High Current Proton Fixed-Field
Alternating-Gradient Accelerator Designs

A thesis submitted to The University of Manchester
for the degree of Doctor of Philosophy (PhD)
in the faculty of Engineering and Physical Sciences

Samuel C T Tygier

School of Physics and Astronomy
2012
## Contents

Abstract .............................. 11  
Declaration ................................ 12  
Copyright ................................ 13  
Acknowledgements ......................... 14  

1 Introduction .......................... 16  
1.1 The Energy Crisis ...................... 16  
1.1.1 Climate Change ...................... 17  
1.1.2 Resource Exhaustion .................. 21  
1.1.3 Energy Supply in the UK .......... 23  
1.1.4 Opposition to Nuclear Power ...... 26  
1.2 The Accelerator Driven Subcritical Reactor .......... 29  

2 Accelerator Driven Subcritical Reactors ........ 31  
2.1 The Energy Amplifier ................. 31  
2.2 Fuel Cycle ............................ 36  
2.2.1 Uranium Fuel ....................... 36  
2.2.2 Thorium Fuel ....................... 37  
2.3 Beam Requirements .................... 39  

3 Fixed-Field Alternating-Gradient Accelerators .... 44  
3.1 Principles ............................ 46  
3.2 Scaling and Non-Scaling .............. 49  
3.3 Beam Species .......................... 51  
3.4 Existing FFAGs ......................... 51  
3.4.1 MURA ............................. 51  
3.4.2 Japanese FFAGs ..................... 53  
3.4.3 EMMA ............................. 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Reliability</td>
<td>57</td>
</tr>
<tr>
<td>3.6</td>
<td>Space Charge</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>Current Machines</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>The PSI</td>
<td>68</td>
</tr>
<tr>
<td>4.2</td>
<td>ISIS</td>
<td>70</td>
</tr>
<tr>
<td>4.3</td>
<td>The SNS</td>
<td>71</td>
</tr>
<tr>
<td>4.4</td>
<td>J-PARC</td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>The ESS</td>
<td>75</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>Tracking Codes</td>
<td>77</td>
</tr>
<tr>
<td>5.1</td>
<td>Tracking With Maps</td>
<td>78</td>
</tr>
<tr>
<td>5.2</td>
<td>Zgoubi</td>
<td>81</td>
</tr>
<tr>
<td>5.3</td>
<td>Space Charge Codes</td>
<td>82</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Zgoubi Plus Space Charge</td>
<td>82</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Space Charge Models</td>
<td>83</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Conclusion</td>
<td>98</td>
</tr>
<tr>
<td>6</td>
<td>Lattices</td>
<td>100</td>
</tr>
<tr>
<td>6.1</td>
<td>Scaling FFAGs</td>
<td>100</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Radial Sector</td>
<td>101</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Spiral Sector</td>
<td>102</td>
</tr>
<tr>
<td>6.2</td>
<td>Non-Scaling FFAGs</td>
<td>104</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Linear Non-Scaling FFAGs</td>
<td>104</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Non-linear Non-scaling FFAG</td>
<td>106</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Isochronous FFAGs</td>
<td>107</td>
</tr>
<tr>
<td>6.3</td>
<td>Cascading Accelerators</td>
<td>110</td>
</tr>
<tr>
<td>6.4</td>
<td>RF Limits</td>
<td>112</td>
</tr>
<tr>
<td>6.5</td>
<td>Lattice for Simulations</td>
<td>113</td>
</tr>
<tr>
<td>7</td>
<td>Multi-Bunch Methods</td>
<td>118</td>
</tr>
<tr>
<td>7.1</td>
<td>Multiple Harmonics</td>
<td>119</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Conditions for Acceleration</td>
<td>119</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Two Beams</td>
<td>120</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Three Beams</td>
<td>125</td>
</tr>
<tr>
<td>7.1.4</td>
<td>More Than Three Beams</td>
<td>127</td>
</tr>
<tr>
<td>7.1.5</td>
<td>Practicality</td>
<td>128</td>
</tr>
<tr>
<td>7.2</td>
<td>Multiple Turn Injection</td>
<td>129</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Simulation</td>
<td>130</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Second Turn Acceptance</td>
<td>132</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Evolution of world total primary energy supply by fuel in Mtoe. 17
1.2 Concentration of CO$_2$ in atmosphere over past 1000 years. . 18
1.3 Changes in temperature, sea level and northern hemisphere
snow cover from 1850. ........................................... 19
1.4 Anthropogenic and natural radiative forcing components. . . 20
1.5 Predictions of temperature increase for various scenarios of
stabilisation of concentration of CO$_2$ and other gases. ....... 21
1.6 Brent crude price from 1987 to present day. ................. 22
1.7 Energy flow chart for UK in 2009. .......................... 24
1.8 Estimated levelised electricity cost ranges for low-carbon power
technologies. ....................................................... 28

2.1 Rubbia’s conceptual design for an energy amplifier. ......... 33
2.2 Ingestion radiotoxicity of spent nuclear fuel ................. 35
2.3 Yield of neutrons for given proton energy. ................. 41
2.4 Yield of neutrons for given proton energy, simulated with
GEANT4. .......................................................... 42
2.5 Current required to generate 1.5 GW thermal power from an
ADSR with a range of $k_{eff}$. ................................. 43

3.1 Schematic of a linac, cyclotron and synchrotron. ............. 45
3.2 MURA Mark II ring. .......................................... 52
3.3 KEK POP layout. ............................................. 54
3.4 EMMA ring showing the 42 doublet cells. .................... 56
3.5 Charge $q$ feels the field from the charge enclosed within the
dotted circle with radius $r$ ................................. 62
3.6 Semi-major and semi-minor axes of the elliptical cross section 63

4.1 Energy and current of several existing and planned high power
proton accelerators. ............................................. 67
4.2 PSI main ring layout. ........................................ 68
4.3 PSI injector 2 layout ................................................. 69
4.4 ISIS site layout ..................................................... 71
4.5 SNS site layout ..................................................... 72
4.6 J-PARC site layout ............................................... 73
4.7 J-PARC 3 GeV synchrotron ..................................... 74
4.8 ESS linac design (1 GeV proposal) ............................. 75

5.1 In reality (black) particles are deflected smoothly; in a simulation (blue) a single step can bring two particles very close. 84
5.2 A Gaussian beam is split into concentric elliptical rings in x-y space ................................................. 89
5.3 The force from each ring is summed ............................... 90
5.4 Electric field in a circular KV beam with radius 1 mm ............................ 91
5.5 Electric field in an elliptical KV beam with \(a = 1\) mm and \(b = 2\) mm ................................................................. 92
5.6 Electric field in a circular Gaussian beam with \(\sigma = 1\) mm ................................................................. 93
5.7 Electric field in an elliptical Gaussian beam with \(\sigma_a = 1\) mm and \(\sigma_b = 2\) mm ................................................................. 93
5.8 Evolution of beam envelope in a drift tube for varying initial \(a'\), by integration of the envelope equation ................................................................. 96
5.9 Evolution of beam envelope in a drift tube for varying initial \(a'\), by Zgoubi with linear method ................................................................. 96
5.10 Evolution of beam envelope in a drift tube for varying initial \(a'\), by Zgoubi with rings method ................................................................. 97
5.11 Evolution of beam envelope in a drift tube for varying initial \(a'\), by Zgoubi with N-body method ................................................................. 98

6.1 Layout of magnets in a radial sector FFAG .......................... 101
6.2 Magnetic field in a spiral sector FFAG ............................. 103
6.3 Spiral sector FFAG at MURA ....................................... 103
6.4 EMMA magnet lattice. Dotted black lines show the magnetic centres, which are offset from the solid black reference line. Vacuum pipe is approximate ................................................................. 105
6.5 Required variation in field with \(\beta\) for isochronicity .................. 109
6.6 Required variation in field with radius for isochronicity .................. 109
6.7 Variation in speed and momentum of protons up to 1 GeV .......... 111
6.8 (a) 500 m linac, (b) 10 m diameter FFAG and (c) 3 x 10 m diameter FFAG, compared to the Sizewell nuclear power station. 114
6.9 NS-FFAG lattice with dipole, showing orbits from 35 MeV to 400 MeV ................................................................. 115
6.10 Twiss parameters for NS-FFAG lattice with dipole at 35 MeV 116
6.11 NS-FFAG lattice with dipole, showing revolution frequency from 35 MeV to 400 MeV ........................................ 116
6.12 NS-FFAG lattice with dipole, showing horizontal and vertical tune from 35 MeV to 400 MeV ................................. 117

7.1 Beam revolution frequencies against time .................... 121
7.2 Cavity frequencies against time ................................ 122
7.3 Energy after 1000 turns for range of RF phase offsets ... 131
7.4 Longitudinal phase space over 1000 turns. Blue particles start earlier than the reference phase, whilst red particles start later. Starting positions are marked with an X. .......................... 131
7.5 Energy after 1000 turns for range of launch phases on the second turn .......................................................... 132
7.6 Longitudinal phase space over 1000 turns, for particles launched on the second turn ................................ ............. 133
7.7 Longitudinal acceptance for injecting on successive turns .... 134
7.8 Survival of realistic bunches injected on successive turns ... 135

8.1 Generalised perveance as a function of current for proton beams at 35 and 70 MeV .................................................. 139
8.2 Maximum stable current for 30 m, 35 to 400 MeV NS-FFAG lattice ........................................................................... 141
8.3 Maximum stable current for 30 m, 35 to 400 MeV NS-FFAG lattice (using linear space charge model) ......................... 141
8.4 Tune shift against current for 10 mm mrad beam in 30 m 35 to 400 MeV NS-FFAG lattice ........................................... 143
8.5 Tune shift against current for 5 mm mrad beam in 30 m 35 to 400 MeV NS-FFAG lattice .............................................. 144
8.6 Horizontal cell tune against beam energy for 10 mm mrad beam in 30 m 35 to 400 MeV NS-FFAG lattice for a range of currents ................................................................. 145
8.7 Beam that has experienced distortion due to strong space charge .......................................................... 146
8.8 Beam horizontal emittance in the absence of space charge ... 148
8.9 Beam horizontal emittance with 1 A current and initial emittance of 10 mm mrad ...................................................... 148
8.10 Beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad .................................................... 149
8.11 Beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad. Using multiple initial seeds ......... 150
8.12 Beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad. With 200k macroparticles .......... 151
8.13 Beam horizontal emittance with 3 A current and initial emittance of 10 mm mrad. ........................................ 152
8.14 Beam horizontal emittance with 8 A current and initial emittance of 10 mm mrad. ........................................ 153
8.15 Beam horizontal width with 8 A current and initial emittance of 10 mm mrad. ........................................... 154
8.16 Beam horizontal width with 10 A current and initial emittance of 10 mm mrad. ........................................... 155
8.17 10 A beam after 8 cells. ........................................... 155
8.18 10 A beam after 13 cells. ......................................... 156
8.19 10 A beam after 29 cells. ......................................... 156
8.20 Beam horizontal emittance with 2 A current and initial emittance of 20 mm mrad. .................................... 158
8.21 Generalised perveance as a function of beam energy for proton beams at 1, 2 and 10 A. .......................... 159

A.1 Phase space ellipse, showing Twiss parameters $\alpha$, $\beta$, $\gamma$ and emittance. ............................................ 165

B.1 Semi-major and semi-minor axes of the elliptical cross section. 173
B.2 Electric field along $R$ at $\theta = 0$ around an ellipse with $a = 2$ and $b = 1$. Black line shows numerical integration. Coloured lines show multipole expansion to increasing order. Grey dashed line is the edge of the ellipse. ......................... 177
B.3 Electric field along $R$ at $\theta = \pi/5$ around an ellipse with $a = 2$ and $b = 1$. Black line shows numerical integration. Coloured lines show multipole expansion to increasing order. Grey dashed line is the edge of the ellipse. ......................... 177
B.4 $E_x$ around an ellipse with $a = 2$ and $b = 1$, as calculated with numerical integration, multipole up to octupole, and the difference between them. All scaled to the maximum field. ... 178
B.5 Electric field along $R$ at $\theta = 0$ around an ellipse with $a = 2$ and $b = 1$. It shows numerical integration (black), quadrupole expansion (blue) and adjusted quadrupole expansion (red). Grey dashed line is the edge of the ellipse. ......................... 179
B.6 Electric field along $R$ at $\theta = \pi/5$ around an ellipse with $a = 2$ and $b = 1$. It shows numerical integration (black), quadrupole expansion (blue) and adjusted quadrupole expansion (red). Grey dashed line is the edge of the ellipse. ......................... 180
B.7 $E_x$ around an ellipse with $a = 2$ and $b = 1$, as calculated
with numerical integration, multipole up to quadrupole with
correction, and the difference between them. All scaled to the
maximum field. . . . . . . . . . . . . . . . . . . . . . . . . . . 181

C.1 Zgoubi reads a zgoubi.dat file, and outputs several files. . . . . 183
C.2 PyZgoubi is controlled by an input file. It then calls Zgoubi,
and collects the results for further processing. . . . . . . . . . 185
C.3 Zgoubi bend definition. . . . . . . . . . . . . . . . . . . . . . 187
List of Tables

1.1 Current UK nuclear reactors. ........................................... 26
1.2 Death rates from various energy sources. ......................... 27

2.1 Estimated thorium reserves (tonnes of Th metal). ............... 37

3.1 MURA FFAG parameters. ............................................... 52
3.2 KEK POP parameters. .................................................. 53
3.3 KEK 150 MeV FFAG parameters. ................................... 55
3.4 KART FFAG parameters. ............................................... 55
3.5 EMMA parameters. ..................................................... 57

4.1 Neutron source parameters. .......................................... 67

6.1 Midpoints in kinetic energy in MeV when splitting acceleration by kinetic energy, momentum and speed. ...................... 111

7.1 Harmonic separation for $n$ beams .................................. 128
7.2 Harmonic separation for $n$ beams, 3× sweep ................... 128
Abstract

To make energy production sustainable and reduce carbon dioxide emissions it is necessary to stop using fossil fuels as our primary energy source. The Accelerator Driven Subcritical Reactor (ADSR) could provide safe nuclear power. It uses thorium as fuel, which is more abundant than uranium, and produces less long lived waste. An ADSR uses neutron spallation, caused by a high power proton beam impacting a metal target, to drive and control the reaction.

The beam needs to have an energy of around 1 GeV and a current of 10 mA with a very high reliability, the combination of which is beyond the capabilities of existing particle accelerators. Cyclotrons and synchrotrons both have trouble producing such a beam, while a suitable linac would be several hundred metres long, and expensive. A more compact accelerator design would allow multiple accelerators to be combined to improve reliability.

This thesis examines the use of a Fixed-Field Alternating-Gradient (FFAG) accelerator as the proton driver. FFAGs are compact, and can simultaneously achieve higher energies than a cyclotron at higher repetition rates than a synchrotron. However, it is still a challenge to reach the high currents required. A 35 to 400 MeV non-scaling FFAG was designed to demonstrate issues encountered at high currents.

Two methods were investigated in order to increase the number of particle bunches that could be simultaneously accelerated. One uses multiple solutions to the harmonic conditions for acceleration, and the second injects bunches after the acceleration has started. Neither was found to give significant practical improvement in current.

Space charge is a destructive force at high currents. Software was developed to simulate the effect of space charge in an FFAG using several models. Space charge tune shifts were measured for a range of energies and currents, and peak currents of above 1 A were found to be unstable. In order to provide 10 mA of average current, acceleration would need to occur in around 100 turns, which will require a very rapid RF sweep.
Declaration

The University of Manchester

Candidate Name: Samuel C T Tygier

Faculty: Engineering and Physical Sciences

Thesis Title: High Current Proton Fixed-Field Alternating-Gradient Accelerator Designs

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DoculInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Librarys regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The Universitys policy on presentation of Theses.
Acknowledgements

I’d like to thank my supervisor Roger Barlow. My project came about due to his desire to see accelerator technology used for the benefit of society, and I am grateful for the direction and guidance that he has given me. Thanks also to Hywel Owen who has helped me throughout my PhD and has taken over as my supervisor after Roger’s move to Huddersfield University.

I have received support from colleagues at the Cockcroft Institute and the Rutherford Appleton Laboratory and collaborators within the BASROC, CONFORM and Thorea groups. Special thanks to Chris Prior who advised me on the space charge simulation, François Méot and Scott Berg for help with Zgoubi, and Neil Marks for answering questions and arranging the Cockcroft lectures.

I’d like to thank the Manchester High Energy Physics group for its constant supply of cake and tea. Thanks to Sabah Salih for keeping the computers running.

I’m grateful to my friends inside and outside physics who kept me in good spirits during the past few years, and especially the past few months. Special thanks to Simon Hammersley, Rob Segrott, Adina Toader and everyone in the accelerator office.

Thank you to my family who has always supported and encouraged me in my work. All the trips to the science museum paid off.

And a special thanks to Vashti; I could not have achieved this without your love, support and (especially) proof reading.
To my grandmothers.
Chapter 1

Introduction

1.1 The Energy Crisis

One of the biggest problems facing the world today is energy supply. As of 2008, 81.3% of the world’s primary energy supply comes from fossil fuels [1]. Global energy consumption has been steadily increasing over time, as shown in figure 1.1 and is showing no sign of stabilising or reducing.

With an increasing population and increasing wealth it is likely that energy demand will continue to grow. The United Nations (UN) estimates that the global population will increase from the current level of just below 7 billion to 9.3 billion by 2050 [2]. There is also likely to be an increase of energy consumption per capita as countries become more developed. If energy consumption in the rest of the world increased to that of the member countries of the OECD (Organisation for Economic Co-operation and Development) today then global consumption would rise from 12,369 Mtoe to
There are two main reasons why it is unsustainable to meet our energy needs with fossil fuels. Firstly, burning of fossil fuels releases carbon dioxide (CO$_2$) and other gases into the atmosphere, which is having a damaging effect on the global climate. Secondly, reserves of fossil fuels are finite and so will not last forever.

### 1.1.1 Climate Change

Since the industrial revolution in the 18th century, the burning of fossil fuels has had a large impact on the concentrations of CO$_2$ in the atmosphere, as shown in figure 1.2. Pre-industrial CO$_2$ levels have been within the range of 180-300 ppm for 650,000 years [5]. As of May 2011 it has surpassed 394.97 ppm as measured at Mauna Loa in Hawaii [4]. Emissions of CO$_2$ have
Figure 1.2: Concentration of CO$_2$ in atmosphere over past 1000 years. From multiple data sets listed in [3], blue points are direct measurements from [4]. James Watt patented his steam engine in 1769.

continued to increase in recent years, to a record 30.6 Gt in 2010, up from 29.3 Gt in 2008 [6].

Since the industrial revolution there have also been significant changes to global temperatures and other indicators of climate change, as shown in figure 1.3 from the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report in 2007. This warming trend has continued to 2010, which was (along with 2005) one of the two warmest years on record [7].

The IPCC found that the increase in CO$_2$ and other greenhouse gases are the largest contributor to the current increase in global temperatures, as shown in figure 1.4 [5]. It concluded:
“Most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations.”

The IPCC also discusses changes to extreme weather events:

“Anthropogenic forcing is likely to have contributed to changes in wind patterns, affecting extratropical storm tracks and temperature patterns in both hemispheres. However, the observed
Figure 1.4: Anthropogenic and natural radiative forcing components [5]. Final column shows Level of Scientific Understanding (LOSU).

Changes in the Northern Hemisphere circulation are larger than simulated in response to 20th century forcing change.”

Climate change is predicted to have negative impacts on water and food supply, ecosystems, human health and to cause damage to property from increased extreme weather and flooding due to sea level rise [8].

The amount of warming seen in the future strongly depends on how soon the concentrations of greenhouse gases can be stabilised and then reduced. Figure 1.5 shows a range of scenarios.

It is obvious that emissions for CO₂ and other greenhouse gases must be rapidly reduced in order to limit further warming. “The primary source of the increased atmospheric concentration of carbon dioxide since the pre-
Figures for the amount of remaining fossil fuel reserves are complex. Surveying can find new reserves, so it is misleading to consider only known reserves. New reserves may differ in qualities, and in difficulty of extraction. However as prices rise new methods of extraction such as hydraulic fracturing...
and gas injection become economically viable.

For example for liquid fossil fuels the U.S. Energy Information Administration (EIA) estimates that there are 1,354 billion barrels of proven global reserves, which at 2007 usage rates of 86.1 million barrels per day [10] would last for around 43 years. Changes in demand, discovery of new reserves and use of new extraction methods could affect this.

It is unlikely that the full depletion of fossil fuels will happen soon enough to prevent catastrophic climate change. However, over the past decades there have been significant increases in fossil fuel price. For example, since the mid 1990s the crude oil price has increased from around $20 to over $100 per barrel, as shown in figure 1.6.

![Brent crude price from 1987 to present day](image)

Figure 1.6: Brent crude price from 1987 to present day [11].
1.1.3 Energy Supply in the UK

The solutions to energy production vary greatly by country. Many renewable sources require or benefit from certain features of local geography such as relief, rainfall, average solar irradiance, shallow offshore areas or access to geothermal energy. Also population density and energy use profiles can have an effect. Fossil fuel mix in many countries is influenced by local reserves. Here only the UK will be considered.

The UK consumed 211.1 Mtoe ($8.838 \times 10^{18}$ J/year or an average rate of 280.3 GW) of primary energy (includes conversion and transmission losses) in 2009 [12]. Figure 1.7 shows the flows of energy from production and import of primary sources, to final usage. It can be seen that transport accounts for 37 % of the UK’s final energy usage, and is almost entirely provided by petroleum. 29 % of energy is used in homes, of this 25 % is in the form of electricity and 75 % in the form of fossil fuels.

As we have seen above there must be a rapid reduction in the use of fossil fuels for energy. The massive energy use reduction required to remove the need for fossil fuels would seem completely unfeasible, although greater energy conservation could help somewhat.

The report *Zero Carbon Britain* [14] by the Centre for Alternative Energy (CAT), suggests that it is possible to reduce UK energy consumption by 50 %, however this requires large changes including: increased insulation and improved thermal design of homes; reduction in temperature of heating and areas heated; improved vehicle efficiency; switching to electric vehicles; 20 % reduction in distances travelled; 66 % reduction in aviation; increase
Figure 1.7: Energy flow chart for UK in 2009. [13]
in car sharing and public transport use; 70-80% decrease in meat consumption; biofuel production. Even with this optimistic reduction of energy use, electricity production would be required to double. However, without a way to enforce these lifestyle changes, it is unlikely that such a big cut to energy consumption can be made.

Therefore new low carbon sources of energy must be developed.

