LHC Beam Stability and Feedback Control

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von

Diplom-Physiker Ralph Steinhagen
aus Bonn

Berichter: Universitätsprofessor Dr. A. Böhm
Universitätsprofessor Dr. T. Hebbeker

Tag der mündlichen Prüfung: 20 Juli 2007

Why lovest thou so this brittle world's joy?
Take all the mirth, take all the fantasies,
Take every game, take every wanton toy,
Take every sport that men can thee devise:
And among them all on warrantise
Thou shalt no pleasure comparable find
To th' inward gladness of a virtuous mind.

Thomas More

Zusammenfassung

Die vorliegende Arbeit befaßt sich mit der Strahlstabilität und der Kontrolle der beiden Protonenstrahlen des LHC Beschleunigers (*Large Hadron Collider*), der sich am CERN im Bau befindet und Ende 2007 in in Betrieb genommen wird. Der Schwerpunkt dieser Arbeit liegt in der Analyse möglicher dynamischer Störungen der mittleren Strahlposition (dem sog. *Orbit*), der Strahlenergie und deren Kontrolle durch rückgekoppelte Regelungssysteme (engl. *feedbacks*). Das Ziel dieser Arbeit ist zu einem guten Start und zuverlässigen Betrieb des LHCs beizutragen.

Im LHC werden zwei Protonenstrahlen auf eine Schwerpunktsenergie von 14 TeV beschleunigt und Protonen-Protonen Kollisionen mit einer nominalen Luminosität von $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ den Experimenten zur Verfügung gestellt. Die Verwendung zweier Strahlen mit sowohl hoher gespeicherter Intensität als auch hoher Energie in einer supraleitenden Umgebung bedarf einer ausgezeichneten Kontrolle der Strahlverlusten, die durch die LHC Collimations und Schutzsysteme gewährleistet werden wird. Die Leistung dieser Systeme hängt kritisch von der Strahlstabilität ab, welche unter Umständen die maximal möglichen LHC Kenngrößen limitieren kann.

Umgebungs und beschleuniger-spezifische Quellen können die Strahlstabilität beeinträchtigen und zur Zunahme von Strahlverlusten in die supraleitenden Apertur führen. Da manuelle Korrekturen auf Grund der benötigten Präzision und erwarteten Zeitskalen nur begrenzt möglich sind, werden automatische Regelungssysteme für die Kontrolle von Orbit, Tune, Chromatizität, Energie und Kopplung daher ein wesentlicher Bestandteil des LHC Betriebes sein.

Der erste Teil dieser Arbeit befaßt sich mit der Abschätzung von Störungen von Orbit und Änderungen des mittleren Strahlimpulses: Die langsamsten Orbitänderungen sind durch Bodenbewegungen, Gezeiten, Temperaturänderungen der Tunnelumgebung und andere Umgebungseinflüsse dominiert: Basierend auf Vergleichsmessungen am Elektron-Positron Ring LEP und weiteren Beschleunigern am CERN, kann der Einfluß von un-korrelierten Bodenbewegungen auf die Stabilität des Orbits über einen Zeitraum von 10 Stunden auf $200-300\,\mu$m abgeschätzt werden. Sonne- und Mondtiden variieren den relativen Strahlimpuls $\Delta p/p$ zwischen $-0.9\ldots0.5\cdot10^{-4}$ auf einer Zeitskala von etwa 6 Stunden. Die Tiden führen zusätzlich zu einer $-170\ldots+100\,\mu$m Änderung des Orbit in Regionen mit Werten der Dispersionsfunktion von etwa 2 m.

Die schnellsten und größten Orbitänderungen werden während der Beschleunigung zur Enerieerhöhung der Protonenstrahlen (sog. 'Rampe') und der Verkleinerung der transversalen Strahlgröße auf etwa $15\,\mu$m in den Interaktionspunkten (dem sog. *beta-squeeze*) erwartet. Während der Rampe können Orbitänderungen mit bis zu $600-700\,\mu$m und Driftgeschwindigkeiten mit $15\,\mu$m/s auftreten. Abhängig von der Präzision der Magnetausrichtung in der Nähe der Interaktionspunkte, können die Orbitänderungen während des *beta-squeeze*'s bis zu 30 mm und Änderungsgeschwindigkeiten bis zu $25\,\mu$m/s erreichen.

Die Änderungen der Strahlposition können um eine Größenordnung die benötigte Strahlstabilität übersteigen: Das *LHC Cleaning System* hat zum Beispiel eine der striktesten Anforderung, und bedarf eine Strahlstabilität, die besser als etwa $15-25\,\mu$m ist. Die LHC Schutzsysteme und andere Systeme haben Anforderungen mit ähnlichen Größenordnungen. Die Anzahl und die örtliche Verteilung der einzelnen Anforderungen im LHC legt es nahe den Orbit nicht nur lokal sondern in dem ganzen Beschleuniger auf dem Niveau von wenigen zehn Mikrometern zu kontrollieren.

Der zweite Teil dieser Arbeit befaßt sich mit der Kontrolle des Orbits und des mittleren Strahlimpulses durch rückgekoppelte Regelungssysteme. Die erreichbare Stabilität dieser Systeme wird im Wesentlichen durch die Qualität und Ausfallwahrscheinlichkeit der 1056 Strahlpositionsmonitore, der mehr als 1060 Dipol-Korrektorkreise und der Auslegung des Steuersystems selbst gegeben.

Die Strahlpositionsmonitore basieren auf dem sogenannten *wide-band time normaliser* Prinzip, einem analogen Schaltkreis, der die transversale Position durch den Zeitabstand zweier Pulse dargestellt und diese über optische Fasern zu einer digitalen Akquisitionskarte übermittelt welche den Orbit durch Mittelung der einzelnen Protonenpaket-Positionen über mehrere Maschinenumläufe.

Es zeigt sich, daß die Positionsauflösung von der Anzahl der Protonen pro Paket und der Paketlänge abhängt. Für die im SPS mit nominellem Strahl getesteten Monitore konnte gezeigt werden, daß die Auflösung pro Strahlpaket und Umlauf etwa 60 μm beträgt, was unter der Annahme eines einzelnen zirkulierenden Strahlpaketes im LHC einer Orbitauflösung von etwa 5 μm entspricht. Ein einfaches Fehlermodel wurde mit der Meßgenauigkeit der LHC Testmonitore im CERN SPS Beschleuniger verglichen. Im Rahmen der spezifizierten Strahlparameter beträgt dabei die (bezüglich der Monitor Halb-Appertur normalisierte) Meßgenauigkeit etwa 1%. Die Auflösung als auch die Meß-Ungenauigkeiten wurden experimentell mit strahlunabhängigen Messungen überprüft und sind in Übereinstimmung mit der Monitor Spezifikation.

Die mittlere Strahlposition wird durch 1060 supraleitende Korrektur-Dipolmagnete gesteuert. Jeder dieser Magnete wird durch jeweils einen unhabhängigen Stromkreislauf gespeist. Die untere Grenze der Strahlstabilität ist durch die gemessene relative Stabilität der Stromversorgung von $5 \cdot 10^{-6}$ gegeben. Diese entspricht einer mittleren Strahlstabilität von $(6 \pm 2)\,\mu$m bei einem Strahlimpuls von 450 GeV/c und verbessert sich wegen der Energieabhängigkeit der Magnetfelder auf auf etwa $(0.4 \pm 0.1)\,\mu$m für einen Strahlimpuls von 7 TeV/c. Die maximalen Stromänderungen sind durch die große Induktanz der Magnete (6 H) und Spannungsbegrenzungen der Stromversorgung auf weniger als 0.5 A/s beschränkt. Folglich kann ein einzelner Dipol-Korrektormagnet eine maximale sinus-förmige Orbitänderung von etwa 200 μm/s bei einem Strahlimpuls 450 GeV/c und 13 μm/s bei einem Strahlimpuls 7 GeV/c erzeugen. Die Begrenzung der maximalen Stromänderungsrate hat zur Folge, daß die maximale effektive Frequenz mit der das Orbit Regelungssystem bei nominellen Betrieb eine externe Störung mit einer Amplitude von 13 μm dämpfen kann auf weniger als 1 Hz limitiert wird.

Fehler oder Ausfälle von Korrekturmagneten sind im Vergleich zu denen der Strahlpositionsmonitoren noch folgenschwerer. Obwohl eine mittlere Zeit von etwa 100000 Stunden zwischen Fehlern pro Stromversorgung erwartet wird, führt die große Anzahl von Dipol-Korrektorkreisen die unabhängig voneinander ausfallen können dazu, daß etwa ein Korrektorkreis Ausfall pro Woche während des LHC Betriebes erwartet wird. Mit mindestens 89% Wahrscheinlichkeit hat der Ausfalls eines einzelnen Korrektorkreises eine Strahlpositionsänderung von mehr als 100 μm zur Folge, welche zusammen mit den engen Toleranzen des LHC Collimations Systems einen direkten Strahlverlust zur Folge haben. In Rahmen diese Arbeit konnte gezeigt werden, daß die Abklingzeitkonstanten der Dipol-Korrektorkreise durch die Entfernung eines Widerstandes, der für die Entladung des Kreises im Fehlerfall vorgesehen war, von ursprünglich 30-35 s auf 60-80 s erhöht werden kann und schlägt zusätzlich eine mögliche aktive Kompensation des sich im Ausfallen befindenden Korrektorkreises vor. Es is geplant den Effekt des ausfallenden Dipol auf den Orbit durch benachbarte Korrektoren abzufangen, was eine erzwungenen Strahlextraktion vermeidet und somit die Verfügbarkeit der Maschine für die Experimente erhöht. Eine realistische Angabe der Ausfallwahrscheinlichkeit kann nur aus den Erfahrungen des LHC Betriebes bestimmt werden.

In dieser Arbeit wird vorgeschlagen die Stabilisierung des Orbits und des relativen Strahlimpulses in eine “raumartige” und “zeitartige” Komponente aufzuteilen. Für kleine Strahlperturbationen kann der Effekt der Dipol-Korrektorkreise auf den Orbit und Impuls durch einen Satz linearer Gleichungen beschrieben und folglich in Matrix-Form gebracht werden. Die 'raumartige' Lösung des Kontrollproblems besteht daher im wesentlichen aus einer Zerlegung der Singulärwerte (engl. *singular-value-decomposition*, SVD) und der darauf folgenden Inversion dieser Matrix. Dabei steuert die Auswahl der verwendeten Eigenwerte die Sensitivität der Korrektur bezüglich Meßfehlern, Meß-Rauschen, Fehlern in der Strahloptik und örtlicher Beschränkung der Korrektur. Es zeigt sich, daß die Verwendung der Modell- im Vergleich zur wahren Strahloptik nicht die Präzision sondern die Konvergenz-Geschwindigkeit der Korrektur beeinträchtigt und somit die effektive Bandbreite des geschlossenen Regelungssystem reduziert.

Die vorgeschlagene 'zeitartige' Orbit Steuerung besteht aus einen Proportional-Integral (PI) Regler und wird durch einen *Smith*-Prädiktor und einen sogenannten *anti-windup* Mechanismus erweitert um die nachteiligen Effekte der nicht-linearen Dynamik der Stromratenlimitierung und der unvermeidlichen Todzeiten zu minimieren. Für den PI Regler und die Prädiktor Erweiterungen wird Youlas affine Parametrisierung aller stabilisierenden Regler verwendet. Die Bandbreite des geschlossenen Regelkreises kann direkt durch einen einzelnen Parameter gesteuert werden, was die Inbetriebnahme und Anpassung des Systems im laufenden Betrieb erleichtert.

Um das Kontrollschema bereits vor der LHC Inbetriebnahme testen zu können, wurde eine Echtzeit-Simulation entwickelt, die die Dynamik der Korrektorkreise und deren Einfluß auf die beiden Protonenstrahlen einschließlich Magnetfehler und anderer externer Störungen berechnet. Dieses System implementiert die gleichen für den LHC geplanten Schnittstellen und kann somit auch zukünftig zur Validierung neuer Orbit Kontrollsysteme und Revisionen eingesetzt werden bevor diese auf den Strahl agieren.

Summary

This report presents the stability and the control of the Large Hadron Collider's (LHC) two beam orbits and their particle momenta using beam-based feedback systems. The LHC, presently being built at CERN, will store, accelerate and provide particle collisions with a maximum particle momentum of 7 TeV/c and a nominal luminosity of $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The presence of two beams, with both high intensity as well as high particle energies, requires excellent control of particle losses inside a superconducting environment, which will be provided by the LHC Cleaning and Machine Protection System. The performance and function of these systems depends critically on the stability of the beam and may eventually limit the LHC performance. The aim of this report is to contribute to a safe and reliable LHC commissioning and machine operation.

Environmental and accelerator-inherent sources as well as failure of magnets and their power converters may perturb and reduce beam stability and may consequently lead to an increase of particle loss inside the cryogenic mass. In order to counteract these disturbances, control of the key beam parameters – orbit, tune, energy, coupling and chromaticity – will be an integral part of LHC operation. Since manual correction of these parameters may reach its limit with respect to required precision and expected timescales, the LHC is the first proton collider that requires feedback control systems for safe and reliable machine operation.

The first part of the analysis gives an estimate of the expected sources of orbit and energy perturbations that can be grouped into environmental sources, machine-inherent sources and machine element failures: the slowest perturbation due to ground motion, tides, temperature fluctuations of the tunnel and other environmental influences are described in this report by a propagation model that is both qualitatively and quantitatively supported by geophone and beam motion measurements at LEP and other CERN accelerators. These confirm that ground motion contribution to the orbit can be negligible in the LHC for frequencies above 1 Hz. For frequencies below 1 Hz, orbit drifts are dominated by random ground motion and can reach $200 - 300\,\mu$m within 10 hours. Solar and lunar tides primarily vary the beam momentum between $-0.9 \ldots 0.5 \cdot 10^{-4}$ on a time scale of about 6 hours and, as a secondary effect, the orbit shifts by $^{+100}_{-170}\ \mu$m in regions with dispersion function values of 2 m. The fastest perturbations are expected during the energy ramp and during the change of the final focus optics (beta-squeeze) in the experimental insertions. During the ramp, these drifts can reach up to $600-700\,\mu$m with drift velocities of about $15\,\mu$m/s and, dependent on the magnet alignment in the experimental insertions, the beta-squeeze induced orbit drifts can reach up to 30 mm with drift velocities of up to $25\,\mu$m/s.

These orbit perturbations exceed the required beam stability by one order of magnitude. The LHC Cleaning System, imposing one of the tightest constraints on beam stability, requires a beam stability in the range of about $15 - 25\,\mu$m at the location of the collimator jaws. LHC Machine Protection and other systems have similar orbit stability requirements. Due to the large number of requirements and their spacial distribution, it is advisable to stabilise the orbit not only in selected local regions but in the entire accelerator.

The second part of this analysis deals with the control of the two LHC beams' orbit and energy through automated feedback systems. Based on the reading of the more than 1056 beam position monitors (BPMs) that are distributed over the machine, a central global feedback controller calculates new deflection strengths for the more than 1060 orbit corrector magnets (CODs) that are suitable to correct the orbit and momentum around their references. The achievable stability is essentially limited by the quality and performance of the individual feedback components. Hence, this report provides an analysis of the BPMs and CODs involved in the orbit and energy feedback.

The BPMs are based on a *wide-band time normaliser* circuit that converts the transverse beam position reading of each individual particle bunch into two laser pulses that are separated by a time delay and transmitted through optical fibres to an acquisition card that converts the delay signals into a digital position. The calculation of the orbit is accordingly computed by averaging the individual bunch position

measurements for a given number of bunches and turns inside the machine. The position measurement resolution is dependent on the bunch length as well as the number of particles per bunch. Based on tests with nominal beam performed in the SPS, the BPMs show a bunch-by-bunch position resolution of about $60\,\mu$m, which after averaging, corresponds to an orbit resolution of about $5\,\mu$m. A simple error model has been tested and compared to the measurement accuracy of LHC type BPMs, obtained through beam-based measurements in the SPS. The measurement accuracy that is normalised to the BPM half-aperture was shown to be in the order of 1% and, within the specified beam parameters, is in accordance to the specification.

The average beam position is controlled through 1060 superconducting and individually powered corrector dipole magnets. The lower beam stability limit is given by the current stability delivered by the power supplies. Based on the measured relative current stability of $5 \cdot 10^{-6}$, the lower orbit stability is given to $6 \pm 2\,\mu$m at a beam momentum of 450 GeV/c and improves to $0.4 \pm 0.1\,\mu$m for a beam momentum of 7 TeV/c. The maximum current changes are limited by the maximum available voltage and corrector inductance (6 H) to less than 0.5 A/s. As a consequence, the maximum sinusoidal orbit drift that can be generated by a single orbit corrector is about $200\,\mu$m/s for a beam momentum of 450 GeV/c and about $13\,\mu$m/s for a beam momentum of 7 TeV/c. The maximum current change rate limit also imposes a limitation on the maximum effective frequency that the orbit feedback can compensate to an external perturbation with an amplitude of $13\,\mu$m to less than 1 Hz.

Errors and failures of the corrector magnets are more severe when compared to those of the beam position monitors. Although the single orbit corrector circuit is expected to have a mean time between failures of about 100000 hours, the large total number of circuits implicates that, in average, about one orbit corrector circuit will fail during operation per week. With at least 89% probability, failure of a single corrector circuit will shift the orbit by more than $100\,\mu$m, which in combination with the tight tolerance of the LHC collimation system will cause an immediate beam loss. It is shown in the line of this thesis that the removal of a resistor from the power converter's 'crow-bar', intended to speed up the discharge of the circuit in case of a converter failure, increases the decay time from 30-35 s to about 60-80 s and thus reduces the severity of power converter failures. In addition to this, a scheme is proposed that compensates the deteriorating effect of the failing circuit by using neighbouring circuits. This scheme may reduce the need for unplanned energy and beam extractions and possibly increase the availability of the LHC for physics. However, a realistic estimate of the single circuit failure rates can only be determined through operational experiences in the LHC.

This thesis proposes to decouple the control into *space-* and *time-domain*. The dipole correctors' response on orbit and beam energy can, for small perturbations, be described by a set of linear equations and accordingly be cast into matrix form. The solution in *space domain* is performed through singular value decomposition (SVD) followed by the inversion of these matrices. The number and type of the eigenvalues used for the inversion steers the precision and locality of the solution as well as the robustness with respect to measurement noise, failures of BPMs and CODs, and the sensitivity to errors on the accelerator lattice. Further, it is shown that the use of an imperfect beam response that differs from the one describing the true beam response for the orbit correction does not affect the precision of the correction but may lead to a reduction of the effective feedback bandwidth.

The proposed correction in 'time-domain' consists of a proportional-integral (PI) controller that is extended by an internal *Smith-Predictor* and so called *anti-windup* mechanism in order to compensate the deteriorating effects of non-linearities due to the current rate limiter and inevitable dead times in the system. The predictor logic has been derived using Youla's affine parameterisation enables to adjust the effective feedback bandwidth through a single parameter. It is thus possible to adapt and optimise the feedback to the given operational scenario.

A real-time 'testbed' complementary to the feedback controller has been developed to test part of the controller logic and its components prior to LHC commissioning. This testbed simulates a realistic open loop orbit response including effects due to delays, BPM and magnet responses and in order to test the controller under nominal operational conditions. The system implements the same for the LHC's planned control interfaces and can thus be used to debug, test and validate future controller implementations prior to their operation on the real beams.

It is believed that during the life cycle of the LHC, the merit of the orbit and energy feedbacks on beam parameters is not limited to nominal machine operation but is also beneficial if used during early commissioning.

Table of Contents

Introduction

The Large Hadron Collider (LHC) is a circular particle accelerator used to investigate aspects of high-energy physics, and is presently under construction at CERN[1]. Extending and using the former Large Electron Positron Collider (LEP) tunnel, the LHC is located at an average depth of 100 metres and has a circumference of 27 km. It will store, accelerate and collide two counter-rotating proton beams up to a maximum particle momentum of 7 TeV/c per beam and provide a nominal peak luminosity of $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in four interaction regions. The particles emerging from these interaction regions will be used to identify new fundamental particles such as the Higgs Boson, explore and search for particles that may provide evidence for super-symmetric extensions of the standard model and other models including, for example, extra dimensions. At a later stage, for the study of the quark-gluon plasma, it is foreseen to collide beams of lead nuclei ($^{208}\mathrm{Pb}^{82+}$) with a momentum of $Z \cdot 7\,\mathrm{TeV/c} \approx 574\,\mathrm{TeV/c}$, with Z being the number of charges per nucleon, and a nominal luminosity of $10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

Accelerators are vital for nuclear as well as high-energy physics, providing particle collisions in a controlled experimental environment. These collisions may lead to elastic scattering of the initial particles or the production of new particles emerging from the interaction point. An interaction of incomming particles detected by a detector is referred to as *event*. Selected events are then used to test elementary constituents that describe matter and verify their properties. The fundamental interactions underlying an event can be derived from the properties of the emerging particles such as type, energy, momentum and the rate at which they are produced. The event rate dN_i/dt is commonly decomposed into the *cross-section* σ_i of a given process and luminosity $\mathscr{L}$ given by the accelerator design:

$$\frac{dN_i}{dt} = \mathscr{L} \cdot \sigma_i$$

The cross-sections are specific for the type of interactions and are usually energy-dependent. Some cross-sections such as the top-quark or the Higgs-Boson production can be very small compared to the unavoidable elastic and inelastic scattering background. In order to obtain significant results and to search for new particles, the design of accelerators is driven by the optimisation of particle collisions at highest energies and highest luminosities possible.

The energy of hadron colliders is defined by the available maximum magnetic field of the main dipole magnets, whereas the luminosity mainly depends on the total number of stored particles and transverse beam size in the interaction region. In the case of the LHC, the main dipole field reaches up to 8.33 T in order to provide the required centre of mass energy of 14 TeV. Nearly all LHC magnets are based on a superconducting Niobium-Titanium (NbTi) alloy and are operated at temperatures between 1.9 and 4 K. In order to meet the luminosity requirement, it is foreseen to store, accelerate and collide two beams with an intensity[2] of up to about $3 \cdot 10^{14}$ particles per beam.

The control of particle loss is an important issue in a superconducting environment.

Though the use of superconducting magnets has an advantage with respect to reduced energy dissipation and consequently reduced overall running costs, the LHC magnets may lose their superconductivity if the energy deposition into the magnets exceeds given limits due to the absorption of particles or failures of the cryogenic system.

The nominal proton intensity of about $3 \cdot 10^{14}$ particles and maximum energy of 7 TeV corresponds to an energy of about 350 MJ stored in the beams. The minimum energy density that may induce a loss of superconductivity inside the magnets is approximately $30\,\mathrm{mJ/cm}^3$. In comparison with the stored energy, it is clear that this energy is sufficient to compromise the function of all magnets if evenly distributed around the machine and is capable of inducing severe damage to accelerator equipment and experiments.

[1] CERN is the European Organisation for Nuclear Research based in Geneva, Switzerland.

[2] The number of stored particles is often referred to as 'intensity', contrary to the classical definition, but however became a common term within the accelerator and high-energy physics community.

A sophisticated Machine Protection and Cleaning System is thus being developed to minimise particle loss into the cryogenic aperture. The required level of particle loss control is about 10^{-7}.

Already during an earlier LHC design stage, both machine protection and beam cleaning systems were foreseen to protect against accidental beam loss and minimise particle loss into the sensitive cryogenic aperture. Among other systems, machine protection and collimation systems require good alignment of the machine and stable beam parameters. The beam cleaning system, for instance, requires a stable beam position at the location of its collimators in order to provide cleaning and protection of the superconducting magnets. The stability is in the order of a fraction of the beam size, which is between about 100 and 300 μm in the LHC dipole magnets. Additional systems require stabilisation on a similar level.

Environmental and accelerator-inherent sources as well as failure of machine devices may perturb and reduce beam stability and may consequently lead to an increase in particle loss inside the machine. The scope of this thesis is the analysis and control of dynamic perturbations of the equilibrium beam position (orbit) and beam momentum. The beam perturbations may exceed the required beam stability by more than two orders of magnitude. Thus, it is foreseen to implement automated control systems to guarantee the proper function of the protection, collimation and other systems that rely on a stable beam. This control system also helps to optimise machine efficiency and increase the availability of the LHC for particle physics.

The orbit and beam momentum are complex beam properties that depend on the magnetic settings, their layout as well as the operational mode of the accelerator. The LHC design, its parameters, as well as issues related to the use of superconducting magnets are summarised in Chapter 1. The *decay* and *snap-back* phenomena, and the necessity for complex two-stage beam cleaning systems are emphasised. Chapter 2 gives a brief summary of the formalisms that are used to describe transverse and longitudinal particle motion in a circular accelerator.

Beam stability can be improved with respect to disturbances if beam-based measurements are used to compute the compensation signals that are sent to the corrector magnets. Beam stability dependes on the type and parameters of the control algorithm used. The control theory that describes feedback systems is summarised in Chapter 3.

Anticipating the basic principles described in Chapter 3, the design of the orbit and energy control system depends on the stability of the beam and vice versa. Beam perturbation effects are discussed in Chapter 4, followed by Chapters 5 and 6 that discuss the measures that stabilise the beam. Beam perturbations can be subdivided into three groups:

- Environmental sources such as un-correlated ground motion, ground motion waves, lunar and solar tides, thermal expansion of magnet girders and tunnel walls.
- Machine-inherent sources such as dynamic multipole field errors, change of the final focus, corrector hysteresis and power converter noise.
- Machine element failures, in particular failures of orbit corrector magnets and their individual power converters.

