, the benefit would be minor. The above information for the remaining five structures are also known, and each cavity is $\sim$ 1 m long, with the number of cells per structure equal to 47, 46, 45, 43, 42, and 41. The input and output group velocity are very similar to the values of the first cavity.

6.4 Conclusion of Power Requirements

This chapter outlined a simulation method that calculated the energy gain of a single cell as a function of input rf power. In order to model the cells as accurately as possible, CST models were created, which allowed for the direct corresponding E_z magnitude for a given rf power. This method was used to calculate the optimal synchronous beta, and required power, for a given energy gain. Both SW and constant gradient TW structures were simulated, allowing for a deep comparison between all possible structure types. Given a conservative constraint on the deposited average power per unit length, and the available input power, the entire parameter space was simulated to find the optimal solutions. The conceptual design will be a six cavity linac, each powered with a 5.5 MW klystron. Each cavity will be a $4\pi/5$ TW structure, with lengths $\sim$ 1 m, allowing the use of the FODO-like focusing scheme. The rf sources are able to run with a RR of 200 Hz and pulse length of 5 μs, which are standard values for proton radiotherapy medical linacs. The average power deposition was limited to a conservative value of $\sim$ 4.5 kW/m, to reduce the risk of operational detuning. The fill time is $<$ 0.8 μs. Another solution was considered at the latter stages of the design choice, however were omitted due to lower rf efficiency and lower real estate gradient. The next stages require the electromagnetic design of the first cavity of six, including the design of the

two power couplers.

Chapter 7

Final Electromagnetic Design

7.1 RF Power Coupler Design

An important aspect of the electromagnetic design involves the rf couplers. As the cavity is TW, both input and output couplers (also called matching cells) need to designed separately. The coupler waveguide is attached to the first/last cell via a 'bottleneck', with an associated width, height and length. The length of the waveguide port is equal to the length of the bottleneck and the first cell length. The waveguide length is tapered such that the width $\times$ length dimensions match that of a standard waveguide port. The taper can be seen in Fig. 7.1. The design of a TW coupler must ensure there are zero reflections from the boundary between the bottleneck and first cell (also refereed to as the drive cell), at the design frequency.

This is achieved by tuning two parameters, the drive cell radius, shown in Fig. 7.1, and either the bottleneck width or height. Changes to either bottleneck parameter can be utilised as both effect the cut-off frequency in the same way. However, it should be noted that increasing the bottleneck height has the opposite effect to increasing the bottleneck width.

The scattering matrix parameters can be found using an electromagnetic solver software, such as CST, after an input port is defined. Contrary to the method of coupler design for a SW cavity, a global minimum in S11 at the design frequency is not a valid approach. This is because reflections within the cavity travel back and destructively interfere with reflections from the coupler port. This method of global matching ensures no rf power is reflected, and is therefore inside the cavity. However, a TW coupler is designed to have zero reflections from the input port. Therefore a method is required to calculate the local reflections inside a cavity, and to tune the coupler dimensions until the reflections are zero [102]. A coupler cell is designed with the aim of achieving an S11 (the reflection coefficient) of less than -20 dB (1 % reflections).

Before the symmetrical method is simulated, a closer approximation can be calculated by matching the external Q of the coupler to the Q of the matching cell. The Q factor of the matching cell have two components, the intrinsic Q_0 and the Q associated with the flow of

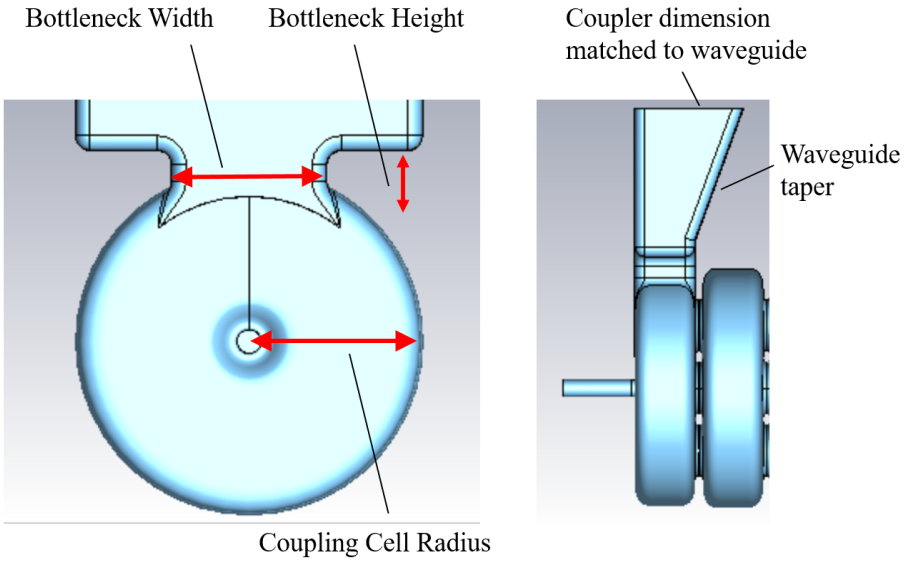


Figure 7.1: Figure showing the coupler dimensions. The waveguide length is tapered to the length of the bottleneck (and single cell).

power due to the finite group velocity, termed Q_{flow}.

$$\frac{1}{Q_{ext}} = \frac{1}{Q_0} + \frac{1}{Q_{flow}} \tag{7.1}$$

Q_0 is the intrinsic Q and is therefore given $Q_0 = \frac{\omega U}{P_{cell}}$. Where U is the deposited energy and P_{cell} is the deposited power. As Q_{flow} is associated withe the flow of power, it is given as the energy lost per cycle due to the total power flow, not just the deposited power, P_{cell}, from Ohmic losses.

$$Q_{flow} = \frac{\omega U}{P_0}. \tag{7.2}$$

The deposited energy is $\frac{P_0 L_{cell}}{v_g}$, therefore;

$$Q_{flow} = \frac{\omega L_{cell}}{v_g}. \tag{7.3}$$

Thus, the matching criteria is;

$$\frac{1}{Q_{ext}} = \frac{P_{cell} L_{cell} v_g}{\omega P_0} + \frac{v_g}{\omega L_{cell}}. \tag{7.4}$$

The second term in Eqn. 7.4 dominates, and therefore

$$Q_{ext} \approx \frac{\omega L_{cell}}{v_g}. \tag{7.5}$$

The phase advance per cell is $\phi = \frac{\omega L_{cell}}{\beta_s c}$, hence

$$Q_{ext} \approx \frac{\phi \beta_s c}{v_g}. \tag{7.6}$$

A simple model can be constructed in CST to calculate the coupler Q_{ext}. By changing the matching cell radius and the coupler width or height, a good initial coupler can be designed. The following method can then be implemented to fine tune the matching cell dimensions.

The following coupler match stage considers a symmetric simulated structure. The structure is comprised of a beam pipe section (with length equal to the cell length), the matching cell, and a certain number of regular cells. The more regular cells, the better the approximation will be. For a tapered structure, at least four cells are required [103]. The simulated structure is then mirrored at the fourth cell so that there are eight cells (including two identical matching cells) and two beam pipes, as shown in Fig. 7.2.