The UK’s most promising renewable energy source is wind. Utilising all of the UK’s territorial water up to a depth of 50 m, an area of 120,000 km² could provide 360 GW of electricity [3]. However this would not be technically or economically feasible. On what scale wind could be used is open to much debate. As of 2010 the UK has 5 GW of wind capacity [15], which with a typical load factor of 30% [3] gives an average of 1.5 GW to the national grid.

In 2009 hydroelectric production provided 0.5 GW in the UK [12]. Hydro power can be very useful to have in the energy supply mix, as controlled at short notice to deliver electricity at times of peak demand. However it would be difficult to expand this as it requires the flooding of large amounts of land to create reservoirs.

In 2009 7.2 GW or 18% of electricity [12] was provided by 18 nuclear reactors on 10 sites, as shown in table 1.1. France, which has a similar population to the UK, produces 49 GW or 78% of their electricity from 58 nuclear reactors [16]. There are plans and proposals in the UK to build approximately 19 GW of new nuclear capacity on existing nuclear sites by the early 2020s [17], although some of this is replacement for existing reactors being decommissioned.
### Plant Type

<table>
<thead>
<tr>
<th>Plant</th>
<th>Type</th>
<th>Current capacity (MWe net)</th>
<th>First power</th>
<th>Expected shutdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oldbury 1</td>
<td>Magnox</td>
<td>217</td>
<td>1967</td>
<td>End 2012</td>
</tr>
<tr>
<td>Wylfa 1 &amp; 2</td>
<td>Magnox</td>
<td>2x490</td>
<td>1971</td>
<td>End 2012</td>
</tr>
<tr>
<td>Dungeness B 1 &amp; 2</td>
<td>AGR</td>
<td>2 × 545</td>
<td>1983 &amp; 1985</td>
<td>2018</td>
</tr>
<tr>
<td>Hartlepool 1 &amp; 2</td>
<td>AGR</td>
<td>2 × 595</td>
<td>1983 &amp; 1984</td>
<td>2019</td>
</tr>
<tr>
<td>Heysham I-1 &amp; I-2</td>
<td>AGR</td>
<td>2 × 580</td>
<td>1983 &amp; 1984</td>
<td>2019</td>
</tr>
<tr>
<td>Heysham II-1 &amp; II-2</td>
<td>AGR</td>
<td>2 × 615</td>
<td>1988</td>
<td>2023</td>
</tr>
<tr>
<td>Torness 1 &amp; 2</td>
<td>AGR</td>
<td>2 × 625</td>
<td>1988 &amp; 1989</td>
<td>2023</td>
</tr>
<tr>
<td>Sizewell B</td>
<td>PWR</td>
<td>1188</td>
<td>1995</td>
<td>2035</td>
</tr>
</tbody>
</table>

Table 1.1: Current UK nuclear reactors [17]. Hinkley Point and Hunterston are currently running at reduced capacity, shown in square brackets.

#### 1.1.4 Opposition to Nuclear Power

There is a widespread opposition to nuclear power from environmental organisations such as Greenpeace International [18] and Friends of the Earth [19], governments including Germany [20] and members of the public. The main arguments against nuclear power are:

- Risk of accidents
- Waste disposal
- Nuclear proliferation
- Cost

There is also an association between nuclear power and nuclear weapons for many people, for example the Campaign for Nuclear Disarmament (CND) campaign against both [21].

Safety is a very important concern for nuclear reactors. Nuclear fuel and waste are radioactive and radiation exposure can be harmful and deadly to
human health. There have been a number of very well known accidents at nuclear reactors that have resulted in deaths and releases of radioactive material. However when comparing deaths per TWh across all energy production, nuclear energy is found to be very safe, as shown in table 1.2 with data from a variety of sources collected by Brian Wang [22].

<table>
<thead>
<tr>
<th>Energy Source</th>
<th>Death Rate (per TWh)</th>
<th>Fraction of world supply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal, world average</td>
<td>161</td>
<td>26%</td>
</tr>
<tr>
<td>Coal, China</td>
<td>278</td>
<td></td>
</tr>
<tr>
<td>Coal, USA</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Oil</td>
<td>36</td>
<td>36%</td>
</tr>
<tr>
<td>Natural Gas</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>Biofuel/Biomass</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Peat</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Solar (rooftop)</td>
<td>0.44</td>
<td>less than 0.1%</td>
</tr>
<tr>
<td>Wind</td>
<td>0.15</td>
<td>less than 1%</td>
</tr>
<tr>
<td>Hydro, Europe</td>
<td>0.10</td>
<td>2.2%</td>
</tr>
<tr>
<td>Hydro, world including Banqiao</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>Nuclear</td>
<td>0.04</td>
<td>5.9%</td>
</tr>
</tbody>
</table>

Table 1.2: Death rates from various energy sources [22]. Nuclear figures include the Chernobyl disaster.

Although nuclear power is a relatively safe way to generate electricity and today’s reactors are safer than earlier models, it is still perceived by many as dangerous. Human risk perception can be inaccurate and depends on many factors other than actual risk [23]. Major new reactor designs may help change perceptions; for example, the Accelerator Driven Subcritical Reactor (ADSR) that is the topic of this thesis has a subcritical core that removes the risk of a supercriticality incident (where the reaction runs away due to the core becoming supercritical), or the pebble bed reactor which has a passive feedback system based on thermal expansion.
There is a widespread belief that nuclear power is expensive. It is true that there is a large upfront cost required to build a nuclear power station. However this is also true of any renewable deployment on the same scale as shown in figure 1.8.

Figure 1.8: Estimated levelised electricity cost ranges for low-carbon power technologies [24]. Based on 10 % discount rate. Nuclear includes decommissioning and waste management. Costs due to intermittency are not considered.

Waste is a difficult issue, and the nuclear industry is held to very high standards on waste disposal. A reduction in long-lived waste would be a large improvement to nuclear power. Part of the nuclear proliferation issue is related to waste: early reactors were designed primarily for the production of plutonium for the manufacture of weapons and modern reactors still produce plutonium. A switch to a thorium-based nuclear fuel cycle vastly reduces plutonium production, and can allow for the use of plutonium as fuel.

Nuclear power is already a viable source of low-carbon energy. Although it is currently unpopular among many people and organisations, there are improvements that can be made to address their concerns.
1.2 The Accelerator Driven Subcritical Reactor

ADSRs offer several advantages over conventional nuclear reactors in areas of safety, waste production and proliferation risk. This may ease public fears of nuclear technology and allow its more widespread adoption. ADSRs are discussed in detail in chapter 2.

A high current proton beam is required to drive an ADSR by producing neutrons in a spallation target. The requirements are beyond what can be provided by today’s accelerators. The Fixed-Field Alternating-Gradient (FFAG) accelerator has been proposed as a design that could provide such a beam. It is described in chapter 3.

There are several accelerators around the world that produce high current proton beams for neutron experiments. Some of these approach the requirements of an ADSR. Chapter 4 compares the world’s most powerful existing proton accelerators.

Accelerator design requires computer modelling. Simulations in this thesis were performed using the existing code Zgoubi and also with new code that added in the effects of space charge. The codes used are described in chapter 5.

Chapter 6 discusses various different designs of FFAGs, including a non-linear non-scaling design used for the simulations in this thesis.

The high current requirements for an ADSR are a challenge. Having a high charge in a bunch causes strong space charge effects. Chapter 7
investigates methods that allow more bunches to be accelerated so that the charge can be spread out among them.

Chapter 8 investigates the effects that space charge has on the beam and places limits on the current that could be provided by an FFAG.
Chapter 2

Accelerator Driven Subcritical Reactors

2.1 The Energy Amplifier

ADSRs, also known as Accelerator Driven Systems (ADSs), are a novel design of nuclear fission reactor, in which the reaction is controlled by an external drive beam.

In a traditional fission reactor, energy is released by a nuclear chain reaction. The fission of a heavy nucleus, typically $^{235}\text{U}$, releases an average of 2.4 neutrons. These neutrons can then cause further fissions. However some will be absorbed without causing a fission, and some will escape from the reactor core. The average number of neutrons that actually go on to cause a further fission is called the effective neutron multiplication factor, $k_{\text{eff}}$.

A traditional reactor for power generation runs with the $k_{\text{eff}}$ very close
to unity: that is, on average each fission causes one further fission with a released neutron, and so the reaction continues at a steady rate. A system with $k_{\text{eff}} = 1$ is said to be in a critical state. The $k_{\text{eff}}$ value can be varied by introducing neutron absorbing material into the core, either with boron control rods or adjusting the concentration of boron in the water used for cooling and moderation. Fully withdrawing the rods takes $k_{\text{eff}}$ above 1, called supercritical, and results in a rapid increase of power output.

For values of $k_{\text{eff}}$ less than 1 the reaction is not self sustaining. Any initial reaction or spontaneous reaction will die away. The principle of a subcritical reactor is to drive this reaction by external means, e.g. an accelerator.

As early as 1952 it was suggested by Lewis [25] that spallation neutrons from an accelerator could be used to drive a reactor for energy production. The idea was further developed by Furukawa [26] and Bowman [27].

In his 1993 paper, Carlo Rubbia proposed a design, shown in figure 2.1, which he called the energy amplifier [28], and he further developed this in his 1995 paper [29]. The design consists of a high intensity proton beam incident on a spallation source which produces a flux of neutrons. These neutrons cause fission in the fuel, and secondary neutrons cause further fissions. The heat produced can be used as in a conventional fission reactor to generate electricity. Some of the energy released can be used to power the accelerator that produces the initial beam.

ADSRs have several advantages over conventional critical reactors:

Safety advantage due to subcriticality of fuel. In any reactor that goes above prompt critical (where there are enough primary neutrons to reach $k_{\text{eff}} = 1$), the reaction will runaway very rapidly. To use a reactor for sig-
significant breeding or transmutation it needs to operate with a fast neutron spectrum, although some thorium breeding is possible in thermal spectrum. This is more difficult to control than a thermal reactor as the reaction rate can change very rapidly due to the shorter lifetime of neutrons. A subcritical reactor gives an additional safety margin as a small increase in $k_{\text{eff}}$ will not take it over one.

Shutdown of a critical reactor requires some action to reduce $k_{\text{eff}}$. This is usually the insertion of control rods. To shutdown a subcritical reactor it is sufficient to switch off, divert or block the drive beam.

There is often a gap between perception of safety and actual safety, and this is especially true for nuclear reactors. The safety of modern reactor de-
signs is increased hugely by ensuring negative void coefficients\textsuperscript{1} and negative thermal feedback. Even so, a portion of the public still regards them as unsafe. The ability to stop an ADSR by switching off the beam may go a long way to improving the perceived safety.

Alternative fuels can be used. Conventional reactors are limited to isotopes that can sustain a chain reaction, normally $^{235}\text{U}$, which makes up 0.72\% of natural uranium deposits. With an external neutron source this condition is relaxed. This makes thorium, specifically $^{232}\text{Th}$, a viable fuel. Thorium is also estimated to be 3 to 4 times as abundant as uranium in the Earth’s crust\textsuperscript{30}.

Radioactive waste production is dependent on the fuel cycle used. In a thorium based ADSR the “production of transuranic or transplutonic elements could be two to three orders of magnitude smaller” than in a current PWR\textsuperscript{31}, as they get ‘burned up’ by the process.

It is possible to use an ADSR for the destruction of nuclear waste, transforming long lived radionuclides into shorter lived ones, generating some power in the process. For fission products this is predominantly by transmuting an unstable nucleus into a stable one by neutron capture. For transuranic nuclides it is mostly due to fission following neutron capture\textsuperscript{31}. Transmutation would normally be combined with partitioning, to prevent the stable isotopes being activated. Figure 2.2 from The European Technical Working Group on ADS shows the reduction of radiotoxicity possible using transmutation.

\textsuperscript{1}A reactor has a negative void coefficient if voids (e.g. bubbles) in the coolant cause the reactivity to decrease.

\textsuperscript{30}A reactor has a negative void coefficient if voids (e.g. bubbles) in the coolant cause the reactivity to decrease.
Figure 2.2: Ingestion radiotoxicity of spent nuclear fuel. With a separation efficiency of 99.9% of the long-lived by-products from the waste, followed by transmutation, reference radiotoxicity levels can be reached within 700 years [32].
2.2 Fuel Cycle

2.2.1 Uranium Fuel

A nuclear chain reaction requires a fissile isotope: that is, one that can be caused to fission by capture of a neutron, and in fissioning will release enough neutrons to make it possible to sustain the reaction. $^{235}\text{U}$ is the only naturally occurring isotope that is fissile. It occurs as 0.72 % of natural uranium [33], the remainder of which is 99.27 % $^{238}\text{U}$ and 0.005 % $^{234}\text{U}$.

Nuclear reactors can use natural uranium, for example the Magnox CO$_2$ cooled, graphite moderated reactors in the UK and the RBMK light water cooled, graphite moderated reactors in the Russia. It is now more common to use enriched uranium, where the $^{235}\text{U}$ concentration has been increased.

The $^{238}\text{U}$ present in a reactor can capture a neutron and then undergo $\beta$ decay twice to become $^{239}\text{Pu}$, another fissile isotope. $^{239}\text{Pu}$ will then also contribute to energy production. The earliest nuclear reactors were designed for plutonium production for military use, rather than power generation. Calder Hall at Sellafield, the first reactor to produce significant quantities of electricity, was designed primarily for plutonium production [34].

The International Atomic Energy Agency (IAEA) estimate that there are 5.5 million tonnes of conventional uranium sources, which would be sufficient for at least 100 years at current consumption rates [35]. This could be extended by using breeder reactors. Currently, there are also large reserves of uranium from military stockpiles that supplement production by mining.
<table>
<thead>
<tr>
<th>Country</th>
<th>Reasonably assured reserves</th>
<th>Estimated additional reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>19 000</td>
<td>-</td>
</tr>
<tr>
<td>Brazil</td>
<td>606 000</td>
<td>700 000</td>
</tr>
<tr>
<td>Canada</td>
<td>45 000</td>
<td>128 000</td>
</tr>
<tr>
<td>Greenland</td>
<td>54 000</td>
<td>32 000</td>
</tr>
<tr>
<td>Egypt</td>
<td>15 000</td>
<td>309 000</td>
</tr>
<tr>
<td>India</td>
<td>319 000</td>
<td>-</td>
</tr>
<tr>
<td>Norway</td>
<td>132 000</td>
<td>132 000</td>
</tr>
<tr>
<td>South Africa</td>
<td>18 000</td>
<td>-</td>
</tr>
<tr>
<td>Turkey</td>
<td>380 000</td>
<td>500 000</td>
</tr>
<tr>
<td>United States</td>
<td>137 000</td>
<td>295 000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1 725 000</strong></td>
<td><strong>2 096 000</strong></td>
</tr>
</tbody>
</table>

Table 2.1: Estimated thorium reserves (tonnes of Th metal) [30].

### 2.2.2 Thorium Fuel

An alternative nuclear fuel is thorium. Natural thorium is considered monoisotopic, consisting entirely of $^{232}$Th [33], which is itself not fissile, but is fertile. Through neutron capture and two $\beta$ decays it becomes $^{233}$U which is fissile.

An ADSR could be started with a pure thorium fuel, and breed the $^{233}$U from scratch. In this set-up there would be little initial energy production until a large enough fraction of $^{233}$U was reached. Alternatively, one could start with a fuel that was seeded with a fissile nuclide. For example $^{239}$Pu could be used as the seed.

The abundance of thorium is estimated to be between 6 ppm [36] and 12 ppm [37] in the Earth’s crust. This compares to 1.8 ppm [36] to 2.7 ppm [38] for uranium. Around 1.7 million tonnes of thorium reserves are known, and an additional 2.1 are estimated; table 2.1 shows this in further detail. Thorium is not used heavily by industry, so it is to be expected that there are more undiscovered reserves than in the case of uranium or fossil fuels. As
with uranium, thorium could be extracted from sea water if prices of ground reserves rose enough to make it commercially viable.

An often quoted ‘fact’ is that thorium reactors do not produce plutonium. This is an oversimplification. Coates and Parks have shown that the ratios of nuclides within a reactor tend towards a steady state [39]. This state is reached with around 50 to 100 years of operating time. If a reactor is preloaded with plutonium above the equilibrium level then it is burned up. However starting with pure thorium would lead to plutonium production up to the equilibrium level. The same is true for other actinides. However, the production rates are lower than with a uranium cycle.

The thorium cycle is considered to be more proliferation resistant: that is it does not produce materials useful for the production of nuclear weapons. Although there can be some \(^{239}\text{Pu}\) production, there is less than in a uranium reactor. \(^{233}\text{U}\) is produced and could be used for making a weapon. It has a higher spontaneous fission rate than \(^{235}\text{U}\), but lower than \(^{239}\text{Pu}\), and could be used in a gun-type or implosion-type bomb. However, it would also would be difficult to separate from the \(^{232}\text{U}\) which is also produced, due to the small mass difference (if one had the facilities to do this, then uranium enrichment would be simpler). \(^{232}\text{U}\) has a half-life of 68.9 years and decays to strong \(\gamma\) emitters. This makes it hard to handle, and easy to detect.

Thorium fuel cycles have been used, though mostly in experimental reactors. The experimental Shippingport light-water breeder reactor in the USA ran for 5 years from 1977, starting with a mixed uranium-thorium fuel. The initial uranium is needed to provide the neutrons to convert the \(^{232}\text{Th}\) to \(^{233}\text{U}\). Analysis of the fuel elements after the 5 years showed 1.39 % increase
2.3 Beam Requirements

Beam requirements are defined by output requirements. An electrical output of 600 MW is a reasonable figure for a small power station. For comparison in the UK nuclear power stations currently operating range from Oldbury at 217 MW [41] to Sizewell B at 1188 MW [42]. Assuming a 40 % thermal to electrical conversion efficiency, then 1.5 GW of thermal power is needed.

The rate of fissions, $N_f$, depends on the rate of incoming neutrons, $N_0$, the effective neutron multiplication factor, $k_{\text{eff}}$ and the number of neutrons released per fission, $v$ (around 2.5), [31]

$$N_f = \frac{N_0 k_{\text{eff}}}{v(1 - k_{\text{eff}})} \quad \text{for} \quad k_{\text{eff}} < 1.$$  \hspace{1cm} (2.1)

The $1/v$ factor is needed to account for the fact that $k_{\text{eff}}$ is defined by the number of secondary fissions caused by the neutrons from a single fission, rather than for a single incident neutron.

It can be noted that as $k_{\text{eff}}$ goes to 1 then $N_f$ goes to infinity, i.e. an endless chain reaction is sustained with no need for input. The energy produced per fission, $E_f$, is around 200 MeV. So the thermal output is,

$$P_{th} = \frac{E_f N_0 k_{\text{eff}}}{v(1 - k_{\text{eff})}}.$$  \hspace{1cm} (2.2)
Rearranged for $N_0$ gives,

$$N_0 = \frac{P_{th} v (1 - k_{\text{eff}})}{E_f k_{\text{eff}}}$$

(2.3)

and with the above values,

$$N_0 = 9.4 \times 10^{19} \frac{(1 - k_{\text{eff}})}{k_{\text{eff}}}.$$

(2.4)

The neutrons are produced by the spallation process. When a high energy proton hits a heavy nucleus, typically lead, tantalum or mercury, it will eject a large number of neutrons.

The number of neutrons produced by spallation increases with the energy of the proton. However, we are more concerned with the neutrons per unit of beam energy, as this will relate to the power gain of the system. Simulation and experiment both put the optimum proton energy at around 1 GeV [43], as shown in figure 2.3 for a cylindrical lead target ($\phi = 20 \text{ cm}, L = 60 \text{ cm}$).

There is also a dependence on the geometry of the target. Too thin and protons will pass through, too thick and neutrons will be absorbed. Figure 2.4 shows neutron yield for a range of target diameters and energies around 1 GeV, as simulated in GEANT4 [44].

If we assume that an optimised target can give 30 neutrons per 1 GeV proton then we can find the required beam current as a function of $k_{\text{eff}}$. This is shown in figure 2.5. For a $k_{\text{eff}}$ of about 0.985 a 10 mA beam would be required. Increasing $k_{\text{eff}}$ reduces the beam current required, but it also reduces the safety margin by moving towards criticality.
The spallation target puts some additional constraints on the beam. If the beam ceases then the target will cool, and when the beam resumes the target will heat. This will cause thermal stress and eventually damage the target. Some gaps in the beam are inevitable.

All accelerators that use radio frequency (RF) cavities for acceleration will have a bunch structure at the accelerating frequency. This is because the field in the cavity must alternate its polarity, so only one half of the cycle will accelerate the beam. This bunching is typically in the 100 MHz to 1 GHz range, fast enough that it should appear to be smooth on the time scale of thermal effects.

Some accelerators work in a continuous mode where low energy particles
Figure 2.4: Yield of neutrons for given proton energy, simulated with GEANT4 [44].

are injected and high energy extracted continuously, for example, the cyclotron. Others have an acceleration cycle: a low energy bunch is injected, accelerated for some time and then extracted. A synchrotron works in this way and so is pulsed. For a synchrotron the length of the acceleration period is limited by the time required to ramp the magnetic fields in the dipoles. The fastest synchrotrons have a pulse rate of around 50 Hz. At this time-scale thermal stress is likely to be significant.

A third reason for the beam stopping would be a technical failure. In existing accelerators beam trips are common. In some cases they only last a few seconds, but they can require repairs to the machine, or replacement of parts that can take hours or days. Even the short trips will allow the target to cool and cause large stresses. The long trips will stop the electrical output of the power plant. Unscheduled shutdowns would be expensive to
Figure 2.5: Current required to generate 1.5 GW thermal power from an ADSR with a range of $k_{\text{eff}}$.

the operator as they would be liable to buy the electricity needed to fulfil their contract. This is discussed further in section 3.5.

In summary, the beam must have an energy of around 1 GeV and an average current of 10 mA to produce a useful power output of 600 MW from an ADSR with a $k_{\text{eff}}$ of 0.98. It must also be smooth at the time-scale of thermal effects. The reliability of the system must be high to avoid target damage and unscheduled shutdowns.

The rest of this thesis addresses the problem of how to achieve such a beam.
Chapter 3

Fixed-Field
Alternating-Gradient
Accelerators

Most particle accelerators fall into one of three types: linear accelerators (linacs), cyclotrons and synchrotrons. They are shown schematically in figure 3.1. Their different properties make them useful in different applications.

A linac has many accelerating cavities in a straight line. The particles make a single pass along the machine. The cavities switch polarity at radio frequency (RF) so that the particles always see an accelerating force. Some quadrupole magnets are used for focusing, but the lattice is dominated by RF cavities.