The predictions are supported by measurements at LEP, SPS and its TI8 transfer line, as well as magnet measurements in CERN's magnet test facilities. The perturbations are compared with the required stability given by the collimation and protection system, as well as other systems.

Besides the environmental and machine-inherent effects, the final quality of the orbit and energy control is determined by the accuracy, resolution and noise of the involved beam position monitors (BPMs), orbit corrector magnets and their individual power supplies. A description of these sub-systems is provided in Chapter 5. Where possible, the analysis is compared against measurements performed at the SPS accelerator and magnet test facilities. As a control strategy, the LHC orbit and energy feedback will be based on a central feedback controller that receives the information from BPMs, evaluates new settings for the corrector dipole magnets (CODs) and sends them to the corresponding front-end computers controlling these magnets. There are in total 1056 BPMs and 1060 CODs. The communication and synchronisation between the sub-systems is a key aspect of the system and is based on the LHC technical network and timing infrastructure.

Chapter 6 discusses the different control and orbit correction strategies. The problem can be split into *space-* and *time domains*:

Assuming steady-state perturbations only, the task in the *space domain* is to find a set of dipole corrector deflections that minimises the residual of the orbit with respect to its reference. The orbit perturbations to be corrected can, in a linear approximation, be written in a simple matrix. The input vector in this approximation is given by the strength δ_i of the N CODs. The output vector $\Delta\underline{z} =$

$\Delta(z_1, \ldots, z_M^T)$ is given by the beam position that is sampled at the M BPMs.

$$\Delta \underline{z}_{error} = \underline{z}_{ref} - \Delta \underline{z} = \underline{x}_{ref} - \mathbf{M}_R \cdot \underline{\delta}$$

with z_{ref} being the reference orbit. This formalism can be further extended to also describe the effects of the main dipoles, other magnets, and momentum and energy shift, respectively. The orbit correction consists of a minimisation of the norm of the difference vector $\varepsilon = \left\| \underline{z}_{ref} + \Delta \underline{z} \right\|$. There are several linear algebraic methods that could be used to solve this least-square problem. However, naive methods such as the classic Gauss elimination are sensitive to uncertainties and noise of the BPM measurements and often fail due to singularities in the beam response matrix $\mathbf{M}_R$. Singularities are a result of over- or under-determined equation systems. The cause for these singular solutions are orbit responses of two or more CODs that may produce beam position distributions that are equal or very similar. For robustness with respect to singularities, noise, and uncertainties on the response matrix, the orbit correction uses an approach based on the Singular Value Decomposition (SVD). The SVD is a eigenvalue decomposition of the response matrix that facilitates the identification and removal of singular solutions. The algorithm and its applications are further discussed in Chapter 6.

The task in the *time domain* is to find an appropriate control law that optimises the transition from the individual initial deflection to steady-state dipole deflection while retaining given constraints such as required bandwidth, minimisation of overshoot and robustness with respect to model and measurement errors. The baseline controller and its architecture is discussed in Chapter 6.

The feasibility of the feedback design is further evaluated and verified using a real-time testbed. This test environment simulates the accelerator response including noise, errors and failures of the BPMs and CODs. It implements similar interfaces that will be used for the LHC. The testbed estimates the performance and robustness of the orbit and energy control loops with respect to perturbations, and their behaviour in case of machine element failures such as failing CODs or BPMs.

Chapter 1

The Large Hadron Collider

The Large Hadron Collider is a separated function synchrotron that will store, accelerate and collide two counter-rotating proton beams with a centre-of-mass energy of 14 TeV and a total number of up to $3 \cdot 10^{14}$ protons per beam. It will provide proton collisions at a nominal luminosity of up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1} = 10\,\mathrm{nb}^{-1}\mathrm{s}^{-1}$ in four experimental insertions occupied by the ATLAS, ALICE, CMS and LHCb detectors. The four additional insertions are used for accelerator equipment such as the Radio Frequency System (RF), the Beam Cleaning System, and the Beam Extraction System. The use of the LEP tunnel and the CERN pre-accelerator chain and infrastructure, consisting of the SPS and its pre-accelerators, was an important design choice reducing the required budget, and it limits some of the LHC parameters such as the injection energy of 0.45 TeV/c, which is given by the Super Proton Synchrotron (SPS) pre-injector and the maximum bunch intensity, which is determined by the space charge limits in the PS Booster.

Tables 1.1 and 1.2 summarise the nominal LHC beam and machine parameter as proposed and described in the LHC Technical Design Report [1].

Beam Parameter		**Injection**	**Collisions**
Proton momentum	[GeV/c^2]	450	7000
Relativistic gamma ($\gamma = \left(1-\beta^2\right)^{-0.5}$)		479.6	7461
Number of particles per bunch		$1.15 \cdot 10^{11}$	
Number of bunches		2808	
Minimum bunch spacing	[ns]	25	
Longitudinal emittance (4σ)	[eVs]	1.0	2.5
Transverse normalised emittance	[μmrad]	3.5	3.75
Circulating beam current	[A]	0.584	
Stored energy per beam	[MJ]	23.3	362
RMS bunch length	[cm]	11.24	7.55
RMS beam size at IP1 and IP5[1]	[μm]	375.2	16.7
RMS beam size at IP2 and IP8[2]	[μm]	279.6	70.9
Form factor due to crossing angle[2]		-	0.836
Peak luminosity in IP1 and IP5 ($\mathscr{L}_{peak}$)	[cm^{-2}s^{-1}]	-	10^{34}
Peak luminosity in IP1 and IP5 per bunch	[cm^{-2}s^{-1}]	-	$3.6 \cdot 10^{30}$
Int. luminosity per bunch crossing ($\mathscr{L}^{int}_{bunch}$)	[b^{-1}bunch^{-1}]	-	317
Total p-p cross section (σ_{pp}) [2]	[mb]	67.5	111
Total events per bunch crossing ($\mathscr{L}^{int}_{bunch} \cdot \sigma_{pp}$)		-	35

Table 1.1: LHC beam parameters. A complete list and more details can be found in [1,3].

[1] The Collision beam sizes assume a for collision squeezed optic with $beta^* = 0.55$ m.
[2] The beam sizes assume a for injection and collisions squeezed optic with $beta^* = 10$ m.
[3] The total crossing angle in IP1 and IP5 is 285 μrad

Storage Ring Parameter		Injection	Collisions
Ring circumference	[m]	26658.883	
Proton energy loss due to synchrotron radiation	[eV]	$1.15 \cdot 10^{-1}$	$6.71 \cdot 10^{3}$
Revolution frequency	[kHz]	≈ 11.2455	
Number of main dipole magnets		1232	
Number of lattice quadrupole magnets		478	
Main dipole field	[T]	0.535	8.33
Main dipole bending radius ϱ	[m]	2803.95	
equivalent bending angle per dipole δ_{MB}	[mrad]	5.1	

Table 1.2: LHC machine parameters. A complete list and more details can be found in [1,3].

1.1 LHC Physics and Parameters

The objective of particle physics, which is to find the elementary constituents that describe matter and its interactions, is best described by the '*Standard Model*'. It encompasses twelve point-like fundamental constituents of matter that can be grouped into light and heavy particles referred to as *Leptons* and *Hadrons*, and five force carriers that describe the electromagnetic, weak and strong interactions between leptons and quarks. The masses of the particles observed up to now are shown in shown in Figure 1.1. The Standard Model postulates an additional massive '*Higgs*' boson that may explain the generation

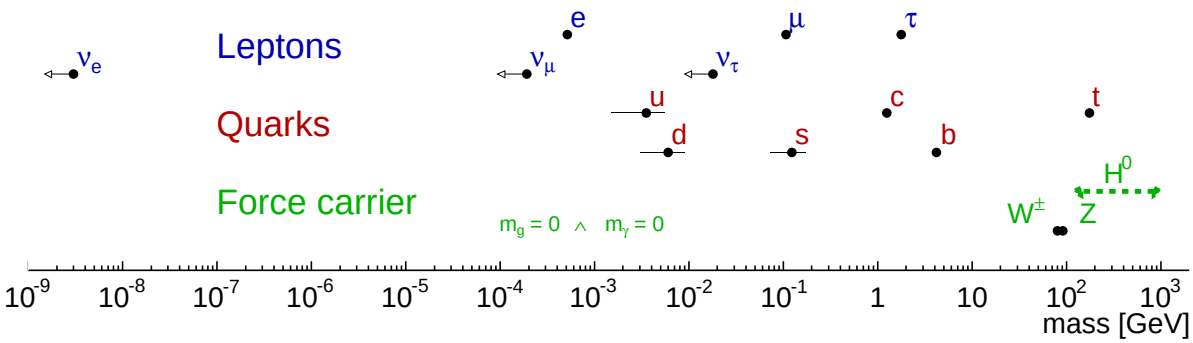


Figure 1.1: Standard Model particle masses and their uncertainties: Photons and gluons are assumed to be mass less. For the neutrino masses (ν_e, ν_μ and ν_τ,) the upper limits are given [4]. A likely range for the expected Higgs mass (H^0) is indicated.

of these masses [5, 6]. With exception of the Higgs, all above particles have been observed, and their interactions are experimentally tested in great detail, for example up to a centre-of-mass energy of about 200 GeV at the Large Electron Positron Collider (LEP) at CERN.

However, there are indications that the Standard Model may be incomplete: The most obvious limitation of the model is the omission of the description of gravitation. Though the Higgs boson may describe the mass generation of the observed particles, it does not explain their actual measured values. Apart from some constraints on the ratios, these masses are part of the 19 free parameters of the Standard Model. Ten additional parameters are required to describe the masses and oscillations of neutrinos.

The Standard Model gives a common description for electro-weak and strong interactions. The coupling constants for these processes do not unify at one high energy. Further observations in astrophysics and cosmology such as the observed matter-antimatter asymmetry in the universe and the galaxy rotation curves are not explained by the model. The latter may indicate either the existence of *dark matter*, *dark energy*, a possibly different behaviour of gravity on large scales, or a new theory of quantum gravity.

Many theories were ruled out by LEP and its detectors as they did not hold up to an energy of about 100 GeV. Apart from a conclusive test of the Higgs-mechanism, there are two promising extentions of the Standard Model, Super Symmetry (SUSY) and theories including extra dimensions that may eventually answer some of the remaining issues. The LHC experiments are expected to provide important measurements for clarifying these ideas. Theoretical physics indicates that the evidence for new possible particles and physics are likely found through collision experiments up to a centre-of-mass energy of 1 TeV.

Due to the strong E^4 and m^4 dependence on the energy E and particle mass m, respectively, of the synchrotron radiation that is emitted by charged particles describing a circular path, the choice of colliding

hadrons in the LHC is motivated by the presently feasible and affordable accelerator technology. In a circular electron-positron collider with similar parameters as LEP but with a beam energy of 500 GeV, electrons would lose a significant amount of energy that would require continuous replenishment each turn through very large radio frequency systems which are considered impractical due to their high operating costs. In contrast, protons with similar machine parameters and energy lose about 10^{-15} times less energy than electrons. Thus, the maximum achievable energy is given by the price of industrially available magnets rather than by the required RF system.

However – unlike leptons – hadrons are not fundamental particles but composed of quarks and gluons (partons), each carrying an unknown fraction of the parent hadron's momentum. The objects which actually 'collide' and that produce an 'event' are therefore uncertain in type, momentum, polarisation, and energy. Thus, only a fraction of the proton energy is available for the inelastic parton interactions that eventually may produce the Higgs or other exotic particles. As a consequence, theory indicates that a much higher proton-proton centre-of-mass energy of about 14 TeV is required to probe the standard model up to about 1 TeV.

The rate of a given event $\dot{N}$ is defined by its cross section σ_i and the luminosity $\mathscr{L}$.

$$\dot{N}_i = \mathscr{L} \cdot \sigma_i \tag{1.1}$$

The luminosity, by definition, is process-independent and can be derived from machine parameters and its optics. Example for the Higgs production cross sections and decay branching ratios are shown in Figure 1.2.

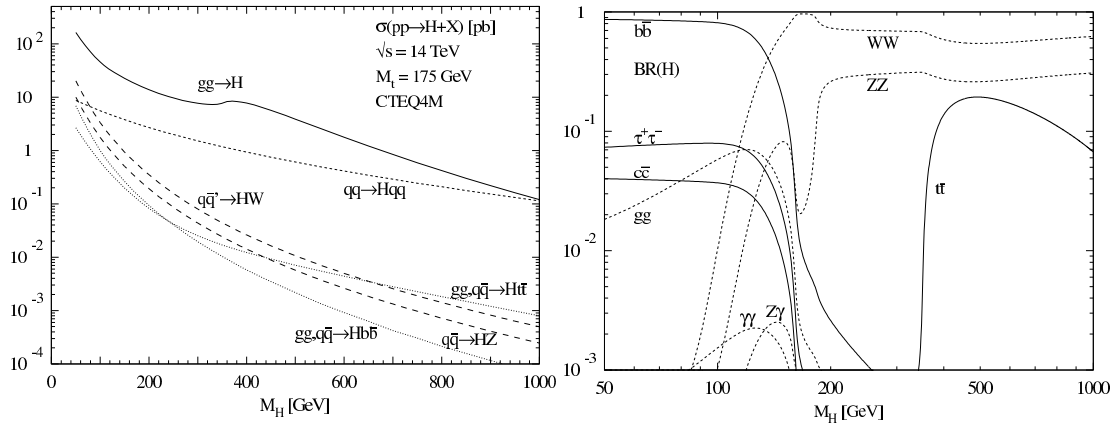


Figure 1.2: Standard Model Higgs boson production cross section in pb (left) and decay branching ratios (right) as a function of the Higgs mass M_H [7]. Since the coupling of the Higgs boson depends on the mass of the particle, it is visible that for higher Higgs masses, the branching ratio favours the production of heavy $W^{\pm}$ and Z boson as well as top-anti-top pairs as soon the production threshold reached twice the particle mass.

The strongest constraints on the Standard Model Higgs mass is given by LEP [8]. The combined experimental excludes directly Higgs masses below 114.4 GeV at 95% confidence level and give an indirect upper limit of less than about 200 GeV. In this range, as seen in Figure 1.2, the total Higgs production cross section is about 20 pb. The total proton-proton cross section at a centre-of-mass energy of 14 TeV is estimated from lower energy data [2] to:

$$\sigma_{pp} = 111 \pm 1.2^{+4.1}_{-2.1}\,\text{mb} \tag{1.2}$$

In comparison with the Higgs production cross section at, for example, 150 GeV ($\sigma_H \approx 20$ pb), only one Higgs boson will be produced for about six billion interactions, which requires very large collision rates to observe some hundred Higgs events per year. For the observation and testing of the Higgs hypothesis, different possible decay channels have to be studied. One of the challenging channels for possible Higgs masses below 150 GeV/c^2 is the Higgs decay into two photons $H^0 \to \gamma\gamma$[1]. Its branching ratio is in the order of 10^{-3} and corresponds to a cross section of a few ten fb according to Figure 1.2. With respect to the total cross section given in equation 1.2 this compares to, on average, one $H^0 \to \gamma\gamma$ decay for

[1]Though the hadronic $q\bar{q}$ decay channels would provide significantly higher cross section, they are difficult to separate from the high expected QCD di-jet background with cross section in the order of a few μb.

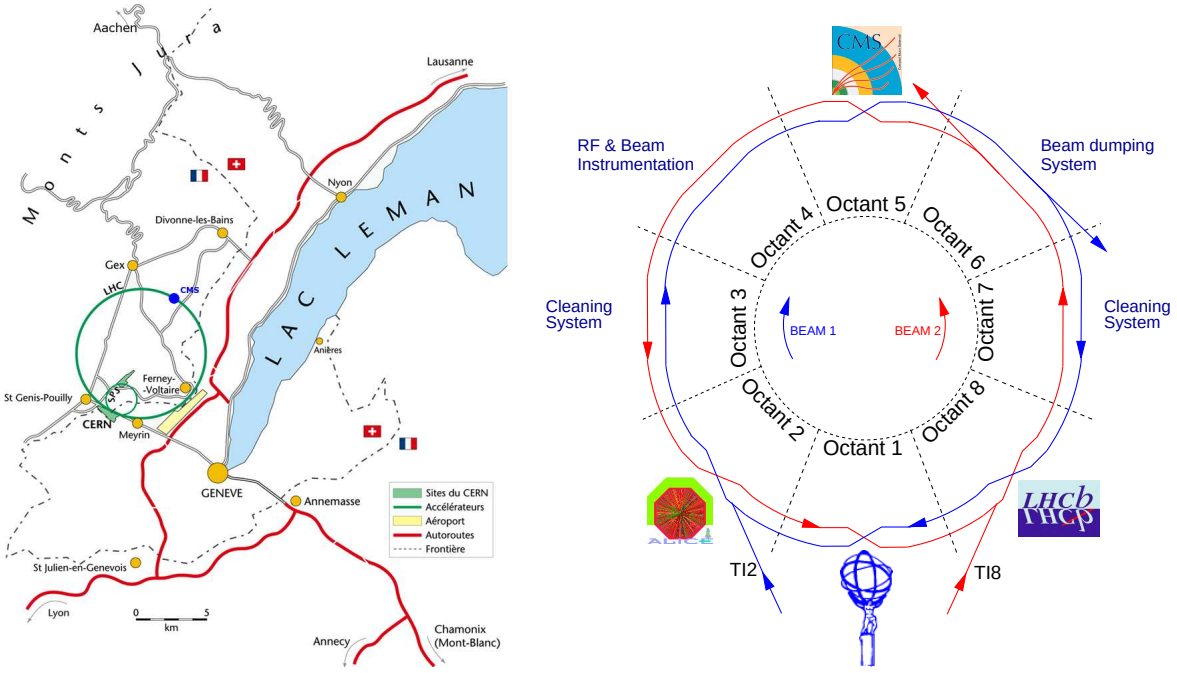
about 10^{13} proton-proton collisions. In other words, it requires a signal-to-noise rejection of more than 13 orders of magnitude. It is clear that this demands best possible particle detectors and high number of collisions.

In order to meet this challenge, CMS, as one of the LHC experiments, requires an integrated luminosity of about $30\,\mathrm{fb}^{-1}$ for a 5 σ significance for the Higgs verification in the $H \rightarrow \gamma\gamma$ channel [9]. Relating this to an approximate accelerator performance, this corresponds to an LHC peak luminosity of about $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ provided over one to two years, assuming a common nominal accelerator efficiency of 50% and assuming that the accelerator is effectively running 200 days per year[2]. Experience with other high energy and high luminosity accelerators shows that usually several years were required until nominal parameters were reached.

Besides the pure existence of the Standard Model Higgs, there are other aspects to be tested, for instance, the dependence of the Higgs coupling on the fermion and boson masses. One must, for example, distinguish the Higgs decay from the decay of other heavy particles of SUSY. In some cases, these potential Standard Model deviations are very small and thus require highest peak luminosities in order to obtain significant results within a meaningful time of less than a decade.

1.2 LHC Accelerator Layout

The LHC is installed in and extends the former LEP tunnel, has a circumference of 27 km and is located at an average depth of 100 metres. The geographical and schematic layout is shown in Figure 1.3.



(a) CERN Geographical Layout

(b) LHC Schematic Layout

Figure 1.3: Schematic layout of the LHC (left) and one octant (right).

The LHC is divided into eight octants separated by eight insertions that are used by experiments and accelerator equipment. To accelerate and collide particles with the same sign of charge, the two hadron beams are kept separately, in most parts of the LHC. The clockwise and counter-clockwise rotating beams are referred to as 'Beam 1' (B1) and 'Beam 2' (B2) and are only brought together in the four interaction regions (IR): IR1, IR2, IR5 and IR8. The ATLAS [10] and CMS detectors [11] are installed in the insertions IR1 and IR5, which will provide high luminosity collisions. Both detectors are general purpose detectors searching for the Higgs boson and particles that may provide evidence for Super-Symmetry. Additionally, these experiments make high precision measurements for testing the Standard Model processes and search for indications of other new physics. IR5 is also used by the TOTEM experiment, which

[2]The efficiency numbers and days per year are rough estimates only, other choices are possible.

studies the total proton-proton cross section, elastic scattering and diffraction dissociation of the particles emerging from the interaction point (IP) at very low scattering angles with respect to the beams [12]. Its particle tracking detectors are installed downstream of IP5 in so-called 'Roman Pots'. The other interaction regions in IR2 and IR8 are used by ALICE and LHCb. The ALICE detector is optimised for high-energy ion collisions [13] and LHCb is specialised in the study of b-physics [14]. The insertions in IR2 and IR8 also contain the injection system for Beam 1 and Beam 2, respectively.

Each octant, with the interaction point (IP) as its centre, is logically sub-divided into 68 cells. The cell layout is symmetrical with respect to the IP and is grouped by its function into 'arc', 'dispersion suppressor' (DS) and 'insertion' (IR), also referred to as 'long straight section' (LSS) as illustrated in Figure 1.4.

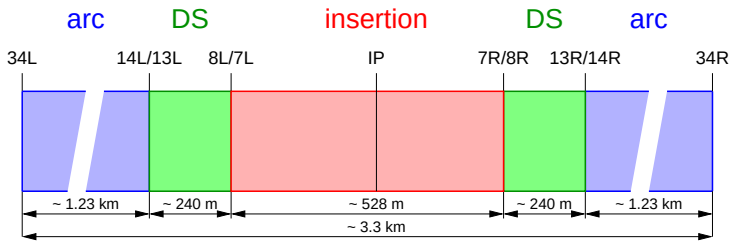


Figure 1.4: Schematic octant layout. Each octant can be subdivided into arc, dispersion suppressor (DS) and insertion. In case of point 1,2,5 and 8, the insertion regions contain the final focus triplets. The naming is illustrated: For example, the seventh cell on the left of IP5 is denoted by '7L5'. A beam position monitor ('BPM') measuring beam 1 ('B1') in this cell is thus named as 'BPM.7L5.B1'.

The naming scheme of the LHC is based on cells that are numbered with respect to the nearest IP. There are 34 cells on each side of the IP. According to LHC naming rules, the third cell on the left ('L') of IP5 is for example named '3L5'. An additional prefix denotes the equipment type inside the cell, for example: 'MB' for main dipoles, 'MCB' for orbit corrector dipole magnet and 'BPM' for beam position monitors. Further characters in the prefix distinguish different equipment class types. The possible suffix, either '.B1' or '.B2' denotes whether the equipment is acting on or measuring 'Beam 1' or 'Beam 2', respectively. For illustration of the naming scheme: A beam position monitor (BPM) measuring the position of Beam 2 in the twelfth cell on the right of IP5 is named 'BPM.12R5.B2'. Further details on LHC naming and conventions can be found in [15].

1.2.1 Arc and Dispersion Suppressor Layout

The LHC arcs have a regular layout consisting of 23 cells with a length of 106.9 m each. Since magnetic quadrupoles focus in one plane and defocus in the other, the quadrupoles are arranged to a 'FODO' cell in order to achieve focusing in both planes. A FODO cell consists of a combination of focusing (F) and defocusing (D) quadrupoles that are separated by a drift space (O). In order to optimise the use of the accelerator circumference, the drift spaces are usually filled up with main dipole magnets (MB) and other multipole magnets. Their use and impact on particle dynamics is discussed in Chapter 2.

Where possible, the LHC magnets have an optimised compact two-in-one design with two bores and opposite fields for Beam 1 and Beam 2, as discussed in Section 1.3. As a consequence of this design choice, a lattice quadrupole magnet (MQ) focusing in the horizontal plane inside the bore of Beam 1 is defocusing for Beam 2 in the other bore. The opposite relation applies for the vertical planes. The schematic layout of the LHC arc FODO cell is shown in Figure 1.5.