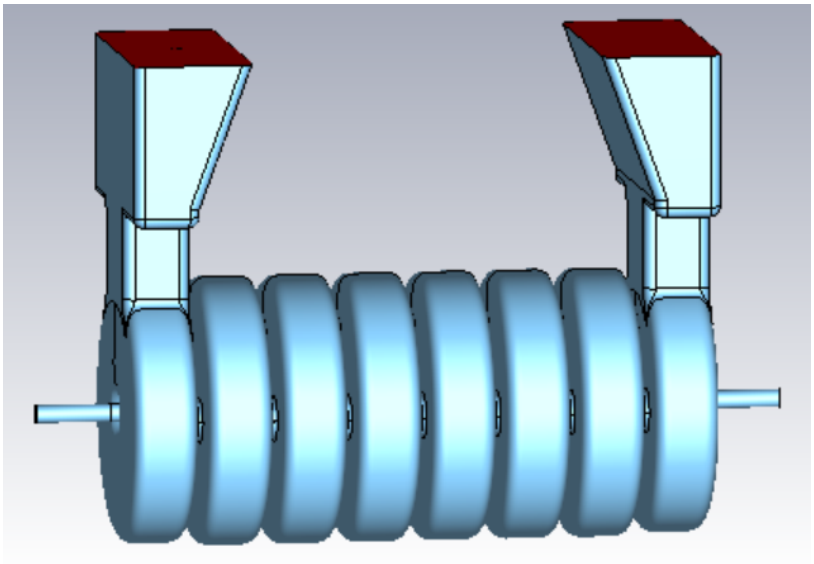


Figure 7.2: Figure showing the simulated structure with eight total cells, two of which are matching cells, and two beam pipes.

The wave that exists inside the cavity in Fig. 7.1 is a periodic traveling wave, that can be written as a propagating wave and a reflected wave, with some reflection coefficient, $R(z)$. The reflection is a function of longitudinal distance, z.

$$E(z,t) = e^{ikz} + R(z)e^{-ikz} \tag{7.7}$$

Where k is the wave vector. As the field is periodic, Floquets theorem states that the field at two locations, separated by a one period, differ only by some complex phase, called the

phase advance. Therefore, the field at a position $z + L$, where L is the period of $E(z, t)$ is,

$$E(z + L, t) = e^{ikz+\phi} + R(z)e^{i(-kz-\phi)}. \tag{7.8}$$

Equivalently, at $z - L$,

$$E(z - L, t) = e^{ikz-\phi} + R(z)e^{i(-kz+\phi)}. \tag{7.9}$$

Summing Eqns. 7.8 and 7.9 and dividing by Eqn. 7.7 gives $\Sigma(z)$;

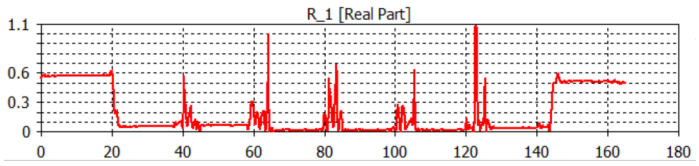
$$\Sigma(z) = \frac{E(z + L, t) + E(z - L, t)}{E(z, t)}. \tag{7.10}$$

Substituting the field profiles produces the following

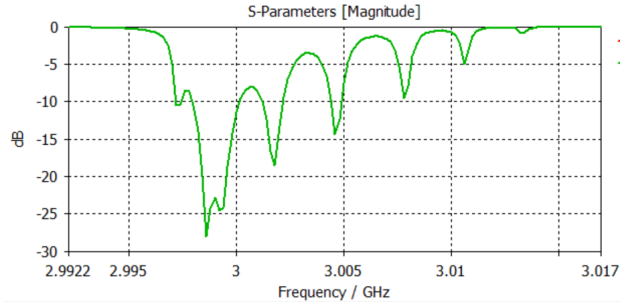
$$\Sigma(z) = \frac{e^{i(kz+\phi)} + R(z)e^{i(-kz-\phi)} + e^{i(kz-\phi)} + R(z)e^{i(-kz+\phi)}}{e^{ikz} + R(z)e^{-ikz}} = 2\cos(\phi). \tag{7.11}$$

Therefore,

$$\phi(z) = \arccos\left(\frac{\Sigma(z)}{2}\right). \tag{7.12}$$



(a) Reflection coefficient



(b) S11

Figure 7.3: Reflection coefficient and S11 for the eight cell model for the output coupler.

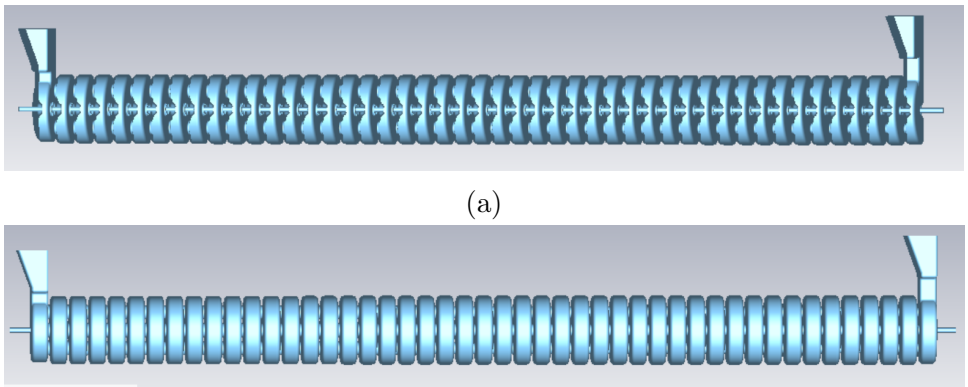
Equation 7.12 describes the phase advance of the field as a function of z. Taking the difference between Eqns. 7.8 and 7.9 and dividing by Eqn. 7.7 gives $\Delta(z)$;

$$\Delta(z) = 2i\sin(\phi)\frac{e^{ikz} - R(z)e^{-ikz}}{e^{ikz} + R(z)e^{-ikz}}. \tag{7.13}$$

Rearranging for the reflection coefficient, $R(z)$,

$$R(z) = \frac{2\sin(\phi) - i\Delta(z)}{2\sin(\phi) + i\Delta(z)}. \tag{7.14}$$

Thus, the phase advance, $\phi(z)$ and local reflection, $R(z)$, can be calculated provided the field $E(z,t)$ is known at three locations separated by the period L. The fields can be calculated in CST for the model shown in Fig. 7.1.



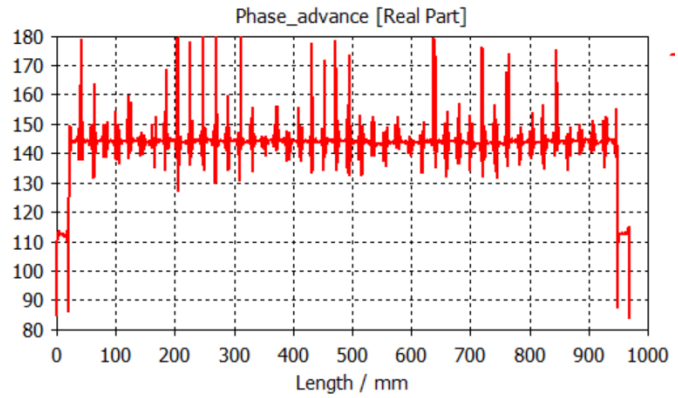
(a)

(b)

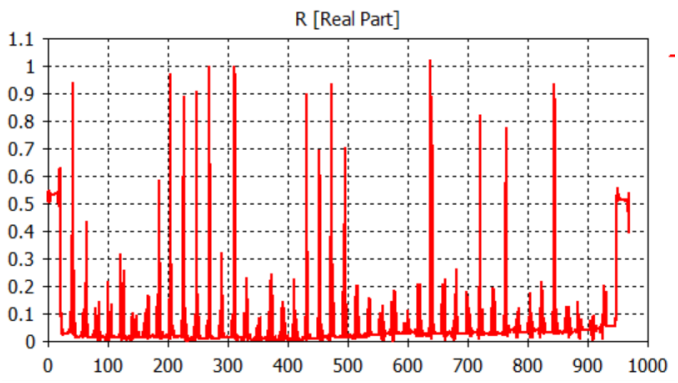
Figure 7.4: Full 47-cell TW structure.