A standard cyclotron has one large dipole magnet, and two large ‘D’ shaped electrodes (dees) that are driven with an RF source. The particles
spiral outwards, moving alternately between the two electrodes. The RF is timed so that when the bunch passes between the dees it always sees an accelerating field. Modern high energy cyclotrons often have several large magnets to produce the field.

A synchrotron is a circular machine, and has small magnets (compared to a cyclotron) arranged in a ring. The particles have a circular orbit due to dipole magnets. The radius is kept constant during acceleration by adjusting the magnet strength in proportion to the particles’ momentum. Alternating quadrupole magnets are used to focus the beam, allowing a small beam pipe. This is known as strong focusing.

An FFAG accelerator has characteristics of both the cyclotron and the synchrotron. It has a ring of magnets, but these are not ramped during
acceleration hence ‘fixed-field’. Instead, the particle orbits increase in radius by a small amount so that they move into an area of increased field. It also uses alternating field gradients to provide strong focusing.

3.1 Principles

The motion of a charged particle is governed by the Lorentz force,

\[ \mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \] (3.1)

where \( q \) is the charge, \( \mathbf{v} \) is the velocity and \( \mathbf{E} \) and \( \mathbf{B} \) are the electric and magnetic fields. The cross product allows the path of the particle to be bent into a circular arc by a uniform magnetic field. The radius of the arc is,

\[ \rho = \frac{p}{Bq} \] (3.2)

where \( p \) is the particle’s momentum transverse to the magnetic field. This is usually rearranged to give the quantity of magnetic rigidity, often just called rigidity,

\[ |\mathbf{B}\rho| = \frac{p}{q} \] (3.3)

In order for a particle to orbit in a circular accelerator it is therefore necessary for either the field or the radius to increase as the momentum increases.

In a classical cyclotron it is only the radius that increases; the field is uniform across the whole system and constant in time. For low energies the speed increases linearly with the momentum and hence the radius, so the
revolution frequency is constant. This simplifies the acceleration system, as a constant RF frequency can be used. As the motion becomes relativistic this relationship breaks down. The highest relativistic $\beta$ achieved in a classical cyclotron is 0.22, in the Oak Ridge 86 inch cyclotron [45], which is around 24 MeV for protons.

There are two methods to extend cyclotrons in energy. One is to sweep the RF frequency to take into account the fact that orbit size is growing faster than the speed, and hence the revolution time is decreasing. This is a synchro-cyclotron, for example the 600 MeV CERN synchro-cyclotron constructed in the late 1950s [46]. Having to change the RF frequency means that the beam must be pulsed. The CERN synchro-cyclotron had a repetition rate of 55 Hz. The RF sweep was achieved by using a variable capacitor with 2 plates that vibrated at 55 Hz like a tuning fork. In modern RF systems the frequency can be adjusted in a more flexible way using digital electronics.

The alternative method is to shape the field so that the orbit time stays constant. If the field is made stronger on the outer edge, then the beam does not need to move to such a large orbit, and so the revolution time will not increase. This is called an isochronous-cyclotron, an example of which is the PSI cyclotron, discussed in detail in section 4.1. A field that increases with radius gives a weak vertical defocusing force, which needs to be countered. This is achieved by splitting the cyclotron’s magnets, and angling the fields to give edge focusing.

A synchrotron takes the opposite approach to balance the increase in rigidity. As the particle is accelerated the magnetic field is increased so that the radius remains constant. Now the beam is confined to a narrow beam
pipe, around which the magnets are placed. The radius of the machine can be made much large than a cyclotron. This allows very high energies to be attained. However, changing the field takes time. Existing rapid cycling synchrotrons have a repetition rates up to about 50 Hz, still too low for an ADSR. It may be possible to increase this, however rapid ramping of magnets causes eddy currents that reduce field quality.

Repetition rate is an important factor in average beam current. It is also important in an ADSR to reduce thermal stress on the spallation target.

The FFAG is a hybrid approach. The bending field increases across the magnet, so that at a larger radius there is a stronger field. As the particle’s rigidity increases, the orbit will move out to a slightly larger radius. Here it will feel a stronger bending force. It does not need to move as far as in a cyclotron, and the smaller orbit excursion allows smaller magnets and vacuum chamber than in a cyclotron. The magnetic fields are constant in time, so they do not limit the acceleration rate.

Another important aspect in an accelerator is focusing. This confines the beam so that it does not diverge. Classical cyclotrons have only a weak focusing due to field fall off towards the edge of the magnet. Synchrotrons use alternating quadrupole gradients to give strong focusing [47]. The alternation is necessary because a quadrupole can only focus in one of the transverse directions at a time.

An FFAG has alternating-gradient strong focusing like a synchrotron. This also keeps the beam small in order to reduce magnet size, which decreases costs.
3.2 Scaling and Non-Scaling

In a synchrotron, the fields in the dipoles and quadrupoles increase with the particle energy so that the normalised fields are constant. This preserves the optics of the system, so that the beam effectively sees the same bending and focusing. The result is that the tune (the number of transverse oscillations around the reference orbit per lap) is constant for all energies, though in reality there are often small changes in tune that need to be corrected. This is important because the tune must be chosen to avoid resonances, which would result in emittance growth or beam loss. In an FFAG the variation of the fields with radius can be chosen to keep the tune constant during acceleration. This is known as scaling. If this were not done the tune would vary and cross through resonances.

The scaling condition can be met by ensuring that fields at different radii have geometric similarity. At a given position around the orbit, the ratio of local curvature, $\rho$, to average curvature, $\rho_0$, must not change with momentum, $p$,

$\frac{\partial}{\partial p} \left( \frac{\rho}{\rho_0} \right) \bigg|_{\theta=\text{const}} = 0 \quad (3.4)$

where $\theta$ is the azimuthal coordinate around the ring. This can be achieved by having the field, $B$, vary, so that

$B(r, \theta) = B_0 \left( \frac{r}{r_0} \right)^k F(\theta) \quad (3.5)$

where $k$ is known as the field index [48]. To achieve these fields, complex non-linear magnets are needed. These fields can be produced by shaping the
magnet face.

Resonances can be devastating to a beam in a synchrotron as the beam will see the same field errors at the same phase many times. They occur when the horizontal and vertical tunes, $\nu_h$ and $\nu_v$, meet the condition,

$$m_h \nu_h + m_v \nu_v = l$$

(3.6)

where $m_h$, $m_v$, and $l$ are integers. $|m_h| + |m_v|$ is the order of the resonance [45]. In general lower order resonances are more destructive.

In 1999, Johnstone proposed the idea of an FFAG that did not follow the scaling law [49]. Muon accelerators require very rapid acceleration as the muon has a lifetime of about 2 µs. An FFAG satisfies this as there is no slow ramping of the magnets. As the tune varies quickly it is possible to cross a resonance without being on it long enough to damage the beam. This allows the scaling conditions to be relaxed and a Non-Scaling FFAG (NS-FFAG) to be designed [50].

The main advantage of NS-FFAGs is that they can be built with linear magnets: only dipole and quadrupole fields are needed. This improves transverse acceptance as there are fewer non-linear field components [50]. They can also be made using smaller magnets, as gradient need not be lower at smaller radii to satisfy equation 3.5.

There are a variety of possible FFAG lattice designs. These are described in more detail in chapter 6.
3.3 Beam Species

In this thesis only protons beams will be considered. Electrons can be used to drive a neutron source [51], but protons give a much better neutron yield and are used in all high power spallation sources as seen in chapter 4.

An H\(^-\) ion, a proton with two bound electrons can also be used. The dynamics of an H\(^-\) are almost identical to a proton, except that it bends in the opposite direction in a magnetic field. This can be taken advantage of during injection or extraction. Consider an H\(^-\) beam circulating in an FFAG lattice, with a foil placed on the outside edge of the beam pipe. The beam is accelerated until its orbit reaches the foil which strips off the extra electrons. The beam will now be bent outwards by the next dipole, and can be extracted. A similar method can be used at injection. This method can be used once in an accelerator chain, for example at ISIS it is used during injection into the synchrotron.

3.4 Existing FFAGs

Only a small number of FFAGs have ever been built: three in the 1950s in the USA, and more recently several in Japan and one in the UK.

3.4.1 MURA

FFAGs were first developed in the 1950s at the Midwestern Universities Research Association (MURA) in the USA [52]. Three electron FFAGs were built at MURA; their parameters are shown in table 3.1.
**Table 3.1: MURA FFAG parameters [53].**

<table>
<thead>
<tr>
<th></th>
<th>Mark II</th>
<th>Mark V</th>
<th>Third FFAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injection Energy</td>
<td>25 keV</td>
<td>35 keV</td>
<td>0.1 MeV</td>
</tr>
<tr>
<td>Maximum Energy</td>
<td>400 keV</td>
<td>180 keV</td>
<td>50 MeV</td>
</tr>
<tr>
<td>Orbit Radius (m)</td>
<td>0.35 - 0.5</td>
<td>0.34 - 0.52</td>
<td>1.2 - 2.0</td>
</tr>
<tr>
<td>Magnet Shape</td>
<td>radial sector</td>
<td>spiral sector</td>
<td>radial sector</td>
</tr>
<tr>
<td>Number of Cells</td>
<td>8</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>Field index</td>
<td>3.36</td>
<td>0.7</td>
<td>9.25</td>
</tr>
<tr>
<td>Particle</td>
<td>electron</td>
<td>electron</td>
<td>electron</td>
</tr>
</tbody>
</table>

Figure 3.2 shows the first FFAG built at MURA, the 400 keV radial sector electron accelerator.

One of the goals of the research was to design an accelerator that could be used for a particle collider, which would give much higher centre of mass energies than the fixed target experiments of the time. Some FFAG designs, such as the third MURA FFAG, can allow the same charge particle beams to orbit in opposite directions simultaneously. However the development
of cascaded synchrotrons by Sands [52] rapidly overtook FFAGs for high energy beams, and the use of synchrotrons as storage rings allowed for high luminosity colliders.

With the limitations of computer power in the 1950s it was difficult for the designers to model the complex magnet requirements and non-linearities. In the following decades no FFAG were built, although some design studies continued, for example an FFAG was proposed by Meads and Wüsterfeld as a driver for a neutron spallation source [54].

### 3.4.2 Japanese FFAGs

The first of the new FFAGs to be built since the 1950s was the Proof of Principle (POP) at the Kō Enerugi Kasokuki Kenkyū Kikō (KEK) in Ibaraki, Japan, built in 2000. This revival was partially triggered by the development of high-gradient high-bandwidth RF cavities using magnetic alloys [55]. It accelerates protons from 50 keV to 500 keV in just 1 ms. The parameters are shown in table 3.2 and the machine layout in figure 3.3. The radius is

<table>
<thead>
<tr>
<th>Type of magnet</th>
<th>Radial sector type (Triplet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of sectors</td>
<td>8</td>
</tr>
<tr>
<td>Field index (k-value)</td>
<td>2.5</td>
</tr>
<tr>
<td>Energy</td>
<td>50 keV → 500 keV</td>
</tr>
<tr>
<td>Repetition rate</td>
<td>1 kHz</td>
</tr>
<tr>
<td>Magnetic field:</td>
<td></td>
</tr>
<tr>
<td>Focus-mag.</td>
<td>0.14 → 0.32 T</td>
</tr>
<tr>
<td>Defocus-mag.</td>
<td>0.04 → 0.13 T</td>
</tr>
<tr>
<td>Radii of closed orbit</td>
<td>0.81 → 1.14 m</td>
</tr>
<tr>
<td>Betatron tune:</td>
<td></td>
</tr>
<tr>
<td>Horizontal</td>
<td>2.17 → 2.22</td>
</tr>
<tr>
<td>Vertical</td>
<td>1.24 → 1.26</td>
</tr>
<tr>
<td>RF frequency</td>
<td>0.61 → 1.38 MHz</td>
</tr>
<tr>
<td>RF voltage</td>
<td>1.3 → 3.0 kV</td>
</tr>
</tbody>
</table>

Table 3.2: KEK POP parameters [56].
small, with an outer orbit of just 1.14 m, but the magnets have a very large horizontal aperture to accommodate the orbit shift.

Following the successful operation of POP a larger 150 MeV proton FFAG was built at KEK in 2003 [57]. Protons are injected from a 12 MeV H⁻ cyclotron situated in the middle of the FFAG ring. The FFAG is notable for its return yoke free magnet design [58], where the focusing and defocusing elements of the triplet act as the yokes for each other. This gives more space for the injection system than was available in the POP machine. It is also the first FFAG to accelerate protons up to an energy where one can start using the beam to drive a neutron spallation source. Its main design parameters are shown in table 3.3.

These prototype FFAGs have led to the development and construction of an ADSR prototype, the Kumatori Accelerator-driven Reactor Test project (KART) at the Kyoto University Research Reactor Institute (KURRI) [59].

Figure 3.3: KEK POP layout [56].
<table>
<thead>
<tr>
<th>Type of Magnet</th>
<th>Triplet Radial (DFD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Sector</td>
<td>12</td>
</tr>
<tr>
<td>Field index (k-value)</td>
<td>7.6</td>
</tr>
<tr>
<td>Beam Energy</td>
<td>12 ( \rightarrow ) 150 MeV</td>
</tr>
<tr>
<td>Average Radius</td>
<td>4.47 ( \rightarrow ) 5.20 m</td>
</tr>
<tr>
<td>Betatron Tune</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Horizontal 3.69 - 3.80</td>
</tr>
<tr>
<td></td>
<td>Vertical 1.14 - 1.30</td>
</tr>
<tr>
<td></td>
<td>Maximum Field Focus 1.63 T</td>
</tr>
<tr>
<td></td>
<td>Defocus 0.78 T</td>
</tr>
<tr>
<td></td>
<td>Repetition 250 Hz</td>
</tr>
</tbody>
</table>

Table 3.3: KEK 150 MeV FFAG parameters [58].

This consists of 3 FFAGs, the third of which is based on the KEK 150 MeV FFAG design. Their main parameters are shown in table 3.4.

<table>
<thead>
<tr>
<th>Focusing Acceleration</th>
<th>Injector</th>
<th>Booster</th>
<th>Main</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>Spiral</td>
<td>Radial</td>
<td>Radial</td>
</tr>
<tr>
<td>E_{inj}</td>
<td>100 keV</td>
<td>2.5 MeV</td>
<td>20 MeV</td>
</tr>
<tr>
<td>E_{ext}</td>
<td>2.5 MeV</td>
<td>20 MeV</td>
<td>150 MeV</td>
</tr>
<tr>
<td>p_{ext} / p_{inj}</td>
<td>5.00</td>
<td>2.84</td>
<td>2.83</td>
</tr>
<tr>
<td>r_{inj}</td>
<td>0.60 m</td>
<td>1.42 m</td>
<td>4.54 m</td>
</tr>
<tr>
<td>r_{ext}</td>
<td>0.99 m</td>
<td>1.71 m</td>
<td>5.12 m</td>
</tr>
</tbody>
</table>

Table 3.4: KART FFAG parameters [59].

The extracted 150 MeV proton beam can then be put into the Kyoto University Critical Assembly (KUCA). This was first achieved in 2009, with a beam of 10 pA pulsed at 30 Hz [60]. A tungsten target was used for the spallation and a neutron rate of \( 1 \times 10^6 \) s\(^{-1} \) achieved. The current is several orders of magnitude below what is needed for power generation.

### 3.4.3 EMMA

The first NS-FFAG, the Electron Model for Many Applications (EMMA), is being commissioned in Daresbury, UK. Its lattice is composed of 42 pairs
of quadrupole magnets, offset from the reference orbit to give a dipole component. The 16.57 m circumference ring is shown in figure 3.4. The main design parameters are shown in table 3.5. The aim of EMMA is to demonstrate the linear non-scaling concept, study resonance crossing, and also to demonstrate a novel acceleration method.

Figure 3.4: EMMA ring showing the 42 doublet cells [61].

EMMA accelerates electrons from 10 to 20 MeV using a novel method called gutter or serpentine acceleration. Rather than the beam being captured in an RF bucket, it moves along a gutter in phase space outside the stable separatrix. This takes the beam across the peak of the RF field multiple times to maximise acceleration. This is possible because EMMA operates around transition energy, the point where the increase in particle speed
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum kinetic energy</td>
<td>10 MeV</td>
</tr>
<tr>
<td>Maximum kinetic energy</td>
<td>20 MeV</td>
</tr>
<tr>
<td>Approximate RF frequency</td>
<td>1.3 GHz</td>
</tr>
<tr>
<td>Lattice cells</td>
<td>42</td>
</tr>
<tr>
<td>RF cavities</td>
<td>19</td>
</tr>
<tr>
<td>Lattice type</td>
<td>Doublet</td>
</tr>
<tr>
<td>Normalized transverse acceptance</td>
<td>3 mm</td>
</tr>
<tr>
<td>Nominal long drift length</td>
<td>210.000 mm</td>
</tr>
<tr>
<td>Nominal short drift length</td>
<td>50.000 mm</td>
</tr>
<tr>
<td>Nominal D magnet length</td>
<td>75.699 mm</td>
</tr>
<tr>
<td>Nominal F magnet length</td>
<td>58.782 mm</td>
</tr>
</tbody>
</table>

Table 3.5: EMMA parameters [62].

is matched by the increase in path length, and so revolution frequency is constant. The revolution time in EMMA only changes by a small amount, parabolically with energy. The change in time of flight is exploited to keep the beam near the crest of the RF, to increase efficiency.

The whole process only takes around 10 turns. There is no need to ramp the RF frequency during acceleration.

EMMA achieved first acceleration on the March 31st 2011 [61]. The beam was accelerated from 12 to 18 MeV/c within 6 turns [63]. The tune rapidly crosses 4 integer resonances during this acceleration. Slow resonance crossing will be tested in the future.

This rapid acceleration method is likely to be important for muon acceleration, but will not be useful for a low energy proton FFAG as the change in velocity, and hence revolution frequency, is too great.

### 3.5 Reliability

For an ADSR to be commercially viable it must have a high availability, the fraction of time that the system is able to output power. As with a traditional
nuclear power station most of the cost of an ADSR will be in the building; the cost of fuel while running is small. While it is outputting electricity it will make money, but while it is off-line it is still costing money to pay back its building loan. This differs from a fossil fuel power station where most of the cost is the fuel, and so while off-line its costs are reduced.

Unscheduled outages are additionally costly. This is because in order to fulfil its contract with the national grid it must buy electricity at a higher ‘system buy price’ until it can arrange a better value contract with another supplier. Steer [64] has modelled costs using historical prices from the UK national grid. For a 600 MW station the mean opportunity cost of an unplanned 24 hour shutdown is £883,000 (2009 money). Accelerators typically have many more short than long outages, and these will cost proportionally more as there is less time to arrange contracts with other suppliers.

It is therefore important to minimise unscheduled shutdowns that prevent power generation. Scheduled shutdowns are less expensive as there would be no contract with the grid to provide power during that time.

Very short beam trips may not stop the system outputting power. The nuclear reaction will take some short time to slow down, and there would be enough residual heat to keep the steam turbines supplied. However they can still cause thermal stress to the spallation target.

The European Technical Working Group on ADS, which is a similar system although designed primarily for waste transmutation, specified of the order of 100 beam trips per year of a duration over 1 second, assuming that the plant is designed for a 40-60 year lifetime [32]. No existing accelerator reaches this level of availability. This is not surprising as it is not normally
such a strong driver in the design. For example; for a research accelerator, it would be cheaper to design a less reliable machine and run it for longer to collect the same amount of data.

It is useful to distinguish mean time between failure (MTBF) and mean time to repair (MTTR) which both contribute to availability. Preventative design and maintenance procedures can increase MTBF. Diagnostics, spare parts on site and a design that allows components to swapped in and out can reduce MTTR.

It is not necessarily true that a component failure results in a system failure. If one RF cavity fails it may be possible to provide a beam at a lower energy or current.

Redundancy can be used to increase availability. This can be at a high level, by having multiple accelerators that feed into the core, or at a lower level with redundant components such as power supplies. Having multiple accelerators allows several options. All the accelerators could run under normal conditions. If one was to fail then either the power station output could be allowed to drop to a fraction of its nominal output, or the remaining accelerators could be run at a higher current to compensate. Another option would be to keep a hot spare that can be switched in at short notice.

Running accelerators produce unsafe levels of radiation, and so accelerator halls must be shielded and sealed during operations. In order for maintenance to be carried out on one accelerator while another is in operation they must be located in separate halls, or the hall must be partitioned with shielding. This rules out a compact solution such as stacking multiple accelerators, and places more emphasis on designing small accelerators.
A study at the PSI cyclotron found that in 1997 11% of planned beam time was lost to long interruption (> 4 hours). Of this about 45% of time was lost to water leaks, 35% to RF problems and 15% to vacuum leaks. Interruptions of less than 4 hours were caused by component failures and some RF tuning problems. They also had typically 600 to 1600 short (< 1 minute) interruptions per week due to sparking on electrostatic injection and extraction systems, sparks in RF cavities and other issues [65]. For experiments at PSI the short trips are not a major problem, but for an ADSR they would be.

SINQ the spallation source project at the PSI reported an improvement in availability from 75% to 98% from the year 1997 to 2003 [66]. This was achieved mostly by reducing water leaks and introducing preventative maintenance. This shows that with effort the reliability of a system can be vastly improved.

### 3.6 Space Charge

Space charge is a collective effect within a charged particle bunch. It becomes important for high intensities, and can be the limiting factor in high current machines.

Consider a proton bunch in an accelerator: it has a large number of positively charged particles confined magnetically into a small space. There will be a repulsive Coulomb force between particles which, in general, will push the bunch apart. It is clear that this effect will be stronger when the charge density is high. The effect is lessened at high energy; this can either be
explained by the magnetic attraction between two currents, or analogously by a relativistic transformation. It can be neglected in the ultra-relativistic limit.

For a uniform cylinder of charge the electric field felt by a particle can be found easily from Gauss’s electric flux law. By symmetry it can be seen that there is no azimuthal component. Figure 3.5 shows the cross section of a bunch with a particle at radius $r$. The radial field depends on the charge enclosed within a circle of radius $r$,

$$E_r = \frac{\rho}{2\epsilon_0} r \quad (3.7)$$

where $\rho$ is the charge density. The corresponding magnetic field is

$$B_\phi = \frac{\rho \; v}{2\epsilon_0 \; c^2} r. \quad (3.8)$$

This is true while $r$ is less that the radius of the bunch.

It can also be expressed in terms of the number of particles per unit length $N$, their charge $q$ and cylinder radius $a$, using

$$\rho = \frac{Nq}{\pi a^2} \quad (3.9)$$

to give

$$E_r = \frac{Nq}{2\pi\epsilon_0 a^2} r \quad (3.10)$$
Figure 3.5: Charge \( q \) feels the field from the charge enclosed within the dotted circle with radius \( r \)

inside the circle \((r \leq a)\), and

\[
E_r = \frac{Nq}{2\pi\epsilon_0 r} \tag{3.11}
\]

outside it \((r \geq a)\). Note that both give the same field for \( r = a \).