The main dipoles as well as the main quadrupoles of one arc are powered in series. The main dipole types 'MBA' and 'MBB' that are interleaved with the quadrupoles (MQ) essentially differ in their available aperture and attached type of corrector spool pieces. The spool piece magnets are used to compensate for local magnet imperfections and for control of long-term beam stability and lifetime. The MBA dipoles, for example, have additional spool pieces for local correction of the main dipoles' sextupole (MSC), octupole and decapole (MCDO) field errors, whereas the MBB has only the MSC spool pieces. Besides the main quadrupoles, the short straight section (SSS) further accommodates the lattice octupoles (MO), skew and trim quadrupoles (MQS and MQT), MSCB magnets consisting of a lattice sextupole, and an orbit corrector. As an important design choice discussed in Chapter 6, each LHC quadrupole is equipped with at least one beam position monitor (BPM) per bore. The additional higher order magnets such as sextupole (MS), octupole (MO) and decapole magnets are required for higher order particle dynamics

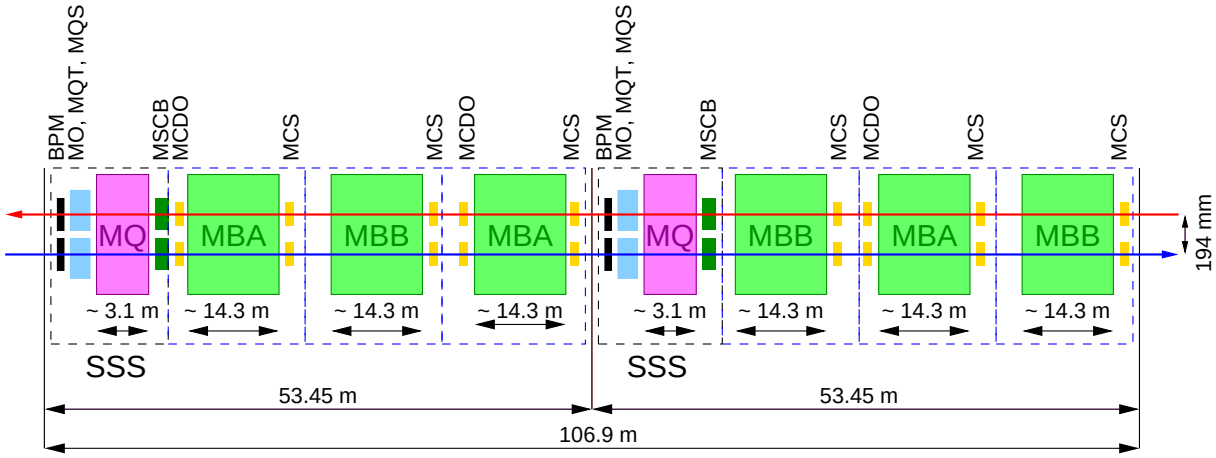


Figure 1.5: Standard LHC arc FODO cell.

due to off-momentum particles. Their use and impact on beam stability is however outside the scope of this thesis.

An intrinsic property of dipole magnets is to sort particles according to their individual momentum, referred to as *dispersion*. The 'dispersion suppressor' is the transition between the regular arc lattice and the long straight section. The function of the dispersion suppressor is to minimise the contribution of the momentum spread to the total beam size of the proton beams inside the long straight section. The common 'missing dipole' principle is used, consisting of a regular FODO cell, similar to the arc, but with one dipole missing per half-cell. The quadrupoles in the dispersion suppressors are powered individually in order to give some additional degree of freedom, required to match the optics functions in the arc to the optics requirements in the long straight section, as these quadrupoles are used to match the final focus in IP.

1.2.2 Insertion Layout

All octants have similar arc-dispersion-suppressor transitions and mainly differ in the layout of their long straight section. Out of the eight LHC insertions, four are used for particle collisions and hence require a crossing of the two beams. The other four are dedicated for accelerator equipment that act individually on each beam. The beams in this sections are further separated with respect to the arc.

Layout of IR1 and IR5

The two high luminosity insertions IR1 and IR5 are occupied by ATLAS and CMS and will provide peak luminosities of up to $= 10 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ during nominal operation. Figure 1.6 shows a schematic overview of the CMS insertion.

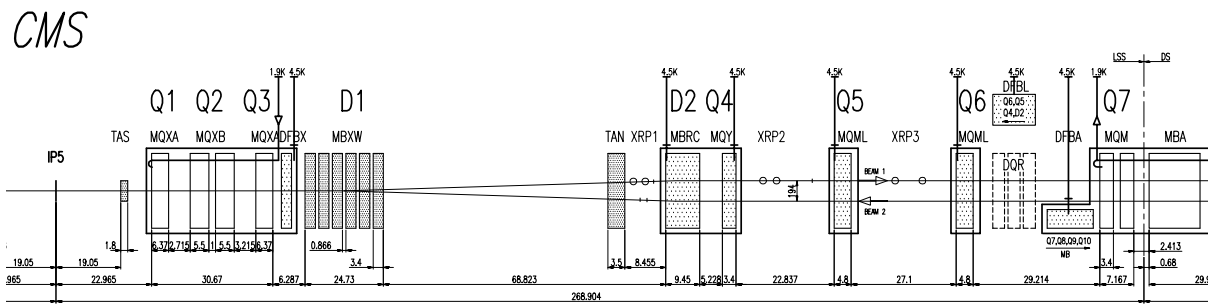


Figure 1.6: Schematic layout of IR5 housing the CMS experiment. The layout of IR1 is similar to IR5. The distances in the longitudinal coordinate are given in meters and the temperatures of the helium line in Kelvin.

The two beams are brought together by a D1-D2 dipole combination located before and after the interaction region. Due to the high expected background and particle loss, the D1 dipoles are normal

conducting magnets. The D1 actually consists of five individual magnets powered in series and is located directly after the last quadrupole magnet Q3. The D2 dipole is found in front of the quadrupole Q4. The particle loss levels are expected to be less for the D2 dipole, and is thus superconducting.

The average beam size in the arc is in the order of $100 - 300\,\mu$m. In order to provide the required high luminosity, a combination of three magnets (quadrupoles Q1, Q2 and Q3) referred to as *final focus triplet* is used to reduce the transverse beam size to about $15\,\mu$m in the IP. In the jargon of accelerator physics, the reduction of beam size is commonly referred to as *squeeze*.

In comparison to LEP for example, the final focus consisted of a quadrupole duplet on each side of the IP. The additional quadrupole triplets in the LHC give an additional degree of freedom which provides the flexibility to independently focus the beam in both planes and thus provide a round beam in the IP[3]. The closest accelerator-specific device to the IP is the TAS absorber that intercepts, together with the TAN absorber that is installed further downstream, the debris, which are mostly elastic proton-proton scattering and neutrons, created by the primary collisions that emerge along the beam axis. If not intercepted, these particles may otherwise quench the superconducting magnets.

TOTEM is an additional experiment in IR5 for measuring the total proton-proton cross section and other physics that involve very low transverse momenta in the order of about $t \approx 10^{-3}\,\mathrm{GeV}^2$. These momenta correspond to an angle of about $10^{-5}\,$rad with respect to the beam axis. Apart from shared detectors within CMS, it uses additional silicon tracker detectors that are installed in so-called *Roman pots* marked as 'XRP1', 'XRP2' and 'XRP3' and that can move very close to the beam centre. The Roman pots are located behind the D1 magnet with an approximate distance of 147 m, 180 m and 220 m to IP5 and IP1, respectively [12].

Layout of IR2 and IR8

The insertions IR2 and IR8 provide particle collisions for ALICE and the LHCb experiments. Furthermore, these insertions are used to inject Beams 1 and 2 coming from the SPS transfer lines TI2 and TI8, respectively, into the LHC. The clock-wise rotating Beam 1 is injected in IR2 into ring 1 and the counter-clockwise Beam 2 is injected in IR8 into ring 2 of the LHC. A schematic drawing is shown in Figure 1.7.

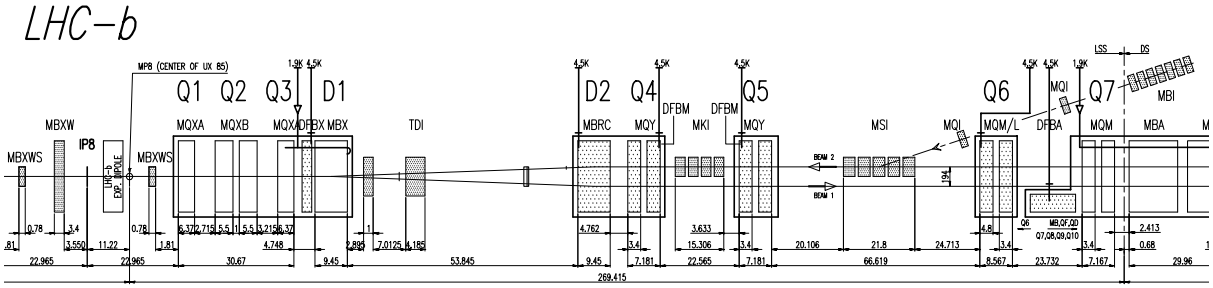


Figure 1.7: Schematic layout of IR8 housing the LHCb experiment and injection devices for Beam 2. The layout of IR2 is similar and symmetric with respect to IR8.

As for IR1 and IR5, for nominal operation the beams are crossed horizontally in IR8 and vertically in IR2 with a nominal crossing angle of $\theta_c = 285\,\mu$rad.

Similar to IR1 and IR5, both insertions are symmetrical with respect to their design interaction point. They consist of a standard dispersion suppressor region and long straight sections housing superconducting final focus triplets and D1-D2 dipole combinations for joining and separating the two beams. A noteworthy particularity of IR8 is the displaced actual collision point required by LHCb, that is shifted by 11.22 m with respect to the symmetry interaction point IP8[4]. ALICE and LHCb use additional spectrometer magnets inside the detectors. The effect of these magnets on the beam is compensated locally.

Since the required luminosity in these experiments is a magnitude lower than in IR1 and I5, the background rates are expected to be significantly lower, and both D1-D2 separation dipoles are superconducting, which provides space for the additional injection system elements.

The injection equipment for Beam 1 is located on the left of IP2 (ALICE) and for Beam 2 on the right of IP8 (LHC). Apart from this left-right asymmetry, both injection chains are similar and consist of

[3] At LEP, the beams in the IP were asymmetric and focused more in the vertical than in the horizontal plane.

[4] The displaced actual collision point in LHCb causes an increased bunch pattern complexity required in order to provide collisions in all IPs.

five horizontally deflecting magnetic injection septa 'MSI', four horizontal fast kicker magnets 'MKI' and an absorber block 'TCDI'. The MSI, located between the quadrupoles Q5 and Q6, magnetically deflect the beam to be injected close to the circulating beam trajectory. The remaining required kick is supplied by the MKI kicker that is located between the quadrupole Q4 and Q5. The TCDI is a passive absorber and is necessary to shadow and protect the superconducting D1 magnet in case of faulty injected beam due to kicker failures. There are further passive absorbers downstream of the TCDI which are discussed in [16].

Layout of IR3, IR4 and IR7

The insertions IR3, IR4 and IR7 are used for momentum cleaning system (IR3), betatron cleaning system (IR7), the LHC radio frequency and damper system and some beam diagnostic systems that measure global beam properties (IR4). Their insertion layout is similar as shown in Figure 1.8. Since particle collisions are not foreseen in these insertions, the two beams are kept separate. Most of the equipment installed in these insertions act on a per-beam basis. Since the radio frequency cavities and collimator devices of the cleaning system require more space, the two beams are separated further using two dipoles on each side of the IP referred to as D3 and D4. The particle losses are expected to be highest in IR3 and IR7 and thus most of the magnets in this region are normal conducting.

Layout of IR6

The only devices that can withstand a full impact of a nominal LHC beam at 7 TeV are the two beam dump absorbers installed in IR6. The schematic layout of the insertion is shown in Figure 1.8. The beams are kept separated in this section and the layout is symmetric for both beams with respect to the IP. For each beam, there is an ultra reliable beam dumping system consisting of 15 kicker magnets labelled as 'MKD' with a rise time of less than 3 μs that kick the beam within a turn from the circulating beam trajectory to the extraction transfer lines. The additional strong septum magnets (MSD) are used to further increase the separation between circulating beam trajectory and the extracted beam. In case of a kicker failure, the septum and the Q4 quadrupole magnet are protected by the additional passive absorbers 'TCDS' and 'TCDQ' for Beams 1 and 2, respectively.

1.3 LHC Magnets

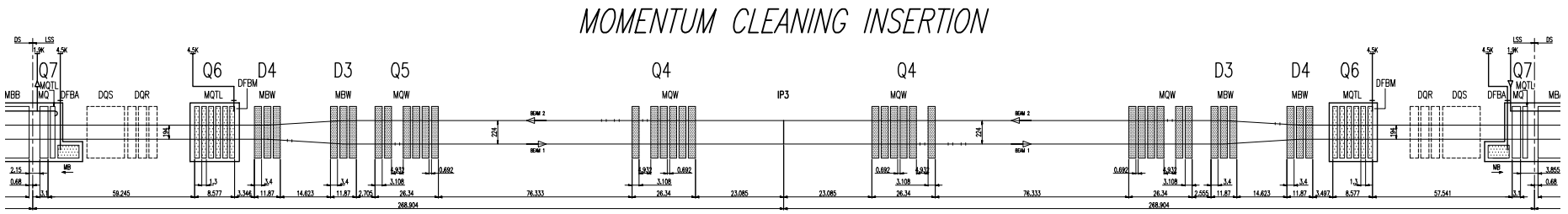
To describe guiding of particle beams, the general magnetic field description $B(x, y, s)$ is commonly expanded into their different multipole components. In order to independently steer the individual multipole components and related beam properties, modern accelerators, such as the LHC, implement most of the multipole fields in dedicated magnets. The magnet's names usually refer to their main multipole field such as *dipole*, *quadrupole*, *sextupole* or *octupole* magnets. Most accelerator magnets have only transverse magnetic components. An exception are the solenoid magnets in the experimental detectors that are used for particle identification and their momentum measurements. The solenoid magnets are designed to have ideally only a longitudinal field component.

Since synchrotron radiation is less an issue for hadrons, the maximum energy is limited by the magnetic field strengths available for bending and focusing the particles on their circular design trajectory. Charged particles describe circles with radius ϱ inside a homogenous magnetic field B. The maximum momentum p relates to magnetic fields and bending radius through

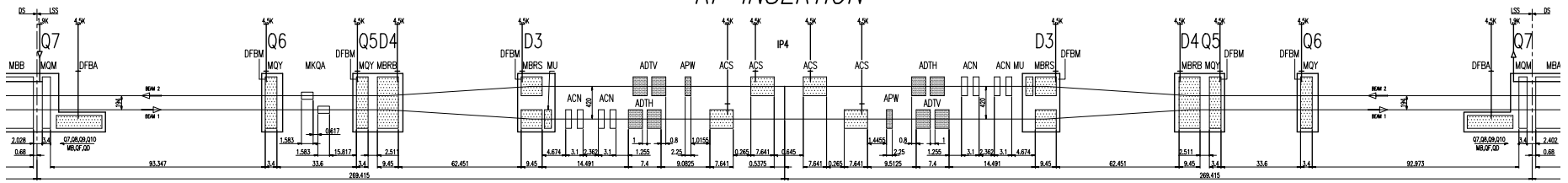
$$p = e \cdot \varrho B_{max}, \tag{1.3}$$

with e being the particle's charge. The expression $(\varrho B) = \frac{p}{e}$ is called *magnetic rigidity* and is commonly used to normalise the equations in order to make them independent from the momentum. In case of the LHC, the magnetic bending radius is determined by the LEP/LHC tunnel to $\varrho \approx 2804$ m. To fulfil the required beam momentum of 7 TeV/c, the maximum required magnetic dipole field has to be at least 8.33 T. This relatively high field can only efficiently be provided by superconducting magnets. Warm magnet technology, due to cooling issues of the resistive coils and saturation of iron yokes, is effectively limited to fields up to about 2 T.

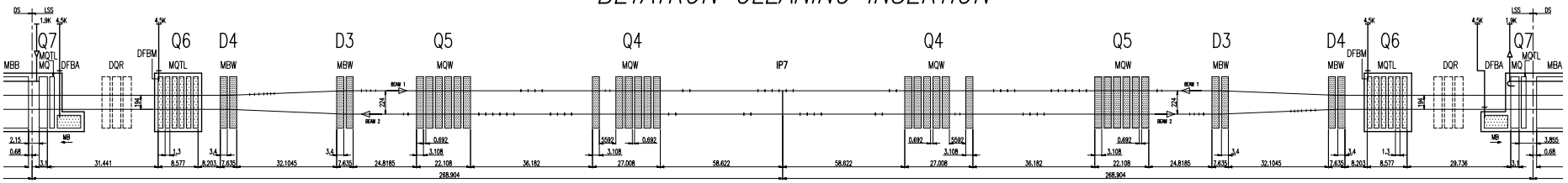
For normal conducting magnets, the field distribution and its quality is determined by the number of poles, their shape, and quality of the magnet pole surfaces. In contrast to normal conducting magnets, the field of superconducting magnets is defined by the current distribution around the bores. In order to



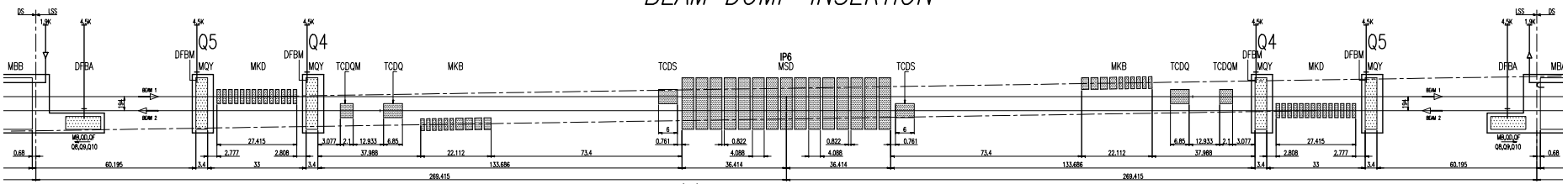
(a) Schematic layout IR3



(b) Schematic layout IR4



(c) Schematic layout IR7



(d) Schematic layout IR6

Figure 1.8: Schematic Layout of the accelerator insertions IR3, IR4, IR6 and IR7.

produce a superconducting magnet with a desired multipole order n ($n = 1$: dipole, $n = 2$: quadrupole, etc.), the coils ideally approximate a

$$I = I_0 \cos(n\phi) \tag{1.4}$$

distribution around the bore, with ϕ being the transverse angle with respect to the symmetry plane and I_0 the maximum current flowing longitudinally along the bore with radius a as illustrated for some low order multipole magnets in Figure 1.9. As shown in [17], the resulting field distribution inside the

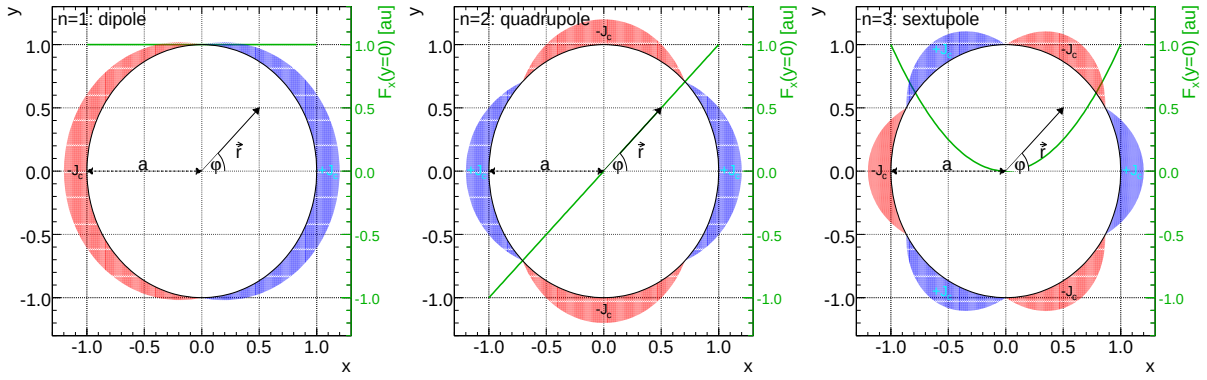


Figure 1.9: The $\cos(n\phi)$ current distributions for superconducting dipoles (dipoles), quadrupoles (middle) and sextuples (right) are shown. The colours ($\pm$ signs) indicate whether the polarity of the current density J_c is going 'in' (red,'-') or coming 'out' (blue,'+') of the drawing plane.

magnets is thus given by:

$$\vec{B}(r,\phi) = -\frac{\mu_0 I_0}{2a}\left(\frac{r}{a}\right)^{n-1} \cdot [\sin(n\phi), \cos(n\phi)] \tag{1.5}$$

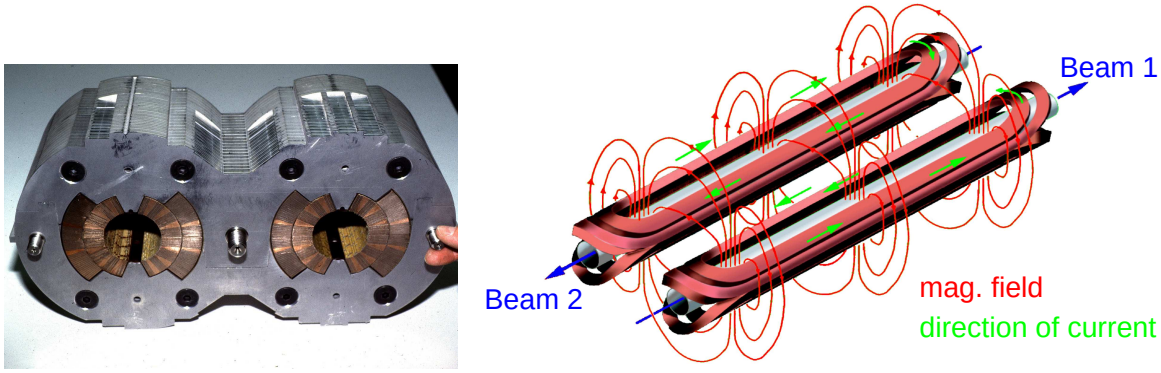
with r being the distance to the magnetic centre and ϕ the angle with respect to the reference (in general the horizontal) plane. Within the framework of magnet design and measurements, the following expansion is more commonly used

$$\begin{aligned} B(x,y) &= B_y + iB_x \\ &= B_m \cdot 10^4 \cdot \sum_{n=1}^{\infty}(b_n + ia_n)\left(\frac{x+iy}{R_{ref}}\right)^{n-1} \end{aligned} \tag{1.6}$$

with n being the index of the multipole and m the main field components of the magnet. The field of the higher multipole usually depends on the position (x,y) inside the bore with respect to the magnetic centre. In order to compare the higher multipole magnets, the fields are commonly normalised to an arbitrary reference radius R_{ref}. In case of the LHC, the reference radius is chosen to be $R_{ref} = 17\,\text{mm}$, less than the aperture. The normal (b_n) and skew (a_n) components of the field expansion are, in principle, without dimension. However, within the field of magnet design, the relative diversion part of a given magnetic field component is commonly referred to as 'one *unit*'. For example, "one unit of b_1" refers to a relative field of 10^{-4} with respect to the main dipole component.

A two-in-one design was chosen for the LHC, making the magnets sufficiently compact to fit in the same cryostat, minimising the occupied space inside the present tunnel and effectively reducing the manufacturing costs. An example of the LHC main dipole cross section and coil geometry is shown in Figure 1.10. The $\cos(n\phi)$ distribution of the coils around the bore is visible. As mentioned earlier, the field distribution inside superconducting magnets is primarily determined by the shape and position of the coils, and the current distribution inside the coils. From the magnetisation point of view, the laminated yoke has a negligible effect on the field inside the two magnet bores and mainly reduces the stray fields outside the magnet. However, the yoke is required and plays an important role in stabilising the superconducting coil. The coils are subject to high stress due to the magnetic repulsion of the strands that can exceed 1.8 MN/m for nominal currents. These forces must be taken by the yoke without changing the coil geometry and modifying the magnetic field.

The LHC magnets are kept in a liquid helium bath to keep them at their required low cryogenic temperatures. The liquid helium cryostats are isolated through liquid nitrogen and finally a vacuum tank, and for reasons of protection and transportation are kept in steel tubes.



(a) dipole cross section (b) dipole coil geometry

Figure 1.10: LHC main dipoles' two-in-one magnet design. The bore separation inside the arc is 194 mm. The approximate $\cos(n\phi)$ of the superconducting dipole coils is visible.

1.3.1 Superconductivity and Phase Transitions

Certain materials undergo a phase transition and may become superconducting once their temperature drops below the critical temperature T_c specific for the material. The virtue of superconducting materials is that they are capable of conducting very high currents within small cross sections while having negligible ohmic losses. Thus they require a significantly reduced cooling power compared to normal conducting magnets. However if the temperature, the maximum current density, or the maximum magnetic field inside the material exceeds given limits, the material undergoes reverse phase transition. In case of metallic superconducting alloys, the material may become normal conducting. This reverse phase transition from superconducting to the normal conducting phase is called as *quench*. In the LHC, the loss of particles inside the magnets is considered the most critical source of magnet quenches, as discussed below.

The LHC magnets are based on the metallic superconducting Niobium Titanium (NbTi) alloy and are operated at temperatures of 1.9 K (superfluid helium) to 4.3 K (normal liquid helium). The NbTi are formed to filaments with a diameter of about 6 μm and are implanted into a copper matrix forming a strand of about 1.4 mm diameter. In case of the main dipole and lattice quadrupole magnets, several strands are grouped and twisted to Rutherford type cables as illustrated in Figure 1.11.

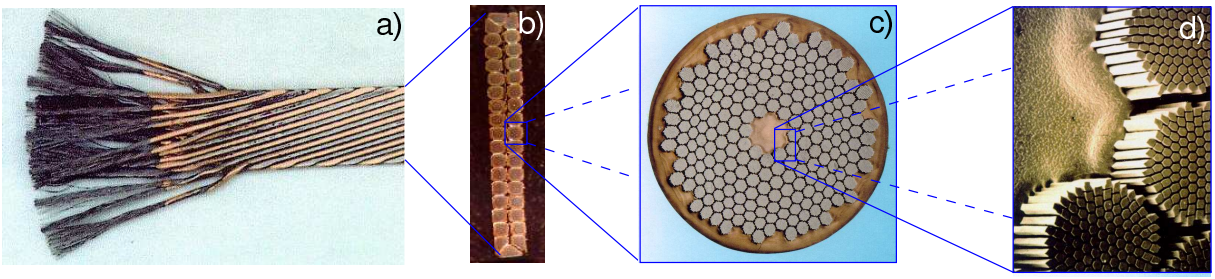


Figure 1.11: Rutherford type cable (a) as used in the LHC main dipoles and lattice quadrupoles. The cable (b), strand (c) and filament (d) cross section are shown. The cable consists of 36 strands and has a width of 15.1 mm. The strand and filament diameters are approximately 1.4 mm and 6 μm, respectively.