Figure 7.3 displays the reflection coefficient and S11 for the output coupler structure. This was achieved by iterating over the matching cell radius and bottleneck height variable in turn, moving toward the minimum in R. The S11 shows all eight modes overlap into one continuous spectrum. The S11 at the design frequency of 2.9985 GHz is lower than -20 dB and is therefore acceptable. The S11 can not be a minimum at 2.9985 GHz due to reflections at the end coupler destructively interfering, as the reflections are very low along the entire structure. As the group velocity is low at the output port (≈ 0.2 % of c) the penultimate cell radius required slight changes to reduce the reflections further. The procedure was repeated for the input matching cell. If the output matching cell is correctly designed, there are very low reflections back to the input port.

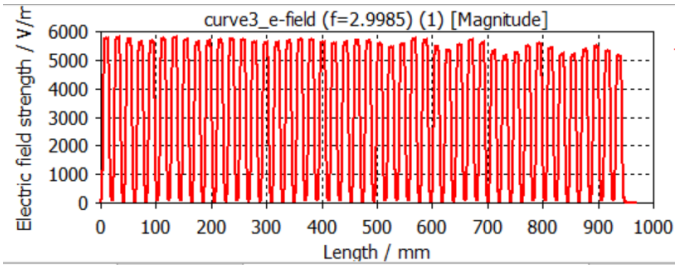
The full structure is shown in Fig. 7.4. The phase advance per cell is shown in Fig. 7.5 (a). The phase advance of individual cells must be corrected via small changes to the cell radii. For large discrepancies in phase advance, individual cells can create small local reflections, which inhibits field flatness. In order to suppress this effect, the phase advance is tuned to be within 1° of the design value (144°) for each cell. Figure 7.5 (b) shows the minimised reflection coefficient along the structure, as required during the design of the output coupler cell. Figure 7.5 (c) shows the on-axis electric field magnitude along the structure. The field is relatively flat, with a small diminish towards the end of the structure. There is a slight



(a) Phase Advance



(b) Reflection coefficient



(c) On-axis electric field

Figure 7.5: Phase advance (a), reflection coefficient (b), and on-axis electric field magnitude (c) along entire structure.

reflection at the output coupler, as the field magnitude oscillates from the output cell to the input, with the reflection decreasing towards to the input port. As the magnitude of the oscillation is small, there was no further optimisation.

Table 7.1: Final electromagnetic Design Criteria

Number of Cells	47
v_g (% of c) [first - last]	[0.82 - 0.18]
Cavity Length (m)	0.96
Input Peak Power (MW) (20% losses)	5.5
Pulse Length (μs)	5
Repetition Rate (Hz)	200
Fill time (μs)	0.75
Gradient (MeV/m)	33.5
P_{out}/P_{in}	0.22
Phase Advance	$4\pi/5$
Z (MΩ/m) [first - last]	[77.5 - 79.5]
E_{peak}/E_{acc}	3.9
$\sqrt{S_c}/E_{acc}$	0.028
B_{peak}/E_{acc} (mT/MV/m) [first - last]	[6.2-5]

7.2 Acceptance of Six Cavity Linac using FC2CT

To fully complete the beam dynamics aspect of this project, all 6 TW cavities were simulated with a realistic input beam.

The lattice parameters for the simulated linac were determined using semi-thin lens approximation. The quadrupole and drift space lengths had predefined values of 0.05 m. The cavity lengths used were approximated during the comparison of different cavity types in Chapter 6, and are ~ 1 m each. Chapter 6 also determined the operating gradient $E_0 \sim 33$ MeV. The value of k_1 was determined such that the maximum beam size is minimised at the cavity entrance/exit analytically using Eqn. 4.107. The calculated value is 10.9 m^{-1}, and thus for an appropriate choice of bore aperture, permanent magnet quadrupoles could be used for the beam line.

Recall, from Chapter 4, the requirement for consecutive $L_{eff,n}$ in Eqn. 4.119 is needed to ensure constant aspect ratio and beam size at each half FODO boundary (in addition to other requirements). As Chapter 6 provided similar lengths for all six structures, it was not possible to satisfy Eqn. 4.119 with increasing cavity length, leaving the alternative of increasing drift length. As increasing the drift lengths reduces the real estate gradient, the FODO-like scheme was not implemented, and the standard FODO scheme was used as the focusing scheme.

As the standard FODO scheme was simulated, all quadrupole strengths and lengths were constant along the lattice. The value of Twiss β_0 was calculated using Eqn. 4.42, which is a function of known lattice parameters. The determined values were $\beta_{x0} = 3.75$ m and $\beta_{y0} = 0.83$ m with $\alpha_{x0} = \alpha_{y0} = 0$. Drift lengths and quadrupoles were simulated as linear

transfer maps, whilst the rf cavities used the TW FC2CT method. Each cavity operated at a gradient of 33 MeV/m, the cell lengths and number of cells were calculated in Chapter 6. The rf phase of each cavity was predefined such that energy gain was maximum. This required running a 1D (to decrease computation time) simulation and fine tuning the rf phase of each cavity such that the ideal particle of the previous cavity had the maximum energy gain in the following cavity. This was required as the optimal rf phase of each cavity was not constant due to relativistic effects. As the protons energy increases, the total rf phase slippage per cavity reduces. Therefore, the optimal input phase shifts by a small amount for each cavity. Figure 7.6 shows the electric field magnitude observed by a particle and the cell entrance and exit along a 40 cell cavity. The optimal rf phase is the value such that the particle returns to the same phase at the cavity exit, maximising energy gain. The particle in Fig. 7.6 is close to the ideal phase, as the start and end phases are near equal.

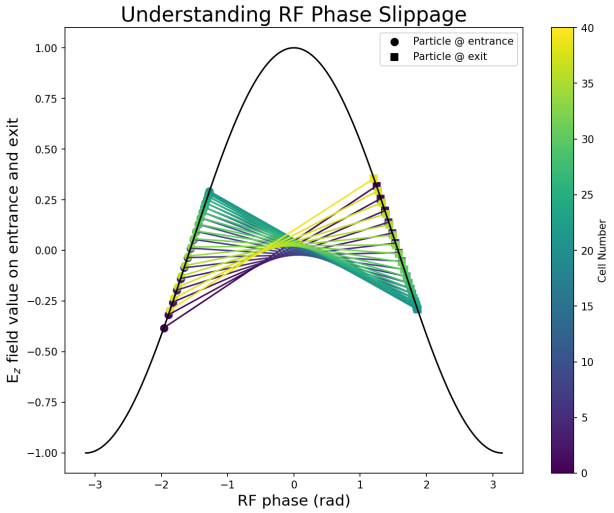


Figure 7.6: Understanding phase slippage.

Figure 7.7 shows the evolution of an initial phase space ellipse with normalised emittance $\varepsilon_N = 0.032$ πm rad. A schematic of the lattice design is also shown. The quadrupole k-strength and length is fixed at 10.9 m^{-1} and 0.05 m, respectively. The drift lengths are 0.05 m. The cavities are TW with phase advance of $\frac{4\pi}{5}$ and lengths $\sim$ 1 m. The longitudinal phase space had an initial rf phase range of 10°, centered on the optimal phase. Each cavity was phased independently as discussed above. The output longitudinal phase space is shown in Fig. 7.8 (a), which shows particles at the final energy of $\sim$ 250 MeV. The simulation assumed an initial longitudinal bunch length of 10°. The majority of particles are located at the head of the bunch. Few particles have started to slip down the neck of the phase space distribution, however still have energy $>$ 249 MeV and would not be lost in the energy selection process.