The Lorentz force on the particle is,

\[
F = q(E + v \times B) = \frac{q\rho}{2\epsilon_0}(1 - \beta^2)r = \frac{q\rho r}{2\epsilon_0\gamma^2} \tag{3.12}
\]

where \( q \) is the particle charge and \( \beta \) and \( \gamma \) are the standard relativistic parameters.

Note that the force is linear in space and so its effect will not change
the distribution of the beam. This would not be true in general for a non-
uniform initial distribution. The beam must be uniform in physical space;
this is a property of the Kapchinskij-Vladimirskij (K-V) distribution [67]. A
K-V beam is stationary under space charge effects.

![Figure 3.6: Semi-major and semi-minor axes of the elliptical cross section](image)

In the case of a bunch with an elliptic cross section with semi-axes $a$ and $b$, as shown in figure 3.6, equation 3.7 can be generalised to [68],

$$E_x = \frac{\rho ab}{\epsilon_0(a+b)} \frac{x}{a} \quad \text{and} \quad E_y = \frac{\rho ab}{\epsilon_0(a+b)} \frac{y}{b} \quad (3.13)$$

or

$$E_x = \frac{Nq}{\pi \epsilon_0(a+b)} \frac{x}{a} \quad \text{and} \quad E_y = \frac{Nq}{\pi \epsilon_0(a+b)} \frac{y}{b}. \quad (3.14)$$

This is still linear as $E$ depends linearly on $x$ and $y$.

Space charge can be seen as a defocusing force by adding it to Hill’s
equation, to give

$$\frac{d^2 x}{ds^2} + k(s)x - \frac{q}{m_0 c^2} E_x = 0 \quad (3.15)$$

where $m_0$ is the proton rest mass, and $k(s)$ is the focusing along the lattice
from the magnetic elements.

In a synchrotron it is typical to derive a parameter called tune shift,
$\delta Q$, to quantify how the space charge depresses the tune. This is given
approximately as

$$\delta Q = -\frac{r_0 R N}{2Q\beta^2\gamma^3 S f}$$

(3.16)

where $r_0$ is the classical proton radius, $R$ is the machine radius, $N$ the number of protons in the bunch, $Q$ the original tune, $S$ the beam cross-section, and $f$ the fill factor [69]. In a synchrotron the tune must be kept away from resonances. To some extent the quadrupole fields can be adjusted during acceleration to counter the space charge tune shift. However in a non-linear system space charge can cause a tune spread, rather than purely a shift. In practice $|\delta Q|$ must be kept below a value around 0.25, in order to maintain the beam.

In the rapid acceleration of an FFAG the situation is different. Resonances are destructive because an identical error is seen for many turns. If acceleration is fast then the tune change from one turn to another means that the error is not identical.

In practice beams are not perfectly uniform. If the beam had a known distribution, e.g. Gaussian, one could modify equation 3.7 to have a different $r$ dependence, as done by Lee [70]. However the non-linear force would modify the distribution over time, as it would not be stationary.

A completely general method would be to consider the force between each pair of particles. This becomes impractical due to the amount of calculation it would require. Various computational methods for simulating space charge and their implementation are discussed in section 5.3.

The proton driver for an ADSR will need to provide a high current, at a fairly low energy by particle accelerator standards ($\gamma \lesssim 2$). This is a regime
where space charge effects will be strong. The maximum beam current will be limited by this. It is necessary to have a good understanding of how space charge will affect the beam’s dynamics and size, as well as particle loss out of the beam. This will be covered later in the thesis.
Chapter 4

Current Machines

There are several high power proton accelerators used around the world today for the production of neutrons by spallation. Their primary purpose is neutron scattering, for analysing materials in fields from biology to aerospace. They have much in common with the accelerator that would be needed for an ADSR, although none currently reach the 10 MW needed for a commercial power station. They use a range of proton accelerators: cyclotrons, synchrotrons and linacs.

Beam power is a function of energy and current. Figure 4.1 shows the energy and current of some of the world’s leading existing and planned high power accelerators. Some of these will be examined in more detail in the chapter; their parameters are shown in table 4.1.
Figure 4.1: Energy and current of several existing and planned high power proton accelerators [71, 72, 73, 74, 75, 76, 77, 78].

Table 4.1: Neutron source parameters [71, 72, 73, 74, 75].
4.1 The PSI

The Paul Scherrer Institute (PSI) in Switzerland hosts what is currently the most powerful proton cyclotron in the world, the 590 MeV Ring Cyclotron [71]. Construction was completed in 1974. Figure 4.2 shows the layout of the accelerator, with the 8 sector magnets, each weighing 245 tonnes, and the 4 main cavities. The radius of orbit increases from 2.1 m at injection to 4.5 m at extraction [79].

Figure 4.2: PSI main ring layout [80].
The beam is pre-accelerated to 870 keV by a Cockcroft-Walton accelerator, and then to 72 MeV in the Injector 2 cyclotron [81]. Injector 2 has 4 sector magnets, and is shown in figure 4.3.

![Figure 4.3: PSI injector 2 layout [80].](image)

The design current was 100 µA but it has since been increased to 2.2 mA, which gives a beam power of 1.3 MW. There are plans to increase the power further to 1.8 MW [82]. The increase has been achieved by increasing the RF cavity voltage so that fewer turns are required. This increases the separation of the penultimate and final laps, leaving more space for the extraction element.

The beam is used to drive the SINQ neutron source, as well some nuclear and pion experiments. It produces a CW (up to the 51 MHz bunch structure) neutron flux of up to $10^{14}$ n/cm$^2$/s [83] using solid zirconium and lead targets. The Megawatt Pilot Experiment (MEGAPIE) has successfully operated a
lead-bismuth eutectic target at a power of 0.77 MW for a period of over 4 months [84].

4.2 ISIS

ISIS is a synchrotron-based pulsed neutron spallation source at the Rutherford Appleton Laboratory (RAL) in Oxfordshire, UK. It delivered its first beam to target in 1984. This accelerator layout is shown in figure 4.4.

The H\textsuperscript{−} beam is initially accelerated to 665 keV by a radio frequency quadrupole (RFQ), and then to 70 MeV by a 55 m drift tube linac [86]. The beam is then stripped as it is injected into the 26 m radius synchrotron. The protons are accumulated for 130 turns, then are accelerated to 800 MeV over 12,250 turns and extracted [87]. The average current is up to 200 $\mu$A which gives a power of 160 kW. The bunches reach the target as 100 ns pulses, with a repetition rate of 50 Hz, rather than the continuous beam at PSI [72]. Neutrons are produced by spallation in a tungsten and tantalum target. A second low power target was added in 2007.

Pulsed neutrons are preferred for some imaging experiments. For example in dynamic processes the short pulse can ‘freeze’ the motion. Also with a short pulse it is possible to measure time of flight information.

ISIS typically delivers beam to users for 180 days per year; the rest is shutdown, maintenance and recovery time, with some time reserved for machine physics. The average availability during run time is 88 % [88].
Figure 4.4: ISIS site layout [85].

4.3 The SNS

The Spallation Neutron Source (SNS) at the Oak Ridge National Laboratory in the USA was completed and had first beam in 2006. It is currently the world’s most powerful pulsed neutron source, having surpassed ISIS [73].
The beam is accelerated in a linac, then accumulated in a storage ring before being sent to the target. The layout of the site can be seen in figure 4.5.

Figure 4.5: SNS site layout [89].

The beam is accelerated from 65 keV to 2.5 MeV in a 3.75 m RFQ. It is then accelerated to 1 GeV by the main linac, which is composed of a drift tube linac, a coupled-cell linac and a superconducting linac, with a total length of 335 m. This beam is accumulated in the 248 m circumference storage ring for 1 ms, 1060 turns. The beam is then extracted in a single turn to deliver a 695 ns pulse to the target.

The pulse of $1.5 \times 10^{14}$ protons is delivered to the target 60 times per second, giving an average current of 1.44 mA or a power of 1.44 MW. The target is made of liquid mercury.
4.4 J-PARC

The J-PARC Laboratory in Tōkai, Japan has several accelerators for a range of experiments, the layout of which is shown in figure 4.6. The 3 GeV synchrotron is designed to provide a beam for a pulsed spallation source which is currently being commissioned, as well as being the booster for the larger 50 GeV synchrotron [90].

The relevant parts for the spallation source are the 400 MeV linac, composed of an RFQ, DTL, separated-type DTL and an annular-coupled structure linac with a total length of 249 m. This injects into the 3 GeV rapid cycling synchrotron, which has a circumference of 348.3 m. The ring layout is shown in figure 4.7. The ring is filled over 308 turns with pulses from the linac. It is then accelerated to 3 GeV in 20 ms, and extracted onto the mercury target with a repetition rate of 25 Hz [74].

As of 2010 the spallation source has been commissioned up to 300 kW [91]. This will be increased to 1 MW.

Figure 4.6: J-PARC site layout [74].
Figure 4.7: J-PARC 3 GeV synchrotron [74].
4.5 The ESS

The European Spallation Source (ESS) will be built in Lund, Sweden and is scheduled to switch on in 2019 [92]. The design is not finalised: the 2008 5 MW Long Pulse proposal has a final proton energy of 1 GeV [93], but the latest parameter list states that this has been updated to 2.5 GeV [75].

The 1 GeV proposal is for a 5 MW source with 2 ms pulses at 16.67 Hz. Earlier proposals included a second target for short pulses, with an accumulator ring similar to the SNS. The main linac, shown in figure 4.8, has several sections with a total length of 633 m: RFQ, drift tube linac, coupled-cell linac, superconducting linac, and then an additional coupled-cell linac for bunch rotation. The target is likely to be mercury, although liquid lead-bismuth and lead-gold eutectics have been considered [93].

The full design for the 2.5 GeV option is not yet published.

Figure 4.8: ESS linac design (1 GeV proposal) [93].
4.6 Summary

High current proton accelerators are well established. The latest machines under construction approach the current and energy requirements for an ADSR. However no current machine has the extreme reliability that is required. This is expect as reliability and availability requirements for physics experiments is much lower than those for an ADSR.

If money was no issue and reliability could be increased, then it would be possible with today’s technology to combine several linacs of a design similar to the SNS or ESS to provide a beam for an ADSR. It would be necessary to make some modifications to give more frequent, smaller pulses.

However, in order for an ADSR to be a viable energy source it is likely the accelerator will need to be smaller and cheaper than a linac, hence the investigation into FFAGs.
Chapter 5

Tracking Codes

Since the 1950s computers have been an essential tool for accelerator design [52]. Without them only the simplest of accelerator designs and effects can be understood.

Accelerator simulation codes are used for a variety of tasks. Given a lattice design they can be used to find properties such as closed orbits, beam parameters, aperture limits, stability and so on. By using a feedback system they can be used to adjust a design to optimise these properties.

Codes can be divided into some broad categories. Tracking codes track particles or macro particles through a system of electric and magnetic fields. Envelope codes simulate the evolution of the envelope of the beam through a system. Also EM solvers can be used to find the fields given by particular geometries, however they have not been used in this work. Some codes combine multiple types of modelling.

At a simple level a particle tracking code will take the initial coordinates
of a particle, and calculate its trajectory through the magnetic and electric fields in the model. The global Cartesian coordinates are usually not very useful, so instead most codes transform into local coordinates relative to the magnets or some reference line. Similarly, it is simpler to describe the magnets as separate localised fields around a reference line, than as a global magnetic field.

All simulation codes make some approximations to simplify the problem and to speed up computation; for example they may completely neglect a small effect, or work with a truncated series expansion of a force rather than the full equation. The model may be idealised; for example it may have infinitely extended perfect magnetic fields and a perfect vacuum in the beam pipe, especially in early stages of design. Some approximations may be valid for some particle accelerators, but not others. For example, in a large synchrotron particles stay close to the reference orbit, and have only a small angular deviation. This allows paraxial approximations. However, in a smaller accelerator, angles might be bigger and so the approximation would not hold. In high energy accelerators, an ultra-relativistic approximation \((v = c)\) can be used, but not at low energies. The result is that there are many different tracking codes, some for very specific tasks, and some more general.

5.1 Tracking With Maps

Transport maps are widely used for particle tracking. Each particle (or macro-particle) is a vector containing a pair of coordinates in each dimen-
sion. In the transverse plane, a position $x$ or $y$ and either a momentum $p_x$ or an angle $x'$, both relative to a reference line often going though the centre of the magnets. In the longitudinal dimension the position relative to the bunch centre $\Delta s$ and the momentum relative to a reference momentum $\Delta p$:

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix} = \begin{pmatrix} x \\ p_x \\ y \\ p_y \\ \Delta s \\ \Delta p \end{pmatrix} \quad (5.1)$$

A map is any function, $F$, that transforms from before the element $Z(0)$ to the coordinates after $Z(1)$,

$$Z(1) = F(Z(0)). \quad (5.2)$$

The simulation code will construct a map for each magnetic element and the drift spaces between them. A particle is tracked around a lattice by applying the maps sequentially. In some cases the maps can be concatenated into a one turn map.

For a linear system it is convenient to express the map as matrix multiplication [45]. The matrix is known as the transport matrix, and often denoted as $R$. The transformation is,

$$Z(1) = RZ(0) \quad (5.3)$$
or,

\[ \mathbf{Z}(1)_i = \sum_{j=1}^{6} \mathbf{R}_{ij} \mathbf{Z}(0)_j. \]  

(5.4)

The matrices for each element can be multiplied to find the matrix for a group of elements or one turn map. The matrix can also be used to find some useful parameters such as tune and Twiss parameters, as shown in section A.4.

For non-linear systems high order terms are needed. Second order terms can be added by using a 3rd-rank tensor, usually called \( \mathbf{T} \),

\[ \mathbf{Z}(1)_i = \sum_{j=1}^{6} \mathbf{R}_{ij} \mathbf{Z}(0)_j + \sum_{j=1}^{6} \sum_{k=1}^{6} \mathbf{T}_{ijk} \mathbf{Z}(0)_j \mathbf{Z}(0)_k. \]  

(5.5)

This is used in the SLAC code TRANSPORT [94] and is the basis for MAD-X [95].

Using a first or second order matrix as a mapping is an approximation. However, by ensuring the mapping is symplectic (that area in phase space remains constant) some important properties are guaranteed to be preserved. A physical system with only conservative forces is symplectic by Liouville’s theorem. This is valid for an accelerator when neglecting effects such as synchrotron radiation and gas scattering. If the transfer map is not symplectic then there will be an artificial damping or exciting of the particles [96].

Some codes, such as MAD-X, can use a Lie algebra method to make its low order maps symplectic even though they are derived from truncated expressions for the Hamiltonian [97].

When tracking for a large number of turns symplecticity can be more important than absolute accuracy of the model. For FFAGs symplecticity is
of less importance as particles are held for a smaller number of turns. However, it is important to check that quantities like emittance remain constant over the time-scale of the simulation. Instead it is more important to use a method that is accurate for particles that enter magnets at large angles, and travel at a large distance from the magnet centres.

Maps can also be constructed to transform the beam envelope around an accelerator.

5.2 Zgoubi

Most of the common accelerator codes are designed for simulating linacs or synchrotrons. FFAGs break some of the assumptions made by these codes: particles do not necessarily pass through the magnet centres, due to the orbit excursion; particles do not necessarily have small angles to the magnet axes; magnet strengths are not scaled with beam momentum. FFAGs can also use complex magnetic elements, with curved faces or non-linear field profiles.

Zgoubi [98] is a charged particle tracking code that was originally developed in the 1970s for the design of spectrometers. Spectrometers have large magnets, and the trajectories of particles of a range of energies must be known. FFAG magnets have similar features, so Zgoubi is widely used to study them.

Rather than constructing a transfer map for each element, Zgoubi tracks each particle and evaluates the fields and forces at that point. This is a numerical integration method which is referred to as ray-tracing in the manual. The magnetic field and its derivatives up to 6th order are calculated
analytically or from interpolation of a field map. These are used in a 6\textsuperscript{th} order expansion of the Lorentz force for the integration step. The method is described in detail in the Zgoubi manual [99].

As this is not an expansion around the magnet centre or a reference energy as in a more traditional code, it is more accurate for FFAGs.

I have developed PyZgoubi [100], a python interface to Zgoubi, that simplifies its use, and works around some of its limitations. It is described in appendix C.

5.3 Space Charge Codes

5.3.1 Zgoubi Plus Space Charge

In order to benefit from Zgoubi’s suitability for FFAG simulation, it was used as the basis of the space charge simulation. However, adding space charge directly into Zgoubi was avoided. Zgoubi takes particles one at a time and tracks them step by step through a whole magnet. Space charge effects could not be calculated at each step, as the positions of the rest of the particles are not known. The beam is only together at the start and end of an element, so the calculation must be done there. To increase accuracy it was necessary to split elements into slices, so that space charge could be applied multiple times per element.

Instead of putting the space charge model inside Zgoubi, it was kept outside in PyZgoubi. PyZgoubi was used to split up magnetic elements into slices, the beam was then passed to Zgoubi, to track through one slice, and
retrieved again by PyZgoubi. Then the effect of space charge was calculated, and applied to the particles as a kick, before tracking them through the next slice.

Communication of particle coordinates between PyZgoubi and Zgoubi was done using data binary files, stored in /dev/shm, a shared memory area of the file system, in order to avoid the overheads of ASCII conversion and writing to disk.

### 5.3.2 Space Charge Models

Computer simulations of physical effects often require a compromise between accuracy and computation time. This is especially true for collective effects for large numbers of particles. A bunch of $10^{12}$ particles, with 6 double precision coordinates would require 48 TB of storage. There would be $10^{24}$ particle pairs whose interactions would need to be calculated. This is impractical on current computers.

It is common practice to create macro-particles each of which represent many physical particles. For example $10^{12}$ protons could be represented by $10^6$ macro-particles, each with $10^6$ times the mass and charge of a proton. It must be recognised that this gives an artificial lumpiness to the simulation, and can produce non-physical effects.

The simplest method for the simulation of space charge is to assume a uniform particle beam, and use the linear equations (equation 3.13). This is quick to calculate, as one only needs to find the width of the beam, and then each particle will feel a force depending on its position.
A simulation of the force between each pair of particles or macro-particles is called an N-body simulation. If the number of macro-particles is limited this can be calculated in a practical amount of time. It can also be used as a check against approximate models. It is, however, susceptible to noise, firstly due to the limited precision of storing and calculating on a computer. It requires the summing of many numbers with a range of magnitudes, which is a common way to lose precision. Secondly, the Coulomb force falls off as $1/r^2$, and so gets very large when 2 macro particles are very close. The stepwise nature of simulation allows particles to approach very close, where in reality they would already have felt a deflective force, as shown in figure 5.1.

![Figure 5.1: In reality (black) particles are deflected smoothly; in a simulation (blue) a single step can bring two particles very close.](image)

The effect in a simulation is that particles will occasionally receive an unphysically large kick. One solution is to modify the distance dependence of the force. Rather than using $1/r^2$, a potential that reaches a maximum value can be used. This can be justified by considering that a macro-particle
represents a cloud of particles rather than a point charge.

Ideally a compromise between the over-simplified linear model, and the full N-body simulation is needed. There are various ways the beam can be decomposed into some representation that can quickly be computed. One method described by Machida [101], is to break it into rings of charge. The number of particles in each ring can be counted. Then for each particle, one can sum the forces produced by each ring. This method assumes rotational symmetry of the beam, but can be generalised to an ellipse.

**Particle distributions**

Space charge is sensitive to the distribution of charge density within the beam. Each particle in the beam has 6 coordinates, 3 of position and 3 of momentum. One must therefore be careful with defining terms such as uniform. For transverse space charge, only the 2 transverse dimensions are considered, giving 4 coordinates $x, x', y$ and $y'$. This is effectively assuming that the beam is long and has no energy spread.

The linear space charge model assumes that the beam has a uniform charge density, i.e. when projected into $x$ and $y$ it forms a uniformly filled ellipse. It will not however be uniform when projected into 1 dimension such as $x$. The $x'$ and $y'$ must be chosen so that this property is preserved as the beam is transported though a focusing lattice. The Kapchinskij-Vladimirskij (KV) [67] distribution achieves this by uniformly distributing particles on the surface of a hypersphere in the 4 dimensions.

To uniformly populate a 4D hypersphere one can take 4 random variables, $r_1, r_2, r_3$ and $r_4$, in the range -1 to 1, discarding sets where $r_1^2 + r_2^2 \geq 1$ or
\( r_3^2 + r_4^2 \geq 1 \). Then the coordinates are [102]

\[
\begin{align*}
  x &= r_1 \\
  x' &= r_2 \\
  y &= r_3 \sqrt{\frac{1 - r_1^2 - r_2^2}{r_3^2 + r_4^2}} \\
  y' &= r_4 \sqrt{\frac{1 - r_1^2 - r_2^2}{r_3^2 + r_4^2}}. \\
\end{align*}
\]

(5.6)

When using the linear space charge algorithm only the width of the beam is needed. So it is unnecessary for the beam to be uniformly filled when projected into 2 dimensions. The simplest distribution that retains its width is the outline of a circle in \( x-x' \) and \( y-y' \). This is generated with 2 uniform random variables, \( r_1 \) and \( r_2 \) in the range 0 to \( 2\pi \). The coordinates are given by:

\[
\begin{align*}
  x &= \sin(r_1) \\
  x' &= \cos(r_1) \\
  y &= \sin(r_2) \\
  y' &= \cos(r_2). \\
\end{align*}
\]

(5.7)

Note that this gives a non-elliptical beam when viewed in the \( x-y \) plane.

A common distribution used in accelerator simulation is a Gaussian distribution. For this each coordinate is give a value directly from a Gaussian random variable.

Another distribution that gives a good representation of a proton beam is a waterbag distribution. This is defined similarly to the KV distribution,
except that a filled hypersphere is used instead of just populating the surface. A simple method to generate this is to uniformly fill a unit hypercube, and reject the coordinates that fall outside of the unit hypersphere.

The methods above give a spherically symmetric beam, with a width of 1. Once this is generated it can be transformed to the required size by multiplying the $x$ and $x'$ coordinates by the square root of the emittance in $x$, and likewise for the $y$ direction.