The critical surface for NbTi that determines the phase transition (quench) from the superconducting to normal conducting state is, apart from the mechanical stress, dependent on the critical temperature T_c, field B_c and current density J_c as shown in Figure 1.12. The maximum field and current density in the magnet are defined by coil, cable and strand geometry. The peak magnetic field in the main dipole can reach up to 8.76 T for a nominal current of 11850 A flowing through the cable cross-section. The energy stored in the magnetic field is about 7 MJ per LHC dipole and is sufficient to melt the coils and induces serious damage if released into the magnet in an uncontrolled manner.

Thus, a sophisticated magnet protection system has been developed to protect the superconducting magnets. This system is able to detect quenches at an early stage and fires so-called *quench heaters*, in

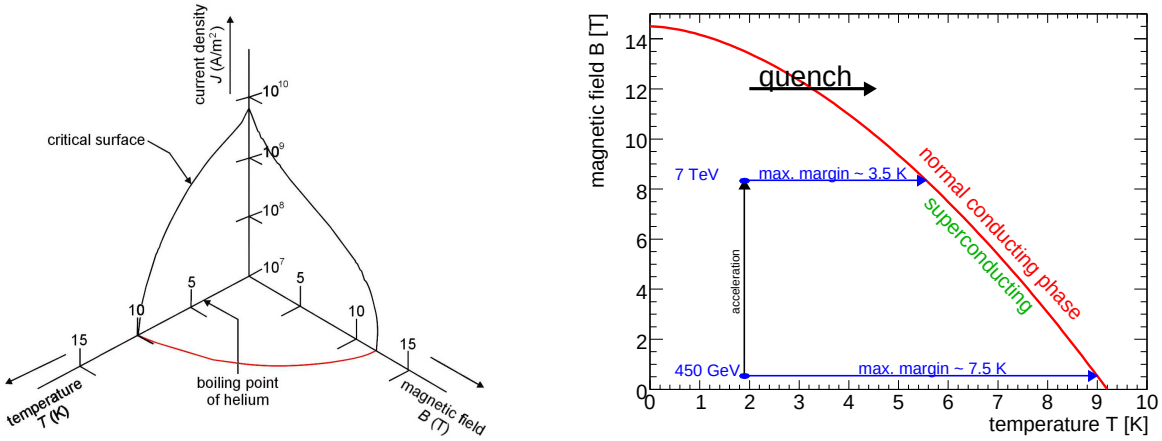


Figure 1.12: Boundary of the superconducting state: (left) the dependence on T_c, B_c and J_c. The superconducting phase is below and normal conducting if above the surface; (right) projection in the T_c-B_c plane for a given current density j_c. The phase transition and maximum temperature margins are indicated. Comparing the injection at 450 GeV and nominal settings at 7 TeV beam momentum, it is visible that the maximum margin reduces by about a factor of two, since the magnetic field is increased from about 0.5 T to 8.33 T.

order to induce a quench of the whole magnet and reduces the energy dissipation at the initial location of the quench. At the same time, a high ohmic resistor is put into series with the dipole circuit, extracting the energy out of the circuit in a safe way by heating a large block of water-cooled copper. Details on the study of quenches of LHC magnets and the required protection systems for the LHC magnets can be found in [18].

1.3.2 Eddy currents, Decay and Snapback of Multipole Components

The control of higher field momenta is critical for machine operation since they may cause particle instabilities and consequently losses in the machine.

Field imperfections can originate from deviations of the coil geometry and – to a little extent – also the effects due to the saturation of the iron yoke. By design, the $\cos(n\phi)$ current distribution around the bore suppresses the even higher multipoles for an odd multipole magnet and vice versa. For example, the dipole ($n = 1$) may have additional sextupole ($n = 3$) and decapole ($n = 5$) components whereas the quadrupole ($n = 2$) may have certain octupole ($n = 4$) and duo-decapole ($n = 6$) moments. In general, the higher multipoles decay fast with their order and are often negligible for $n > 6$.

Additional field imperfections arise due to the very small resistivity of the superconducting material and cause stationary and time dependent multipole fields. As described by the Maxwell equation

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1.7}$$

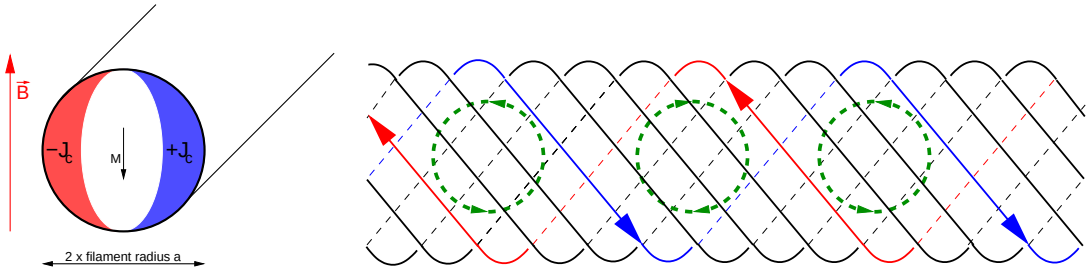
a magnetic field $\vec{B}$ that changes in time induces an electric field $\vec{E}$ that creates so-called *eddy currents* inside a conductor. These currents shielding the main field may create additional higher multipole components. Superconducting materials are, in this respect, not different to normal ones.

For example, while ramping the main dipole magnets, additional *coupling current* loops are created between filaments and strands along the coils. This is – among others – the reason to embed the filaments into a copper matrix. The interconnections between the filaments are then resistive and these types of eddy currents are damped fast as soon as the ramping stops.

However, since the resistance within the filaments itself is very small, the eddy currents may prevail much longer with decay times of the order of several days to several weeks. The currents inside the filaments are commonly referred to as *persistent currents* and are parallel with the current, creating the main field component. Since the maximum current is limited by the critical current density J_c of the superconductor, the persistent current decreases if the nominal current increases, and is thus more dominant for low fields.

In case the magnetic field is reversed, the initial persistent current loops may prevail and additional loops with opposite field are created inside the remaining filament cross section. This effect yields a

hysteresis of the persistent currents inside the filaments and is hence also referred to as *filament magnetisation* as first described in [19]. It is important to note: the hysteresis of superconducting magnets is mainly due to the above-explained superconducting property and less due to the hysteresis of the iron yoke. The hysteresis of the iron can normally be neglected since most superconducting magnets operate beyond the full saturation of the yoke. The schematic effect of an initial filament magnetisation is illustrated in Figure 1.13(a).



(a) Filament Magnetisation (b) Current loops in the LHC superconducting cables

Figure 1.13: Filament magnetisation and current loops in superconducting Rutherford type cable. Decay mechanism (right) is driven by the imbalance with respect to the main current between two strands (red and blue) inside the cable and causes various smaller current loops (green). The size of the additional loops can vary between half the twist pitch up to the full cable length.

Another effect related to superconductors is the *Decay and Snap-back* phenomenon. It is believed ([20]) that this effect is caused by imbalances of currents in different strands and filaments, illustrated in Figure 1.13(b). These imbalances are a result of earlier ramping and thus depend on the magnet ramping history and the duration the magnet stayed at the maximum field. In addition to the filament magnetisation the imbalances can cause a reduction of the main field and cause higher multipole field errors. These additional loops may *decay* due to the finite contact resistance while the main current is kept constant. This effect has time constants in the order of thousand seconds and is dominant during injection, while the main dipole current is kept constant. However, as soon as the magnet is ramped up after injection, these current loops 'regenerate' and the initial main field is restored to its pre-decayed value. The time constant of this so-called *snapback* is dependent on the magnitude of these current loops before the start of the ramp and thus depends on the magnet powering history similar to the decay. The time scale of the snapback is in the order of about hundred seconds. Details on quantitative analysis of filament magnetisation, magnet measurements, the decay and snapback phenomenon can be found in [19–22]. An example of the decay and snapback of the main dipole's sextupole component is shown in Figure 1.14. The decay during the injection plateau and the fast snapback is visible once the main dipole is ramped up to nominal settings.

The decay and snapback among other effects, has a significant impact on particle stability and thus on losses inside the machine. For example, the b_3 field components causes a swing of the chromaticity, which is the focusing of the machine for off-momentum particles, of up to about 150 units, whereas a stable machine operation requires a chromaticity in the order of 2 to 10 units. If un-compensated, this large swing of chromaticity would inevitably lead to instabilities and particle losses. The same effect also affect the other multipole field components and beam parameters.

In order to guarantee a stable beam and long beam lifetimes, the level of control of magnetic field errors depends on the specific component but is usually on the percent to per-mille level.

1.4 Luminosity and Control of Particle Losses

The maximum luminosity in the LHC is determined by the ability to control particle losses inside the machine. For design purposes, the following simplified luminosity approximation can be used

$$\mathscr{L} = \frac{N n_b^2 f_{rev}}{4\pi \cdot (\sigma^*)^2} \cdot F, \tag{1.8}$$

where n_b is the number of protons per bunch. The LHC revolution frequency $f_{rev} \approx 11\,\text{kHz}$ is determined by the machine circumference and the speed of light. F is a generic reduction factor including effects due to the non-zero crossing angle, hour-glass effect, non-perfect beam overlap and other effects. In the LHC,

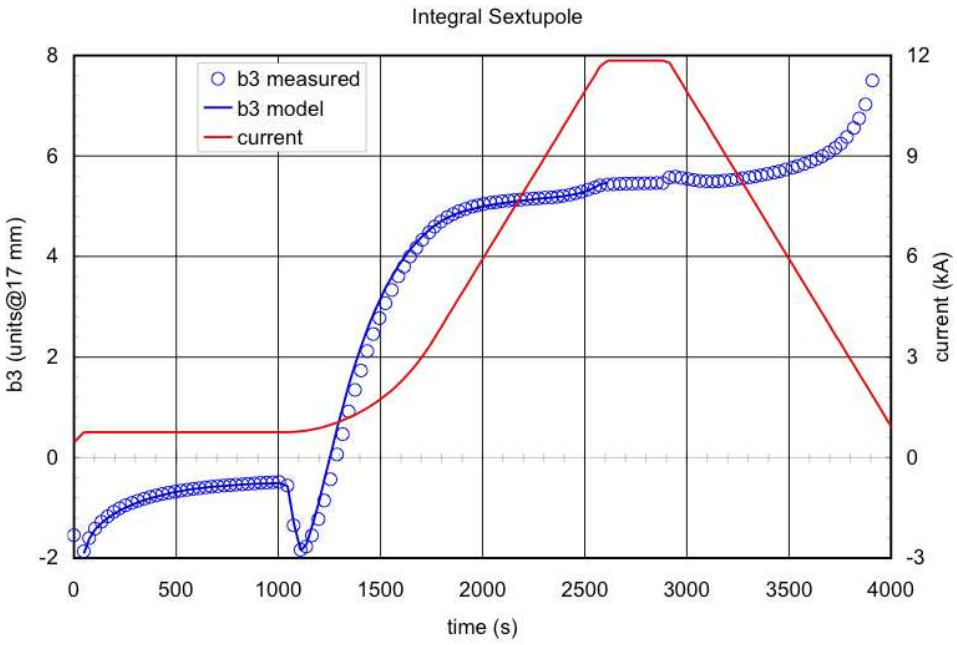


Figure 1.14: Example of the main dipole's sextupole field (b_3). The decay and snap-back are visible [23].

the most dominant reduction factor is due to the non-zero crossing angle θ_c that causes an imperfect overlap of the colliding bunches in the IP. The geometric reduction factor due to crossing can be written as

$$F_{crossing} = 1/\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2}, \tag{1.9}$$

with σ_z the bunch length ($\sigma_z \approx 1\,\text{ns} \cdot c \approx 30\,\text{cm}$).

As visible in equation 1.8, there are three possibilities to increase total luminosity: reducing the beam size σ^* at the IP, increasing the number of protons per bunch n_b and increasing the number N of colliding bunches in the machine.

The focus to smaller beam sizes is limited by the maximum available strength of the final focus quadrupole magnets and the *hour-glass* effect, which describes the effect of luminosity reduction as soon as the transverse beam size in the IP is much smaller than the bunch length. The present luminosity optimisation and most of the other proposed luminosity upgrade plans essentially establish or increase the total stored intensity $N \cdot n_b$ per beam. In its nominal configuration, the LHC will store up to $N = 2808$ bunches with each about $n_b = 1.15 \cdot 10^{11}$ protons. This corresponds to a total stored intensity of about $3 \cdot 10^{14}$ protons or a beam current of about 0.5 A per beam.

1.4.1 Stored Energy

With an increase of the stored intensity, particle loss becomes more critical. For nominal beam parameters the $3 \cdot 10^{14}$ protons with a particle momentum of 7 TeV yield a total stored energy of about 350 MJ per beam. Figure 1.15 shows a Livingston-style plot that relates present and previous particle accelerators with respect to their stored beam energy and particle momentum. The maximum momentum as well as the stored beam energy of the LHC exceeds all prior and present colliders. It is worthwhile to note that the LHC is also the first machine where even an injected nominal beam is sufficient to cause damage to accelerator or detectors. The damage threshold for copper that has been experimentally verified in [24,25] is indicated. It is important to note that the threshold is very sensitive to the input geometry. The so-called *pilot bunch* is a single bunch consisting of about $5 \cdot 10^9$ protons, and is the lowest foreseen intensity that will be injected and can be observed by the LHC beam instrumentation. The analysis presented in [24, 25] indicates that even this very low intensity may possibly cause material damage if accelerated to 7 TeV and lost due to machine failure or beam instability. It was realised at an early stage that the detection, prevention and capture of failures that may release large parts of the energy, stored in the magnet and in the beam itself, in an uncontrolled manner may potentially limit the maximum possible LHC performance. Hence an intensive research and development programme was dedicated to these issues.

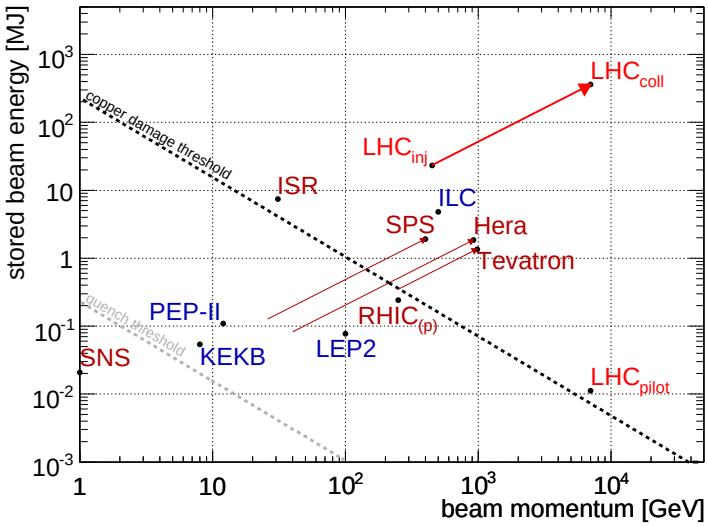


Figure 1.15: Livingston-style plot of the stored energy versus particle beam momentum of selected prior, present and future lepton (blue) and hadron (red) colliders. The equivalent LHC quench and damage level for copper are indicated. The arrows mark the operational range for circular colliders for injection to collision energy. The LHC exceeds all prior and present colliders not only by maximum momentum but also by stored beam energy. It is also the first machine where even an injected nominal beam is sufficient to cause damage.

At least two classes of particle losses are distinguished for the LHC:

1. fast or accidental beam losses
2. slow particle losses

The protection against the first type of loss is treated by the LHC Machine Protection System whereas the second is handled by the LHC Cleaning System.

1.4.2 Machine Protection System

The LHC Machine Protection System is designed to minimise the impact of accidental or fast proton losses into the machine that could cause severe damage and potentially compromise the function of the accelerator and experiments. The system contains a large number of redundant passive and active devices that, in combination, ensures safe operation. The machine protection consists of two classes of systems. The sensors of both classes are used as indicators to predict whether the system is under verge or already subject to a failure that could lead to a quench and potential damage of the machine. In case of a warning signal, these systems either directly or indirectly trigger a fast beam energy extraction that would otherwise damage the machine.

The first class is related to the protection of individual magnets. The stored energy in the magnetic field is about 11 GJ and has to be extracted on a timescale of tens of seconds in order to avoid damage to the magnets themselves. The magnet inherent protection consists of quench detection sensors on the superconducting coils and a large ohmic resistance. In case of a failure, this high ohmic resistance is switched into series with the magnets and that quickly and safely extracts the stored energy in the magnet that would otherwise lead to melting of the coils. See [18] for details. For redundancy and reliability of the protection system, the function of certain critical systems is complemented by active monitoring, such as the current in critical magnets or the correct functioning of the beam extraction system. For example, the warm D1 magnets in the experimental insertion are equipped with fast magnet current decay monitors, that in case of a current drop may directly trigger a beam dump. The measurement system is described in [26].

The second class of protection systems is beam related and deals with the fast extraction of the stored energy in the beam of about 350 MJ per beam, which has to be removed as fast as in one turn ($\approx 89\,\mu$s). It consists of a combination of passive and active devices. The passive devices are particle absorbers placed in critical locations such as the TDI in the injection insertions IR8 and IR2 that are likely to prone to intercepting unplanned particle loss. In case the speed of the failure is beyond the scope of the active

protection devices (e.g. one turn failures), these devices would sacrificially absorb the beam in favour of more sensitive accelerator equipment that they shadow. The active system relies on beam measurements using a wide range of sensors such as more than 4000 beam loss monitors (BLM) detecting increased particle losses outside the magnets, measurement of the beam current and its distribution, beam position monitors and other beam diagnostics connected to Beam Interlock Controller. These devices collect and transmit the triggers for beam extraction to the beam dumps installed in IR6.

The design of these systems is very demanding and requires a thorough understanding of the physics of accelerators, its hardware and operation. Though the beam stability and control issues described in this thesis are closely related to machine protection, a detailed description of the whole machine protection system is in its complexity, beyond the scope of this thesis. Further details can be found in [27, 28].

1.4.3 Beam Cleaning System

The LHC Beam Cleaning System minimises slow particle losses and consists of a set of collimators installed close to the beam intercepting particles with large amplitudes that may otherwise be lost into the cryogenic aperture, and induce quenches or damage the machine.

It has been shown that a single stage system, consisting of only a few primary collimators with very high density material that would intercept nearly all particles diffusing into large amplitudes, may not be able to survive a direct impact or high beam losses into the collimator jaws [29]. In addition, such a system does not provide a sufficiently high absorption. Thus it was decided to implement a two-stage cleaning system. In the two-stage collimation approach, each stage either absorbs or scatters a given number of particles to higher amplitudes creating a secondary beam halo that is dealt with the successive stages as illustrated in Figure 1.16. In the LHC, the primary collimators consist of about 0.2 m long

are

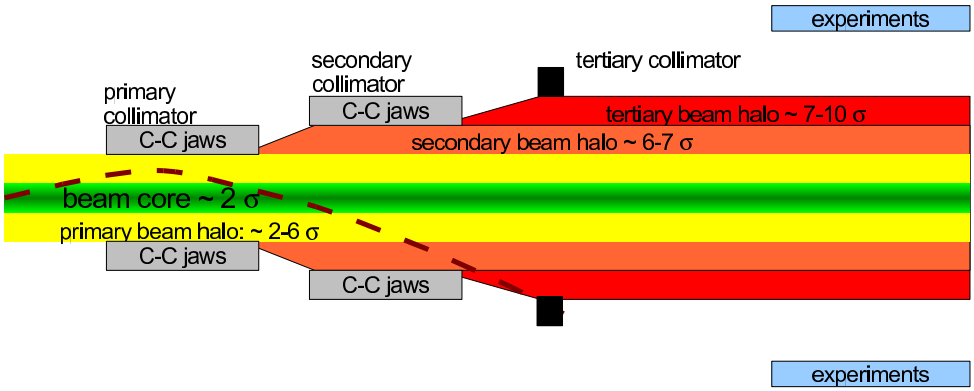


Figure 1.16: LHC beam cleaning system schematic: the system design aims to minimise the particles loss into the cryogenic mass. Beam particles that diffuse to high transverse amplitudes may either be intercepted or scattered by the primary collimators. The scattered particles may reach the secondary collimators and again are either absorbed or scattered and eventually captured by the tertiary collimator.

carbon-carbon compound jaws that are as close as about $5.7\,\sigma$ to the beam, with σ being the local beam size. The 1 m long secondary collimators are retracted by about one beam size and located at $6.7\,\sigma$. They either absorb the secondary halo particles or scatter them into a tertiary halo. The system has to ensure that the remaining particles escaping the cleaning system are below the number creating a quench of the magnet. Two sub-categories are distinguished: the momentum cleaning with collimators installed in IR3 that intercept un-captured and off-momentum particles, and the beta cleaning system located in IR7 that intercepts particles with large betatron oscillation amplitudes. Additional tertiary collimators are installed in certain regions in order to improve the local cleaning for special sensitive equipment (such as the final focus triplet) and to provide additional protection against accidental beam loss by acting as passive absorbers.

The performance of the LHC collimation system is given by its cleaning inefficiency η as defined by the ratio between the particles escaping the cleaning system and the number of particles impacting on to the primary collimator jaws:

$$\eta = \frac{\#\text{ particle escaping the cleaning system}}{\#\text{ particle impacting onto the primary collimator } (\sim 5.7\sigma)} \tag{1.10}$$

As discussed in Section 4.5.1, the cleaning inefficiency of the LHC Collimation System depends critically on the alignment of the primary and secondary collimator jaws with respect to beam and machine optics.

1.5 Beam Stabilisation

Both machine protection and collimation require active control of beam parameters such as beam orbit, energy and other parameters, for proper functioning of the LHC. To counteract disturbances due to decay and snapback of multipole moments, misalignments, ground motion and other dynamic effects, the control of the key beam parameters – orbit, tune, chromaticity and energy – will be an integral part of LHC operation. Manual correction of these parameters may soon reach their limit with respect to required precision and expected time-scales of the expected perturbation.

The LHC orbit feedback system, as the largest and most complex feedback system, consists of 1056 monitors grouped into 64 beam position monitor (BPM) front-end crates with each up to 18 monitors, a central feedback controller and 1060 individually powered orbit correction dipole magnets controlled by about 50 power converter gateways. The individual systems that are distributed over the 27 km ring circumference are connected through the LHC technical network which is based on a fast non-blocking Gigabit Ethernet network technology with fast store-and-forward switches. The orbit feedback architecture is shown in Figure 1.17. Based on this analysis, it was decided to implement a global

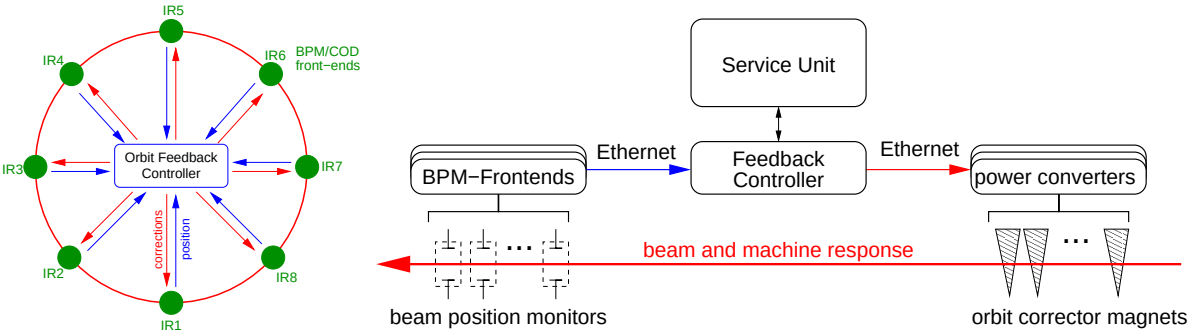


Figure 1.17: Schematic Feedback Overview

orbit correction with local constraints. The pre-processed beam position information is sent through the technical LHC network to the central controller that computes the appropriate corrections in time and space that will stabilise the beam on its reference position and send the new corrector strengths to the converter powering the correction dipole magnets. This automated measurement and correction procedure is repeated periodically at 10 Hz up to a maximum rate of 50 Hz. The orbit feedback system depends on beam parameters and conditions such as bunch systematics, machine imperfections and the state and dynamics of the over 3000 devices involved. In order to provide a high availability and benefit of this system during operation, the reliability and robustness was an important consideration for the design choices of this system.

This thesis presents an analysis of the baseline requirements and systematics to machine imperfections and element failures of the proposed orbit and energy feedback systems.

Chapter 2

Acceleration and Guiding of Charged Particle Beams

This chapter summarises the basic principles and formalism required to describe longitudinal and transverse particle dynamics and its control in a circular accelerator. Further information can be found in [17, 30–33].