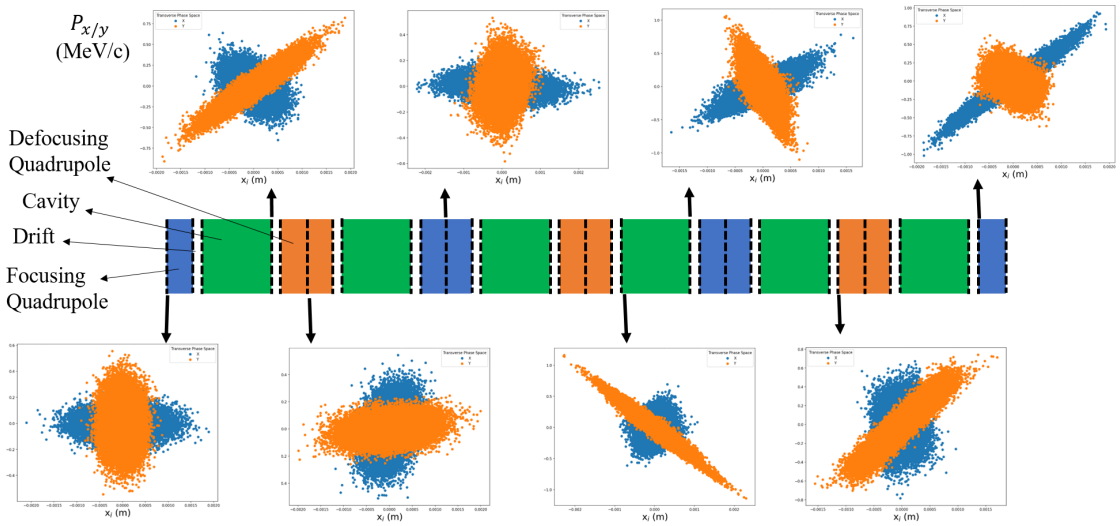
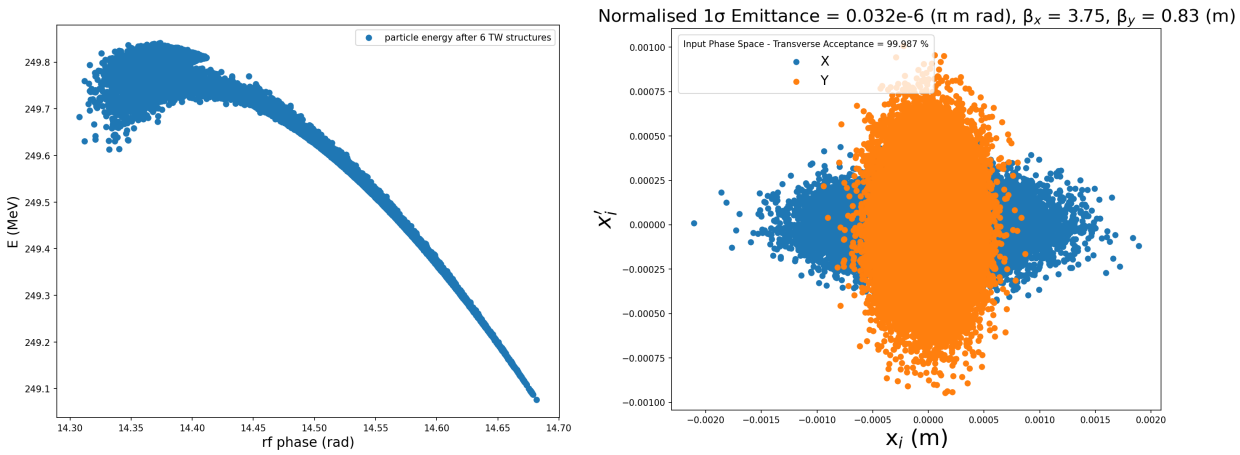


Figure 7.7: Schematic of entire linac lattice design with associated phase space plot at discrete sections. Focusing scheme is the standard FODO with optimised quadrupole k-strength. 30,000 macro-particles.



(a) (b)

Figure 7.8: (a) Particle energy after traversing six TW cavities in standard FODO scheme. (b) Acceptance ellipse of the 6 TW linac solution. Initial beam ellipse taken from analytical FODO method.

Figure 7.8 (b) shows the acceptance ellipse of the 6 cavity lattice. The acceptance is high, with the linac losing less than 0.2% of the initial beam. Particles were lost if the radial displacement exceeded the cavity aperture, 2.5 mm. The simulated ellipse was a 5σ beam ellipse.

7.3 Conclusion to Final Electromagnetic Design

This chapter discussed the design procedure of the two end power coupler cells. The end cell radii and bottleneck width/height were used as the degrees of freedom. An initial approximation for the matching cell was achieved by approximating the required Q factor of the end cells given the rf group velocity. Once an initial matching cell had been designed, a method that exploits Floquets' theorem was used to provide a fine tuning of the matching cell dimensions. The matching cells are tuned to minimise all local reflections and the individual cells are optimised for the correct phase advance per cell. Lastly, a simulation was completed using FC2CT to asses the transmission of a realistic beam for a given initial phase space ellipse. The results of Fig. 7.8 (b) show a 5σ beam can be accepted by the linac with transmission over 99%. The initial phase space ellipse and the quadrupole k- strength were both defined analytically from results discussed in Chapter 4. The focusing scheme used is the standard FODO scheme. The FODO-like scheme could not focus the beam longitudinally, due to the increasing drift lengths in order to maintain constant aspect ratio and beam size at the half-FODO cell boundaries. The cavity lengths in this simulation are all $\sim$ 1 m as determined in Chapter 6, and therefore the FODO-like scheme is not appropriate. Further work could explore the FODO-like focusing scheme for a set of short cavity lengths, where the increase in drift length (in order to satisfy FODO-like conditions) is small, and the real estate gradient remains unchanged.

Chapter 8

Conclusions

This project concerned the electromagnetic design of a linear proton booster from 150 MeV to 250 MeV for proton radiotherapy. The booster section is the final accelerator in an all-linac solution. Protons are initially accelerated to 5 MeV with an RFQ, from 5 MeV to 37.5 MeV a side-coupled drift tube linac (SCDTL) is used. From 37.5 MeV to 250 MeV a coupled cavity linac is used. As the entire accelerator is linear, the beam emittance is kept low, which allowed for small cavity apertures in the CCL. Small cavity apertures produce high shunt impedance, however can limit the length of structures due to transverse focusing requirements.

8.1 Summary

Single Cell Geometry

The geometry of the single cell nose cone, alongside the side-coupled cells and coupling slots was discussed in Chapter 3.

The single cell nose cone optimisation was completed by producing a set of objectives, that change as a function of inputs. The objectives used are as follows,

$$\text{Obj} = [Z, \frac{E_{pk}}{E_{acc}}, \frac{B_{pk}}{E_{acc}}, \frac{\sqrt{S_c}}{E_{acc}}].$$

The inputs are the relative locations of Non-uniform ration basis spline (NURBS) points, which define the geometry of the single cell. In order to maximise the objectives with respect to the inputs variables Objective Genetic Algorithms (MOGAs) were used. MOGAs produce solutions based on survival of the fittest, inspired by nature. Each generation, or iteration, produces a number of solutions, and the best are taken based upon the objective values.

The optimal single cell geometries are the set of Pareto optimal solutions. The Pareto optimal space is defined as the set of solutions that can not be made superior in any one

objective without making another objective worse. In order to select a final single cell geometry, the Pareto optimal solutions must be compared visually. This was completed using MO visualisation techniques, that preserve independent objective values and plot each solution separately. Preserving independant objective values is vital during the selection stage as dimensional reduction methods may plot combined objective functions, losing all information of individual objective values. It was found the Parallel Coordinate Plot (PCP) and Radar Chart (RC) were highly effective visualisation tools. Each solution is plotted as a line that crosses an objective axis at a position relating to the objective value.