It is also usually necessary to transform the beam so that it has the required Twiss parameters. This can be done by multiplication by a matrix $T$ [103],

$$T = T_b T_a$$

(5.8)

where

$$T_b = \begin{pmatrix}
\sqrt{\beta_x} & 0 & 0 & 0 & 0 & 0 \\
0 & 1/\sqrt{\beta_x} & 0 & 0 & 0 & 0 \\
0 & 0 & \sqrt{\beta_y} & 0 & 0 & 0 \\
0 & 0 & 0 & 1/\sqrt{\beta_y} & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

(5.9)

and

$$T_b = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
-\alpha_x & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -\alpha_y & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

(5.10)
Space charge algorithms

Linear space charge is simple to implement, but only useful for KV beams. All that needs to be known about the beam is the semi-major axes of the cross section. This can be used to find the electric field \( E_x \) and \( E_y \), as functions of \( x \) and \( y \) by equation 3.13. A kick can then be applied to each particle. This does not require much CPU time to calculate, and scales linearly with the number of particles.

As only the extent of the beam is considered, the inside of the beam need not be tracked. A halo beam can be used. This gives a smoother edge compared to a KV for the same number of particles, and therefore reduces statistical noise.

The linear method is quite limited as it does not exhibit tune spread, and will not show any non-linear effects.

A very general method is to do an N-body simulation. Here the distance between each pair of macro-particles is measured, and from this the force calculated. There is no assumption about the particle distribution. However, it is very susceptible to noise so requires a large number of particles. Also CPU time scales with the square of the particle number. For typical size beams it is significantly slower than the other methods.

For the initial beam distributions discussed above there is a symmetry that can be taken advantage of to speed up the calculation. They are all built from a distribution that has a density depending only on \( r \). They are then deformed by a linear transformation, so contours of equal density will form ellipses. The beam has an elliptical symmetry, The beam can therefore
be decomposed into a set of concentric elliptical rings of equal aspect ratio, as shown in figure 5.2.

![Figure 5.2: A Gaussian beam is split into concentric elliptical rings in x-y space.](image)

The number of macro-particles in each ring is counted, and so the ring’s charge density measured and stored. Then for each particle the force contribution from each ring is measured, as shown in figure 5.3. Rings outside the particle, 5.3.a, can be ignored as they give no net force to the particle [104]. Then working from the outside inwards each ring is considered. The field due to a uniform filled ellipse with the charge density and outer dimensions of the ring on the particle is calculated. For the next most outer ring, the difference in density used: if the density is the same then the contribution is included; if the inner ring is more dense, then contribution is added; and if less dense then contribution is subtracted.

If the particle is within the ring, as in 5.3.b, then the field can be calculated using the linear formula for the ellipse. If the particle is outside the ring, 5.3.c and 5.3.d, then the field outside an ellipse is needed. This is calculated using a multipole field with an adjustment to account for the field.
Another method is to place a grid over the beam and count the particles in each segment. This density map can then be used to calculate electric field at any point, using a method such as finite element analysis. This would remove any assumption of symmetry of the beam. This has not been implemented as it would take significant effort, and elliptical symmetry is a good assumption for the beam. However it could be done as future work.

Each of the methods described above gives you the electric field $E_x$ experienced by each particle. One can compare the field between the models by measuring the kick given to some witness particles by a beam. Figure 5.4 shows the field for the three models for a circular KV beam with $10^4$ macroparticles. All models agree within the beam. The linear model is only valid within the beam, whereas the ellipse and the N-body methods give the expected fall off of the field away from the beam. The linear model is only useful in a case where it is known that the beam will remain perfectly uni-
form for the duration of the simulation, and that there will be no particles forming a halo. Some noise is visible in the N-Body model as expected.

Figure 5.4: Electric field in a circular KV beam with radius 1 mm.

The same is shown for an elliptical KV beam in figure 5.5. There is good agreement between the models inside the beam, and the ellipse and N-body methods agree well outside the beam. A numerical integration computed using SciPy [105] is also shown; it agrees well with the other methods. The numerical integration covers a pole when measuring the field inside the beam, as there is a charge density at the witness position. This generated warnings from the integrator, however the result agrees with other methods.

Figure 5.6 shows the electric field for a Gaussian beam. As well as a numerical integration, the second order analytical solution from Lee [70] is plotted. The linear model is obviously no use when the beam is not uniform.
Figure 5.5: Electric field in an elliptical KV beam with $a = 1$ mm and $b = 2$ mm.

The N-body and ellipse models agree well with each other and the numerical integration. Lee’s curve is good for small amplitudes where the bulk of the particles are, but falls off too fast on the edge of the beam.

Figure 5.7 shows the electric field for a Gaussian beam with a different width in $x$ and $y$. Again, there is good agreement between N-body and ellipse models throughout the beam.

The linear model is clearly very limited. Although it is quick, it is not useful if there is any deviation from a linear beam. The ellipse method is useful whenever the beam has a cylindrical or elliptical symmetry. This the case for the standard distributions such as Gaussian, KV and waterbag. The N-body makes no assumptions of symmetry, but has issues of noise and computation time for beams modelled with a large number of macro-particles.
Figure 5.6: Electric field in a circular Gaussian beam with $\sigma = 1$ mm.

Figure 5.7: Electric field in an elliptical Gaussian beam with $\sigma_a=1$ mm and $\sigma_b=2$ mm.
It is however very useful for benchmarking while developing other models.

Once the electric field is calculated the kick given to each particle is,

\[
\Delta x' = \frac{q_{mp}}{m_{mp} \gamma^3 \beta^2 c^2} E_x \Delta s
\]  \hspace{1cm} (5.11)

where \( q_{mp} \) and \( m_{mp} \) are respectively the charge and rest mass of the macro-particle.

In some case it can be useful to bring in a quantity perveance, \( K \), which depends only on the beam energy and current [68]. It is defined as,

\[
K = \frac{I}{I_0} \frac{2}{\beta^2 \gamma^3}
\]  \hspace{1cm} (5.12)

with \( I_0 \) being the characteristic current,

\[
I_0 = \frac{4\pi \epsilon_0 m_0 c^3}{q}.
\]  \hspace{1cm} (5.13)

In the uniform charge density case, we can substitute equation 3.14 into equation 5.11, giving,

\[
\Delta x' = \frac{q_{mp}}{m_{mp} \gamma^3 \beta^2 c^2} \frac{Nq}{\pi \epsilon_0 (a + b)} \frac{x}{a} \Delta s
\]  \hspace{1cm} (5.14)

and re-express in terms of \( K \),

\[
\Delta x' = K \frac{2}{(a + b)} \frac{x}{a} \Delta s.
\]  \hspace{1cm} (5.15)
Benchmarking

It is necessary to compare a new code against analytic models and existing codes to ensure that it is correct.

A very simple case is a uniform beam in a drift tube, an area of zero magnetic field. In the absence of space charge a particle will continue in a straight line. The beam envelope $a$ will evolve according to [68],

$$\frac{d^2 a}{ds^2} + ka - \frac{\epsilon^2}{a^3} - \frac{K}{a} = 0 \quad (5.16)$$

where $k$ is the focusing (zero in a drift) and $\epsilon$ is the emittance.

In the simplest case $\epsilon$ is set to 0, and the beam has no emittance. This is known as a laminar beam, as there will be no crossing of particle trajectories. Starting with a KV beam the $x'$ and $y'$ coordinates for each particle were adjusted to be $xa'/a$ and $ya'/a$, so that the coordinates in $xx'$ space and $yy'$ space form a straight line.

It is useful to create normalised coordinates

$$R = \frac{a}{a_0} \quad Z = \sqrt{2K} \frac{s}{a_0} \quad (5.17)$$

for plotting, where $a_0$ is the initial radius.

This can now be integrated for various initial values of $a'$. The envelopes were calculated using the odeint function in SciPy and are shown in figure 5.8.

Figure 5.9 shows, in crosses, the envelope of a beam tracked with Zgoubi plus the linear space charge kicks. The analytic result is shown again as
Figure 5.8: Evolution of beam envelope in a drift tube for varying initial $a'$, by integration of the envelope equation. $Z$ and $R$ are normalised coordinates, see equation 5.17.

Figure 5.9: Evolution of beam envelope in a drift tube for varying initial $a'$, by Zgoubi with linear method.
solid lines. This was done with $10^4$ macro-particles, and arbitrary values of current, 10 A, beam energy, 50 MeV, and $r_0$, 1 m. The beam was tracked for 420 m, with a step size of 4.2 m, to give 100 points in the range $0 < Z < 2.5$. The agreement is very good.

Figures 5.10 and 5.11 show the comparison again, this time using the ring method, and N-body method. The ring method gives a good answer, as it should reproduce the results of the linear method for a uniform beam.

The N-body method gives a large envelope for several cases. This is because $R$ is calculated from the full width of the beam. The N-body method will sometimes give one particle an unphysically large kick, and so kick it out of the beam.

Figure 5.10: Evolution of beam envelope in a drift tube for varying initial $a'$, by Zgoubi with rings method.

Zgoubi was also compared to an existing space charge code KVBL [106]
Figure 5.11: Evolution of beam envelope in a drift tube for varying initial $a'$, by Zgoubi with N-body method

which solves the above envelope equations for a lattice.

5.3.3 Conclusion

I have shown that Zgoubi with the addition of space charge models is suitable for simulating space charge. For a uniform beam the quick linear method can be used. For other distributions that still have cylindrical symmetry the rings method can be used. The N-body method is not useful, as it is too computationally slow, and creates artefacts.

Currently only transverse space charge effects have been considered. This is equivalent to considering an infinitely long bunch. In the case of a high current machine it will be beneficial to have long bunches in order to spread the charge over a large volume, so this is a reasonable approximation.
The methods currently used assume that the particles within the bunch are only moving slowly relative to each other. This allows the forces to be calculated in the rest frame of the bunch. In some FFAG scenarios there would be multiple bunches with different energies simultaneously in the accelerator. These would be moving at large speeds relative to each other; however, they would still affect each other’s motion. A full electrodynamic approach would be needed to calculate space charge in this case.

The extension of Zgoubi to include space charge is likely to be useful in a range of applications. FFAGs have been proposed for a range of uses such as providing beams for proton and ion therapy and for the rapid acceleration of muons.
Chapter 6

Lattices

There are a wide variety of FFAG designs. The characteristic that they all share is that the magnetic fields remain fixed in time, even though the energy changes, and there is strong focusing. The beam moves from a small orbit in a low field area of the magnets, to a larger orbit with stronger fields. The original FFAGs satisfied the scaling condition, constant tune with energy, but some of the modern designs have relaxed this. This relaxation gives the non-scaling FFAGs which have more freedom in design.

6.1 Scaling FFAGs

The FFAG project at MURA developed two FFAG designs, radial sector and spiral sector. The magnetic field shapes and profiles are tightly defined in order to preserve the scaling condition.
6.1.1 Radial Sector

The radial sector FFAG is made up of radial sector magnets with a radial field gradient. The gradient of the magnets alternates to provide strong focusing. Within each magnet the field at the median plane at any azimuth is [48],

\[ B = B_0 \left( \frac{r}{r_0} \right)^k \]  

(6.1)

where \( r \) is the distance from the centre of the machine and \( k \) is the field index for that machine.

Figure 6.1 shows the layout of the magnets. The MURA radial sector FFAGs alternated between focusing and defocusing elements. The KEK designs combine 3 magnets into a defocusing-focusing-defocusing (DFD) triplet unit, with drift spaces and RF cavities between them as show in figure 3.3.

![Figure 6.1: Layout of magnets in a radial sector FFAG [48].](image)
If the focusing and defocusing elements are identical in strength apart from the sign of their magnetic field then the lattice can stably hold two same charge beams travelling in opposite directions. This would allow construction of a particle collider; to build a collider with a synchrotron the beams must be opposite sign, for example proton on anti-proton, or the beams must go through separate magnets.

If the focusing elements (which provide the positive bending) are made larger or stronger then the overall radius of the machine can be reduced.

### 6.1.2 Spiral Sector

The spiral sector FFAG has a more complex field. The guide field on the median plane is [48],

\[
B = B_0 \left( \frac{r}{r_0} \right)^k \left( 1 + f \cos \left( N\theta - N \tan \zeta \ln \left( \frac{r}{r_0} \right) \right) \right)
\]

(6.2)

where \( \theta \) is the azimuthal angle, \( f \) the flutter factor, \( N \) the number of sectors and \( \zeta \) the spiral angle between the locus of maximum field and the radius.

This gives the field shown in figure 6.2. The alternation of the field is now seen as the particles cross at an angle from regions of high to low field, and back again.

Figure 6.3 shows the spiral sector FFAG at MURA.

Spiral sector FFAGs can have a smaller radius than a radial sector FFAG, as there can be a bending field applied for most of the orbit.

There is some overlap between spiral sector FFAGs and sector cyclotrons. Cyclotrons however are always designed such that the revolution frequency
is constant with energy, whereas for the scaling FFAGs the main design constraint is to keep the tune constant.
6.2 Non-Scaling FFAGs

Relaxing the scaling condition opens up a huge range of possible lattices. The constraint now is only that the lattice must be stable for particles over a range of energies.

6.2.1 Linear Non-Scaling FFAGs

An accelerator lattice that contains only dipole and or quadrupole components, either as separate elements or combined function magnets is described as linear. This is because, with the paraxial approximation, the transfer maps for such a lattice are linear. For non paraxial systems such as FFAGs this is not true, but the name sticks.

A linear non-scaling FFAG has 2 main advantages over the scaling designs. Firstly the magnets can be much simpler. Only magnets up to quadrupole are needed, and they can be rectangular instead of sectors or with curved edges. Secondly the beam excursion with energy can be made smaller. This is because it is no longer necessary to use the field shown in equation 6.1, instead a high gradient can be applied across the whole magnet aperture.

EMMA, described in section 3.4.3, has a very simple lattice composed of 42 straight sections. In each of these there are 2 quadrupoles. These are square to the reference line of the straight section, but with their magnetic centres offset to provide a dipole component. This is shown in figure 6.4. The long drift contains a RF cavity in 19 of the 42 cells and is also used for the injection and extraction septum and kickers.
As EMMA is a prototype each magnet is mounted on a transverse horizontal slider so that a range of lattice parameters can be investigated. In a production machine this would not be necessary.

EMMA has a large radius given its energy, a 20 MeV electron beam has a rigidity of 0.068 Tm. This is because only a small fraction of the length is used for bending. A 1 GeV proton has a rigidity of 5.66 Tm, and so an EMMA lattice with the same magnet strengths would be around 100 times larger.

EMMA has very little space between magnets. This was a problem for the design of the injection and extraction section. The beam has to make a tight 70° bend so that it avoids having to pass though the previous magnet. A more practical design would need to leave more space for injection.

There are possible variations to a linear non-scaling design. For exam-
ple adding a dedicated dipole for bending removes the need to offset the quadrupoles, simplifying their design. A lattice that looks like a traditional synchrotron FODO cell will work. In a synchrotron dispersion is the measure of orbit shift due to relative momentum. For an FFAG this shift is the separation of orbits for different momenta. Increasing the quadrupole strength reduces this orbit excursion, as the particles don’t need to move as far to experience the higher fields.

6.2.2 Non-linear Non-scaling FFAG

While linear non-scaling FFAGs simplify magnet design, it may still be desirable to minimise tune shift during acceleration. There are several ways to modify a non-scaling lattice to reduce the tune shift. The lattice is now non-linear, but it still does not follow the scaling law. This non-linear non-scaling FFAG design has been proposed for the PAMELA medical proton accelerator.

Johnstone proposed using quadrupoles with wedge angles to adjust the tune [108, 109]. Although the radial field profile is linear, this effectively adds some higher order terms to the lattice.

PAMELA takes a different approach to stabilise the tune. A lattice is designed starting with the $r^k$ field, then the field is expanded into multipoles, and higher order terms discarded [110]. For PAMELA multipoles up to octopole would give sufficient tune flattening. The magnets can have parallel entrance and exit faces, which simplifies the design. The multipole components are provided by superconducting magnets with multiple helical
windings [111].

### 6.2.3 Isochronous FFAGs

An accelerator is isochronous if its revolution frequency is constant at all energies. The revolution frequency depends on the speed of the particles and the length of the orbit.

At high energies the change in speed of the particles is very small as $\beta$ is close to 1. To make an isochronous lattice you need to keep the path length constant. At low energies ($\gamma \lesssim 2$) the speed can vary significantly and so one must make the path length follow this speed change.

At low energies the particle’s speed is proportional to its momentum. In a classical cyclotron with a constant magnetic field, the radius will also be proportional to the momentum. This allows the speed change and orbit change to cancel perfectly, and the revolution frequency to be constant.

For protons this relation begins to break down around 10s of MeV. To get to higher energies in an isochronous cyclotron the field profile can be shaped so that the change in orbit radius continues to cancel. The angular frequency of particles in a cyclotron will be

$$\omega = \frac{\beta c}{\rho}$$ \hspace{1cm} (6.3)

and is constant as long as $\beta$ and the radius $\rho$ are proportional. Using the rigidity equation

$$B\rho = \frac{p}{q}$$ \hspace{1cm} (6.4)
which relates momentum, magnetic field and radius, and the relativistic definition of momentum,

\[ p = \gamma m_0 \beta c = m_0 c \frac{\beta}{\sqrt{1 - \beta^2}} \]  

we can find the necessary shape of the magnetic field equation. In terms of \( \beta \), the field must be,

\[ B(\beta) = \frac{m_0 \omega}{q \sqrt{1 - \beta^2}} \]  

and more usefully, in terms of radius,

\[ B(\rho) = \frac{m_0 \omega}{q \sqrt{1 - \omega^2 c^2 \rho^2}} \]  

A field that increases with radius gives a weak vertical defocusing force, so isochronous cyclotrons must vary their azimuthal field. In this case the above formula are approximations for the average field and radius. Figure 6.5 shows how for low \( \beta \) a constant field gives isochronicity, but as \( \beta \) rises above 0.2 the field must also rise.

Figure 6.6 shows the average bending field when the parameters for the PSI are put into equation 6.7. The actual field in the PSI magnets goes up to 2 T [79]. As the magnets fill about half of the lattice, to leave room for RF and drift spaces, the field averaged over the orbit will be close to that plotted.

The same technique can be applied to an FFAG. If the field is shaped correctly then the lattice can be made isochronous.

It is possible to use optimisation to achieve an isochronous design by
Figure 6.5: Required variation in field with $\beta$ for isochronicity.

Figure 6.6: Required variation in field with radius for isochronicity.

tailoring an arbitrary radial field profile [112].

An isochronous FFAG could use a much simpler RF system for acceleration as there would be no need to sweep the RF frequency. This would allow
multiple beams at different energies to be accelerated simultaneously, giving
a continuous beam similar to a cyclotron. However it requires a large change
in path length, and therefore radius, for a machine with a large change in $\beta$. This requires magnets with larger physical apertures.

In a continuous accelerator more care needs to be taken not to disturb the other orbiting beams while doing injection and extraction. When there is just a single bunch orbiting it is possible to use kicker magnets, which rapidly switch on to provide a field that bends the beam in or out of the accelerator. In the absence of kickers one needs to provide a field that varies in space, so that when the beam reaches the outer edge of the accelerator it is bent into an extraction septum. It will be necessary to have a large separation between orbits to allow this to be done cleanly.

6.3 Cascading Accelerators

The task of designing an accelerator gets greater as its energy range is increased. It is therefore normal to cascade multiple accelerators together to cover the required range. Chapter 4 showed examples of large accelerator complexes combining various types of accelerator.

Typically the ion source will create a beam from tens to hundreds of keV, which is then accelerated to a few MeV in a cyclotron or RFQ. Commercial cyclotrons are available up to energies of 70 MeV [113].

One could then use a chain of 2 FFAGs to bring the beam up to 1 GeV. There are a few possible methods to choose the injection energy for the second FFAG. One could take half way in energy, to maximise sharing of
acceleration, half way in momentum as that defines the orbit shapes, or half way in speed which defines the RF frequency. These each vary non-linearly with each other, as shown in figure 6.7, so no choice satisfies them all.

![Figure 6.7: Variation in speed and momentum of protons up to 1 GeV](image)

Depending on the injection energy of the first FFAG, one can select a mid point. For example if you inject at 35 MeV, which gives $\beta = 0.27$, and extract at 1 GeV where $\beta = 0.88$ then the speed midpoint is $\beta = 0.57$ which corresponds to an energy of 204 MeV. These midpoints for a selection of energies are shown in table 6.1.

<table>
<thead>
<tr>
<th>Injection energy</th>
<th>KE midpoint</th>
<th>momentum midpoint</th>
<th>speed midpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.0</td>
<td>517.5</td>
<td>416.55</td>
<td>204.05</td>
</tr>
<tr>
<td>50.0</td>
<td>525.0</td>
<td>435.32</td>
<td>228.63</td>
</tr>
<tr>
<td>70.0</td>
<td>535.0</td>
<td>456.92</td>
<td>258.26</td>
</tr>
</tbody>
</table>

Table 6.1: Midpoints in kinetic energy in MeV when splitting acceleration by kinetic energy, momentum and speed.
6.4 RF Limits

The acceleration of the particles is governed by the RF system. Each time a particle passes through an RF cavity at the correct phase it receives energy. In a linac the beam passes through each cavity only once, and so many cavities with high gradients are needed. In a circular accelerator the beam can make multiple passes, with a small amount of energy added each time.

In a isochronous or near isochronous accelerator the RF system can run at a constant frequency. However in a non-isochronous accelerator the frequency needs to be swept.

In a synchrotron the acceleration period is limited by the time taken to ramp the magnets, which is from 10 ms in a rapid cycling synchrotron, up to tens of minutes. In a low energy synchrotron the RF frequency will be adjusted during this sweep. For a fixed field accelerator the limit is set by the number of passes through the RF cavity required and the time required to sweep the RF frequency. RF frequencies can be adjusted faster than magnetic fields.

An RF cavity is a resonator. Often, for efficiency, they are designed to have a narrow bandwidth, that is a high quality value (high Q). However they can also be wideband, or dynamically tunable. With high beam loading, due to a high intensity beam the Q value would be reduced anyway.

For example the KEK POP accelerator uses a low Q cavity with a FINEMET magnetic alloy core with a resonant frequency of 5 MHz. They drive it from 0.61 to 1.38 MHz with a 1 kHz repetition rate [114]. In the 150 MeV FFAG at KEK a similar cavity is used with a sweep from 1.5 to 4.6 MHz with a
repetition rate of 100 Hz [115].

The 150 MeV FFAG uses a single cavity to provide 6 kV. It has a gradient of 35 kV/m and takes up 0.4 m of a straight section [115]. This would require 23000 laps to accelerate from 12 to 150 MeV.

6.5 Lattice for Simulations

For the purpose of simulations in this thesis a simple lattice was designed, that exhibits the challenging features of a non-scaling FFAG. It accelerates from 35 to 400 MeV, and has a footprint of approximately 10 m diameter. Large enough to give space for many RF cavities, but still small compared to a linac option, and the overall size of a nuclear power facility as shown in figure 6.8.