2.1 Maxwell Equations

The acceleration, circular bending, and focusing of charged particles is described by interactions of charged particles with electro-magnetic fields. The dynamics of particles with velocities close to the speed of light is well described by the Maxwell equations that are given in their co-variant form as:

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \quad \text{and} \quad \partial_\mu \mathcal{F}^{\mu\nu} = 0 \ , \tag{2.1}$$

with $J^\nu = (\varrho, \vec{j}) = \varrho_0 U^\nu$ being the four-current, ϱ the charge density, $(U^\nu) = \gamma(c, \vec{v})$ the four-velocity and $F^{\mu\nu}$ and $\mathcal{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta}$ are the (dual) electromagnetic field tensors ($\epsilon_{\alpha\beta\mu\nu}$: *Levi-Civita* tensor):

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix} \quad \text{and} \quad (\mathcal{F}_{\mu\nu}) = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \tag{2.2}$$

B_i and E_i are the components of the magnetic flux density and the electric field, respectively. The metric tensor $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ is chosen. In this representation the co- and contravariant four-gradient are given as:

$$(\partial_\alpha) = (\frac{1}{c}\partial_t, \nabla) \quad \text{and} \quad (\partial^\beta) = (\frac{1}{c}\partial_t, -\nabla) \tag{2.3}$$

2.1.1 Lorentz Transformation

An intrinsic property of equation 2.1 is the invariance under Lorentz transform which in a general form can be written as

$$\Lambda(\vec{\beta}) = \begin{pmatrix} \gamma & -\gamma\beta_1 & -\gamma\beta_2 & -\gamma\beta_3 \\ -\gamma\beta_1 & 1 + \frac{(\gamma-1)\beta_1^2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} \\ -\gamma\beta_2 & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & 1 + \frac{(\gamma-1)\beta_2^2}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} \\ -\gamma\beta_3 & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} & 1 + \frac{(\gamma-1)\beta_3^2}{\beta^2} \end{pmatrix} \tag{2.4}$$

with the boost vector $\vec{\beta} = (\beta_1, \beta_2, \beta_3)^T = \vec{v}/c$ from the initial to the boosted frame of reference. The electromagnetic fields observed in the boosted frame of reference are given by the electro-magnetic field tensor $\mathbf{F}$, defined in equation 2.2 that transforms inside the new reference system to ($\mathbf{F}'$):

$$\mathbf{F}' = \Lambda \mathbf{F} \Lambda^T \tag{2.5}$$

After rearranging and regrouping the equation, it can be shown ([31]) that in the boosted reference system, the magnetic flux density $\vec{B}$ and electric field $\vec{E}$ are observed as:

$$\vec{E}' = \gamma(\vec{E} + \vec{\beta} \times \vec{B}) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}(\vec{\beta} \cdot \vec{E}) \tag{2.6}$$

$$\vec{B}' = \gamma(\vec{B} - \vec{\beta} \times \vec{E}) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}(\vec{\beta} \cdot \vec{B}) \tag{2.7}$$

Applying this relation to the field of a single charge travelling close to the speed of light, it is visible that in the laboratory frame of reference the longitudinal electrical field component of the moving charge appears squeezed into the transverse plane as illustrated in Figure 2.1.

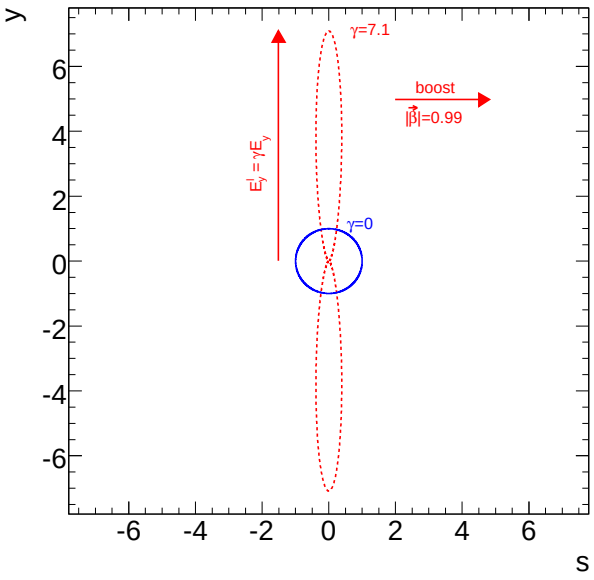


Figure 2.1: Equipotential surface in charged particle's frame of rest (blue) and as observed in $\beta = v_s/c = 0.99$ boosted frame of reference. The transverse field enhancement by $\gamma = 7.1$ is visible. For LHC energies ($\gamma > 400$), the electrical field is compressed into discs travelling along the individual protons.

In the LHC, the relativistic γ is above 480, which corresponds to an opening angle of the radiation cone of less than 2 mrad. One can assume that the longitudinal electrical field of the protons is effectively compressed into a disc with only transverse electric field components.

2.1.2 Lorentz Force

Generally, forces are not necessarily preserved under Lorentz transformation and must be handled with care when changing the frame of reference. The Lorentz force $\mathbf{F}'_L$ describes the force on a charged particle inside an electromagnetic field $F^{\mu\nu}$ and can be written as

$$\mathbf{F}'_L = \frac{d\mathbf{p}}{d\tau} = (q \cdot F^{\mu\nu} U_\nu), \tag{2.8}$$

with U_ν being the (four) velocity, q the charge, and τ the proper time of the particle. Since $\mathbf{p} = (E/c, m\gamma\vec{v})$, the change of total energy $\mathrm{d}E/\mathrm{d}\tau$ of the particle is given by the first component of the Lorentz force:

$$\frac{\mathrm{d}E}{\mathrm{d}\tau} = c \cdot F^0_L = qc \cdot \gamma\, \vec{v}\vec{E} \tag{2.9}$$

From equation 2.9, it is visible that the particle's total energy only increases by the electric field $\vec{E}$ that acts along the direction of flight $\vec{v}$ in the particle's frame of reference. A constant magnetic flux $\vec{B}$ leaves the energy constant. However, it may change the direction of momentum of the particle.

There are two common ways to apply electric fields to charged particles for acceleration: The first method is directly through radio-frequency (RF) cavities that provide a longitudinal acceleration voltage across a gap that the particle traverses. The second method is through induction as first introduced in the BETATRON [34]. The latter principle has been revived under the name '*induction devices*' and is intended for future use in high current accelerators under study at the KEK laboratory [35, 36].

These devices are magnets that generate a transverse fast-changing magnetic flux density $\left|\partial_t \vec{B}\right| \neq 0$ that causes an electric field along the beam axis, as described by the Maxwell equation $\nabla \times \vec{E} = -\partial_t \vec{B}$. To first order, the acceleration voltage dU across the induction device per turn is given by:

$$dU = -\int_{dS} \vec{E} \mathrm{d}\vec{s} = -\iint_A \frac{\mathrm{d}\vec{B}}{\mathrm{d}t} \mathrm{d}\vec{A} \quad \Rightarrow \quad \Delta U \approx -\frac{\Delta \vec{B}}{\Delta t} \cdot A \tag{2.10}$$

where dS is the border of the surface $A = l_{mag} \cdot r$ that is penetrated by the magnetic flux density, l_{mag} being the effective magnetic length and r the aperture of the magnet.

In contrast to the BETATRON that simultaneously uses the changing magnetic field to store and to accelerate particles, this effect is negligible for the LHC. Though the LHC will ramp the 1232 main dipole magnets from about 0.53 T to about 8.3 T within 20 minutes ($l_{mag} = 14.3$ m and $r \approx 28$ mm), the cumulative effect of all main dipole magnets would yield only an accelerating voltage

$$\Delta U_{ind.} \approx 3\,\mathrm{V} \tag{2.11}$$

per turn, which is very small compared to the energy gain of 485 keV applied by the LHC RF cavities [37] and compared to the energy loss due to synchrotron radiation of about 7 keV per turn for a proton of 7 TeV. However, this voltage has a significant effect on creating eddy currents. Due to the very low resistivity inside superconducting magnets, these eddy currents can become very large and drive imperfections of the magnetic fields, as discussed in Section 4.4.1. It is further assumed that the energy and velocity of the particles are constant inside magnetic fields.

2.2 Transverse Dynamics

2.2.1 Co-Rotating Coordinate System

Though charged particle dynamics is fully described by the Maxwell Equations and Lorentz Force (equation 2.8) respectively, a direct solution is rather cumbersome in the context of circular accelerators. Commonly, one introduces a more convenient local co-rotating coordinate system. This system uses the ideal particle's trajectory as reference for its origin, illustrated in Figure 2.2. Using this coordinate system the movement of the individual particles is thus treated as a small perturbation around the specified design trajectory.

By definition, the local axis x always points outwards and the axis y points upwards, with respect to the closed trajectory. The local longitudinal axis s is parallel to the reference particle's propagation direction. This coordinate system assumes implicitly the periodicity of the local coordinate x and y, the local bending radius ϱ and various multipole strengths $b_n(s)$

$$x(s+C) = x(s), \quad y(s+C) = y(s), \quad \varrho(s+C) = \varrho(s), \quad \text{and} \quad \mathrm{b_n(s+C) = b_n(s)} \tag{2.12}$$

with C being the accelerator circumference. Unless otherwise stated, primed notation x' is used to indicate a differentiation with respect to the longitudinal coordinate s and $\dot{x}$ denotes a differentiation in time t.

$$x' := \frac{\partial x}{\partial s} \qquad \text{and} \qquad \dot{x} := \frac{\partial x}{\partial t} \tag{2.13}$$

In the most general form, the full particle state is given as a coordinate in the six-dimensional phase space including the transverse local coordinates (x,y), their derivatives (x',y'), the particle's energy E and longitudinal phase ψ with respect to the reference particle.

$$\underline{x} = (x, x', y, y', E, \psi)^T \tag{2.14}$$

2.2.2 Hill's Equation

Using this local co-rotating coordinate system and keeping only leading terms, it is possible to reduce equation 2.8, which describes the particle dynamics in an accelerator by a Hill's type[1] equation that can

[1] The homogenous part of equation 2.15 was first introduced by HILL ([38]) to describe the three body problem of the orbital trajectory of the moon around the sun under the periodic influence of Earth.

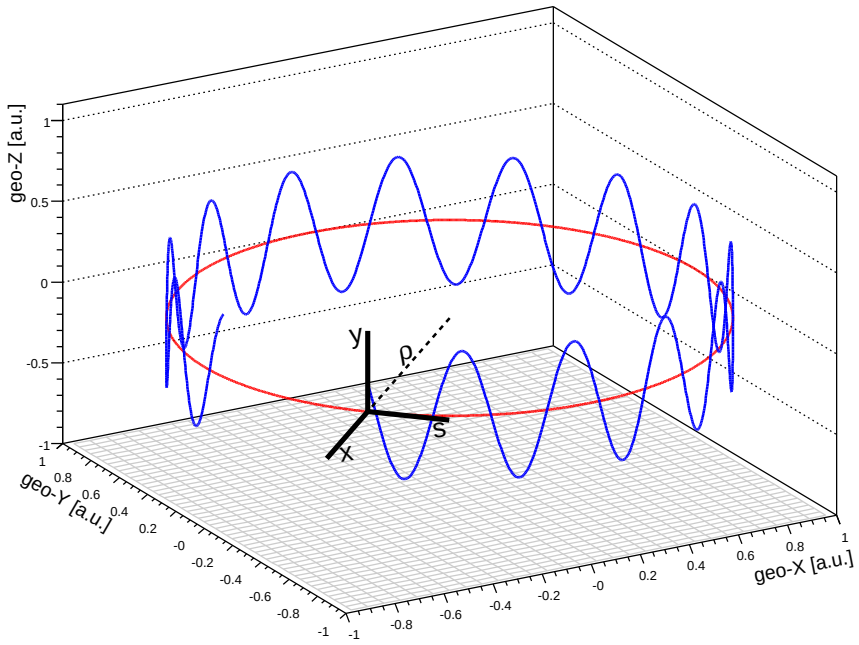


Figure 2.2: Co-rotating local coordinate system. The circular design trajectory (red) and an exemplary orbit perturbation (blue) is shown. The local bending radius $\varrho(s)$ is a function of the path length s inside the accelerator.

be written as:

$$\begin{cases} x'' & +(k(s)+\dfrac{1}{\varrho^2(s)})\cdot x & = & \dfrac{1}{\varrho(s)}\dfrac{\Delta p}{p} \\ y'' & -k(s)\cdot y & = & 0 \end{cases} \tag{2.15}$$

with $\varrho(s)$ being the local bending radius of the main dipoles and

$$k(s) = \frac{dB/dx}{p} = \frac{dB/dx}{(B\varrho)} = -\frac{dB/dy}{(B\varrho)} \tag{2.16}$$

the normalised local quadrupole strength and $\Delta p/p$ the particle momentum offset with respect to the reference particle. The quadrupole strength $k(s)$ is positive for focusing quadrupoles and negative for defocusing. The multipole strengths are commonly normalised by the *magnetic rigidity* $(B\varrho) = p/q$ in order to make the equations independent from the design particle's absolute momentum p and charge q. As a reminder, a positive $k(s)$ quadrupole in one plane is intrinsically negative in the other plane but with the same magnitude. Thus the quadrupoles are usually arranged to so called *FODO* cells with alternating focusing (F) and de-focusing (D) quadrupoles in order to provide focusing in both planes, as illustrated in Figure 1.5.

Betatron Oscillation

The homogeneous part of equation 2.15 is commonly solved using *Floquet's Theorem* ([39]) through the following oscillatory ansatz

$$z_\beta(s) = \sqrt{\epsilon_z \beta_z(s)}\cdot\cos(\Delta\mu_z(s)+\phi_z), \tag{2.17}$$

with z being either the x or y coordinate and ϵ_z and ϕ_z the initial particle conditions. Where required the optical functions and parameters are marked with subscripts (e.g. β_x and β_y) to distinguish between planes, if not they are omitted. These types of particle oscillations were first observed in the Betatron accelerator [34, 40] and are thus called *betatron oscillations* and the positive periodic function $\beta(s)$ accordingly referred to as *betatron function*. The betatron function is, by definition, always positive and – in a regular FODO structure – always maximum at the focusing magnets and minimum at defocusing magnets. The *betatron phase advance* $\Delta\mu(s)$ is dependent on the betatron function and defined as

$$\Delta\mu(s) = \int_{s_0}^{s} \frac{1}{\beta(s')}\mathrm{d}s' \tag{2.18}$$

with s_0 being an arbitrary reference location in the ring and C its circumference. By definition $\Delta\mu$ is monotonic increasing. As a measurable observable, one defines the number of transverse oscillations a particle describes per turn as *tune* Q of the machine, which is related to the total phase advance of the machine through:

$$Q = \frac{1}{2\pi}\int_0^C \mu(s)\mathrm{ds} \tag{2.19}$$

In order to avoid resonances due to magnetic imperfections and consequently amplitude growth of the particles' oscillations inside the machine, the tune is usually a non-integer and ideally an irrational non-fractional number. Additionally, the horizontal and vertical tunes usually differ in order to avoid coupling resonances between planes. One defines $\Delta = Q_x - Q_y - p$ as the fractional unperturbed tune-split, with p an integer. In the LHC, the tune-split is for example $\Delta = 64.28 - 59.31 - 5 = -0.03$ during injection and $\Delta = 64.31 - 59.32 - 5 = -0.01$ during collision. As shown in [33], it is further required to avoid the following resonance condition

$$mQ_x + nQ_y = p \tag{2.20}$$

with m, n, and p being arbitrary integers. The order of the resonance is given by $|m + n|$.
Generally, with exception of a few special cases, the betatron function and phase advance cannot be solved analytically but can be numerically estimated for given locations in the accelerator using tools such as MAD or others described in [41–43]. The numeric values are referred to as *twiss parameter*[2] or *lattice* of the accelerator. The LHC arc lattice is shown in Figure 2.3. The FODO periodicity, horizontal-vertical and maxima-minima asymmetry at the quadrupole locations are visible. In addition to the betatron

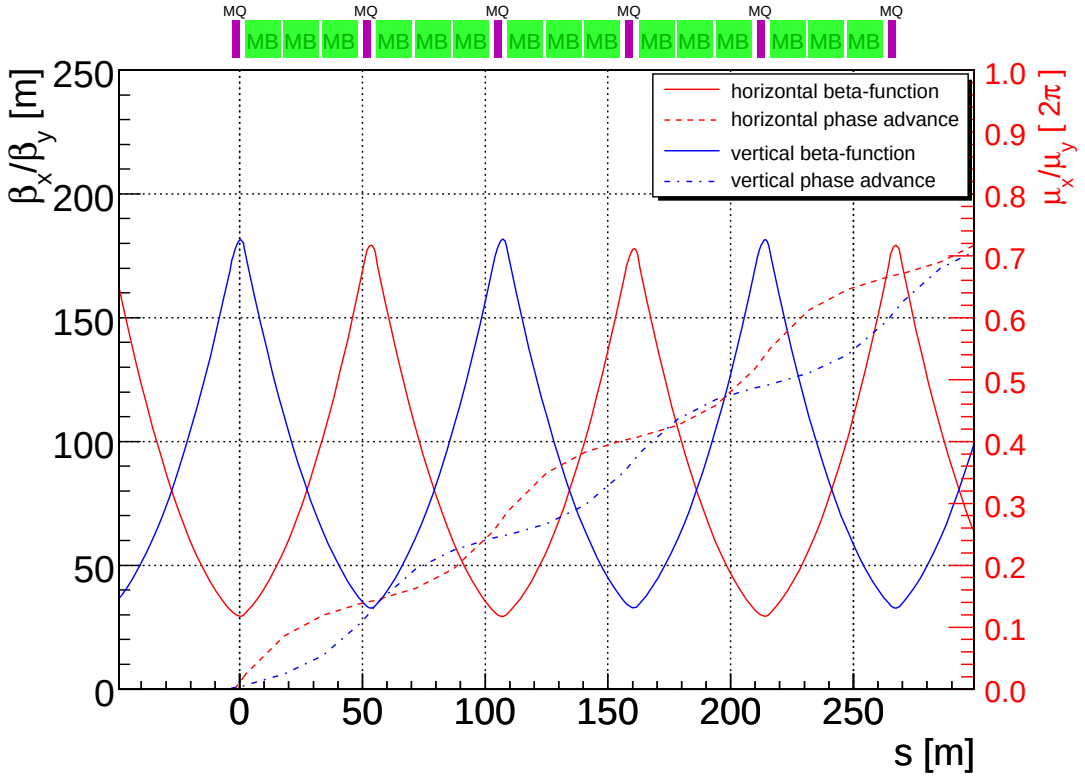


Figure 2.3: Detail of the LHC arc optics: The betatron function and phase advance of the horizontal and vertical plane are plotted as a function of longitudinal position in the ring. The periodicity of the FODO structure and asymmetry with respect to horizontal and vertical plane is visible. The betatron function is maximal (minimal) at the location of the quadrupoles.

function and phase advance, one commonly defines the following two functions:

$$\alpha(s) := -\frac{1}{2}\beta'(s) \qquad \text{and} \qquad \gamma(s) := \frac{1+\alpha^2(s)}{\beta(s)} \tag{2.21}$$

[2]The 'twiss parameters' are named after R.Q. Twiss, who first introduced them for describing particle motions in proton synchrotrons [44].

Intrinsic to the quadrupole focusing, $\alpha(s)$ vanishes inside the quadrupoles. The definitions of $\beta(s)$ and $\gamma(s)$ differ from the relativistic $\beta_{rel} = \frac{v}{c}$ and $\gamma_{rel} = (1-\beta^2)^{-0.5}$ function definitions. Where required, the relativistic parameters are denoted through subscripts. Using this definition, the first derivative of equation 2.17 can thus be written as

$$z'_\beta(s) = -\frac{\sqrt{\epsilon}}{\sqrt{\beta(s)}}\left[\alpha(s)\cos(\Delta\mu(s)+\phi) + \sin(\Delta\mu(s)+\phi)\right] \tag{2.22}$$

Momentum Spread and Dispersion

The above expressions are suited for the dynamics of a single charged particle with reference momentum p. When describing multiple particle dynamics, it is important to take the beam momentum distribution into account. The momentum-wise sorting of the charged particles is an intrinsic property of dipoles and quadrupoles. One distinguishes two momentum effects, the momentum spread of the beam distribution that defines the beam width and the systematic average beam momentum shift.

The shift of the average beam momentum yields a shift of the equilibrium beam position, which is determined by the bending radii of the main dipole magnets as expressed by the inhomogenous part of equation 2.15. The periodic so-called *dispersion function* $D(s)$, usually defined for a normalised momentum error of $\Delta p/p = 1$, is a special solution of the inhomogenous equation. The general solution of Hill's equation for a particle with momentum offset $\Delta p/p$ is thus given by the superimposition of the dispersion orbit $x_d(s)$

$$z_d(s) = D(s)\cdot\Delta p/p \tag{2.23}$$

and the betatron oscillation given in equation 2.17:

$$z(s) = z_\beta(s) + D(s)\cdot\frac{\Delta p}{p} \tag{2.24}$$

Figure 2.4 shows the LHC dispersion function as a function of position in the machine. The horizontal dispersion function reaches values between 1 and 2 m in the arcs and is maximum in the momentum cleaning insertion located in IR3, with values up to about ±2.7 m. As described in Section 1.2.1, the minimising effect of the dispersion suppressors inside the insertion regions is visible. Due to the absence of strong bending magnets in the vertical plane, the vertical dispersion function is usually very small, and in case of the LHC, in the order of a few centimetres.

Chromaticity

A second energy-related effect is due to the beam momentum spread. In equation 2.16, one sees that the quadrupole focusing and thus the transverse oscillations and tune depend on the individual particle momentum. In addition to the momentum shift and dispersion function related shift, individual particles with larger individual momentum experience less deflection and consequently have a smaller tune than the reference particle. Consequently particles with a smaller momentum will have a larger tune. The tune spread due to the momentum dependant focusing is in first order proportional to the momentum spread. The *chromaticity* Q', another observable inside an accelerator, is dimensionless and is commonly defined as the ratio between the relative tune spread $\Delta Q/Q$ and relative momentum spread $\Delta p/p$

$$Q' = \frac{\Delta Q/Q}{\Delta p/p} \tag{2.25}$$

with Q and p being the tune and momentum of the reference particle. The same definition applies to systematic momentum and tune shift, respectively. The *natural chromaticity* Q'_{nat} refers to the chromaticity the machine would have without any sextupole or higher multipole magnets contributions and is given by the integral:

$$Q'_{nat} := \frac{\Delta Q}{\Delta p/p} = -\frac{1}{4\pi}\oint k(s)\beta(s)ds \tag{2.26}$$

Since the quadrupole strength is positive at focusing quadrupoles, which coincides with large values of the betatron function in contrast to small betatron functions at defocusing quadrupoles that have negative focusing strengths, the integral is always usually positive and thus the natural chromaticity always negative. The LHC for example has a natural chromaticity of about:

$$Q'_x \approx -140 \tag{2.27}$$

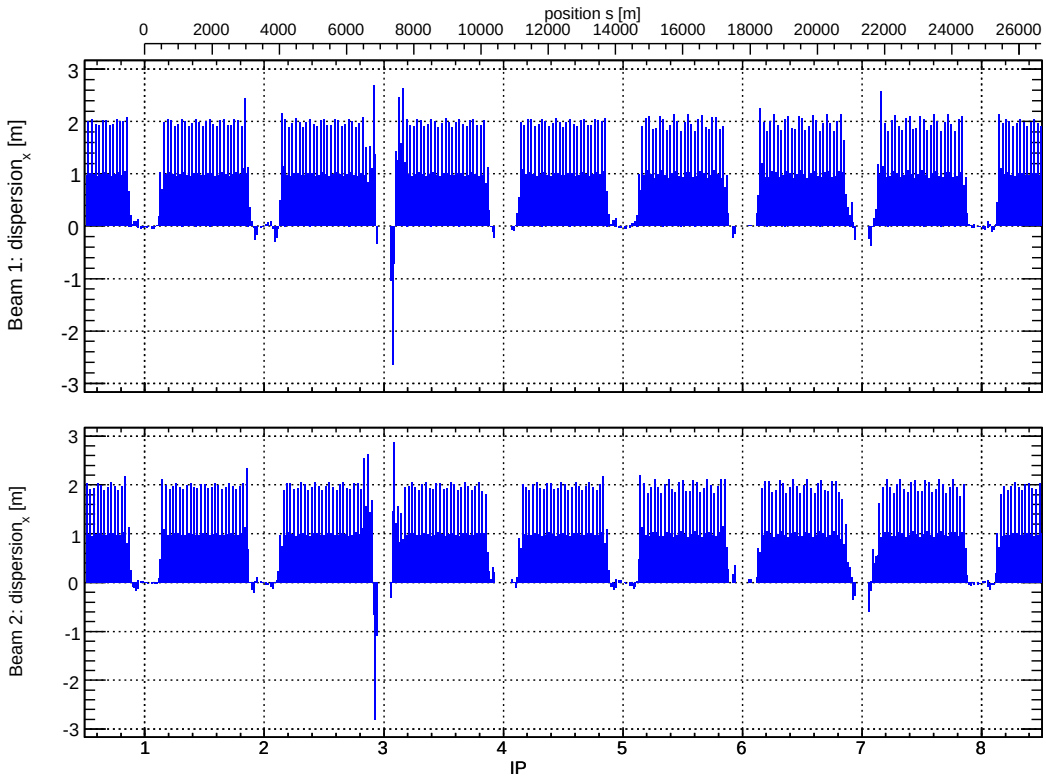


Figure 2.4: Horizontal LHC dispersion. The maximum horizontal dispersion of about $\pm 2.7\,\mathrm{m}$ is reached in the momentum cleaning insertion. With exception of the momentum cleaning in IR3, the effect of the dispersion suppressors on the dispersion function inside the insertion regions is visible. The vertical dispersion is (not plotted) usually in the order of a few centimetres.