As MOGAs can produce a very large set of Pareto optimal solutions, the PCP and RC can become over-crowded, therefore similar solutions are clustered together using a K-means clustering algorithm. Simulations can be clustered together as the definition of Pareto optimal allows to solutions being essentially indistinguishable. The number of clusters is defined by the user, for many clusters, the visualisation plots are still overcrowded. For too few clusters, the individual clusters become too diverse, and some solutions are represented by a cluster that is not appropriate. In this work, 30 - 50 clusters achieved a strong compromise between these conflicting arguments.

A solution can then be accurately chosen by selecting a cluster that best aligns with the type of cavity being designed, as different accelerators have different uses and therefore different objective requirements. The benefit of the clustering allows the entire Pareto space to be observed in one single plot. A candidate was selected that achieved a modest trade-off between all objectives. The process must be repeated for both SW and TW cavity cells, as the cell length changes.

The next step required the design of the side-coupled cells for the SW cavity. The side-coupled cell must be designed such that the resonant frequency is identical to the $\pi/2$ normal mode and confluence is achieved. In order to reduce the transverse size of the cavity, the radius of the side-coupled cells must be limited, which increases the frequency of the $\pi/2$ mode. From Slater's perturbation theory, the frequency of the side-coupled cell can be decreased by removing stored electrical energy from the $\pi/2$ mode. This is achieved by increasing the capacitance of the side-coupled cell, by creating larger 'noses'.

A side-coupled SW cavity experiences end cell phase shift and power flow droop. These are attributed to the flow of power along the structure (due to the excitation of the drive cell), and to perturbation effects from frequency errors in individual oscillators. In order to suppress the phase shift and power droop, the inter-cell coupling constant must be increased. The coupling constant is related to the width of the passband on the dispersion curve. Increasing the slot width between the accelerating and side-coupled cell increases the coupling constant, however reduces the shunt impedance. As the number of cells increases, the required coupling constant increases for a given end cell phase shift and power flow droop. The geometry of the coupling slot between the accelerating and side-coupled cell was altered and the effect of shunt impedance with coupling constant was assessed.

As the cavity aperture was small, the TW cell requires coupling slots in order to produce

a large enough group velocity, v_g. The number and position of coupling slots was determined with respect to the effect on certain objectives. As the surface electric and modified Poynting vector fields peak on the nose cone, they are not influenced by the coupling slot geometry. In order to minimise B_{pk}/E_{acc} and maximising the v_g for a given Z, four slots was chosen. However, six slots also produced a good solution.

Analytical Transverse Beam Dynamics

The transverse beam dynamics (TBD) of the CCL was discussed in Chapter 4. The possible cavity length, as defined by limits in the transverse beam dynamics, is an important design specification. If longer cavities are possible, TW structures become optimal with respect to rf efficiency. However, at shorter cavity lengths, SW cavities are superior. The length at which both structures are equally as efficient is not a defined point, but a fairly broad function of many variables.

Chapter 4 developed an analytical framework that incorporated longitudinal acceleration into the TBD. Two focusing schemes were discussed and compared, the FODO-like scheme (cavities sandwiched between quadrupoles of alternating polarity to produce strong focusing) and Minimum Aperture Scheme (MAS, the scheme defines the case where the acceptance ellipse of the cavity matches the beam ellipse, in order to maximise the cavity length for a given aperture).

The FODO-like scheme incorporated acceleration and was defined such that the aspect ratio and beam size was constant at the boundary between FODO cells. The scheme was analytically solved such that the maximum beam size was minimised at the cavity entrance. It was found that various lattice parameters were functions of the Lorentz factor. In order to satisfy the constraints on the system, consecutive quadrupole parameters must change as follows,

$$l_{q1} = \frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}} l_{q2}, \;\; k_1 = \frac{k_2}{(\frac{\gamma_{r0}\beta_{r0}}{\gamma_{r1}\beta_{r1}})^2}.$$

The value of k_1 is analytically calculated such that the maximum beam size is minimised by solving for the minimum of β_{xc0}, the Twiss β at the cavity entrance, with respect to k_1. The calculation was carried out in the semi-thin lens approximation, defined by taking the second order expansion (with respect to $\sqrt{k_1} l_{q1}$) of the Trigonometric and Hyperbolic functions in the quadrupole transfer map elements. The resultant equation to solve was a cubic in k_1 which was solved. This approach was valid due to the insensitivity in β_{xc0} with respect the quadrupole length.

In order to concatenate multiple FODO-like cells together, there are further constraints on consecutive cavity drift lengths. It was shown that for increasing Lorentz factor, either the cavity or drift lengths must increase. This result shows a limitation for the FODO-like focusing scheme for cases were the cavity lengths are long (~ 1 m), as the drift lengths become long, and the real estate gradient is reduced. In this case, the standard FODO scheme has

a higher real estate gradient, however the optics are not perfectly periodic. For short cavity lengths, however, the FODO-like scheme is a highly achievable scheme, that ensures constant aspect ratio and beam size at FODO cell boundaries.

With the value of the normalise emittance given (defined by the RFQ), 0.032 π m rad, the FODO-like scheme can focus a 5σ beam for cavity lengths up to $\sim$ 1 m, for the cavity aperture of 2.5 mm. This allowed the possibility of considering TW cavities. The MAS was discussed factoring in longitudinal acceleration. It was shown that for a 2.5 mm aperture, the normalised emittance allows cavity lengths of $\sim$ 10 m. For a certain lattice, the real estate gradient of the MAS can become superior to the FODO-like scheme. However, for cavity lengths around 10 m, the rf power and manufacturing capabilities become limiting factors. This chapter therefore produced an upper limit on the cavity lengths possible in different focusing schemes, allowing for the possibility of longer TW structures.

Fast Tracking Code - FC2CT

In order to accurately track the phase space of a proton beam through a cavity, a fast cell to cell tracking code was developed, called FC2CT. This was discussed in Chapter 5. FC2CT uses the on-axis E_z field component to analytically determine the longitudinal momentum change of a particle over a single rf cell, with the approximation of constant particle velocity. The E_z field is described analytically by treating the field as a periodic field and calculating the Fourier series coefficients. The analytical form for the change in energy is intrinsically a function of the Fourier series coefficients, cell length, phase advance per cell, ratio β/β_s, and the input rf phase.

The 1D result showed strong agreement with well trusted tracking code ASTRA for the case of a SW field. In addition, there are fewer calculations to be made using FC2CT over conventional time-step tracking. Thus, a given simulation can be completed more quickly, whilst still producing accurate results. As a result, the method was adapted to complete tracking in 3D. The non-zero field components, B_θ, E_r were expanded as a Taylor series in r, up to, and including terms of the order r^3. This allowed the field components to be expressed as a function of the on-axis E_z field, and its derivatives. The 3D result also showed strong agreement with the ASTRA tracking code, for SW fields.

The E_z field in a TW cavity is different from a SW cavity, as the field is not strictly symmetric about the centre of the cell at all times. The E_z field can be written as the real part of a complex field, and is shown to be mathematically identical to the sum of two standing waves. The standing waves are the real and imaginary component of the TW on-axis E_z field, and can again be described with a Fourier series. As the Fourier series is defined with the same periodicity as E_z (defined by the phase advance per cell), the field being constructed may incorporate many cells (a $4\pi/3$ TW cavity has a periodicity of 5 cells). Therefore, both a_n and b_n Fourier coefficients are required. This resulted in many more calculations per single cell, and the advantage of FC2CT over conventional tracking was reduced.

The well trusted tracking code RF-Track was used to compare the FC2CT method for TW cavities. The results showed good agreement. A approximate TW tracking method was implemented that treated the TW field as if it was SW (called the fast TW method). For this method, the on-axis E_z component was only required in one cell. The field is the following cell was negated to produce a π-mode with periodicity of two cells (as in the case of the SW FC2CT method). The Fourier series only requires b_n coefficients. The imported field is maximally excited, and therefore is symmetric about the centre of the cell at all times. The fast TW FC2CT method showed good agreement with RF-Track, showing good accuracy whilst retaining the key benefit of short computing time.