The lattice is composed of 30 cells, each 1 m long. Each cell, shown in figure 6.9, has 2 quadrupoles and a dipole. The magnets are 0.2 m long with parallel entrance and exit faces. There is a 0.05 m gap between the magnets and the bends in the reference line are either side of the dipole. The rest of the cell is left clear for RF cavities.

The dipole is tuned to midpoint momentum. The quadrupoles have equal strength but opposite sign, and are tuned to give as much focusing as possible. To choose the magnet values an injection and extraction energy must be given. The injection energy is set slightly below the real injection energy, other injection would be right on the edge of stability. Using 35 MeV as the injection and 400 MeV as the extraction energy, the dipole has a field of 2.11 T and the quadrupoles have gradients of 15.95 T/m. These are a bit
Figure 6.8: (a) 500 m Linac, (b) 10 m radius FFAG and (c) $3 \times 10$ m radius FFAG, compared to the Sizewell nuclear power station [116].

beyond the fields that can be achieved with warm magnets. This is not a problem for this simulation work as the effects studied will be common to all similar lattices. For some earlier work the injection energy was set to 70 MeV.

The lattice is implemented in Zgoubi using the \texttt{MULTIPOLE} element for the magnets, \texttt{DRIFT} for the drift spaces and \texttt{CHANGREF} for the bends.

Figure 6.10 shows the Twiss parameters for the lattice at injection energy. Discontinuities are due to the reference changes at the edge of the dipole. At injection the maximum of the $\beta$ function is 1.74 m and 2.13 m horizontally and vertically respectively.

Figure 6.11 shows the large swing in revolution frequency of the lattice, from 2.65 to 6.55 MHz, due to the large change in $\beta$ from 35 to 400 MeV.
Figure 6.9: NS-FFAG lattice with dipole, showing orbits from 35 MeV to 400 MeV

Figure 6.12 shows the how the cell tune changes rapidly during the acceleration.

The issues of frequency change and space charge would be smaller for the second FFAG that accelerates from 400 MeV to the 1 GeV spallation energy, so it is not considered in this thesis.
Figure 6.10: Twiss parameters for NS-FFAG lattice with dipole at 35 MeV

Figure 6.11: NS-FFAG lattice with dipole, showing revolution frequency from 35 MeV to 400 MeV
Figure 6.12: NS-FFAG lattice with dipole, showing horizontal and vertical tune from 35 MeV to 400 MeV
Chapter 7

Multi-Bunch Methods

A defining feature of an FFAG accelerator is that its magnetic field does not change as a function of time. The field provides stable, strongly focused orbits for the whole range of particle energies. This implies a possibility of running an FFAG with continuous injection and extraction as in a cyclotron. This is in contrast to a synchrotron where the magnetic field is ramped up during acceleration, proportional to the particle momentum.

For an accelerator working below ultra-relativistic energies, for example the protons in an ADSR driver which are below 1 GeV, the speed of the particles changes during acceleration. Unless a large orbit shift is allowed, this causes the revolution frequency to vary with energy, which in turn requires the RF frequency to be swept during the acceleration cycle.

With the RF frequency changing with time, the accelerator cannot always hold particles at any energy. All injection must occur before the RF frequency sweep starts, and extraction can not occur until the maximum frequency is
reached. This will limit the repetition rate of the accelerator.

This chapter examines two methods that may allow multiple bunches of different energies to orbit and be accelerated simultaneously. The first takes advantage of the multiple solutions to the harmonic condition for acceleration. The second exploits the wide tolerance for accelerating a bunch that does not have the correct energy.

There was some investigation into similar acceleration schemes in the early days of FFAGs. Symon described a ‘bucket lift’ method and use of multiple oscillators [117], however it seems that this idea was not developed or tested [107].

These methods are distinct from and complementary to the common method of using many RF buckets by running an RF frequency that is a multiple of the revolution frequency.

For simplicity a ring with a single RF cavity will be considered. However, in the first method the cavity needs to have multiple simultaneous frequencies.

### 7.1 Multiple Harmonics

#### 7.1.1 Conditions for Acceleration

To accelerate to energies above a few MeV it is necessary to use alternating electric fields. These must be synchronised to the passing of the beam so that the beam always sees a field that causes it to accelerate. In a circular machine this is achieved by making the cavity’s RF frequency, $C$, an integer
multiple, $h$, of the beam’s revolution frequency, $B$, so that the beam always sees the same phase in the cavity:

$$C = hB. \quad (7.1)$$

The revolution frequency is given by the speed divided by the path length. In an accelerator where the revolution frequency changes with time, $C$ and $B$ vary, but $h$ remains constant,

$$C(t) = hB(t). \quad (7.2)$$

An exception to this is to use a harmonic number jumping scheme, where $h$ can also change, but this is not considered in this thesis.

### 7.1.2 Two Beams

For a given revolution frequency there are an infinite number of harmonic frequencies at which the RF cavities can be run, although only some of those will be practical. Also, multiple revolution frequencies can have the same cavity frequency as a different harmonic.

This means that under some conditions it is possible to have multiple beams, of different energies, and multiple cavities, at different frequencies, and still satisfy equation 7.1, for each beam-cavity pair.

For example, a beam with a revolution frequency of 3 MHz could be driven by cavities at 3 MHz, 6 MHz, 9 MHz and so on. Beams with revolution frequencies 3 and 4 MHz could both be driven by a cavity at 12 MHz, with harmonic numbers 4 and 3 respectively.
A broadband cavity can have multiple RF frequencies simultaneously; such a system is considered in this section.

The following shows a system with 2 beams being accelerated at a time; a third is injected as the first is extracted. The acceleration of each beam takes time $T$. A beam is injected every $T/2$.

\[ \text{Figure 7.1: Beam revolution frequencies against time} \]

Figure 7.1 shows the revolution frequency against time for three beams, $B_A(t)$, $B_B(t)$ and $B_C(t)$, and figure 7.2 shows the same for three cavity frequencies, $C_A(t)$, $C_B(t)$ and $C_C(t)$.

The frequency of cavity A is a fixed integer multiple of the revolution frequency of bunch A. The same is true for cavity B and bunch B, and cavity C and bunch C, and remains true at all times,

\[ \frac{C_A(t)}{B_A(t)} = \frac{C_B(t)}{B_B(t)} = \frac{C_C(t)}{B_C(t)} = h_1 \quad (7.3) \]

where $h_1$ is a constant integer.
Everything repeats in time, hence

\[ B_A(t) = B_B(t + T/2) = B_C(t + T) \]  \hspace{1cm} (7.4)

and

\[ C_A(t) = C_B(t + T/2) = C_C(t + T). \]  \hspace{1cm} (7.5)

Also bunches must be harmonic with the other cavities giving

\[ \frac{C_A(t)}{B_B(t)} = \frac{C_B(t)}{B_C(t)} = h_2 \]  \hspace{1cm} (7.6)

and

\[ \frac{C_B(t)}{B_A(t)} = \frac{C_C(t)}{B_B(t)} = h_0. \]  \hspace{1cm} (7.7)

All the harmonics, \( h_0, h_1 \) and \( h_2 \), must be integer. If we can find a set of harmonics that satisfies the conditions then multiple bunches can be accelerated simultaneously.
From
\[ \frac{C_A(t)}{B_B(t)} = h_2 \quad \frac{C_B(t)}{B_B(t)} = h_1 \] (7.8)
we get
\[ \frac{C_A(t)}{C_B(t)} = \frac{h_2}{h_1}. \] (7.9)
So the frequencies in cavities A and B are always separated by a constant fraction. Combining the time constraints from equation 7.4, we get
\[ B_A(t) = B_B(t) \frac{C_A(t)}{C_B(t)} = B_B(t) \frac{h_2}{h_1}. \] (7.10)
The only non-trivial solution for equation 7.4 and 7.10 is an exponential,
\[ B_A(t) = x e^{yt+z}. \] (7.11)
At \( t = 0 \), \( B_A = f_a \), and at \( t = T/2 \), \( B_A = f_b = f_a h_2/h_1 \), so constants are
\[ B_A(t) = f_a e^{-\log(h_2/h_1)T}. \] (7.12)
The revolution frequency at injection and extraction depends on the particle speed, but also it must satisfy
\[ B_B(T) = B_B(0) \left( \frac{h_2}{h_1} \right)^2 \] (7.13)
so if the revolution frequency were to double, then we need to satisfy,
\[ \left( \frac{h_2}{h_1} \right)^2 = 2. \] (7.14)
The harmonics $h_1$ and $h_2$ must be integers which restricts us to approximations, e.g.

$$\left(\frac{10}{7}\right)^2 = 2.04081 \quad \text{or} \quad \left(\frac{3}{2}\right)^2 = 2.25. \quad (7.15)$$

The approximation is not a problem, as it can be accounted for by making a small change to the injection or extraction energy.

The harmonics $h_0$, $h_1$ and $h_2$ are each separated by the same fraction,

$$h_2 = \left(\frac{h_2}{h_1}\right) h_1 = \left(\frac{h_2}{h_1}\right)^2 h_0. \quad (7.16)$$

The lowest integer solutions look like

$$h_0 = ka^2 \quad h_1 = kab \quad h_2 = kb^2 \quad (7.17)$$

where $a$, $b$ and $k$ are positive integers and

$$\frac{b}{a} = \frac{h_2}{h_1}. \quad (7.18)$$

So the smallest harmonics to satisfy the $10/7$ ratio are

$$h_0 = 49 \quad h_1 = 70 \quad h_2 = 100. \quad (7.19)$$

As an example we can consider an accelerator with an injection revolution frequency of 1 MHz, and an extraction revolution frequency of $1 \times (10/7)^2 \approx 2$ MHz, and a cavity that can sweep from 70 to 142 MHz.

At t=0 there is a bunch, A, already half way through its acceleration, with a revolution frequency of $1 \times (10/7) \approx 1.43$ MHz, and a new bunch,
B, has just been injected with a revolution frequency of 1 MHz. The cavity is generating the frequencies 70 and 100 MHz. It is clear that bunch B has harmonics 70 and 100. Bunch A has harmonics $70/(10/7) = 49$ and $100/(10/7) = 70$. These are the integers $h_0, h_1$ and $h_2$.

At $t=T/2$ the RF cavity has swept its frequencies up, from 70 to 100, and from 100 to $100 \times (10/7) \approx 143$ MHz. This causes bunch A to accelerate to an energy with a revolution frequency of $1 \times (10/7)^2 \approx 2$ MHz, and bunch B to an energy with a revolution frequency of $1 \times (10/7) \approx 1.43$ MHz. As the frequencies have all increased by the same proportions the harmonics have remained constant. Bunch A is now at the extraction energy and can be extracted. Now the 143 MHz frequency is switched off, and a new 70 MHz frequency is switched on. A new bunch, C, is now injected with a revolution frequency of 1 MHz. The beams and frequencies are back to the situation at $t=0$, and the cycle can begin again.

Although some frequencies above were given approximately, the harmonics were always exact.

### 7.1.3 Three Beams

In the two beam case a given beam shares the accelerator with a higher energy beam for the first half of the acceleration, and a lower energy beam for the second half. This requires three integer harmonics, a basic one $h_1$, a lower one $h_0$ and a higher one $h_2$. For a three beam case, at the start there will be two beams of higher energy, and by the end two beams of lower energy. This will require five integer harmonics. This changes the requirements in
For the same overall speed change, the harmonics will need to be closer together. So the $h_2/h_1$ needs to be raised to the fourth power,

$$B_B(T) = B_B(0) \left( \frac{h_2}{h_1} \right)^4. \quad (7.20)$$

For a factor of two frequency change we can use

$$\left( \frac{6}{5} \right)^4 = 2.0736. \quad (7.21)$$

We also need more harmonics,

$$a^4 \ a^3b \ a^2b^2 \ ab^3 \ b^4. \quad (7.22)$$

So the smallest harmonics are

$$625 \ 750 \ 900 \ 1080 \ 1296. \quad (7.23)$$

If we were to choose $a$ and $b$ to improve the approximation, eg

$$\left( \frac{119}{100} \right)^4 = 2.00533921 \quad (7.24)$$

it results in larger harmonics

$$100000000 \ 119000000 \ 141610000 \ 168515900 \ 20053921. \quad (7.25)$$

In accelerators with a circumference of 10 m, the revolution frequency will
be of the order of MHz. The highest frequencies commonly used in acceler-
ators are of the order of GHz. Therefore the maximum practical harmonics
are of the order 1000. It is clear that $a$ and $b$ must be kept small.

### 7.1.4 More Than Three Beams

For a fixed frequency shift, as the number of bunches increase the harmonics
will be closer together, hence $b/a$ will need to be closer to 1. For $n$ bunches
there will need to be a main harmonic, plus $n - 1$ above and below. So the
number of needed harmonics,

$$n_h = 1 + 2(n - 1) = 2n - 1. \quad (7.26)$$

There will need to be $n_h - 1$ steps between the harmonics, so $a$ and $b$ must
satisfy;

$$\left( \frac{b}{a} \right)^{n_h - 1} = \left( \frac{b}{a} \right)^{(2n-2)} \approx 2. \quad (7.27)$$

The highest harmonic will be

$$h_{\text{max}} = b^{2n-2} \quad (7.28)$$

so $b$ must be minimised. Table 7.1 shows the $b/a$ with the smallest $b$ that
satisfies equation 7.27 to within 10% for increasing $n$.

So as $n$ increases, $a$ and $b$ must increase, and hence $h_{\text{max}}$ will increase
rapidly.

For larger frequency sweeps, the values for $a$ and $b$, shown in table 7.2,
are smaller, but they still grow quickly.

127
Table 7.1: Harmonic separation for $n$ beams

<table>
<thead>
<tr>
<th>Beams</th>
<th>$a/b$</th>
<th>Frequency change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$7/5$</td>
<td>$(7/5)^2 = 1.96$</td>
</tr>
<tr>
<td>3</td>
<td>$6/5$</td>
<td>$(6/5)^4 = 2.0736$</td>
</tr>
<tr>
<td>4</td>
<td>$9/8$</td>
<td>$(9/8)^6 = 2.02728652954$</td>
</tr>
<tr>
<td>5</td>
<td>$11/10$</td>
<td>$(11/10)^8 = 2.14358881$</td>
</tr>
<tr>
<td>6</td>
<td>$14/13$</td>
<td>$(14/13)^{10} = 2.09819976337$</td>
</tr>
<tr>
<td>7</td>
<td>$16/15$</td>
<td>$(16/15)^{12} = 2.16942521297$</td>
</tr>
<tr>
<td>8</td>
<td>$19/18$</td>
<td>$(19/18)^{14} = 2.13174543623$</td>
</tr>
</tbody>
</table>

Table 7.2: Harmonic separation for $n$ beams, 3× sweep

<table>
<thead>
<tr>
<th>Beams</th>
<th>$a/b$</th>
<th>Frequency change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$5/3$</td>
<td>$(5/3)^2 = 2.7777777778$</td>
</tr>
<tr>
<td>3</td>
<td>$4/3$</td>
<td>$(4/3)^4 = 3.16049382716$</td>
</tr>
<tr>
<td>4</td>
<td>$6/5$</td>
<td>$(6/5)^6 = 2.985984$</td>
</tr>
<tr>
<td>5</td>
<td>$8/7$</td>
<td>$(8/7)^8 = 2.91028536805$</td>
</tr>
<tr>
<td>6</td>
<td>$9/8$</td>
<td>$(9/8)^{10} = 3.24732102547$</td>
</tr>
<tr>
<td>7</td>
<td>$11/10$</td>
<td>$(11/10)^{12} = 3.13842837672$</td>
</tr>
<tr>
<td>8</td>
<td>$13/12$</td>
<td>$(13/12)^{14} = 3.06668725033$</td>
</tr>
</tbody>
</table>

7.1.5 Practicality

This method places strict requirements on the modulation of the RF frequency. It must increase as an exponential with time, in order that the harmonic conditions are maintained. Although complex modulation is possible with modern RF systems, it may be an awkward constraint. It may also require that the RF voltage is modulated in a complex way to ensure that energy gain per turn is kept in sync with the RF frequency.

As we have seen, with three beams the smallest solution that allows a factor two change in revolution frequency gives a $h_{\text{max}}$ of 1296. This is already a large harmonic number for a small accelerator. For more than three beams the situation becomes more difficult. Increasing the frequency
sweep is not a solution, as this will likely increase the sweep time, slowing the acceleration cycle.

So this is not a practical method for significantly increasing the number of bunches in an FFAG with protons in the region of a few 100 MeV.

However, this does not rule out methods where the beams are not accelerated by all the RF frequencies. If the revolution frequency of a bunch is not an integer factor of an RF frequency then it will sample different phases of the RF waveform on each lap. On average the kicks should cancel, but they could also decelerate the bunch, thus wasting energy.

At KEK two bunch acceleration has been shown in an FFAG [114]. Each bunch had an RF frequency matched to it. As long as the frequencies were kept far enough apart acceleration was achieved.

7.2 Multiple Turn Injection

Given that finding multiple exact solutions to accelerating multiple bunches is not very successful we can look at some inexact solutions.

One possibility is to keep injecting bunches during the RF frequency sweep. Normally one would inject the beam with RF frequency exactly matched to the revolution frequency for the injection energy times the harmonic number. Then one would stop the injection, and sweep up the RF. However if one was to keep injecting during the sweep there would be some time when the frequencies would be close enough to still capture the bunch and accelerate it.

This can also work the other way. If the sweep starts below the matched
frequency then a beam could be injected early.

7.2.1 Simulation

To find how large this window is one needs to do a simulation that includes transverse and longitudinal dynamics. The simulation was carried out using the NS-FFAG lattice described in section 6.5 with injection at 70 Mev and extraction at 500 MeV.

A proton bunch was created that had particles on the injection energy reference orbit at 70 MeV with a time spread equivalent to 1° of RF phase between them, covering a full 360°. The bunch was tracked for a lap through the lattice with Zgoubi. Then in PyZgoubi each particle in the bunch was given an RF kick dependent on its accumulated time of flight through the lattice. The bunch was then tracked through the next lap of the lattice. The RF frequency was swept from 3.7 to 7.3 MHz over the course of 1000 turns, causing some of the particles to accelerate up to 500 MeV.

In a real machine particles would be injected in a bunch timed to correspond to the phase region that has acceleration. The larger this is the more particles can be put into the bunch.

Figure 7.3 shows the final energy of particles after 1000 turns for the range of initial phases around the reference particle. Particles still in the same RF bucket as the reference are shown in red. These have been fully accelerated, whereas others have only been partially accelerated. The longitudinal acceptance is 130° wide.

Figure 7.4 shows the particles in longitudinal phase space around a ref-
Figure 7.3: Energy after 1000 turns for range of RF phase offsets.

Figure 7.4: Longitudinal phase space over 1000 turns. Blue particles start earlier than the reference phase, whilst red particles start later. Starting positions are marked with an X.
ference particle. Particles within the 130° around the reference particle, from -47 to +82, are trapped, and oscillate in phase and energy similar to the motion in a synchrotron. Particles outside the longitudinal acceptance are progressively retarded in phase and energy compared to the reference particle.

### 7.2.2 Second Turn Acceptance

To compare, we show a similar bunch launched with an initial energy still at 70 MeV, but injected at the time that the first bunch would have made one turn. By now the sweep has started and the RF frequency is slightly too high compared to the revolution frequency. This imperfectly matched beam is then tracked as before, but for only 999 turns.

![Energy vs. Phase Graph](image)

Figure 7.5: Energy after 1000 turns for range of launch phases on the second turn.
Figure 7.5 shows the final energy of particles after the remaining 999 turns. Particles still in the same bucket as the second turn reference particle are shown in red. The region of acceptance is reduced 120° as shown in Figure 7.6, due to the imperfect matching. Still a large number of particles are fully accelerated.

![Figure 7.6: Longitudinal phase space over 1000 turns, for particles launched on the second turn.](image)

### 7.2.3 Injecting on Many Turns

This method can now be extended forwards and backwards; bunches were launched before and after the first lap and the longitudinal acceptance measured. Figure 7.7 shows how the acceptance varies for different injection laps. The region in which particles can be accelerated falls to zero degrees about
7 turns before and after the first lap. This means particles could be injected for 14 turns from the start of the RF sweep, or for example the 12 turns with an acceptance above 60°.

![Figure 7.7: Longitudinal acceptance for injecting on successive turns.](image)

### 7.2.4 A Realistic Cyclotron Beam

To see the acceptance of this accelerator for realistic beams, Gaussian-distributed bunches with vertical and horizontal emittance of 1 π mm mrad, and longitudinal emittance of 0.5 % dE/E × RF phase were injected. These values are similar to those of the output beam of the PSI Injector 2 Cyclotron [118]. This beam was injected during the sweep on each of several laps before and after the nominal start lap, and then tracked for 1000 turns as before.

Figure 7.8 shows the survival of the realistic bunch injected on successive turns.
laps. For bunches injected on laps -5 to 6 the survival is above 99%. It then falls off rapidly.

Figure 7.8: Survival of realistic bunches injected on successive turns.

Particles that don’t get caught in the accelerating buckets stay close to the injection energy. This is good, as if there are losses, it is better that they are at low energy which will cause less activation of the accelerator and less energy waste. For a real machine the bunches would be shaped and chopped to fit the acceptance, so beam loss would be less.

7.2.5 Practicality

The FFAG lattice shown in this section has a wide longitudinal acceptance, without any specific optimisation of the RF sweep or lattice. Simulations with a realistic injector output show that it should be possible to inject
before and after the first turn even though the RF frequency will not be
perfectly matched to the revolution frequency at that point. This would
increase the duty factor by over an order of magnitude, in this case from
1/1000 to 12/1000.

This seems to be a practical method for increasing the average current of
an ADSR driver. However, it may increase the complexity of the injection
and extraction systems. Also, the bunches injected in successive turns will be
physically close to each other during the acceleration. There may be beam-
beam effects that require advanced techniques to model [119], and so were
not studied in this thesis.
Chapter 8

Current Limits

An ADSR requires a high current beam to drive the spallation. As shown in section 2.3, our baseline design with a \( k_{\text{eff}} = 0.985 \) requires 10 mA of average current at 1 GeV in order to provide 600 MW electrical power. This is currently beyond that of the world’s largest neutron spallation accelerators discussed in chapter 4, so achieving it will be a challenge.

To provide 10 mA of average current the peak current must be higher. With conventional RF the bunch length must be significantly less than half of the ring circumference, so that there is no beam in the cavities while they are in their decelerating phase. A third harmonic cavity can be used to maximise this filling. A filling factor of one third means that bunches fill one third of the ring. This would give a bunching factor, the ratio of average to peak current of approximately one third.