The natural chromaticity is usually compensated using sextupole magnets that give a quadratic amplitude dependent focusing. These magnets are installed in regions with large dispersion functions in order to make use of the sorting in amplitude of the particles according to their momentum. The total chromaticity including the sextupole magnets is given by

$$Q' = \mp \frac{1}{4\pi} \oint \beta(s) \left[k(s) + m(s)D(s)\right] ds \tag{2.28}$$

with $m(s) = \frac{e}{p}\frac{\partial^2 B}{\partial x^2}|_{0,0}$ being the normalised sextupole strength, $D(s)$ the dispersion and $\beta(s)$ the betatron function. The sign depends on the plane: '-' for the horizontal and '+' for the vertical plane as quadrupoles ($k(s)$) and sextupoles ($m(s)$) are focusing in one and at the same time defocusing in the other plane

The relative beam momentum spread $\Delta p/p$ is usually in the order of 10^{-4}, which combined with the usually large natural chromaticity, leads to a relative tune spread $|\Delta Q/Q|$ in the order of 10^{-2}. This rather large tune spread would inevitably force some particles to oscillate on some high-order resonances, eventually leading to growth of amplitude and particle loss. Thus for reasons of stability, the chromaticity is usually set to be around zero and slightly positive to avoid head-tail instabilities above transition. The total chromaticity Q' is, as visible in equation 2.28, controlled through the sextupole magnets that are installed into dispersive regions and distributed over the ring. In the LHC, a likely working point for the chromaticity is between 2 and 10 and is required to be controlled on the level of ± 1 for nominal operation [45].

Beam Size

Another interpretation of the parameter ϵ is given, using equations 2.17 and 2.22, as the *transverse emittance* of the beam described by the surface in phase space occupied by the particle or beam, respectively:

$$\gamma(s)z^2(s) + 2\alpha(s)z(s)z'(s) + \beta(s)z'(s) = \epsilon \tag{2.29}$$

Further, the emittance is proportional to the total transverse momentum of the particle. Using this emittance definition the beam size can be written as:

$$\sigma = \sqrt{\epsilon\beta(s)} = \sqrt{\frac{\epsilon^*\beta(s)}{\gamma_{rel}}} \tag{2.30}$$

with $\epsilon_z^* = \epsilon \cdot \gamma_{rel}$ being the normalised beam emittance, which ideally in first order is a constant for protons passing through different accelerators.

2.2.3 Matrix Formalism

An alternative description to Hill's equation is the use of transfer functions that map the individual particle state given in equation 2.14 from one accelerator element to the next. This method is also referred to as *particle tracking*. Discussing transverse fields only, the general field description $\mathbf{B}(x, y)$ can be expanded using a two-dimensional Taylor series around $(0, 0)$. The terms normalised to the energy of up to an order of two are given in Table 2.1:

	regular		**skew**	
	$\frac{e}{p}B_x(x,y)$	$\frac{e}{p}B_y(x,y)$	$\frac{e}{p}B_x(x,y)$	$\frac{e}{p}B_y(x,y)$
Dipole	0	ϱ_x	ϱ_y	0
Quadrupole	$-k \cdot y$	$+k \cdot x$	$+\kappa \cdot x$	$-\kappa \cdot y$
Sextupole	$m \cdot xy$	$\frac{1}{2}m(x^2 - y^2)$	$\frac{1}{2}m_s(y^2 - x^2)$	$-m_s \cdot xy$

Table 2.1: Magnetic multipole field expansion. The lowest terms up to an order of two are given and normalised to the reference momentum.

Here, k and κ are the normalised normal and skew quadrupole, and m and m_s the normalised normal and skew sextupole strengths, respectively:

$$k = \frac{e}{p} \cdot \frac{\partial B_x}{\partial y}, \quad \kappa = \frac{e}{p} \cdot \frac{\partial B_x}{\partial x}, \quad m = \frac{e}{p} \cdot \frac{\partial^2 B_x}{\partial y^2} \quad \text{and} \quad m_s = \frac{e}{p} \cdot \frac{\partial^2 B_y}{\partial x^2} \tag{2.31}$$

Commonly the field integral

$$\frac{e}{p}Bl_{mag} := \frac{e}{p}\int_s^{s+l} B(s)ds \tag{2.32}$$

is approximated by the magnetic length l_{mag} as an effective length which is defined so that the product of the average field B and magnetic length equals the integral of the real magnetic field $B(s)$ distribution over the geometric length l of the magnet.

Using equation 2.8 and the magnetic field decomposition in Table 2.1, the beam deflection angle δ while passing through a short magnet can be written as:

$$\sin(\frac{\delta_{x/y}}{2}) = \frac{l_{mag}}{2\varrho} = \frac{l_{mag}}{2} \cdot \frac{e}{p}B_{y/x} \tag{2.33}$$

For large accelerators, the deflection angle due to an individual magnet is usually very small and in the order of at most a few mrad. Hence, it is convenient to use a thin lens assumption and approximate $\sin(x) \approx x$ which simplifies equation 2.33 to:

$$\delta_{x/y} \approx \frac{l_{mag}}{\varrho} = l_{mag} \cdot \frac{e}{p}B_{y/x} \tag{2.34}$$

The magnetic bending in the horizontal plane is proportional to the vertical magnetic field and vice versa. It is important to note that, with exception of the dipole magnet, the field and thus the deflection is position-dependent. For example, it is visible that the deflection inside a quadrupole is given by

$$\delta_{quad}|_H = kl_{mag} \cdot x \tag{2.35}$$

$$\delta_{quad}|_V = kl_{mag} \cdot y \tag{2.36}$$

and for a normal sextupole given by:

$$\delta_{sext}|_H = ml_{mag} \cdot \frac{1}{2}(x^2 - y^2) \tag{2.37}$$

$$\delta_{sext}|_V = ml_{mag} \cdot xy \tag{2.38}$$

For the lower order dipole and quadrupole magnets, the deflection and particle transfer can in good approximation be cast into matrix form

$$\underline{x} = \mathbf{M}_i \underline{x}_0 \tag{2.39}$$

with $\underline{x}_0$ the initial and $\underline{x}$ the final particle state and $\mathbf{M}_i$ a 6x6 matrix describing the specific magnet or accelerator structure. Since magnets, apart from changing the path length discussed later, do not significantly affect the energy or phase of the particle, these matrices can generally be reduced to 4x4 matrices. Ignoring the coupling between planes, the matrices can be further reduced to a 2x2 size. In this approximation, the matrix describing a drift space of length l can thus be written as:

$$\mathbf{M}_{drift} = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \tag{2.40}$$

Neglecting the edge focusing effects, the transfer matrix of a dipole deflecting the beam by an angle δ is given by

$$\mathbf{M}_{dipole} = \begin{pmatrix} \cos\delta & \varrho \sin\delta \\ -\frac{1}{\varrho}\sin\delta & \cos\delta \end{pmatrix} \tag{2.41}$$

and of a focusing quadrupole can be written as

$$\mathbf{M}_{quadrupole} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{f} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & +\frac{1}{f} & 1 \end{pmatrix} \tag{2.42}$$

with the focusing length $f = \frac{1}{kl_{mag}}$ of the quadrupole. A 4×4 notation was exceptionally used in equation 2.42 to show the assymetry between the horizontal and vertical plane. For linear particle transfer, this matrix formalism is equivalent to Hill's equation and thus both can be related using a generic beam transfer matrix $\mathbf{M}$ ([17]) describing the beam transfer from any given point A (twiss: α_0, β_0) to another point B in the machine (twiss: α, β)

$$\mathbf{M}_{AB} = \begin{pmatrix} \sqrt{\frac{\beta}{\beta_0}}\left[\cos(\mu) + \alpha_0 \sin(\mu)\right] & \sqrt{\beta\beta_0}\sin(\mu) \\ \frac{(\alpha_0 - \alpha)\cos(\mu) - (1 + \alpha_0\alpha)\sin(\mu)}{\sqrt{\beta\beta_0}} & \sqrt{\frac{\beta_0}{\beta}}\left[\cos(\mu) - \alpha\sin(\mu)\right] \end{pmatrix} \tag{2.43}$$

where μ is the positive phase advance between A and B. An important case of equation 2.43 is the matrix describing one turn inside the machine that can be written as:

$$\mathbf{M}_{turn} = \begin{pmatrix} \cos(2\pi Q) + \alpha\sin(2\pi Q) & \beta\sin(2\pi Q) \\ -\gamma(s)\sin(2\pi Q) & \cos(2\pi Q) - \alpha\sin(2\pi Q) \end{pmatrix} \tag{2.44}$$

A full particle state tracking would then consist of a computation of the transfer matrix using only the twiss parameters of the origin to the point of interest and then a multiplication of the initial state. Higher multipole magnets can be incorporated through non-linear transfer functions interleaved with the regular matrix multiplication segments.

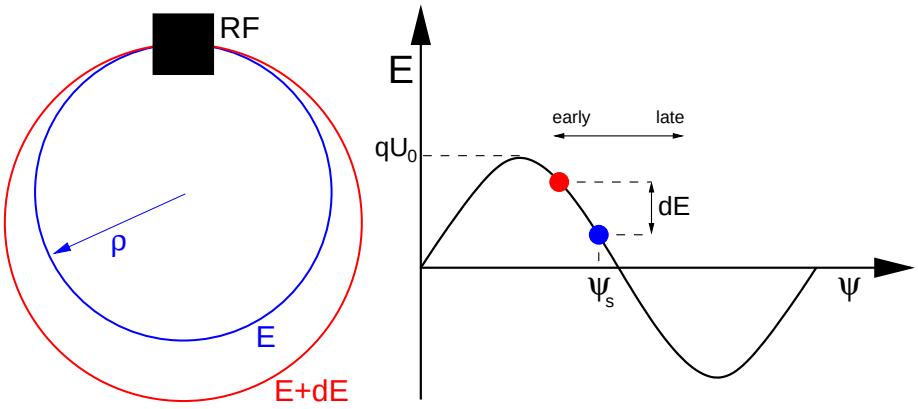
2.3 Longitudinal Dynamics

In contrast to linear accelerators where the charged particle passes only once through each cavity, circular accelerators make use of multiple traversing of the particle through the same cavity. For the following discussion, a simplified synchrotron model consisting only of a homogenous constant magnetic field B for bending and one radio frequency (RF) system is assumed, as illustrated in Figure 2.5.
The magnetic bending radius of the particle trajectory is thus given by:

$$\varrho = \frac{p}{q} \cdot \frac{1}{B} \tag{2.45}$$

Without RF system, the relative change of circumference $\Delta C/C$ is proportional to the relative momentum change $\Delta p/p$

$$\frac{\Delta C}{C} = \alpha_c \frac{\Delta p}{p} \tag{2.46}$$



(a) simplified accelerator model (b) RF energy gain

Figure 2.5: Simplified accelerator (left) and radio-frequency energy gain (right)

with α_c the momentum compaction factor which relates to the dispersion function:

$$\alpha_c = \frac{1}{C_0} \oint \frac{D(s)}{\varrho(s)} \mathrm{ds} \tag{2.47}$$

In case of the LHC, the momentum compaction factor is for example about $\alpha_c = 3.22 \cdot 10^{-4}$.

The revolution frequency f_{rev} defines the number of turns a particle circulates per second in the machine and depends only on the machine circumference C and particle velocity $v = \beta c$. Expanding the relative frequency change to

$$f_{rev} = \frac{v}{2\pi R} = \frac{\beta c}{C} \quad \Rightarrow \quad \frac{\Delta f}{f} = \frac{\Delta \beta}{\beta} - \frac{\Delta C}{C} \tag{2.48}$$

with $\Delta C/C$ being the relative change of machine circumference and $\Delta\beta/\beta$ the relative change of the particle velocity. From the relativistic momentum and energy definition one can similarly derive that:

$$p = \gamma m v = \beta\gamma \frac{E_0}{c} \quad \Rightarrow \quad \frac{\Delta p}{p} = (1-\beta^2)^{-1} \frac{\Delta \beta}{\beta} \tag{2.49}$$

$$E = \frac{c}{\beta} p \quad \Rightarrow \quad \frac{\Delta E}{E} = \beta^2 \frac{\Delta p}{p} \tag{2.50}$$

Inserting equations 2.46 and 2.49 into equation 2.48 yields:

$$\frac{\Delta f}{f} = \left(\frac{1}{\gamma^2} - \alpha_c \right) \cdot \frac{\Delta p}{p} = \eta \cdot \frac{\Delta p}{p} \tag{2.51}$$

or in terms of relative energy change (using equation 2.50) as :

$$\frac{\Delta f}{f} = \frac{1}{\beta^2} \left(\frac{1}{\gamma^2} - \alpha_c \right) \cdot \frac{\Delta E}{E} \tag{2.52}$$

In the ultra-relativistic limit ($\gamma^{-2} \approx 0$ and $\beta^{-2} \approx 1$) the revolution frequency changes linearly with the momentum in order to preserve a constant circumference. The zero-crossing of equation 2.52 is referred to as *transition* energy and only relevant for low energy machines. Since $\gamma_{tr} = \frac{1}{\sqrt{\alpha_c}} \approx 55.7 \Longleftrightarrow p \approx 52\,\mathrm{GeV/c}$ is smaller than the minimum injection momentum ($p_{inj} \approx 450\,\mathrm{GeV/c}$), the LHC operates always above transition.

The acceleration of the charged beam particles that is made by the LHC Radio Frequency System consists of two cryo-modules per beam with four RF cavities each, which in total can supply a maximum accelerating voltage U_0 of 16 MV. The main harmonic of the cavities is $f_{RF} = 400\,\mathrm{MHz}$ and is a multiple (harmonic number $h = 35640$) of the LHC revolution frequency $f_{rev} = 11245.5\,\mathrm{Hz}$. The harmonic number defines the maximum number of longitudinally stable regions, further referred to as (RF) *buckets*, that can be used to capture, store and accelerate protons. The collectivity of protons inside a bucket is called *bunch*. In the LHC, only about every tenth bucket will capture a bunch, which accordingly corresponds to a bunch spacing of about 25 ns or 7.5 m [3].

[3] The extraction kicker requires a gap of about 3 μs inside the bunch pattern which reduces the actual maximum possible number to 2808 bunches per LHC beam.

The energy gain per turn of a particle is dependent on its phase ψ with which it arrives at and passes through the RF system. By definition, in a *synchrotron*, the magnetic field of the main dipole magnets and the RF frequency are matched in order to maintain a constant circumference. The reference particle at nominal momentum arrives synchronously with a phase ψ_s.

As shown in Figure 2.5, the energy gain dE and phase advance per turn of any other non-synchronous particle having a phase ψ with respect to the synchronous particle can thus be written as:

$$\frac{\mathrm{dE}}{\mathrm{dt}} = f_{rev} \cdot qU_0 \cdot [\sin(\psi_s + \psi) - \sin(\psi_s)] \tag{2.53}$$

$$\frac{\mathrm{d}\psi}{\mathrm{dt}} = 2\pi h \Delta f_{rev} \tag{2.54}$$

Inserting equation 2.51 into 2.54 yields

$$\frac{\mathrm{d}\psi}{\mathrm{dt}} + 2\pi h f_{rev} \frac{1}{\beta^2} \left(\frac{1}{\gamma^2} - \alpha_c \right) \frac{\mathrm{dE}}{E} = 0 \tag{2.55}$$

and differentiation and inserting 2.53 results into

$$\frac{d^2\psi}{dt^2} + \left(\frac{2\pi h f_{rev}^2 qU_0}{\beta^2 E} \right) \cdot \left(\frac{1}{\gamma^2} - \alpha_c \right) \cdot [\sin(\psi_s + \psi) - \sin(\psi_s)] = 0 \tag{2.56}$$

This equation is non-linear in ψ and thus in general can not be solved analytically. However, for small phase deviations $\psi \ll 1$, one can linearise and approximate $\sin(\psi) \approx \psi$ which simplifies the equations to

$$\sin(\psi_s + \psi) - \sin(\psi_s) = \sin(\psi_s) \underbrace{\cos(\psi)}_{\approx 1} + \cos(\psi_s) \underbrace{\sin(\psi)}_{\approx \psi} - \sin(\psi_s) \approx \cos(\psi_s) \cdot \psi \tag{2.57}$$

Using this approximation, the longitudinal movement is reduced to a harmonic energy-phase oscillation around the reference particle

$$\frac{d^2\psi}{dt^2} + \Omega_s^2 \cdot \psi = 0 \quad \text{and} \tag{2.58}$$

$$\frac{d^2E}{dt^2} + \Omega_s^2 \cdot E = 0 \tag{2.59}$$

with the synchrotron frequency Ω_s and synchrotron tune Q_s given by:

$$\Omega_s = \pm 2\pi f_{rev} \cdot \sqrt{-\frac{qU_0 h \cos(\psi_s)}{2\pi\beta^2 E} \cdot \left(\alpha_c - \frac{1}{\gamma^2} \right)} = \pm 2\pi f_{rev} \cdot Q_s \tag{2.60}$$

The synchrotron tune Q_s and thus the synchrotron frequency depends on the square root of the maximum RF voltage and the phase for the synchronous particle

$$\Omega_s \sim \sqrt{-U_0 \cdot \cos(\psi_s)} \tag{2.61}$$

For the LHC, the expected synchrotron frequencies are between 30 and 60 Hz depending on the maximum RF voltage (Injection: 7 MV; Collision: 16 MV) and the operational working point ψ_s for the synchronous particle, which is dependent on whether the protons are accelerated ($\psi_s > 0$) or stored ($\psi_s = \pi$). In order to have stable solutions it is required that Ω_s is positive and real and for acceleration it is required that the energy gain per turn $\Delta E \sim \sin(\psi_s)$ is positive. Therefore the allowed stable synchronous phases above transition are within $\frac{\pi}{2} < \psi_s \leq \pi$. The LHC operates at a phase $\psi = \pi$ during injection and collision.

Though a general solution of equation 2.56 is not possible, one can derive the separatrix defining the RF bucket size and phase space in which stable particle oscillation can occur, which is in the end the outer limit for the bunch sizes.

Combining equations 2.53 and 2.54, one can find that:

$$\Delta E = \pm \sqrt{-\frac{\beta^2 qU_0 E}{\pi q \left(\alpha_c - \frac{1}{\gamma^2} \right)} \cdot [\cos(\psi_s - \psi) + \cos(\psi_s) + (2\psi_s + \psi - \pi) \cdot \sin(\psi_s)]} \tag{2.62}$$

The possible phase space separatrices (filled blue) together with some instable trajectories (dotted red) are plotted in Figure 2.6. The maximum energy acceptance $|\Delta E|_{max}$ is given by the maximum stable

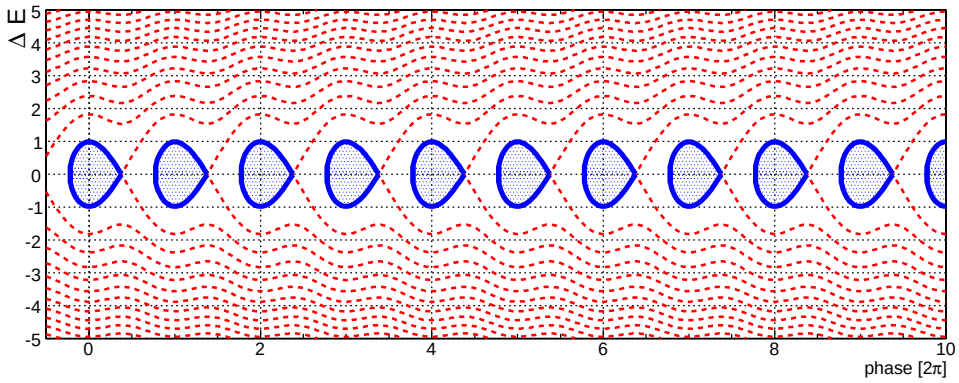


Figure 2.6: Longitudinal phase diagram for an accelerating RF system. The stable phase space, given by the separatrix (blue), and some instable trajectories (dotted red lines) are shown. The RF buckets are periodically spaced by 2π.

phase for the synchronous particle ($\psi = 0$) in equation 2.62 and can be written as:

$$|\Delta E|_{max} = \sqrt{-\frac{2\beta^2 q U_0 E}{\pi q \left(\alpha_c - \frac{1}{\gamma^2}\right)} \cdot \left[\cos(\psi_s) + (\psi_s - \frac{\pi}{2}) \cdot \sin(\psi_s)\right]} \tag{2.63}$$

The maximum acceptance and thus the bunch length is proportional to $|\Delta E|_{max} \sim \sqrt{U_0}$.

2.4 Small Perturbation Treatment

The above mechanism describes single particle dynamics and calculation of single trajectories around the design trajectory. However due to magnetic and mechanical imperfections discussed for the LHC in Chapter 4 the true trajectory and optical function may differ from its design.

2.4.1 Dipole Deflections and Closed Orbit

By definition, an ideal particle with zero position and momentum offset would oscillate on the design trajectory. However in the practical application, there are numerous sources that may induce additional dipole kicks onto the beam and perturb the ideal design trajectory. These dipole kicks δ_i may be caused by:

- magnetic field errors:
$$\delta_i = \frac{e}{p}(B_0 - B_{real})l_{mag} \tag{2.64}$$
- Feed-down dipole due to misaligned and shifted quadrupoles: As visible in equation 2.36, a quadrupole shifted by Δx causes a dipole kick:
$$\delta_{xi} = k l_{mag} \cdot \Delta x \tag{2.65}$$
- Feed-down effects due to gradient errors and shifts Δk and off-centre beams in the quadrupoles: The gradient change induces dipole kick which is given by the change of quadrupole strength Δk_i and off-centre beam position x_s inside the magnet i:
$$\delta_{xi} = \Delta k l_{mag} \cdot x_s \tag{2.66}$$
Commonly, this effect is negligible for the arc quadrupoles, since the ideal trajectory is usually designed to go through the centre of the quadrupoles. Further, the relative field errors are usually very small and in the order of less than a few per-mille with respect to their nominal gradient. However, this effect is relevant for the insertion quadrupole magnets. There, both conditions are met: the quadrupole gradients are deliberately changed from their initial settings to settings in order to provide a reduced betatron function at the location of the IP referred to β^*-squeeze, and the two counter-rotating beams are separated to minimise parasitic encounters outside the IP and thus go off-centre through the insertion magnets.

- Tilted magnets: By design, each LHC main dipole magnet deflects the beam by an angle of about $\delta_{x0} \approx 5.1\,\mathrm{mrad}$ in order to bend the beams onto their circular reference trajectory. However, if the magnet is tilted along its longitudinal axis, part of its deflection angle is projected into the vertical plane as given by:

$$\delta_x = \delta_{x0} \cdot [1 - \cos(\alpha)] \tag{2.67}$$
$$\delta_y = \delta_{x0} \cdot \sin(\alpha) \tag{2.68}$$

  A roll error of $\alpha = 1\,\mathrm{mrad}$ would cause a dipole kick of about $5.1\,\mu\mathrm{rad}$ that is fed-down into the vertical plane at the location of the main dipole.

The orbit corrector dipole magnets (CODs) are usually installed at the quadrupoles in order to correct for the above effects. These additional – usually small – dipole fields deflect the particle onto a new trajectory with respect to the reference. Using the dipole transfer matrix M_{dipole_i} (equation 2.17) and the beam transfer matrix $M_{i\to s}$ (equation 2.43)

$$\begin{pmatrix} z(s) \\ z'(s) \end{pmatrix} = M_{i\to s} \cdot M_{dipole_i} \cdot \begin{pmatrix} z_i \\ z'_i \end{pmatrix} \tag{2.69}$$

the new trajectory is given by

$$\Delta z(s) = \sqrt{\beta_i \beta(s)} \sin \Delta\mu(s) \cdot \delta_i \tag{2.70}$$

with δ_i being the dipole deflection angle, β_i the betatron-function at the dipole and $\beta(s)$ the betatron-function at the point of observation.