Power Requirements

The fast nature of FC2CT allows it to be used to calculate the energy gain of different structure types (SW versus TW) as a function of cavity length quickly. In Chapter 6 the energy gain of different cavity structures was calculated as function of input power and cavity length. The core differences between SW and TW structures was incorporated into the simulations as required. SW structures require increased coupling as the number of cell lengths increases for the same end cell phase shift and end cell field magnitude. Input power is distributed evenly amongst the accelerating cells as the shunt impedance is constant. As the SW cavity length increases, Z decreases and the rf efficiency drops.

TW structures were simulated as constant gradient structures. As a result, the group velocity requires tapering along the structure, using the radius of circular coupling slots in the cell wall. The single cell radius and coupling slot radius were swept and the effect on cell figures of merit - group velocity, Z, Q and frequency - were determined. Polynomial functions were fit so that for a given cell and slot radius, the correct cell parameters could be calculated. For TW structures, the input group velocity was swept and the cell parameters are calculated for the given v_g and correct frequency. For each cell, Z and Q are determined, which allows for more accurate computations of the cell power and therefore cell gradient. A large parameter scan of cavity length and input power showed that high power shorter structures perform better in SW, whilst longer cavities are superior when operated as TW. The MAS was omitted, as the real estate gradient could not compete with the FODO focusing scheme, for this project.

After discussions with industrial supervisors, the design parameter space was constrained. Two klystron sources were considered with 3.5 and 5.5 MW peak power. Cavity lengths under 1 m were considered, as this is the limit for the FODO scheme. The average deposited power per unit length was limited to 4.5 kW/m, to greatly mitigate the risk of operational detuning. Normal conducting cavities with average power values exceeding this have been previously designed and tested. The average power was calculated with a repetition rate of 200 Hz and pulse length (flat top) of 5 μs. TW structures with phase advances of $2\pi/3$ and $4\pi/5$ were simulated, in addition to a side-coupled SW structure.

The results showed that both SW and TW structures produced structures of similar gradients, however SW cavities have much larger fill times, and therefore were omitted from the design space. For lower repetition rates, the average power loss is limited, allowing for higher powers within a given cavity length. In this case, short SW cavities become attractive solutions. The two TW structure types performed similarly, with higher v_g structures being slightly optimal as $2\pi/3$ and lower v_g structures optimal as $4\pi/5$ structures. The design choice simplified to a decision between an eight structure linac of 5.5 MW power per structure, and a six structure linac with 5.5 MW power per structure. A short list of four viable solutions was displayed, varying with respect to total power and operational group velocity. Higher power solutions offered increased v_g, thus are less sensitive to machining errors and have faster filling times. Lower power solutions operate with lower group velocity, however have high rf efficiency. All solutions had gradients of $\sim$ 30 - 33 MeV/m, and were vastly under the limits placed by the peak surface fields.

A choice was made to select a solution with low group velocity, to maximise rf efficiency and minimise initial and running costs. In addition, the higher power solution was comprised of eight structures, and therefore required additional non-accelerating elements, dropping the real estate gradient. The input/output group velocity is 0.83/0.18 % of c. Structures with similarly low v_g have been previously designed. Individual cavity lengths are $\sim$ 1 m and the cavity is operated in the $4\pi/5$ mode. The simulation results provided the cell and coupling slot radii for all cells in the first cavity, in addition to the synchronous beta. At gradients of $\sim$ 30 MeV/m, it was shown that changing the cell length at some point along the structure would have negligibly increased the effective gradient. The total power requirement is 32 MW peak power, with an assumed power losses of 20% due to the rf network.

Final Electromagnetic Design

The final step required designing the coupling/matching cells for the TW structure. This process fundamentally requires minimising the reflection at the input/output coupling port. An additional match was approximated by matching the external Q of the coupler to the Q associated with the flow of power through the cell, due to the group velocity. The coupler dimension was subsequently fine tuned using results derived from Floquets theorem, which allows a calculation of the local reflection at the coupler cell. The full structure was simulated, and individual cell radii were altered to produce the correct phase advance per cell. Lastly, FC2CT was used to track a beam through six TW structures in the standard FODO scheme. The Twiss parameters describing the beam, and the lattice parameters, were calculated analytically from results obtained in Chapter 4, such as the optimum quadrupole k-strength. The quadrupole length and drift length were predefined. The rf cavity lengths were taken from the results obtained in Chapter 6. The results showed high transmission ($>$ 99.5 %) was possible.

8.2 Future Work

This project produced a conceptual design for a 3 GHz Linac booster for protons accelerated from 150 to 250 MeV. As the design is not going to be built, a prototype was not prioritised in this work. The next steps, however, would construct and test a prototype design. The testing phase would perform bead-pull tests on the cavity to measure the resonant frequency and the field levels in individual cells. Whilst the electromagnetic design of the first cavity has been completed in this work, the following five structures would require the design of the input/output coupler cells. The cell length changed between structures, to maintain synchronicity between the beam and accelerating field. Optimisations of the individual single cells in subsequent structures would require simulating, to ensure the objectives were optimal.

Further future work can explore the conceptual design of a cavity capable of accelerating protons from 37.5 to 150 MeV. Due to the different relativistic effects for protons lower energies (37.5 to 150 MeV), the conceptual design may be different to the one designed in this work. Once this energy stage has been designed, there will be a conceptual design for the entire all-linac concept.

References

[1] *Our World In Data - Causes of Death*. https://ourworldindata.org/causes-of-death. Accessed: 12-12-2023.

[2] Geoff Delaney et al. "The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines". In: *Cancer: Interdisciplinary International Journal of the American Cancer Society* 104.6 (2005), pp. 1129–1137.

[3] *IAEA - Rays of Hope: Widening Global Access to Cancer Care*. https://www.iaea.org/newscenter/news/rays-of-hope-widening-global-access-to-cancer-care.. Accessed: 12-12-2023.

[4] Xiufang Tian et al. "The evolution of proton beam therapy: Current and future status". In: *Molecular and clinical oncology* 8.1 (2018), pp. 15–21.

[5] Graham Farmelo. "The discovery of X-rays". In: *Scientific American* 273.5 (1995), pp. 86–91.

[6] Allen Pusey. "Roentgen-rays in the treatment of skin diseases and for the removal of hair. 1". In: *Archives of Dermatology* 119.2 (1983), pp. 162–175.

[7] DJ Th Wagener. *The history of oncology*. Bohn Stafleu van Loghum, 2009.

[8] Richard Francis Mould. "The discovery of radium in 1898 by Maria Sklodowska-Curie (1867-1934) and Pierre Curie (1859-1906) with commentary on their life and times." In: *The British Journal of Radiology* 71.852 (1998), pp. 1229–1254.

[9] Kenneth S Davis. "The history of radium". In: *Radiology* 2.5 (1924), pp. 334–342.

[10] Yunfei Jiao, Fangyu Cao, and Hu Liu. "Radiation-induced cell death and its mechanisms". In: *Health Physics* 123.5 (2022), pp. 376–386.

[11] Jiamin Zhu et al. "Molecular mechanisms of lncRNAs in regulating cancer cell radiosensitivity". In: *Bioscience Reports* 39.8 (2019), BSR20190590.

[12] AR Bekker. "Proton stopping power and range calculation using effective atom number and effective electron density from dual energy CT". PhD thesis. Faculty of Science and Engineering, University of Groningen, 2015.

[13] Andrew Brown and Herman Suit. "The centenary of the discovery of the Bragg peak". In: *Radiotherapy and Oncology* 73.3 (2004), pp. 265–268.