In the case of a pulsed accelerator the peak current is even greater. No beam arrives at the target during the acceleration time; then a short pulse
is delivered. If the duty factor is 1/1000 then the peak current is 1000 times the average current. In this chapter ‘current’ is used to refer to peak current unless otherwise stated.

Increasing the circumference of the machine will spread the bunch out over a longer distance. However it will also increase the time per lap and so lower the repetition rate for acceleration with the same number of laps. These effects cancel, so that the required peak current is independent of machine size. A larger machine does make it easier to accelerate in fewer turns. If the accelerator size is doubled then the time needed to execute the sweep for the same number of laps is doubled. It would also allow space for more RF cavities, so the voltage and power of each cavity could be reduced.

The current in an accelerator is limited by effects that are not obvious at early stages of machine design. Collective effects of the beam (the beam’s interaction with itself) are neglected in most particle accelerator design codes. By adding space charge modelling to Zgoubi, as described in section 5.3.1, some of these collective effects can be investigated.

8.1 Space Charge

Space charge is strongest at lower energies due to its being suppressed by high relativistic $\gamma$ values. I will therefore only consider its effect in the first FFAG, and mostly concentrate on its effect around the injection energy.

Space charge is disruptive to the beam in a number of ways. At the most extreme currents space charge can prevent the beam passing through the lattice; at lower currents space charge has subtler effects.
Delivering 1 mA average current with a 1 kHz repetition rate requires bunches of 1 \( \mu C \) or \( 6.24 \times 10^{12} \) protons. This bunch could be divided into the sub bunches of a bunch train, but as they would need to be physically smaller the peak density and hence transverse space charge effects would be the same.

Figure 8.1 shows the generalised perveance of proton beams at 35 and 70 MeV.

![Figure 8.1: Generalised perveance as a function of current for proton beams at 35 and 70 MeV.](image)

### 8.1.1 Focusing

Finding the point at which the current is great enough to prevent the magnets from focusing the beam is fairly simple, and puts a very hard limit on the maximum current. This is the point at which there is no longer a periodic
solution to the Twiss parameters, and can also be interpreted as the point at which the tune is suppressed to zero.

One can check whether a lattice is stable by taking the trace of the transfer matrix. If the trace of the 2 by 2 matrix for either of the transverse planes is greater than two then it is not stable [120].

Given a lattice, one can search for the limit of stability by increasing the beam current until the lattice is no longer stable, and then searching back to find the limit.

A KV beam of 50k particles was used for the simulation. It was tracked through one cell of the lattice along with a test bunch of 11 particles, (a reference particle and pair particles offset in $x, x', y, y'$ and $p$). The test bunch feels the space charge force from the main beam, but does not contribute to it. The magnets were split into 10 slices and tracked with Zgoubi. Between each slice a space charge kick was calculated and applied with the concentric elliptic rings method, using 128 rings as described in section 5.3.2. The transfer matrix is found using the 11 particle test bunch and the method described in appendix A.3.

The limit will depend on the injection energy and the beam emittance. Taking the 30 m NS-FFAG with quadrupoles and dipoles, and using an injection energy of 35 MeV, the maximum stable charge for a given emittance is shown in figure 8.2. For comparison, the same was done using the linear space charge model, shown in figure 8.3. The results are close, as would be expected given that the beam initially has a uniform cross section.

The emittance dependence is strong and linear. If the beam is physically large the space charge force is reduced due to the lower charge density. It
Figure 8.2: Maximum stable current for 30 m, 35 to 400 MeV NS-FFAG lattice.

Figure 8.3: Maximum stable current for 30 m, 35 to 400 MeV NS-FFAG lattice (using linear space charge model).
appears that with a large enough beam pipe any current could be transported. However, in reality the physical size of the magnet apertures is limited.

The transfer matrix is only a linear description of the dynamics, and so while it is useful for assessing some dynamic properties, its stability does not guarantee that the beam really is stable over a large number of turns.

8.1.2 Tune Shift

Above, we found the point at which space charge reduced the tune to zero and made the transfer matrix unstable. We can also look at the amount of tune depression caused as a function of current.

In a synchrotron any tune shift must be kept small to avoid moving the beam onto a resonance which would cause rapid beam loss. This is less of an issue in a rapid-accelerating FFAG. However with rapid acceleration it is not possible to correct the tune by adjusting the quadrupole strengths as can be done in a slower machine.

A KV beam and test bunch were tracked through one cell as before to find the cell transfer matrix. From this, the cell tune was calculated using the method in appendix A.4.

Figure 8.4 shows how the tune is depressed as the current in the beam is increased. We have already found that the stability limit for this lattice at 10 mm mrad is around 10 A. Here the horizontal tune is seen to drop suddenly at around 9 A as the beam becomes unstable.

Figure 8.5 shows the same for a 5 mm mrad beam. The shape is the same but the tune is reduced to zero at a lower current. Again this is consistent
with the stability limits above.

These plots are smooth and free from noise as they only show the effects that take place in the first cell. There is not enough time for complex instabilities to take effect. At low currents the space charge effect is linear.

In a linear NS-FFAG the tune already varies with energy. At higher energies the tune is lower as the beam is less affected by the magnetic fields, and so focusing is reduced. Space charge is suppressed at higher energy, so the tune depression will be less there. This is shown in figure 8.6. As the current is increased the tune shift is lessened. This is due to space charge partially compensating for the tune shift due to optics changes.

There is some complex behaviour around the stability limits. The behaviour of these high current beams is shown later.
Figure 8.5: Tune shift against current for 5 mm mrad beam in 30 m 35 to 400 MeV NS-FFAG lattice.

As the energy increases the lattice becomes stable for higher energies. So where the lattice is not stable at 35 MeV for beams above 10 A, a 13 A beam can be injected at 65 MeV.

Tune flattening may give some leeway, but has its problems. One could imagine a lattice that had very strong focusing at low energy to overcome space charge. However, during commissioning one starts with low current beams, to lower the risk of damage from beam loss. These low current beams would not have sufficient space charge to compensate for the extra focusing, and would be over-focused. The result would be a lattice that was only stable for beams above a critical current. In this case it may be possible to commission with shorter bunches to increase the peak current.
8.1.3 Beam Growth

Space charge can cause the beam to grow in physical size. Firstly it affects the focusing of the lattice and so increases the $\beta$ functions. This effect is constant with current and energy; the $\beta$ functions do not continue to rise for a circulating beam.

Space charge can also affect the beam size by causing emittance growth. This could cause a continued blow up in the beam size, which is best seen by tracking a beam for many turns.

Beam size can be characterised either by emittance or width. Space charge can cause large distortion of the beam, so it is necessary to use a range of measurements. Figure 8.7 shows such a beam that has been tracked
for 240 cells with 10 A of current. This has caused distortion; its shape is no longer an ellipse, the density is no longer uniform and there is a distinct large halo surrounding the core.

Figure 8.7: Beam that has experienced distortion due to strong space charge. Top left shows beam in real space, top right and bottom left show vertical and horizontal phase space

Emittance is a very useful measurement as it remains constant (in the absence of non-conservative forces) throughout a lattice. The width varies through the lattice as the beam is focused in different planes. However, if the width is taken as the same point in each cell then it would stay constant under the same conditions that keep the emittance constant, and so could be considered a proxy for emittance.

The width can be easier to calculate. To find the emittance from a particle distribution one needs to find its area by finding a shape, usually an ellipse,
that encloses the beam. With a beam that is not a regular shape, or that has a halo is not simple. In such cases measuring the area of an ellipse that encompasses the whole beam, 100 % emittance, will be strongly affected by the halo. Equally, an RMS emittance can give an incomplete picture for a non-gaussian beam.

The width can be found just by looking at the distribution along the axis of interest. It is simple to define and calculate a width that includes a certain fraction of particles, for example 68 %, 95 % and 99 %. Methods for measuring these quantities from a beam are described in section A.2.3.

KV beams of 50k particles were tracked in the space charge model as before, for 1000 lattice cells. Their emittance and width were recorded after each cell. The initial beam was matched to the distorted Twiss parameters for currents below the stability limit. For unstable current values the Twiss parameters found at 9 A were used.

In the absence of space charge, when the current is zero, the beam stays a constant size over time as shown in figure 8.8. This is a good indicator that the simulation does not noticeably suffer from the non-symplecticity of the tracking code. There is some noise in the 100 % emittance as it can be affected by the location of a single particle on the edge of the beam.

At a low current, here 1 A, the beam size is still stable after 1000 cells. This is shown in Figure 8.9.

As the current is increased the beam growth becomes an issue. With a 2 A beam the size is stable for around 200 cells, and then grows to around 30 to 40 mm mrad. Figure 8.10 shows the growth of the 2 A beam. Once the instability sets in the growth is rapid over a few tens of cells, and then
Figure 8.8: Beam horizontal emittance in the absence of space charge

Figure 8.9: Beam horizontal emittance with 1 A current and initial emittance of 10 mm mrad.
becomes gentle though noisy for the remainder of the simulation. A similar growth is also seen at 1.5 A.

![Graph showing beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad.](image)

**Figure 8.10:** Beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad.

It is important to verify that the instability is not just an artificial artefact of the simulation. If the simulation is repeated with several different starting random seeds, a similar growth is seen at around the same time. This is shown in figure 8.11.

If the number of particles is increased up to 200k (from 50k) the growth still occurs, at around the same time, as shown in figure 8.12. This reproducibility shows that it is unlikely that the instability is due just to graininess of using a limited number of macro particles to model the beam.

While a significant growth in the beam is observed, the beam is still held in the lattice and the growth is bounded. It is still possible to accelerate the
beam if the beam pipe is made large enough to hold it. It is worth continuing with higher currents to see how the growth changes.

At 3 A of current a similar pattern is seen, but with a faster onset of growth that reaches a larger size. This is shown in figure 8.13.

As the current is increased towards the stability limit the growth in the emittance becomes very rapid. Figure 8.14 shows the emittance growth for a 8 A beam. Previously 9 A was found to be the stability limit for 10 mm mrad. Here, the difference between RMS values and 100 % values is very obvious. The growth of a halo has a large effect on the 100 % emittance. However the RMS emittance is more sensitive to the core of the beam.

In this case, it is informative to look at the fractional widths of the beam as shown in figure 8.15. From this it is possible to see that 95.5 % of the
Figure 8.12: Beam horizontal emittance with 2 A current and initial emittance of 10 mm mrad. With 200k macroparticles.

beam stays within a full width of 15 mm, about 0.3% of the beam is beyond 25 mm and that the full halo stretches out to 30 mm.

Beam growth happens quickly. Most of the growth happens within a few tens of cells or about one lap. Once the beam size is increased the space charge force is reduced and the beam can stabilise. If the growth was slower then it might be suppressed by the fast acceleration. Here, growth is too fast to take advantage of this.

If the current is increased further there is even more growth. Figure 8.16 shows the growth in width with 10 A, beyond the stability limit. This increases the size of the beam and the halo. It is notable that no particles escape far beyond the beam; the halo is still transported by the lattice.

It is interesting to see how the beam growth happens by looking at the
Figure 8.13: Beam horizontal emittance with 3 A current and initial emittance of 10 mm mrad.

beams. Figure 8.17 shows the transverse phase space of the 10 A beam after eight cells, there is already some distortion visible, and the density is less uniform. Figure 8.18 shows the beam after the 13th cell; distortion has grown in the phase space plots.

Figure 8.19 shows the beam after the 29th cell: filaments have formed in phase space which give a halo in real space.

Given the distortion of the beam after strong space charge effects the symmetry assumption of the elliptic rings space charge model breaks down. This limits confidence in later stages of the simulations. However, during the initial stages of stability and growth the beam is still symmetrical so one can be confident about the onset of the growth.
8.1.4 Physical Size Limits

As we have seen, the effect of space charge for a given current, is reduced by having a larger beam. This is due to the charge density being lower. However a larger beam requires a larger beam pipe and magnet apertures. At 10 mm mrad the beam has a maximum physical width of \( \approx 5 \) mm. The beam pipe needs to be larger to give sufficient clearance. An FFAG inherently has a large horizontal physical aperture to accommodate the orbit excursion, but the vertical aperture will be limited so that strong field can be produced.

A continuous accelerator would have tighter constraints on emittance to prevent overlap of bunches on consecutive laps. This would cause complex interactions between the bunches. Also it would make clean injection and
Figure 8.15: Beam horizontal width with 8 A current and initial emittance of 10 mm mrad.

extraction impossible, as there would be no gap for the septum. In a pulsed accelerator a kicker magnet could be used to charge the orbit on the first and final lap to move the beam completely past the septum. A kicker could not be used in a continuous accelerator as it would disturb the other circulating bunches.

The 30 m 35 to 400 MeV NS-FFAG design has an orbit excursion of 550 mm. A 1000 turn continuous accelerator with a similar orbit excursion would need beam widths below 0.5 mm.
Figure 8.16: Beam horizontal width with 10 A current and initial emittance 10 mm mrad.

Figure 8.17: 10 A beam after 8 cells.
Figure 8.18: 10 A beam after 13 cells.

Figure 8.19: 10 A beam after 29 cells.
8.1.5 Conclusion

The limits on peak current put constraints on the possible average current. A 1 A peak current will give 0.3 mA average with a 1:1000 duty factor and $\frac{1}{3}$ filling factor. If the acceleration can be done in 100 turns, then a 1:100 duty factor could be achieved giving 3 mA. Combining three FFAGs of 3 mA gets very close to the required current.

Reducing the number of turns gives other advantages related to the increased orbit separation. It makes clean injection and extraction easier, as there is more separation between the beam and the septum. Also if multiple bunches can be accelerated simultaneously there will be more separation between them.

Rapid acceleration requires more RF acceleration per turn.

Peak currents could be increased by allowing a larger beam size. This would require larger magnet apertures. If the initial emittance is increased to 20 mm mrad then it can stably hold a 2 A peak current, as shown in figure 8.20.

Larger currents could also be allowed by raising the injection energy. Figure 8.21 shows how the perveance reduces as the energy is increased. Changing the injection energy from 35 to 70 MeV reduces perveance by a factor of 2.8, which may allow for an increase in current by a similar amount. This would also increase the speed of the beam at injection, reducing the sweep requirements.

For 100 turn acceleration the RF system would need to provide 3.6 MeV per turn to accelerate from 35 to 400 MeV. A 30 m circumference would have
Figure 8.20: Beam horizontal emittance with 2 A current and initial emittance of 20 mm mrad.

sufficient space for these voltages at low gradients. A more difficult issue is providing the rapid RF sweep. Running in harmonic 1 the RF would need to sweep from 2.65 to 6.55 MHz in 20 µs. A higher harmonic would raise the frequencies, e.g. harmonic 10 would require a 26.5 to 65.5 MHz sweep in the same time. If the machine was made larger there would be more time available to execute the RF sweep.
Figure 8.21: Generalised perveance as a function of beam energy for proton beams at 1, 2 and 10 A.
Chapter 9

Conclusion

Nuclear power is and will continue to be an essential part of the UK electricity supply. Although nuclear power compares favourably to other energy sources in terms of cost and safety there is still widespread opposition to it. ADSRs offer several advantages over conventional nuclear reactors: increased safety due to subcritical operation; ability to use the more abundant thorium fuel; less long lived waste; and greater resistance to nuclear proliferation.

However, an ADSR requires a more powerful proton beam than any currently in existence. A reactor with an electrical output of 600 MW and a $k_{\text{eff}}$ of 0.985 needs a 10 MW beam of 1 GeV or higher. There are some high power linacs that approach the requirements for current, although they have low repetition which would be unsuitable for an ADSR. Cyclotrons are unsuitable for an ADSR as their energy reach is limited by relativistic effects. Synchrotrons are unsuitable due to their inherently low repetition rate. Also, advances in reliability need to be made.
An FFAG accelerator offers a solution. It is much more compact than a linac, and can reach high energies with high repetition rates.

FFAGs are typically non-isochronous and have a large change in revolution frequency with energy in the region needed for an ADSR. Although the magnets do not need to be ramped in an FFAG, if there is a change in revolution frequency then the RF will need to be swept. This means that an FFAG will be a pulsed machine, with injection, acceleration and then extraction. This gives a low duty cycle, and hence the peak currents are much higher than the average current.

A lattice that contains the features typical in an NS-FFAG, resonance crossing and large revolution frequency shift, was chosen for simulation work. Resonance crossing is less of a problem during rapid acceleration, but designs are published that flatten the tunes to remove resonance crossing.

Two methods for injecting more bunches during the acceleration cycle were investigated. By taking advantage of multiple solutions to the harmonic condition it is possible to accelerate multiple bunches with multiple RF frequencies. This interlacing puts strong constraints on the harmonics that are used. To accelerate more than a small number of bunches requires impractically high harmonic numbers.

It is possible to inject bunches in the region around the correct RF frequency and for them to still be accelerated. This allows several bunches to be injected while the RF is already sweeping. However, these bunches will be very close in space or even overlapping during acceleration. This means that from a space charge point of view they will behave as a larger bunch.

Space charge is a strong and destructive force at the energies and inten-
sities required. I have developed extensions to the Zgoubi particle tracking software to allow modelling of space charge.

For reasonable bunch sizes at 35 MeV peak currents are limited to 1 A. In order to achieve the required average currents, duty cycles of around 1:100 would be needed which corresponds to high acceleration rates, and 100 lap acceleration. This requires an RF system that can sweep more than a factor of two in frequency in tens of microseconds. These demands could be lessened by increasing the physical dimensions of the bunch which would require larger magnet apertures, or by increasing the injection energies.

The ability of RF systems to sweep rapidly will be one of the main limiting factors for a high current FFAG.

Switching to an isochronous design allows continuous acceleration and so reduces peak currents. There are some designs for isochronous FFAGs, but these require large horizontal magnet apertures to allow the orbit radius to increase. In these designs the distinction between an FFAG and a cyclotron are blurred.

It is currently an exciting time for FFAG research. The commissioning of EMMA will give many interesting results such as the effects of slowly crossing resonances. These results will affect future NS-FFAG designs. There are many possible variations in FFAG design that can be explored.
Appendix A

Methods and Definitions

A.1 Closed Orbit Search

In an FFAG the closed orbit does not necessarily fall on the centre of the magnets, and it changes as a function of energy. Therefore, it is usually the first step in any simulation to find the closed orbit.

A particle’s transverse coordinates are transformed as it is tracked through an accelerator lattice. An initial point in phase space is mapped to a new point. A well-behaved lattice will have a fixed point in phase space, where a particle returns to the same point in phase space after a lattice cell, or full orbit. This fixed point is the closed orbit. Around the fixed point is a stable region. Particles in the stable region will oscillate around the fixed point.

In a perfect linear system the oscillation will be an ellipse in phase space, with the fixed point at the centre. In a non-linear system the ellipse will be distorted, but the fixed point will be near the centre.
To find the fixed point one first needs to find a stable orbit. This can either be an arbitrary start point \((x = 0; x' = 0\) is often stable), a guess based on knowledge of the lattice, or a previously found point (e.g. the closed orbit for a similar configuration). In difficult cases a large number of initial coordinates can be tried until a stable one is found.

This stable coordinate can now be tracked for several turns. The coordinates after each turn will draw a path in phase space. If this initial value was within the linear region of the system and enough turns were made, the path will be an ellipse and the centre will be the fixed point. If not, then taking the mean of the coordinates will give a new point that is closer to the fixed point than the initial coordinates. This method can be iterated until it converges.

A.2 Transverse Beam Parameters

A.2.1 Definitions

It is convenient to describe a beam by an ellipse that encloses the beam in phase space. This is parametrised by \(\alpha, \beta, \gamma\), known as the Twiss parameters or Courant-Snyder parameters, and the emittance, \(\epsilon\). These describe the ellipse, as shown in figure A.1.

In an envelope simulation code, these parameters themselves are propagated through the lattice, but in a particle tracking code we must find them from the beam.
A.2.2 100 % Twiss Parameters

The 100 % Twiss parameters are those of an ellipse in phase space that fully encloses the beam. In the case of an elliptical phase space with a hard edge, this ellipse is well defined. First one must find the emittance, $\epsilon$. The area of an ellipse, $A$ is,

$$A = ab\pi$$  \hspace{1cm} (A.1)

where $a$ and $b$ are the semi-major axes. To find $a$ and $b$ the ellipse must first be rotated to align it with the x-axis. This can be done by converting the coordinates of the particles to polar, and finding the particle with the greatest $r$ value. Its $\theta$ can then be used to rotate the coordinates. Back in Cartesian coordinates the width and height of the phase space can be measured. These are $a$ and $b$. 

\hspace{1cm} Figure A.1: Phase space ellipse, showing Twiss parameters $\alpha$, $\beta$, $\gamma$ and emittance. [121]
On the un-rotated phase space, the points with maximum and minimum values in $x$ and $x'$ can be then found. From these points $\alpha$, $\beta$ and $\gamma$ can be found with the formulae in figure A.1.

This method is vulnerable to noise. Averages of the parameters from both sides of the ellipse can be used to reduce this a small amount. As only the particles on the edge of the ellipse are considered using a distribution with just the halo rather than a filled ellipse, can save simulation time.

### A.2.3 RMS Twiss Parameters

For more realistic beams the root mean squared (RMS) properties can be more useful.

The beam is first translated so that its mean in each coordinate is on the origin. The RMS emittance is given by,

$$\tilde{\epsilon}_x = \sqrt{\overline{x^2} \overline{x'^2} - \overline{xx'}}$$  \hspace{1cm} (A.2)

where $\overline{x}$ is the arithmetic mean of the x coordinates.

Similarly, the RMS values can replace the widths in the formulas for the Twiss parameters to give,

$$\tilde{\beta}_x = \frac{\overline{x^2}}{\tilde{\epsilon}_x}$$  \hspace{1cm} (A.3)

and

$$\tilde{\alpha}_x = -\frac{\overline{xx'}}{\tilde{\epsilon}_x}. \hspace{1cm} (A.4)$$
A.3 Finding the Transfer Matrix with a Tracking Code

The transfer matrix is a first order mapping as described in section 5.1. It can also be used to calculate various properties of the lattice. It is usually calculated analytically from the lattice and then used to track particles.