One special solution is the equilibrium state that is given if one requires that the particle state after one turn given in equation 2.69 equals the initial state:

$$z(s) = z(s + C)$$
$$z'(s) = z'(s + C)$$

The trajectory one obtains if this equilibrium state is propagated through the ring is called *closed orbit*. Using the twiss parameters the orbit and angluar response is given as:

$$\Delta z_{co}(s) = \frac{\sqrt{\beta(s)\beta_i}}{2\sin(\pi Q)} \cdot \cos(\Delta\mu_i(s) - \pi Q) \cdot \delta_i \tag{2.71}$$

$$\Delta z'_{co}(s) = -\sqrt{\frac{\beta_i}{\beta(s)}} \cdot \frac{2 \cdot \alpha(s) \cdot \cos(\Delta\mu_i(s) - \pi Q) + \sin(\Delta\mu_i(s) - \pi Q)}{2\sin(\pi Q)} \cdot \delta_i \tag{2.72}$$

The closed orbit depends on the tune Q of the machine in addition to the betatron function β_i, $\beta(s)$ and positive phase advance $\Delta\mu(s)$ between the dipole magnet and point of observation. In order to provide the equilibrium beam position, the tune Q has to be a non-integer as visible in the denominator term $\sin(\pi Q)$, which would make equation 2.71 otherwise singular. Including this dipole perturbation, equation 2.24 that describes the trajectory of an individual particle can be rewritten as:

$$z(s) = z_{co}(s) + z_\beta(s) + D(s) \cdot \frac{\Delta p}{p} \tag{2.73}$$

2.4.2 Dipole Deflections and Beam Energy

A secondary effect of the closed orbit perturbations, in the absence of an accelerating voltage, is the change of the closed orbit pathlength with respect to the reference trajectory. In the presence of a RF voltage the absolute pathlength is defined by the RF frequency and the pathlength difference created by the dipole kick is, according to equation 2.46, converted into a momentum offset $\Delta p/p$. As visible in Figure 2.7, the local change of the pathlength of the perturbed dl with respect to the design trajectory ds is given by the following geometric relation ($r = \varrho + z$)

$$\begin{aligned} dC &= dl - ds \\ &= \sqrt{(r d\varphi)^2 + dz^2} - ds \\ &= \sqrt{r^2 + \left(\frac{dz}{d\varphi}\right)^2}\, d\varphi - \varrho d\varphi \\ &= \left[\sqrt{(\varrho + z)^2 + \left(\frac{dz}{d\varphi}\right)^2} - \varrho\right] d\varphi \end{aligned} \tag{2.74}$$

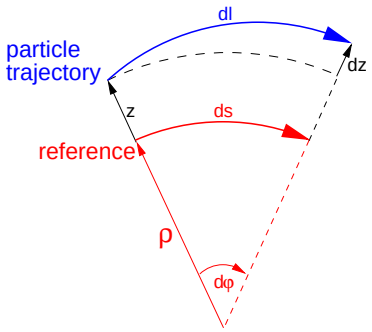


Figure 2.7: Pathlength difference of the perturbed orbit with respect to the design orbit over a short angular segment $d\varphi$ of the accelerator.

with ϱ the local bending radius and z the horizontal coordinate of the perturbed orbit in the local co-rotating coordinate system. As described in [46, 47], using the expansion of the square-root and taking only the lower order terms into account, equation 2.74 can be approximated to:

$$dC \approx \left[\frac{z}{\varrho} + \frac{1}{2}\left(\frac{dz}{ds}\right)^2\right] ds \tag{2.75}$$

The total beam circumference change is given by the integral of equation 2.75

$$\Delta C = \oint dC \tag{2.76}$$

and yields the following total change in pathlength ([46]):

$$\Delta C \approx \sum_{i=0}^{n} D_z(s_i)\delta_i + \frac{8Q_z^2\hat{z}^2}{C_0} \tag{2.77}$$

with $D_z(s_i)$ the dispersion function and δ_i the dipole kick and C_0 the unperturbed circumference of the machine. For the integration of the quadratic term in equation 2.75, the betatron oscillations were approximated by a saw-tooth function with maximum excursion $\hat{z}$ and machine tune Q_z. The momentum change is given according to the definition of the momentum compaction factor (equation 2.46) to:

$$\frac{\Delta p}{p} = \left(\frac{1}{\gamma} - \frac{1}{\alpha_p}\right)\frac{\Delta C}{C} \tag{2.78}$$

In addition to these path lengthening effects, particles with a large oscillation amplitude receive deflections while passing through the sextupoles. The additional momentum shift due to the sextupoles can, according to [47], be approximated to:

$$\left.\frac{\Delta p}{p}\right|_{sext.} \approx \frac{\sum_j \frac{1}{2} m_j z_j^2}{2\pi} \tag{2.79}$$

The presented linearisation is a rough estimate, as the true energy perturbation depends on the actual orbit perturbation in the machine and chromaticity (strength of the sextupoles) inside the accelerator. More precise estimates may be obtained through numerical simulations such as MADX that perform the full tracking through all individual magnets and RF cavities. The modelling error for the LHC due to the linearisation is further discussed in Section 6.1.6.

2.4.3 Machine Tune and Beta-Beat

Quadrupole errors change the machine tune and the betatron function that defines the machine optics. Using the same formalism as for the dipole perturbations, one can show ([17]) that the tune perturbation ΔQ for small quadrupoles errors $\Delta k(s)$ is given by the integral

$$\Delta Q = \frac{1}{4\pi}\oint \beta(s)\Delta k(s) ds \tag{2.80}$$

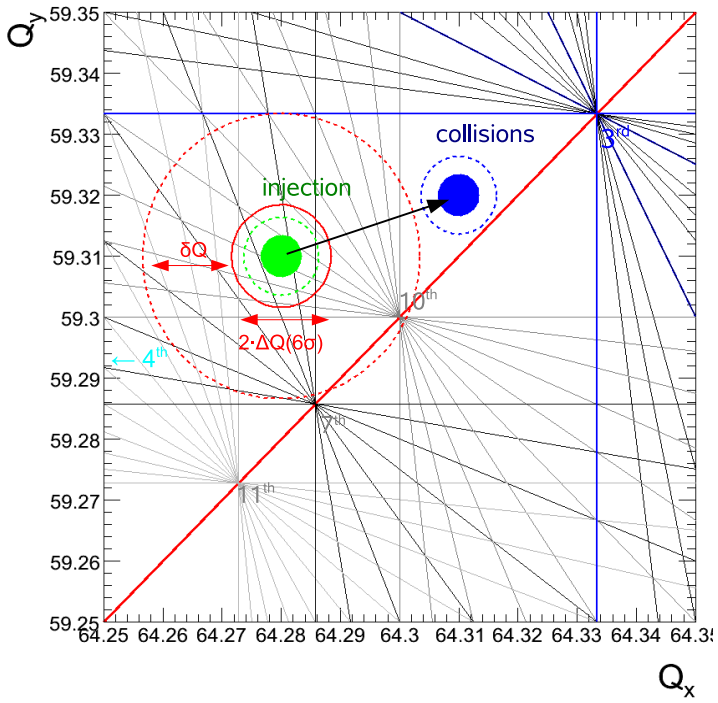


Figure 2.8: LHC tune diagram: The tunes for injection ($Q_x = 64.29$,$Q_y = 50.31$) and for collisions ($Q_x = 64.31$,$Q_y = 59.32$) are indicated. The tune width (filled surface) is the 6σ envelope based on the expected chromaticity and beam momentum spread: $\Delta Q = Q' \frac{\Delta p}{p}$. δQ denotes the available space of the tune distribution without a particle on the 6σ envelope hitting the third order resonance.

The tune is a critical parameter for machine operation and has to be kept constant within tight tolerances in order to avoid that the reference or off-momentum particle oscillate on one of the possible resonances. Figure 2.8 shows exemplarily the LHC tune diagram with the injection and collision tunes. As described in [48,49], for the LHC, it is required to avoid at least the first three orders during injection and up to the twelfth order while colliding the two beams in order to provide long beam life times. Taking the natural tune spread due to the chromaticity and momentum distribution (equation 2.25), beam-beam and other effects into account, the tunes have to be controlled to a level better than $\Delta Q = 0.003$ during injection and better than $\Delta Q = 0.001$ while colliding the beams ([49]). The correction of the tune is an important aspect of machine operation. However, it is not in the scope of this thesis. Further information on tune measurement and control can be found in [49–51]. For further discussions, it is assumed that the tune is already corrected or is within reasonable limits prior to correcting the orbit or energy.

In addition to the tune shift, a quadrupole error causes a perturbation of the betatron functions referred to as *beta-beat*. This effect has an influence on the control of the orbit, energy and other parameters, since their responses are strongly dependent on the value of the betatron function and phase advance, as, for example, visible in equation 2.71.

For small perturbations $\Delta k(s)$ of the quadrupole strength the change of the betatron function can be approximated to

$$\Delta\beta(s) = -\frac{\beta_0}{2\sin(2\pi Q)} \oint \beta(s)\Delta k(s) \cdot \cos\left(2[\mu(s) - \mu_0] - 2\pi Q\right) ds \tag{2.81}$$

The beta-beating creates, similar to the orbit perturbation, an oscillatory wave that propagates with twice the phase advance.

Chapter 3

Control System Design

This chapter summarises concepts of control theory needed for the description and design of the proposed feedback control systems. More information can be found in [52–55].

The LHC orbit and energy control are both large Multiple-Input-Multiple-Output (MIMO) systems. The input consists of over 1000 beam position monitors and over 1000 orbit corrector dipole magnets. For the discussion of control system theory and its concepts, it is favourable to reduce the complexity first to a Single-Input-Single-Output (SISO) system. The issues related to MIMO systems are discussed in Chapter 6. There are two classes of control systems that are distinguished: *open* and *closed loop*, also referred to as *feed-forward* and *feedback* systems. For the practical application, the vast majority of controlled systens use feedbacks. The importance of pure feed-forward systems is given by their simplicity while discussing fundamental concepts of control theory or, in some cases, the inaccessibility of given processes that cannot or are difficult to be measured.

The following example illustrates the concept: The simplified model system (process) steers the position at a single beam position monitor using a single dipole as output (SISO) as illustrated in Figure 3.1. The corrector magnet is located at a distance L in front of the monitor. Ignoring the complexity of geom-

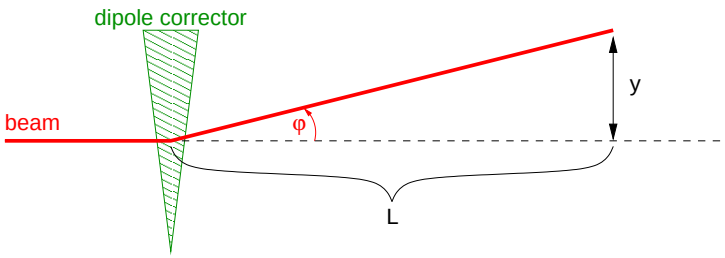


Figure 3.1: Simplified one-to-one beam steering control system consisting of a single position monitor and dipole correction magnet spaced by L.

etry and current-deflection angle conversion of the power converter driving the magnet, it is assumed that the corrector is adjusted by simply specifying the final deflection φ of the magnet. The space between corrector and monitor is assumed to be empty drift space. Using this simple model, the beam position y at the monitor can thus, using the matrix formalism and drift space transfer matrix, be approximated to

$$y = G(s) \cdot \varphi \approx L \cdot \varphi \tag{3.1}$$

with both the initial beam position and angle being zero at the location of the dipole. In order to control the beam position at the monitor, a simple first order control process would simply invert the process transfer function $G(s)$. Since the system is assumed to be time-invariant, this inversion is simply a constant calibration factor $K_P = \frac{1}{L}$ that is multiplied with the requested reference position r prior to sending and applying this angle to corrector dipole converter in order to fulfil:

$$y = K_p G(s) \cdot r \stackrel{!}{=} r \tag{3.2}$$

The control systems' function and differential equation are commonly described using block diagrams, as shown for the above system in Figure 3.2, which are the control theories' equivalent to Feynman graphs. The signal flow and causality are indicated through arrows. Boxes indicate transfer functions and constants that are convoluted with the input signals pointing into them. Circles with either a sum (Σ) or product (Π) symbol indicate whether the inputs are added or multiplied. Minus symbols at the input may indicate whether the signal is to be inverted prior to addition or multiplication with other signals. The system performance and sensitivity is given by the full response and sensitivity function:

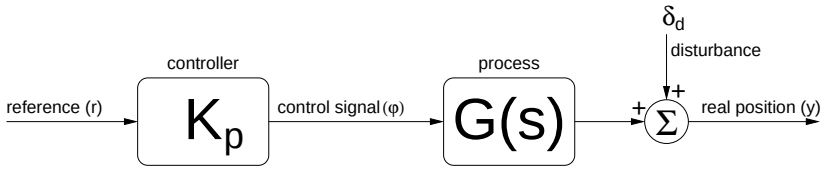


Figure 3.2: Open loop block diagram: The system consists of a proportional controller with gain K_p, a linear time-independent process G and external perturbation source δ_d that is added to the process output.

The transfer function $T(s)$ of the system is defined as the ratio between true y to requested r output:

$$T(s) := \frac{y}{r}$$

In this simple model, the open loop response ideally equals one. In real life, the real state is, however, perturbed by other additional non-accounted external sources such as shifted quadrupole magnets inducing dipole kicks, power converter calibration errors and ripple, machine and magnet imperfection. Without loss of generality, these external sources are commonly summarised to a single perturbation source δ_d and added to the system output. The *disturbance rejection* also referred to as *first sensitivity function* S of the above open loop system is for a non-vanishing disturbance δ_d given by:

$$S := \frac{y}{\delta_d} = 1 + DG \cdot \frac{r}{\delta_d} \tag{3.3}$$

The sensitivity for large perturbations ($\delta_d >> 1$) equals one. Often it is implicitly assumed that the requested reference (change) is zero while evaluating the disturbance rejection. The sensitivity function of the given model illustrates the lack of disturbance rejection of open loop systems and susceptibility to systematic errors. The stability of the parameter (position) to be controlled is thus essentially given by the magnitude of the perturbation sources.

Open loop control systems depend on the precision of the proportional gain (also referred to as calibration) and on the disturbance prediction. They are limited to systematic sources and have limited performance or even fail when dealing with unknown or random perturbations. However, such systems are still applicable and used if the model uncertainties, calibration factors and perturbation sources are precisely known or small with respect to the required stability. In case of the control of orbit and energy in the LHC the expected perturbation exceeds, by far, the required stability margins as discussed in Chapter 4.

3.1 Laplace Transform

The above open loop model system is time invariant. However – nearly always – the system (process) to be steered depends on the perturbation's rate of change, its magnitude and may contain significant non-linear terms. Many physics control problems can be reduced to a first or second order type system. Their differential equation describing the changing state y due to a driving force $f(t)$ can be written for example, as:

$$\tau \dot{y}(t) + y(t) = f(t) \tag{3.4}$$

with τ the natural time constant for a first order system and

$$\ddot{y}(t) + 2\zeta\omega_0 \cdot \dot{y}(t) + \omega_0^2 \cdot y(t) = f(t) \tag{3.5}$$

for a second order system with ζ the relative damping and ω_0 the system's eigenfrequency. Though these differential equations can be solved analytically, it is convenient to use the *Laplace* representation of the

differential equations. The *Laplace* transform ($\mathcal{L}$ -transform) is commonly given in its unilateral form as

$$\mathcal{L}(f(t)) = F(s) := \int_{0^-}^{+\infty} f(t)e^{-st}dt \tag{3.6}$$

with 0^- being an abbreviation for the limes $\lim_{-\infty\to 0}$. The complex parameter s defines the $\mathcal{L}$-space to (i being the imaginary value $i := \sqrt{-1}$):

$$s = \sigma + i\omega \tag{3.7}$$

The abscissa and ordinate of $\mathcal{L}$-space are given in unit of s^{-1}. From the above definition, the inverse $\mathcal{L}$ -transform can be written as

$$\mathcal{L}^{-1}(F(s)) = f(t) = \frac{1}{2\pi i}\int_{\sigma_c - i\infty}^{\sigma_c + i\infty} F(s)e^{st}ds \tag{3.8}$$

with σ_c being an arbitrarily selected value that has to be on the right of all singularities of $F(s)$ in the $\mathcal{L}$-space . It can be shown that the Laplace transform has the following important properties:

Linearity:	$\mathcal{L}(\alpha f_1(t)) = \alpha\mathcal{L}(f_1(t)) = \alpha F_1(s)$	(3.9)
Superposition:	$\mathcal{L}(\alpha f_1(t) + \beta f_2(t)) = \alpha F_1(s) + \beta F_2(s)$	(3.10)
Time Delay:	$\mathcal{L}(f(t' \to t - \lambda)) = e^{-s\lambda}F(s)$	(3.11)
Time Scaling:	$\mathcal{L}(f(t' \to at)) = \frac{1}{\lvert a\rvert}F(\frac{s}{a})$	(3.12)
Differentiation:	$\mathcal{L}(\dot{f}(t)) = -f(0^-) + sF(s)$	(3.13)
Integration:	$\mathcal{L}\left(\int f(t)dt\right) = \frac{1}{s}F(s)$	(3.14)
Convolution:	$\mathcal{L}(f_1(t) * f_2(t)) = F_1(s)\cdot F_2(s)$ and	(3.15)
	$\mathcal{L}(f_1(t)\cdot f_2(t)) = F_1(s) * F_2(s)$	(3.16)

The advantage of the Laplace transform is its flexibility while studying dynamic systems using arbitrary excitations. The convolution relation (equation 3.15) is particularly useful when many sub-processes are involved, since the total response function of successively connected processes in $\mathcal{L}$-space is performed through a simple multiplication, while in time domain the total response has to be calculated using the more complex convolution integral. Using the Laplace transform and its differentiation property, the second order system equation 3.5 can be written as:

$$\begin{aligned} \mathcal{L}\{\ddot{y}(t) + 2\zeta\omega_0\cdot\dot{y}(t) + \omega_0^2 y(t)\} &= \mathcal{L}\{f(t)\} \\ s^2Y(s) + 2\zeta\omega_0 sY(s) + \omega_0^2 Y(s) &= F(s) \qquad (3.17) \\ (s^2 + 2\zeta\omega_0 s + \omega_0^2)\cdot Y(s) &= F(s) \qquad (3.18) \end{aligned}$$

It is visible that the $\mathcal{L}$-transformed response $Y(s)$ is decoupled from the $\mathcal{L}$-transformed excitation source $F(s)$. The process transfer function $H(s)$ is defined by:

$$Y(s) = H(s)\cdot F(s) \quad \Longleftrightarrow \quad H(s) := \frac{Y(s)}{F(s)} \tag{3.19}$$

The $\mathcal{L}$-transformed transfer function of the first and second order system response $H(s)$ can thus be written as:

$$H_{1^{st}}(s) = \frac{K_0}{\tau\cdot s + 1} \tag{3.20}$$

$$H_{2^{nd}}(s) = \frac{K_0\omega_0^2}{s^2 + 2\zeta\omega_0\cdot s + \omega_0^2} \tag{3.21}$$

The factor K_0 is introduced to normalise and generalise the equation.

3.1.1 Effect of Zeros and Poles on Transfer Functions

Generally, any response function $H(s)$ can be written using the *Padé* expansion to a fraction of two polynomials of degrees n and m

$$H(s) := \frac{\prod_{i=0}^{n}(s+b_i)}{\prod_{i=0}^{m}(s+a_i)} \tag{3.22}$$

with b_i being the so-called *zeros* and a_i the *poles* of the transfer function. The system's response in time domain is described by the location of the poles and zeros of equation 3.22 in $\mathcal{L}$ -space. Figure 3.3 gives an example of pole locations and their corresponding solutions in the time domain. The imaginary part of the pole defines the oscillation frequency and the real part the amplitude growth or damping, respectively. In a system with several poles, the poles closest to the imaginary axis define the main behaviour of the system and are hence referred to as *dominant poles*. The poles describe stable dynamics on the left half plane of the $\mathcal{L}$ -space, and instable dynamics on the right half plane. The system is termed as *minimum*

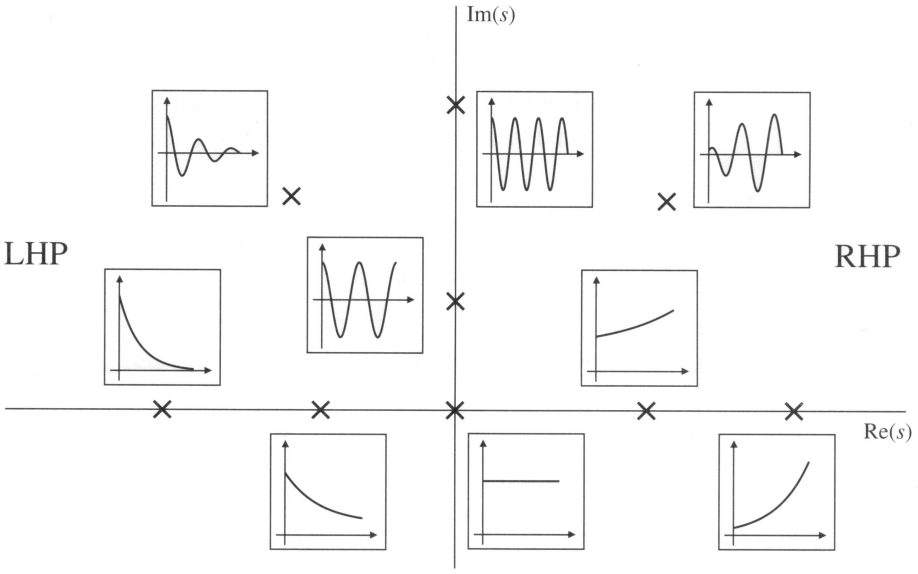


Figure 3.3: Time domain step response associated with certain pole location (X) in the $\mathcal{L}$-space . The imaginary part of the pole determines the frequency of the oscillatory part and the real part (horizontal axis in $\frac{1}{s}$) the damping. Stable systems are required to have poles located in the left half plane (LHP), whereas instable systems (increasing amplitude) are found in the right half plane (RHP). Poles with pure damping correspond to first order systems. Figure taken from [55].

phase system, if all poles and zeros are located on the left-half-plane of the $\mathcal{L}$ -space. The effect of zeros is more subtle and depends on whether they are on the left or right half plane and whether they are in the proximity or potentially cancel poles.
The design of linear control system is based on the location of these poles. If necessary, the controller compensates the effect of the plant by introducing zeros, which may cancel the plant's poles and/or move the plant's poles to a desired location. Usually the poles are chosen so that the loop systems matches the required response.

3.1.2 Step Response

For a general classification and description of linear processes one commonly studies the system's *step response* to a Heaviside step function excitation: The reference is set to zero for times $t < 0$ and to an arbitrary but constant reference y_∞ for times $t \geqslant 0$. Figure 3.4 shows an exemplary response of a second order system. In order to describe the process and performance of the control loop, one commonly defines the following properties:

- the *rise time* t_r as the time the process needs to reach 90% from 10% of the reference signal y_∞ (here: $y_\infty = 1$),

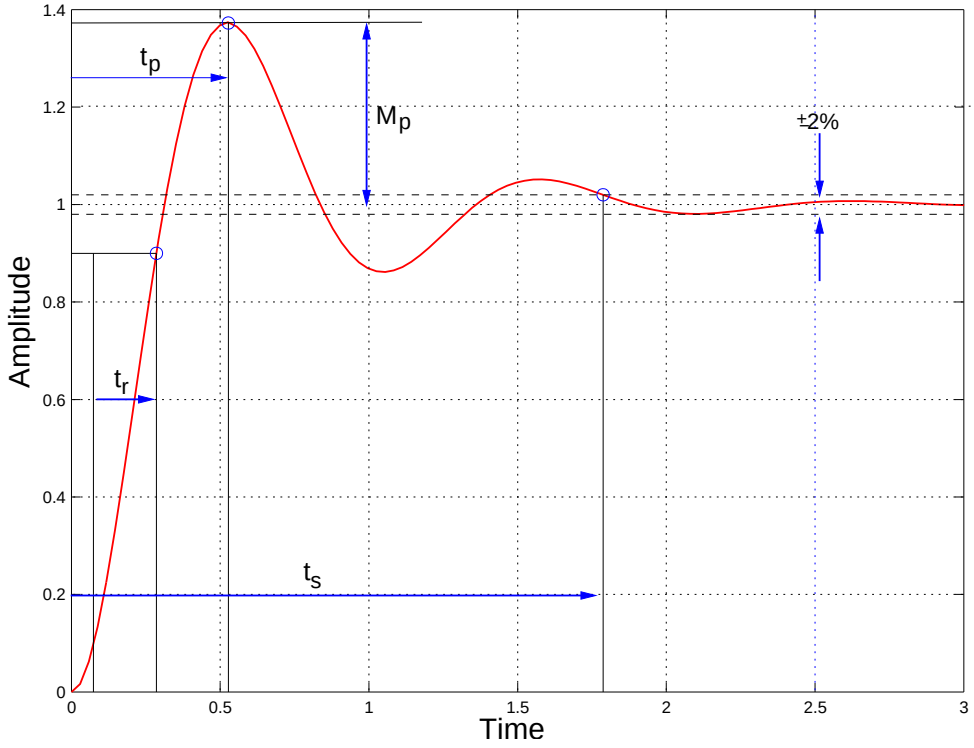


Figure 3.4: Exemplary step reference response of a second order system. The system characteristics rise time t_r, peak time t_p, settling time t_s and relative overshoot M_p are indicated.

- the *peak time* t_p as the time from the start of the excitation until the maximum response,
- the *overshoot* M_p as the difference between the peak amplitude and final steady state and
- *settling time* as the time till the transient settles within 2% with respect to the final steady state value.

In general, the true final state may differ from the desired final state referred to as *steady-state error* and is required to be minimised through the feedback controller.