[14] Harald Paganetti. *Proton beam therapy*. IOP Publishing Bristol, 2017.

[15] David Jette and Weimin Chen. "Creating a spread-out Bragg peak in proton beams". In: *Physics in Medicine & Biology* 56.11 (2011), N131.

[16] Antje-Christin Knopf and Antony Lomax. "In vivo proton range verification: a review". In: *Physics in Medicine & Biology* 58.15 (2013), R131.

[17] Claire Barker, Matthew Lowe, and Ganesh Radhakrishna. "An introduction to proton beam therapy". In: *British Journal of Hospital Medicine* 80.10 (2019), pp. 574–578.

[18] *Proton Beam Therapy - Gleneagles Hospital*. `https://www.gleneagles.com.sg/tests-treatments/proton-beam-therapy`. Accessed: 12-12-2023.

[19] P. J. Bryant. "A Brief history and review of accelerators". In: *CERN Accelerator School: Course on General Accelerator Physics*. 1992, pp. 1–16.

[20] Robert J Van de Graaff. "A 1,500,000 volt electrostatic generator". In: *Phys. Rev* 38 (1931), p. 1919.

[21] R Wideröe. "On a new principle for the production of higher voltages". In: *Archiv für Elektrotechnik* 21 (1928), pp. 387–391.

[22] DW Kerst. *Development of the Betatron*. 1942.

[23] Ernest O Lawrence and M Stanley Livingston. "The production of high speed light ions without the use of high voltages". In: *Physical Review* 40.1 (1932), p. 19.

[24] Andrew Robert Steere. *A timeline of major particle accelerators*. Michigan State University, 2005.

[25] Vladimir Isaakovich Veksler. "A new method of acceleration of relativistic particles". In: *J. Phys.* 9 (1945), pp. 153–158.

[26] Edwin M McMillan. "The synchrotron - a proposed high energy particle accelerator". In: *Physical Review* 68.5-6 (1945), p. 143.

[27] Karel Strijckmans. "The isochronous cyclotron: principles and recent developments". In: *Computerized medical imaging and graphics* 25.2 (2001), pp. 69–78.

[28] Ernest D Courant, M Stanley Livingston, and Hartland S Snyder. "The strong-focusing synchroton - A new high energy accelerator". In: *Physical Review* 88.5 (1952), p. 1190.

[29] Andrzej Wolski. *Beam dynamics in high energy particle accelerators*. World Scientific, 2014.

[30] William C Barber et al. "The first colliders: AdA, VEP-1 and Princeton–Stanford". In: *IN THE XXI CENTURY* (2015), p. 61.

[31] R Schmidt. "Introduction to machine protection". In: *arXiv preprint arXiv:1608.02433* (2016).

[32] *Diamond Light Source*. `https://www.diamond.ac.uk/Home/About.html`. Accessed: 12-12-2023.

[33] Luis W Alvarez et al. "Berkeley proton linear accelerator". In: *Review of Scientific Instruments* 26.2 (1955), pp. 111–133.

[34] HA Schwettman. "Electron Linear Accelerators". In: *Pure and Applied Physics*. Vol. 40. Elsevier, 1974, pp. 129–147.

[35] Thomas P Wangler. *RF Linear accelerators*. John Wiley & Sons, 2008.

[36] Jens Bruér. *Open midplane designs based on sector coils in superconducting dipole magnets*. 2008.

[37] *RFQ accelerator schematic*. `https://commons.wikimedia.org/wiki/File:RFQ_accelerator_schematic.svg`. Accessed: 12-12-2023.

[38] Hasan Padamsee, Jens Knobloch, and Tomas Hays. *RF superconductivity for accelerators*. John Wiley & Sons, 2008.

[39] Brendan Whelan et al. "MRI Linac systems". In: *MRI for radiotherapy: planning, delivery, and response assessment* (2019), pp. 155–168.

[40] RW Hamm, KR Crandall, and JM Potter. "Preliminary design of a dedicated proton therapy linac". In: *Proc. PAC90* 4 (1991), pp. 2583–2585.

[41] Ugo Amaldi et al. "LIBO-a linac-booster for protontherapy: construction and tests of a prototype". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 521.2-3 (2004), pp. 512–529.

[42] C De Martinis et al. "Acceleration tests of a 3 GHz proton linear accelerator (LIBO) for hadrontherapy". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 681 (2012), pp. 10–15.

[43] U Amaldi et al. "Design of a fast-cycling high-gradient rotating linac for protontherapy". In: *Proc. IPAC13* (2013).

[44] Stefano Benedetti. "High-gradient and high-efficiency linear accelerators for hadron therapy". PhD thesis. Ecole Polytechnique Fédérale de Lausanne, 2018.

[45] C Ronsivalle et al. "First acceleration of a proton beam in a Side Coupled Drift tube Linac". In: *Europhysics Letters* 111.1 (2015), p. 14002.

[46] Maurizio Vretenar et al. *A compact high-frequency RFQ for medical applications*. Tech. rep. 2014.

[47] Silvia Verdú Andrés. "High-gradient accelerating structure studies and their application in hadrontherapy". In: (2013).

[48] AE Geisler et al. "Commissioning of the ACCEL 250 MeV proton cyclotron". In: *Proc. of 18th Int. Conf. on Cycl. and their Appl., Giardini Naxos, Italy*. 2007, pp. 9–14.

[49] H Röcken et al. "progress at Varian's superconducting cyclotrons: a base for the Probeam platform". In: *in Cyclotrons, 2013.* (2013).

[50] *UCLH Treatment Room*. `https://www.ucl.ac.uk/medical-physics-biomedical-engineering/news/2022/feb/uclh-proton-beam-therapy-centre-treats-first-patients`. Accessed: 12-12-2023.

[51] Sam Pitman. "Optimisation studies for a high gradient proton Linac for application in proton imaging: ProBE: Proton Boosting Linac for imaging and therapy". PhD thesis. 2019.

[52] Alberto Degiovanni et al. "Status of the Commissioning of the LIGHT Prototype". In: *9th Int. Particle Accelerator Conf.(IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*. JACOW Publishing, Geneva, Switzerland. 2018, pp. 425–428.

[53] Weisstein, Eric W. *Cylindrical Coordinates*. From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/CylindricalCoordinates.html.

[54] David J Griffiths. *Introduction to Electrodynamics Fourth Edition*. 2021.

[55] Frank Bowman. *Introduction to Bessel functions*. Courier Corporation, 2012.

[56] Dassault Systemes. *CST STUDIO SUITE, R2021x*. `https://www.3ds.com/products-services/simulia/products/cst-studio-suite/`. Computer code. 2021.

[57] Reinhard Klatt et al. "Maxwell's grid equations". In: *Frequenz* 44.1 (1990), pp. 9–16.

[58] SC Joshi, V Paramonov, A Skassyrskaya, et al. "The complete 3-D coupled RF-thermal-structural-RF analysis procedure for a normal conducting accelerating structure for high intensity hadron linac". In: *Proc. of the 2002 Linac Conference*. 2003, p. 218.

[59] Rob Appleby et al. *The science and technology of particle accelerators*. Taylor & Francis, 2020.

[60] Yves Renier. "Implementation and validation of the linear collider final focus prototype: ATF2 at KEK (Japan)". PhD thesis. Paris 11, 2010.

[61] Arakaparampil M Mathai. *Jacobians of matrix transformations and functions of matrix argument*. World Scientific, 1997.

[62] Donsub Rim. "An elementary proof that symplectic matrices have determinant one". In: *arXiv preprint arXiv:1505.04240* (2015).

[63] *Quadrupole Field Profile*. `https://en.wikipedia.org/wiki/Quadrupole`. Accessed: 12-12-2023.