It is also possible to calculate the transfer matrix by tracking a prepared set of particles with any tracking code. Each term in the matrix gives the dependence of one coordinate on another coordinate. It is the Jacobian matrix of the transfer.

\[
R = \begin{bmatrix}
\frac{\partial Z(1)}{\partial Z(0)} & \cdots & \frac{\partial Z(1)}{\partial Z(0)} \\
\frac{\partial Z(0)}{\partial Z(0)} & \ddots & \vdots \\
\vdots & \ddots & \vdots \\
\frac{\partial Z(1)}{\partial Z(0)} & \cdots & \frac{\partial Z(1)}{\partial Z(0)}
\end{bmatrix} \tag{A.5}
\]

One can take two particles on the reference orbit and displace one up and one down in \( x \) by a small \( \delta x \). Then these particles can be tracked through a lattice. At the end of the tracking their displacement in each coordinate, divided by \( \delta x \) is the corresponding term in the transfer matrix.

Zgoubi has a built-in method to do this. One can make a special bunch using the KOBJ=5 option to the OBJET bunch generator, and get the transfer matrix MATRIX element. This creates a bunch with 11 particles; a central one on the reference orbit, and five pairs with a positive and negative displacement in the \( x, x', y, y' \) and momentum coordinates respectively. In a case without RF this is sufficient to calculate the matrix.
In a non-linear system the size of the initial displacements will affect the final matrix.

A.4 Lattice Parameters from the Transfer Matrix

With the transfer matrix for a periodic lattice, or for a cell that makes up one period of the lattice, various parameters can be found. From the solutions to Hill’s equation, one can write the matrix for each of the transverse planes as [69]:

\[ R = \begin{bmatrix} \cos \mu + \alpha \sin \mu & \beta \sin \mu \\ -\gamma \sin \mu & \cos \mu - \alpha \sin \mu \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \] (A.6)

One can then solve for \( \alpha, \beta, \gamma \) at the start of the periodic section and the phase advance \( \mu \) over the periodic section [69].

\[ \mu = \arccos \left( \frac{a + d}{2} \right) \] (A.7)

\[ \alpha = \frac{a - d}{2 \sin \mu} \] (A.8)

\[ \beta = \frac{b}{\sin \mu} \] (A.9)

\[ \gamma = -\frac{c}{\sin \mu}. \] (A.10)

It is worth noting that this is only valid for a whole periodic cell; it cannot, for example, be used to find the phase advance through part of a lattice.
A.5 Space Charge Parameters

There are a number of parameters and factors that are frequently used in literature on space charge. They are collected and defined here for convenience.

The classical particle radius, \( r_c \), is defined by equating the rest energy \( mc^2 \) and the potential energy \( q^2 / 4 \pi \varepsilon_0 r_c \) of a particle,

\[
\rho_c = \frac{q^2}{4 \pi \varepsilon_0 m_0 c^2}.
\] (A.11)

It is \( 2.8180 \times 10^{-15} \) m for electrons and \( 1.5347 \times 10^{-18} \) m for protons [68].

The characteristic current, \( I_0 \), is used in several of the definitions, and is defined as

\[
I_0 = \frac{4 \pi \varepsilon_0 m_0 c^3}{q}.
\] (A.12)

It is approximately 17 kA for electrons and 31 MA for protons [68].

In an accelerator, a useful definition of beam current is,

\[
I = \rho_L \beta c
\] (A.13)

where \( \rho_L \) is the charge density per unit of beam length.

The Budker parameter is defined as [68],

\[
\nu_B = \frac{\rho_L r_c}{q} = \frac{I}{I_0 \beta}.
\] (A.14)
The generalised perveance is defined as [68],

\[
K = \frac{I}{I_0} \frac{2}{\beta^3 \gamma^3} = \frac{2\nu_B}{\beta^2 \gamma^3}. \tag{A.15}
\]

It is a dimensionless quantity that measures how great the effect of space charge is at a given current and energy.
Appendix B

2D Multipole Expansion

The transverse field around a long bunch can be considered as the field around an infinitely long wire charged or equivalently a charge distribution in 2D. The electric field at a radius $R$ around a point in 2D is,

$$E_r = \frac{q}{2\pi\epsilon_0} \frac{1}{R}$$  \hspace{0.5cm} (B.1)

where $q$ is the charge per unit length. The electric potential is

$$\phi = -\frac{q}{2\pi\epsilon_0} \log(R) + c$$  \hspace{0.5cm} (B.2)

where $c$ depends on an arbitrary choice of where the potential is zero, and will be dropped.

For a collection of charges $q_i$ at position $r_i$ the potential is given by

$$\phi = -\sum \frac{q_i}{2\pi\epsilon_0} \log(|\mathbf{R} - \mathbf{r}_i|)$$  \hspace{0.5cm} (B.3)
or in polar coordinates, with $\theta_i$ as the polar angle for the $i^{th}$ charge,

$$
\phi = - \sum_i \frac{q_i}{2\pi \epsilon_0} \log \left( \sqrt{R^2 + r_i^2 - 2Rr_i \cos(\theta_i)} \right)
$$

$$
= - \sum_i \frac{q_i}{2\pi \epsilon_0} \log \left( R \sqrt{1 + \frac{r_i^2}{R^2} - \frac{2r_i \cos(\theta_i)}{R}} \right)
$$

$$
= - \sum_i \frac{q_i}{2\pi \epsilon_0} \left( \log(R) + \frac{1}{2} \log \left( 1 + \frac{r_i^2}{R^2} - \frac{2r_i \cos(\theta_i)}{R} \right) \right). \quad (B.4)
$$

Using the expansion

$$
\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \quad (B.5)
$$

with

$$
x = \left( \frac{r_i^2}{R^2} - \frac{2r_i \cos(\theta_i)}{R} \right) \quad (B.6)
$$

gives

$$
\phi = - \sum_i \frac{q_i}{2\pi \epsilon_0} \left( \log(R) - \frac{r_i \cos(\theta_i)}{R} - \frac{r_i^2 \cos(2\theta_i)}{2R^2} - \frac{r_i^3 \cos(3\theta_i)}{3R^3} - \frac{r_i^4 \cos(4\theta_i)}{4R^4} - \frac{r_i^5 \cos(5\theta_i)}{5R^5} - \frac{r_i^6 \cos(6\theta_i)}{6R^6} + \ldots \right) \quad (B.7)
$$

where we can identify the $\log(R)$ term as the monopole, the $1/R$ term as the dipole, the $1/R^2$ as the quadrupole, and so on. The potential can then be written with a set of coefficients, $C_n$, for the contributions from each

172
multipole,

$$\phi = \frac{q}{2\pi\epsilon_0} \left( -C_0 \log(R) + \frac{C_1 \cos(\theta)}{R} + \frac{C_2 \cos(2\theta)}{R^2} \right.$$ 
$$\left. + \frac{C_3 \cos(3\theta)}{R^3} + \frac{C_4 \cos(4\theta)}{R^4} + \frac{C_5 \cos(5\theta)}{R^5} + \frac{C_6 \cos(6\theta)}{R^6} + \ldots \right) \quad \text{(B.8)}$$

where \(q\) is the total charge.

The coefficients can be found by integrating the charge distribution, and the form of the multipole. For a uniform ellipse of charge we take the number density \(\rho = \frac{1}{\pi ab}\) inside the ellipse and 0 outside. The boundary of an ellipse with a semi-major axis, \(a\), and a semi-minor axis, \(b\), as shown in B.1, is in Cartesian coordinates,

$$y = \pm b \sqrt{1 - \frac{x^2}{a^2}} \quad \text{(B.9)}$$

and in polar coordinates,

$$r = \frac{ab}{\sqrt{b^2 \cos^2(\theta) + a^2 \sin^2(\theta)}}. \quad \text{(B.10)}$$

Figure B.1: Semi-major and semi-minor axes of the elliptical cross section.
For example the monopole, which must be equal to 1, can be found with,

\[
C_0 = \int_{\theta=0}^{\pi} \int_{r=0}^{a b} \rho(r, \theta) r \, dr \, d\theta \tag{B.11}
\]

or

\[
C_0 = \int_{x=-a}^{a} \int_{y=-b \sqrt{1 - \frac{x^2}{a^2}}}^{b \sqrt{1 - \frac{x^2}{a^2}}} \rho(x, y) \, dy \, dx \tag{B.12}
\]

and the dipole, which must be equal to 0, with,

\[
C_1 = \int_{\theta=0}^{\pi} \int_{r=0}^{a b} \rho(r, \theta) \cos(\theta) r^2 \, dr \, d\theta \tag{B.13}
\]

or

\[
C_1 = \int_{x=-a}^{a} \int_{y=-b \sqrt{1 - \frac{x^2}{a^2}}}^{b \sqrt{1 - \frac{x^2}{a^2}}} \rho(x, y) \cos(\tan^{-1}(y/x)) \sqrt{x^2 + y^2} \, dy \, dx \tag{B.14}
\]

and so on for high terms.

Using Mathematica [122] the coefficients can be found, the first few are:

\[
C_0 = 1 \tag{B.15}
\]

\[
C_1 = 0
\]

\[
C_2 = \frac{1}{4} (b^2 - a^2)
\]

\[
C_3 = 0
\]

\[
C_4 = \frac{1}{8} (b^2 - a^2)^2
\]
$C_0$ is just the area of the ellipse times the number density, this makes sense as the monopole should only depend on the total charge of the bunch. The dipole, $C_1$, would represent an offset of the bunch so it is expected to be zero, likewise with the other odd terms. The quadrupole and octopole, $C_2$ and $C_4$, depend on the shape, and vanish for a circle, when $a = b$.

In general

$$C_n = \begin{cases} \frac{A_n}{2} (b^2 - a^2)^{\frac{n}{2}} & \text{if } n = 0, 2, 4 \ldots \\ 0 & \text{if } n = 1, 2, 3 \ldots \end{cases}$$ \hspace{1cm} (B.16)

where $A_n$ are the sequence of coefficients in the expansion of $\sqrt{1-x}$ multiplied by 2,

$$A = \left[ \frac{1}{4}, \frac{1}{8}, \frac{5}{64}, \frac{7}{128}, \frac{21}{512}, \frac{33}{1024}, \frac{429}{16384}, \frac{715}{32768}, \frac{2431}{131072} \right].$$ \hspace{1cm} (B.17)

Hence the full expansion of the potential is:

$$\phi = \frac{q}{2\pi\epsilon_0} \left( -\log(R) + \frac{1}{4} \frac{(a^2 - b^2) \cos(2\theta)}{R^2} \right) + \frac{1}{8} \frac{(a^2 - b^2)^2 \cos(4\theta)}{R^4} + \frac{5}{64} \frac{(a^2 - b^2)^3 \cos(6\theta)}{R^6} + \ldots \right).$$ \hspace{1cm} (B.18)

The electric field is

$$E = -\nabla \phi.$$ \hspace{1cm} (B.19)
The field around an ellipse is found to be,

\[ E_x = \frac{q}{2\pi \varepsilon_0} \left( \frac{x}{R^2} + \frac{1}{4} \frac{(a^2 - b^2)}{R^4} (x \cos(2\theta) - y \sin(2\theta)) \right. \]
\[ + \frac{1}{8} \frac{(a^2 - b^2)^2}{R^6} (x \cos(4\theta) - y \sin(4\theta)) \]
\[ + \frac{5}{64} \frac{(a^2 - b^2)^3}{R^8} (x \cos(6\theta) - y \sin(6\theta)) + \ldots \right). \]

A multipole expansion is valid only in regions where the terms form a convergent series. When \( R \) is large only the coefficients fall rapidly and only a few terms are needed. Near the surface of an ellipse \( R \) is comparable in size to \( a \) and \( b \), the series does not converge for \( R < \sqrt{|a^2 - b^2|} \). In practice there is a region close to the ellipse on the broad side where the truncated series gives a poor result. This is an important region for space charge calculations.

Figure B.2 shows how the multipole expansion converges well up to the edge of the ellipse at \( \theta = 0 \). Figure B.3 shows how the multipole expansion does not converge close to the ellipse at \( \theta = \pi/5 \).

Figure B.4 shows the field in 2D space around the ellipse. On the broad sides of the ellipse the non-convergent expansion gives errors up to 60% of the maximum field.

## B.1 Correction to the Multipole Expansion

The multipole formulation is valuable as it gives a very accurate field model at large radius. The field inside the ellipse is given by the simple expression
Figure B.2: Electric field along $R$ at $\theta = 0$ around an ellipse with $a = 2$ and $b = 1$. Black line shows numerical integration. Coloured lines show multipole expansion to increasing order. Grey dashed line is the edge of the ellipse.

Figure B.3: Electric field along $R$ at $\theta = \pi/5$ around an ellipse with $a = 2$ and $b = 1$. Black line shows numerical integration. Coloured lines show multipole expansion to increasing order. Grey dashed line is the edge of the ellipse.
in equation 3.13,

\[ E_x = \frac{\rho ab}{\epsilon_0 (a + b)} \frac{x}{a}. \]

But for some values of theta there is a region between where the multipole expansion is a poor model.

As the field is known at the surface of the ellipse it is possible to construct a correction term to adjust the field given by the multipole model. This adjustment must fall off such that at large \( R \) the field is given by the multipole expansion. Also, the field must be continuous at the surface of the ellipse. If the multipole field is evaluated at a point on the surface and compared to
Figure B.5: Electric field along $R$ at $\theta = 0$ around an ellipse with $a = 2$ and $b = 1$. It shows numerical integration (black), quadrupole expansion (blue) and adjusted quadrupole expansion (red). Grey dashed line is the edge of the ellipse.

linear field at that point then a $\theta$ dependent difference $\Delta E_x$ can be found. The term,

$$\frac{\Delta E_x}{\left(\sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2}}\right)^4}$$  

was found to give a good field when the multipole was calculated up to quadrupole, it falls off approximately with $1/r^4$. Higher multipole terms give larger errors on the broad side, which the correction term is less able to correct.

Figure B.5 shows how the correction improves a quadrupole field at $\theta = 0$. Here the correction is effectively replacing the higher multipole terms. Figure B.6 shows how the correction improves a quadrupole field at $\theta = \pi/5$. Here the correction is quite large and much more effective than adding further
Figure B.6: Electric field along $R$ at $\theta = \pi/5$ around an ellipse with $a = 2$ and $b = 1$. It shows numerical integration (black), quadrupole expansion (blue) and adjusted quadrupole expansion (red). Grey dashed line is the edge of the ellipse.

multipole terms as show above in B.3. Inside the ellipse, on the left of the grey dashed line, the linear formula is used.

Figure B.7 shows the field in 2D space around the ellipse. The error, as compared to the numerical integration, along the broad side has been reduced considerably. The largest difference is now around 2%. Note scale difference compared to figure B.4.

An alternative method would be to derive the electric field in ellipsoidal coordinates using the integrals presented in [104].
Numerical integration

Mulpole up to quadrupole with adjustment

Difference

Figure B.7: $E_x$ around an ellipse with $a = 2$ and $b = 1$, as calculated with numerical integration, multipole up to quadrupole with correction, and the difference between them. All scaled to the maximum field.
Appendix C

PyZgoubi

PyZgoubi [100] is an interface to Zgoubi [98] that I have written in order to make it more practical and easier to use.

C.1 Motivation

Zgoubi is a particle tracking code, which is widely used for FFAGS and has been used for most of the simulation work in this thesis. Its ray tracing method is described in section 5.2.

Although it is a powerful code it has several limitations in its interface that make it difficult to use.

Input files have a rigid format: the element keyword must be followed by a list of numbers that specify the parameters. These parameters must be given in a specific order, as given in the Zgoubi manual [99]. There are no parameter names in the input file.
Parameters use a range of different units; for example metres, centimetres, degrees and radians are all used in various places.

Zgoubi will only read and write to specific file names for its input (zgoubi.dat) and most of its output files (zgoubi.fai, zgoubi.plt, zgoubi.res), as shown in figure C.1.

$ zgoubi

$ zgoubi

Figure C.1: Zgoubi reads a zgoubi.dat file, and outputs several files.

It is not possible to do calculations in the input file. For example, you cannot define an angle as $2\pi/42$, instead you must put 0.14959965017094254 into the input file.

Looping is limited to rerunning the entire line. One cannot define a repeated cell, or reuse a magnet definition without explicitly repeating them in the input file.

I am aware of several users who have written custom scripts to control Zgoubi using C++ and MATLAB to get around these issues. PyZgoubi grew out of my own scripts written in Python and now has several users.
C.2 Implementation

PyZgoubi is written in Python [123] and makes use of the extension libraries NumPy [124], SciPy [105] and MatPlotLib [125]. Python is a dynamic interpreted programming language. This provides several advantages compared to C/C++ or FORTRAN: programs do not need to be compiled; data types are handled automatically and flexibly; there is bound checking on all arrays; and there is no risk of common errors such as using uninitialised values and memory leaks. NumPy and SciPy provide fast routines for arrays and numerical calculations.

Rather than writing an input file, one writes a python script. PyZgoubi is a python module that provides several classes that let one build a lattice, run Zgoubi and analyse the results. One can either load this module, or use the `pyzgoubi` command that loads the required modules, does some set up, and runs the python script. This is shown in figure C.2.

When PyZgoubi runs Zgoubi it does so by first creating a temporary folder, usually in `/tmp/` or `/dev/shm/`. Into this it writes a `zgoubi.dat` file. Then it runs Zgoubi and checks that it returns without error. A `Results` object is then created which can be used to access any of Zgoubi’s output files.

Zgoubi can be run many times from a single input file, for example to allow feedback from one run to another.
Figure C.2: PyZgoubi is controlled by an input file. It then calls Zgoubi, and collects the results for further processing.

C.2.1 Structure

The Line object holds a list of elements. It can also hold other Line objects to simplify the description of a lattice that contains repeated sections. It has several methods for manipulating the lattice, by adding, removing and swapping elements. It has a Run() method that performs the setting up of a temporary folder, running Zgoubi and returns the Results object.

Zgoubi has many magnetic, electrostatic and other beam line elements. It also has some special elements; for example, to create beams, define particle properties and add markers. In PyZgoubi these map to classes which are derived from a zgoubi_element class. Most of these classes are defined in a
.defs file, which contains a list of the required parameters.

For example, a sector bend magnet has 29 parameters which are defined in the Zgoubi manual, as shown in figure C.3.

A bend with a length of 1 m, skew angle of 0.1 radians, field of 50 kG and no fringe fields, would be entered into a zgoubi.dat as follows:

'BEND'
2
100  0.1  50
0  0  0
0  0  0  0  0  0
0  0  0
0  0  0  0  0  0
0.1
0  0  0  0

In PyZgoubi there is a BEND object that is built from the following definition:

BEND
IL : I
XL, Sk, B1 : 3E
X_E, LAM_E, W_E : 3E
N, C_0, C_1, C_2, C_3, C_4, C_5 : I,6E
X_S, LAM_S, W_S : 3E
NS, CS_0, CS_1, CS_2, CS_3, CS_4, CS_5 : I,6E
XPAS: X
KPOS, XCE, YCE, ALE : I,3E

Each line shows the names of the parameters, and the data types (integer or real).

In your input file, you would create the equivalent with the following:

b1 = BEND(IL=2,
    XL=100, Sk=0.1, B1=50,
    XPAS=0.1)

Undefined values are set to zero. Also variables and calculations can be used, for example for unit conversion:
**BEND**  
**Bending magnet, Cartesian frame**

**IL**  
$I$L = 1, 2 : print field and coordinates along trajectories (otherwise $IL = 0$)

**$XL$, $Sk$, $B1$**  
Length ; skew angle ; field cm, rad, kG

**Entrance face** :

**$XE$, $λE$, $WE$**  
Integration zone extent ; fringe field extent (normally $≃$ gap height ; zero for sharp edge) ; wedge angle cm, cm, rad

**$N$, $C0$–$C5$**  
Unused ; fringe field coefficients : $B(s) = B1 F(s)$ with $F(s) = 1/(1 + \exp(P(s)))$ and $P(s) = \sum_{i=0}^{5} C_i (s/\lambda)^i$ unused, 6*no dim.

**Exit face** :

**$XS$, $λS$, $WS$**  
See entrance face cm, cm, rad

**$N$, $C0$–$C5$**  
unused, 6*no dim.

**$XPAS$**  
Integration step cm

**$KPOS$, $XCE$, $YCE$, $ALE$**  
$KPOS=1$ : element aligned, 2 : misaligned ; shifts, tilt (unused if $KPOS=1$) 1-2, 2*cm, rad

$KPOS = 3$ :
entrance and exit frames are shifted by $YCE$
and tilted wrt. the magnet by an angle of
• either ALE if ALE≠0
• or $2 \arcsin(\frac{BRX}{2BO})$ if ALE=0

Geometry and parameters of BEND in its Cartesian frame : $XL$ = length, $θ$ = deviation, $WE$, $WS$ are the entrance and exit wedge angles.

**Figure C.3: Zgoubi bend definition [99].**
\begin{verbatim}
b1_len = 1 * m
b1 = BEND(IL=2,
    XL=b1_len*_cm, Sk=radians(6), B1=50,
    XPAS=0.1)
\end{verbatim}

This allows better readability and flexibility. The full Python language is available so loops and branching can be used to build a complex line; data can be loaded from files or databases; and other programs can interfaced to. For example, PyZgoubi has been interfaced to the EPICS [126] control system as part of the on-line software for EMMA [127].

Once a line has been built by adding several elements, it can be run. If this is successful then a \texttt{Results} object will be returned. This has methods to directly access or save the standard Zgoubi output files. Often it is more useful to use the \texttt{Results} object to read those files, and return the data as a NumPy array. This can then be analysed or plotted.

The results of one run can be used to create a new lattice, or modify the lattice before running it again. This can be used to optimise a lattice.

\section*{C.2.2 Other features}

PyZgoubi also contains a set of utility functions in its \texttt{utils} sub-module. These include a closed orbit finder that iteratively tracks a particle through a lattice, finds the centre of its ellipse in phase space, and uses that as a new initial coordinate. It can start the search with a large grid of particles over phase space, to find a stable orbit.

There are also functions to calculate Twiss parameters and transfer matrices, a function to measure tune using a fast Fourier transform, a function
to misalign elements, and a few functions for plotting.

There is a Bunch class that can hold a bunch of particle coordinates. It has methods for generating various useful particle distributions, e.g. Gaussian, KV or water-bag. It also has methods for measuring bunch parameters such as widths, emittance and Twiss values.

Bunches can be loaded by Zgoubi using Zgoubi’s OBJET mode 3, accessed through PyZgoubi’s OBJET_bunch element.

PyZgoubi can split a large bunch, run several instances of Zgoubi from separate threads and then combine the bunches back together. This is currently limited to cases where you only need the final bunch, and not any other analysis with Zgoubi, or tracks. There is some overhead to this, but in cases where Zgoubi would take a long time to run the speed up is significant.

C.3 Availability

PyZgoubi is open source, and available from the SourceForge.net website. The source code is stored in the Bazaar distributed revision control system [128]. Packages are available and work on Linux, Mac OS X and Microsoft Windows.
Bibliography