[64] Jim Norem, Zeke Insepov, and I Konkashbaev. "Triggers for RF breakdown". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 537.3 (2005), pp. 510–520.

[65] Valery A Dolgashev et al. "RF breakdown in normal conducting single-cell structures". In: *Proceedings of the 2005 Particle Accelerator Conference*. IEEE. 2005, pp. 595–599.

[66] B Bonin. *Field emission in RF cavities*. Tech. rep. CEA Centre d'Etudes de Saclay, 1996.

[67] Kai Nordlund and Flyura Djurabekova. "Defect model for the dependence of breakdown rate on external electric fields". In: *Physical Review Special Topics-Accelerators and Beams* 15.7 (2012), p. 071002.

[68] Alexej Grudiev, S Calatroni, and W Wuensch. "New local field quantity describing the high gradient limit of accelerating structures". In: *Physical Review Special Topics-Accelerators and Beams* 12.10 (2009), p. 102001.

[69] Kyrre Ness Sjobak, Alexej Grudiev, and Erik Adli. "New Criterion for Shape Optimization of Normal-Conducting Accelerator Cells for High-Gradient Applications". In: *27th International Linear Accelerator Conference*. 2014, MOPP028.

[70] A Degiovanni et al. "High gradient RF test results of S-band and C-band cavities for medical linear accelerators". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 890 (2018), pp. 1–7.

[71] S Benedetti et al. "High gradient performance of an s-band backward traveling wave accelerating structure for medical hadron therapy accelerators". In: *Proc. IPAC18* (2018).

[72] Nicholas Shipman. "Experimental study of DC vacuum breakdown and application to high-gradient accelerating structures for CLIC". PhD thesis. 2015.

[73] Olivier L De Weck. "Multiobjective optimization: History and promise". In: *Invited Keynote Paper, GL2-2, The Third China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems, Kanazawa, Japan*. Vol. 2. 2004, p. 34.

[74] R Timothy Marler and Jasbir S Arora. "The weighted sum method for multi-objective optimization: new insights". In: *Structural and multidisciplinary optimization* 41.6 (2010), pp. 853–862.

[75] Hisashi Tamaki, Hajime Kita, and Shigenobu Kobayashi. "Multi-objective optimization by genetic algorithms: A review". In: *Proceedings of IEEE international conference on evolutionary computation*. IEEE. 1996, pp. 517–522.

[76] Seyedali Mirjalili and Seyedali Mirjalili. "Genetic algorithm". In: *Evolutionary Algorithms and Neural Networks: Theory and Applications* (2019), pp. 43–55.

[77] Julie Jacques et al. "Conception of a dominance-based multi-objective local search in the context of classification rule mining in large and imbalanced data sets". In: *Applied Soft Computing* 34 (Sept. 2015), pp. 705–720. DOI: 10.1016/j.asoc.2015.06.002.

[78] T Luo et al. "RF design of APEX2 two-cell continuous-wave normal conducting photoelectron gun cavity based on multi-objective genetic algorithm". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 940 (2019), pp. 12–18.

[79] R Bartolini, M Apollonio, and IPS Martin. "Multiobjective genetic algorithm optimization of the beam dynamics in linac drivers for free electron lasers". In: *Physical Review Special Topics-Accelerators and Beams* 15.3 (2012), p. 030701.

[80] Sam Smith et al. "Multiobjective optimization and Pareto front visualization techniques applied to normal conducting rf accelerating structures". In: *Physical Review Accelerators and Beams* 25.6 (2022), p. 062002.

[81] CA Coello Coello. "Evolutionary multi-objective optimization: a historical view of the field". In: *IEEE computational intelligence magazine* 1.1 (2006), pp. 28–36.

[82] Kalyanmoy Deb et al. "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II". In: *Parallel Problem Solving from Nature PPSN VI: 6th International Conference Paris, France, September 18–20, 2000 Proceedings 6*. Springer. 2000, pp. 849–858.

[83] Santosh Tiwari et al. "AMGA: an archive-based micro genetic algorithm for multi-objective optimization". In: *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. 2008, pp. 729–736.

[84] SW Shin, JS Chai, et al. "Optimization of the RF cavity of the medical purpose electron linac by using genetic algorithm". In: *Generations* 9200 (2014), p. 9250.

[85] Mamdouh Nasr and Sami Tantawi. "New Geometrical-Optimization Approach using Splines for Enhanced Accelerator Cavities' Performance". In: *9th International Particle Accelerator Conference*. June 2018. DOI: 10.18429/JACoW-IPAC2018-THPMK049.

[86] Sam Smith. "Currently Novel Low Energy Linacs for 3D X-ray Scanning Applications". PhD thesis. School of Engineering - Lancaster University - In Preparation, 2024.

[87] Dassault Systemes. *ISIGHT*. https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/SIMULIA/RESOURCES/simulia-isight-brochure.pdf. Computer code.

[88] S Benedetti, A Grudiev, and A Latina. "High gradient linac for proton therapy". In: *Physical Review Accelerators and Beams* 20.4 (2017), p. 040101.

[89] Abdelrahman Elewah et al. "3D-RadViz: Three Dimensional Radial Visualization for Large-Scale Data Visualization". In: *2021 IEEE Congress on Evolutionary Computation (CEC)*. IEEE. 2021, pp. 1037–1046.

[90] John A Hartigan and Manchek A Wong. "Algorithm AS 136: A k-means clustering algorithm". In: *Journal of the royal statistical society. series c (applied statistics)* 28.1 (1979), pp. 100–108.

[91] James Rosenzweig and Luca Serafini. "Transverse particle motion in radio-frequency linear accelerators". In: *Physical Review E* 49.2 (1994), p. 1599.

[92] RWD Nickalls. "Viete, Descartes and the cubic equation". In: *The Mathematical Gazette* 90.518 (2006), pp. 203–208.

[93] Guido Van Rossum and Fred L Drake Jr. *Python tutorial*. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

[94] Veliko Dimov et al. "Beam commissioning of the 750 MHz proton RFQ for the LIGHT prototype". In: *9th International Particle Accelerator Conference (IPAC2018), Vancouver, BC, Canada, JACoW, Geneva, Switzerland*. 2018, pp. 658–660.

[95] Alessandra M Lombardi. "Overview of Linacs". In: *arXiv preprint arXiv:1804.08521* (2018).

[96] Klaus Flottmann. *ASTRA-A Space Charge Tracking Algorithm*. Tech. rep. DESY, Hamburg, 2000.

[97] Andreas Adelmann et al. "OPAL a versatile tool for charged particle accelerator simulations". In: *arXiv preprint arXiv:1905.06654* (2019).

[98] Andrea Latina et al. "RF-TRACK: Beam tracking in field maps including space-charge effects. Features and benchmarks". In: *Proceedings of the 28th Linear Accelerator Conference-LINAC16*. 2016.

[99] *K100 - ScandiNova*. `https://scandinovasystems.com/pulse-modulator/k-series/k100/`. Accessed: 12/05/2023.

[100] Alberto Degiovanni et al. "Linac booster for high energy proton therapy and imaging". In: *Physical Review Accelerators and Beams* 21.6 (2018), p. 064701.

[101] Xian-Cai Lin et al. "Design, fabrication, and testing of low-group-velocity S-band traveling-wave accelerating structure". In: *Nuclear Science and Techniques* 33.11 (2022), p. 147.

[102] WenCheng Fang et al. "Design and experimental study of a C-band traveling-wave accelerating structure". In: *Chinese Science Bulletin* 56 (2011), pp. 18–23.

[103] Christopher Nantista, Sami Tantawi, and Valery Dolgashev. "Low-field accelerator structure couplers and design techniques". In: *Physical Review Special Topics-Accelerators and Beams* 7.7 (2004), p. 072001.