3.1.3 Frequency Response

Another classification is given by the process' *frequency* or *bode response*. Here the system response is given with respect to a sinusoidal excitation with constant amplitude and phase but varying frequency. The response's *magnitude* $M(f)$ is defined by the ratio between the maximum oscillation amplitude of the excitation and the maximum oscillation amplitude of the process response. The phase $\phi(f)$ describes the lag of the response behind the excitation signal. The following equations relate the magnitude and phase to the process' Laplace transformed response function:

$$M(f) = 20\log(|G(s)||_{s=i2\pi f}) \tag{3.23}$$

$$\phi(f) = \tan\left(\frac{\Re[G(s)]}{\Im[G(s)]}\right)\Bigg|_{s=i2\pi f} \tag{3.24}$$

The magnitude $M(f)$ and phase $\phi(f)$ are usually plotted as a function of frequency and measured in units of *decibels* (dB), also referred to as *gain*, as shown for a second order system in Figure 3.5. The abscissa axis is logarithmic in frequency with the base 10. Often the unit '*decade*' is used to describe the distance between the two frequencies f_1 and $10 \cdot f_1$. The maximum phase lag depends on the order n of the system and is given by:

$$\Delta\varphi|_{max} = n \cdot \frac{\pi}{2} \tag{3.25}$$

Another indication for the order of the system is the slope indicated in the Bode plot which is proportional to $n \cdot 20\,\mathrm{dB/decade}$. One defines the following system observables:

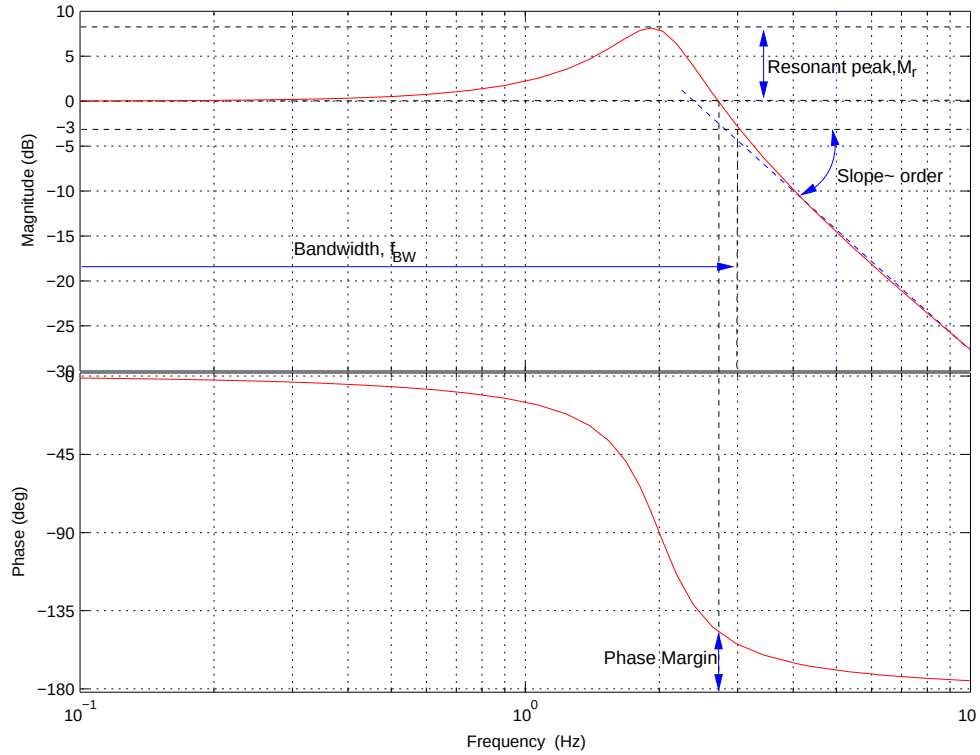


Figure 3.5: Example for the frequency response of a second order system (Bode plot). The bandwidth f_{BW}, resonant peak M_r, phase margin and slope of the system are indicated.

- the *Bandwidth* f_{BW} which is defined by the frequency where the gain drops below 3 dB,
- the *resonant peak* M_r which is the highest value of the frequency response and an indication for the system's damping coefficient: the lower the damping, the higher the resonant peak.
- the *phase margin* $\Delta\phi$ is given as the difference between the systems phase lag and π (180^O) at the frequency when the gain drops below 0 dB.
- similarly one defines the *gain margin* $\Delta G(s)$ as the difference of the magnitude with respect to zero in the logarithmic representation at a frequency where the phase margin is $\Delta\phi = 0$.

Phase and gain margin are important properties that define the stability of the closed loop discussed in the next section.

3.2 Closed Loop Systems

An improvement with respect to disturbance rejection can be achieved using the measured state of the process y in addition to the reference setting r in order to compute the control signal. These feedback systems thus have a cyclic dependence between control and plant process in a way that the controller affects the system output and the system output depends on the controller action. A simple feedback block diagram is shown in Figure 3.6. The loop consists of a process modelled by the transfer function $G(s)$, a

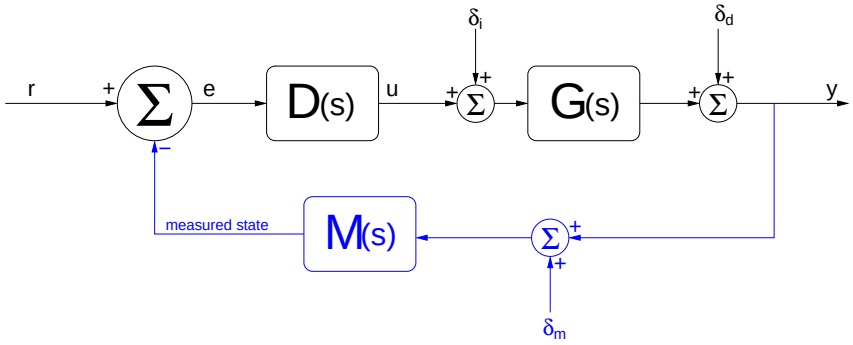


Figure 3.6: Closed loop block diagram: Different to the open loop system, the controller uses the error signal e which is the difference between the reference signal r and the measured state coming from the feedback path. The measurement device transfer function $M(s)$ and superimposed measurement noise δ_m signal are indicated.

controller $D(s)$ and a monitor response function $M(s)$. The loop disturbances can be summarised into three groups: external perturbations δ_d, internal perturbations δ_i and measurement noise δ_n. Depending on the loop representation, often the monitor response is incorporated into one of the other responses and the loop normalised in order to fulfil $M(s) = 1$. The equivalent equations, without disturbances, described by Figure 3.6 are given by:

$$y = D(s)G(s) \cdot (e + \delta_i) + \delta_d \tag{3.26}$$

$$e = r - M(s) \cdot y \tag{3.27}$$

The definition of the closed loop transfer function and disturbance rejection is similar to the open loop case. In addition, one defines the following set of process sensitivity functions

$$T_0(s) := \frac{y}{r} = \frac{D(s)G(s)}{1 + D(s)G(s)} \tag{3.28}$$

$$S_{d0}(s) := \frac{y}{\delta_d} = \frac{1}{1 + D(s)G(s)} \tag{3.29}$$

$$S_{i0}(s) := \frac{y}{\delta_i} = \frac{G(s)}{1 + D(s)G(s)} \tag{3.30}$$

$$S_{u0}(s) := \frac{u}{\delta_d} = \frac{D(s)}{1 + D(s)G(s)} \tag{3.31}$$

where $T_0(s)$ is the transfer function (sometimes also referred to as *nominal complementary sensitivity*), $S_d(s)$ the *nominal sensitivity* or disturbance rejection, $S_i(s)$ the *input-disturbance sensitivity* and $S_u(s)$ the *control sensitivity*. For the sensitivity functions, it is implicitly assumed that $r(s) = 0$. The subindex '0' indicates that the functions refer to the unperturbed plant. The effect of process model errors is further discussed below. As visible in equation 3.28, the closed loop stability is fundamentally determined by the closed loop pole

$$1 + D(s)G(s) \tag{3.32}$$

which for reasons of stability should avoid that the function $D(s)G(s)$ encircles the '-1' point in the complex Laplace space. Thus, in terms of frequency domains, the combined control and process function should provide

$$|G(s)| \cos(2\pi) \overset{!}{<} -1 \tag{3.33}$$

which is equivalent to the requirement that a stable closed loop has to have a positive phase margin or a negative gain at a phase of 180^O as indicated in the bode plot. Thus, in case the gain margin is not

sufficient to avoid an encirclement of the '-1', the control function $D(s)$ is designed to deliberately reduce the systems bandwidth and to roll-off the total response in order to provide the required stability.

Further it is visible in equations 3.28 and 3.29, that the sum of the transfer function and nominal sensitivity equals one:

$$T(s) + S_d(s) = 1 \tag{3.34}$$

Hence, in order to provide a good disturbance rejection, one has to optimise the tracking properties of the transfer function $T(s)$ based on the optimisation of the step or frequency response described earlier. One could in principle further define the sensitivity of the control loop with respect to measurement noise $S_n(s)$, which is given by:

$$S_{n0}(s) := \frac{y}{\delta_n} = \frac{D(s)G(s)}{1 + D(s)G(s)} \tag{3.35}$$

However equation 3.35 is essentially the same as the transfer function T_0.

Thus, in combination with equation 3.34, it is visible that there is an intrinsic trade-off between disturbance rejection and sensitivity to measurement noise: A control loop design that is robust with respect to measurement noise would thus reduce the system's bandwidth (or rise time) $|T_0(i\omega)|$, which automatically increases the sensitivity to external noise $S_{d0}(s)$ and vice versa.

Usually, most processes have low-pass characteristics which rolls off for high frequency. For example: A simple converter powering a simple superconducting dipole circuit consisting of a single magnet with $6H$ inductance and a resistance of about $0.1\,\Omega$ would require a steady-state voltage of about 1 V to supply a constant current of 10 A. However, if switched on with only one volt, the required current would, due to the large inductance, reach 95% of its required target only after about 180 seconds (rise-time $\tau = \frac{L}{R} = 60\,\mathrm{s}$). Thus all controlled power converters have some additional voltage margin in order to momentarily increase the voltage while the magnet current is far from its required steady-state value to boost the rise-time. As visible in 3.31, the controller needs to provide larger excitation signals in order to provide larger closed loop bandwidth. However, the maximum excitation signals that can be sent to the process are usually limited by, in case of power converters, the maximum available voltage or available power that drive the process. The minimum rise time (or maximum bandwidth) is given by the constraints on the control sensitivity $S_u(s)$ that should be constrained and respect those limitations.

Summarising: The optimisation of control systems is thus finding a compromise and adequate trade-off between:

- high disturbance rejection, in particular the minimisation of steady state errors,
- optimising the control system requirements and the total cost has to be taken into account, and
- robustness against model errors and measurement noise.

3.3 Proportional-Integral-Derivative (PID) Control

The proportional controller is the simplest kind of controller. Using the controller described in equation 3.2 for the closed loop case yields the following closed loop transfer and first sensitivity function:

$$T_0(s) = \frac{K_p \cdot G(s)}{1 + K_p \cdot G(s)} \tag{3.36}$$

$$S_{d0}(s) = \frac{1}{1 + K_p \cdot G(s)} \tag{3.37}$$

Compared to equation 3.3, the sensitivity $S_{d0}(s)$ can be made smaller than for the open loop case by increasing the proportional gain K_p. However, the closed loop creates a so-called steady-state error e_{ss} given by

$$e_{ss} = \lim_{t \longrightarrow \infty} r(\theta(t)) - y(t) \tag{3.38}$$

with $\theta(t)$ the Heaviside step function. Though an increase of the proportional gain K_p may reduce the steady-state error, it does not fully minimise it. Further, large proportional gains often cause large and thus infeasible excitation signals which generally should be avoided. Thus nearly all closed loop controllers combine at least proportional (P) with integral (I) action in a so-called PI controller $D_{PI}(s)$ that can be written in the Laplace domain as

$$D_{PI}(s) = \frac{u(s)}{e(s)} = K_p + K_i \cdot \frac{1}{s} \tag{3.39}$$

with $u(s)$ the excitation, $e(s)$ the error signal and K_i the integral gain. As visible in equation 3.14, the integral term is proportional to the integral over error signal e:

$$u_{PI}(t) = K_p \cdot e(t) + K_i \cdot \int_{-\infty}^{t} e(t')dt' \tag{3.40}$$

The integral unlike the proportional term increases the controller signal $u(t)$ till the error signal $e(t)$ is zero and thus minimises any steady state error. However, the integral term also introduces an additional phase lag

$$\Delta\phi \approx -\frac{K_i}{2\pi f} \tag{3.41}$$

that for large integral gains may lead to loop instability. A further extension is to add a derivative term with gain K_d that acts on the first derivative of the error signal. The derivative term, sometimes also referred to as *predictive* control, is proportional to s, as visible in equation 3.13. It introduces a positive phase lag

$$\Delta\phi \approx +K_d \cdot 2\pi f \tag{3.42}$$

that can be used to stabilise the phase lag due to the integral control. The combination of proportional, integral and derivative action, commonly referred to as *PID controller*, is one of the most widely used controllers in feedback systems. One possible Laplace domain representation is

$$D_{PID}(s) = K_p + K_i \cdot \frac{1}{s} + K_d \cdot s \tag{3.43}$$

with K_p the proportional, K_i the integral and K_d the derivative gain. This representation is unsuitable for large frequency f, since the derivate term ($\sim s = i2\pi f$) would yield unfeasibly large excitation signals. Thus, in order to obtain a proper controller representation, the derivative term is often combined with a high frequency cut-off

$$D_D(s) = K_d \frac{s}{\tau_D s + 1} \tag{3.44}$$

with τ_D the derivative time constant which specifies the range and exactness of the derivative control. Usually τ_D is chosen to be short for rounding off the response and reducing the controller's sensitivity to measurement noise. The smaller it is, the larger the frequency range over which the filtered derivative action approximates the exact derivative. The PID controller using this modified derivative term is thus given as:

$$\boxed{D_{PID}(s) = K_p + K_i \cdot \frac{1}{s} + K_d \cdot \frac{s}{\tau_D s + 1}} \tag{3.45}$$

Though – strictly – the effect of proportional, integral and derivative terms are related and the parameter dependent on each other, it is possible to give a few qualitative effects on the closed loop response characteristics which are summarised in Table 3.1.

PID term	rise time	bandwidth	overshoot	settling time	steady-state-error
proportional	decrease	increase	increase	small change	decrease
integral	decrease	increase	increase	increase	eliminate
derivative	small effect	small effect	decrease	decrease	small effect

Table 3.1: Qualitative effects of the PID terms on closed loop response.

One possibility to find the optimal settings is to use the qualitative relations given in Table 3.1 and to empirically optimise the parameter. A first optimisation study has been performed for several types of processes by Ziegler and Nichols, and later Cohen and Coon [56, 57]. Their optimisation methods are presented in the Annexe D. These empirical tuning rules should only be considered as a fall-back option. A more rigorous PID tuning should rely on the model based approach described in the next section.

3.4 Model Based Feedback Controller Design

This thesis focuses on YOULA'S *affine parameterisation method* for optimal controllers, which is based on the analytic process inversion, first described in [58, 59].
Youla found ([58]):

Lemma:
For an open-loop stable process $G(s)$, the nominal closed-loop transfer function is stable if and only if $Q(s)$ is an arbitrary stable proper transfer function and $D(s)$ parameterised as:

$$D(s) = \frac{Q(s)}{1 - Q(s)G(s)} \qquad \Longleftrightarrow \qquad Q(s) = \frac{D(s)}{1 + D(s)G(s)} \tag{3.46}$$

The stability of the closed loop system follows immediately out of the definition (equation 3.46) when inserted into equations 3.28 to 3.35. The sensitivity functions in the $Q(s)$ form are given as:

$$T_0(s) = Q(s)G(s) \tag{3.47}$$
$$S_{d0}(s) = 1 - Q(s)G(s) \tag{3.48}$$
$$S_{i0}(s) = (1 - Q(s)G(s))G(s) \tag{3.49}$$
$$S_{u0}(s) = Q(s) \tag{3.50}$$

Assuming $G(s)$ is stable, the only requirement for closed loop stability is for $Q(s)$ to be stable. The same principle can be extended to open-loop unstable systems, as described in [52, 58]. The strength of this design method is the explicit controller design with respect to required closed loop performance, as visible in equation 3.47, and required stability (equations 3.48 to 3.50).

The controller design is carried out on the transfer function $Q(s)$ satisfying the required performance and stability criteria and then converted to the controller response function using equation 3.46. For optimal control, the function $Q(s)$ would ideally invert the process response $G(s)$. However, as visible in equation 3.50, this would lead to improper controller implementation that requires infinite control excitation signals at high frequencies. Thus, one typically decomposes Q into

$$Q(s) = F_Q(s) \cdot G_0^i(s) \tag{3.51}$$

with $F_Q(s)$ the roll-off function and $G_0^i(s)$ the exact inversion ($G_0^{-1}(s)$), if $G_0(s)$ contains no unstable zeros. In case the process contains unstable zeros, $G_0(s)$ can be parameterised as

$$G_0(s) = \frac{B_{os}(s)B_{us}(s)}{A_0(s)} \tag{3.52}$$

with $B_{os}(s)$ and $B_{us}(s)$ being the stable and unstable zeros, and $A_{os}(s)$ and $A_{us}(s)$ being the stable and unstable poles of the plant. Using this decomposition $G^i(s)$ can then be written as:

$$G^i(s) = \frac{A_{os}(s)}{B_{os}(s)} \tag{3.53}$$

The choice of $F_Q(s)$ determines the stability transfer function. In case of a process with no unstable zeros, $G_0^i(s)$ is given by the exact inversion and the closed loop response (equation 3.47)

$$T_0(s) = F_Q(s) \tag{3.54}$$

reduces the roll-off function $F_Q(s)$ and thus one could in principle achieve an arbitrary closed loop response which, as visible in equation 3.50, is limited by a required feasible control sensitivity in order to provide constraint controller excitation signals. For processes with non-stable zeros ($B_{ou}(s)$) the transfer function is given by:

$$T_0(s) = F_Q(s)B_{ou}(s) \tag{3.55}$$

Common unstable zeros are, for example, systematic delays or transportation lags that are intrinsic to sampled systems and discussed in section 3.5.

Up to now, the design was based on a perfect process description $G_0(s)$. In the real application, the process $G(s)$ usually contains uncertainties and errors that can be written as

$$G(s) = G_0(s) + G_\epsilon(s) = G_0(s)(1 + G_\Delta(s)) \tag{3.56}$$

with G_ϵ the *additive model error* and G_Δ the *multiplicative model error*. After inserting this error definition into equations 3.47 to 3.50, the sensitivity functions including the process model errors can be written as:

$$\begin{aligned} T(s) &= T_0(s)(1 + G_\Delta(s))S_\Delta(s) && (3.57)\\ S_d(s) &= S_0(s)S_\Delta(s) && (3.58)\\ S_i(s) &= S_{i0}(1 + G_\Delta(s))S_\Delta(s) && (3.59)\\ S_u(s) &= Q(s)S_\Delta(s) && (3.60) \end{aligned}$$

with

$$S_\Delta(s) = \frac{1}{1 + Q(s)G_\epsilon(s)} \tag{3.61}$$

Model errors usually occur at higher frequencies. Thus, in order to provide robust control, the function $F_Q(s)$ is rolled off before the contribution of the unstable zeros $B_{ou}(s)$ or model errors become dominant and in order to provide that the product in equations 3.57 to 3.60 stay bounded. This is equivalent to limiting the desired process response $F_Q(s)$ bandwidth, where the reduced bandwidth depends on the expected magnitude of the perturbations and model errors.

3.4.1 Example: First Order System

This formalism is illustrated using a simple first and second order system. Assuming a first order plant

$$G(s) = \frac{K_0}{\tau \cdot s + 1} \tag{3.62}$$

it is possible, due to the absence of unstable zeros, to perform an inversion of the full process and write the controllers Q-function as

$$Q(s) = F_Q(s) \cdot G^{-1}(s) \tag{3.63}$$

with

$$G^{-1}(s) := [G(s)]^{-1} = \frac{\tau \cdot s + 1}{K_0} \tag{3.64}$$

In order that $Q(s)$ and thus $D(s)$ are biproper, $F_Q(s)$ must have a degree of one, and can, for example, be written as:

$$F_Q(s) = \frac{1}{\alpha s + 1} \tag{3.65}$$

Inserting it into equation 3.46 yields:

$$D(s) = \frac{\tau \cdot s + 1}{K_0 \alpha s} = \frac{\tau}{K_0 \alpha} + \frac{1}{K_0 \alpha s} \tag{3.66}$$

The PID control structure given in equation 3.45 is visible and the gains are identified by:

$$K_p = \frac{\tau}{K_0 \alpha} \qquad \text{and} \qquad K_i = \frac{1}{K_0 \alpha} \tag{3.67}$$

The parameter α can be seen as a tuning parameter between a high noise sensitivity or high disturbance rejection (small α) and low noise sensitivity or low disturbance rejection (large α). Another interpretation of α, as visible in equations 3.47 and 3.63, is as feedback bandwidth of the closed loop system. Figure 3.7 shows the bandwidth/sensitivity trade-off as a function of the model parameter α. The range for α is given by the maximum possible control signal the actuator can supply for a given disturbance, as described by equation 3.50. Often, the processes to-be-steered contain internal feedback loops themself that optimise the process bandwidth. Thus the parameter or bandwidth α of the closed loop has to be smaller in order to deliver feasible excitation signals.

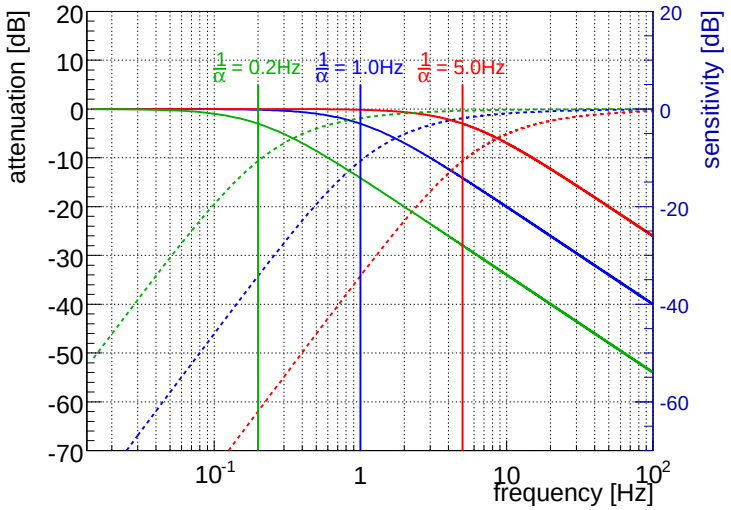


Figure 3.7: Effective closed loop bandwidth trade-off parameterisation α.

3.4.2 Example: Second Order System

In case of a second order system, defined in equation 3.21, one can similarly define a second order roll-off function F_Q:

$$F_Q(s) = \frac{K_0\omega_{cl}^2}{s^2 + 2\zeta_{cl}\omega_{cl}s + \omega_{cl}^2} \tag{3.68}$$

Re-inserting equation 3.68 into equation 3.46 and comparing with the PID controller equation 3.45, it is possible to derive the following PID gain parameterisation

$$K_p = \frac{4\zeta_0\zeta_{cl}\omega_0\omega_{cl} - \omega_0^2}{4K_0\zeta_{cl}^2} \tag{3.69}$$

$$K_i = \frac{\omega_0^2\omega_{cl}}{2K_0\zeta_{cl}} \tag{3.70}$$

$$K_d = \frac{4\zeta_{cl}^2\omega_{cl}^2 - 4\zeta_0\omega_0\zeta_{cl} + \omega_0^2}{8K_0\zeta_{cl}^3\omega_{cl}} \tag{3.71}$$

$$\tau_d = \frac{1}{2\zeta_{cl}\omega_{cl}} \tag{3.72}$$

with ζ_0 the damping coefficient, ω_0 the eigenfrequency and K_0 the steady-state gain of the open loop process, and ζ_{cl} and ω_{cl} the desired damping and eigenfrequency of the closed loop.

The first and second order PID parameterisations are useful for gain scheduling since the gains are decoupled with respect to real process parameters and desired closed loop response. Of course, as visible in equation 3.50, the closed loop parameters have to respect the control sensitivity and provide excitations that are within process control capabilities and limits.

For example, in order to minimise the overshoot of the closed loop one may choose a critical damping $\zeta_{cl} = 1$ and vary the desired closed loop bandwidth ω_{cl} while meeting the other controls requirements such as the maximum bound of the control effort and maximum sensitivity to measurement noise.

3.5 Sampled Systems

Most physics processes have to be controlled. However, nowadays it is common to exploit digital signal processors and to convert continuous process signals in the digital domain. Digitisation gives the advantage of higher flexibility with respect to possible controller design and higher reproducibility of the controller action with respect to environmental influences and disturbances that would influence analogue circuits. For the digitisation process, the continuous function f are evaluated for discrete times

$$f(k) := f(T_k) = f(T_0 + k \cdot T_s) \tag{3.73}$$

with T_0 an initial time offset, $k \in \mathbb{N}$ the sampling index and T_s the sampling interval. Strictly speaking the sampling, using for example, an analogue-digital-converter (ADC), provides only values at the time of the sampling measured functions. Thus commonly and often implicitly, it is assumed that the measured value is kept constant till the next sampling period, referred to as zero-order-hold (ZOH) as illustrated in Figure 3.8. Due to the sampling and ZOH the approximated function lags behind the original function.

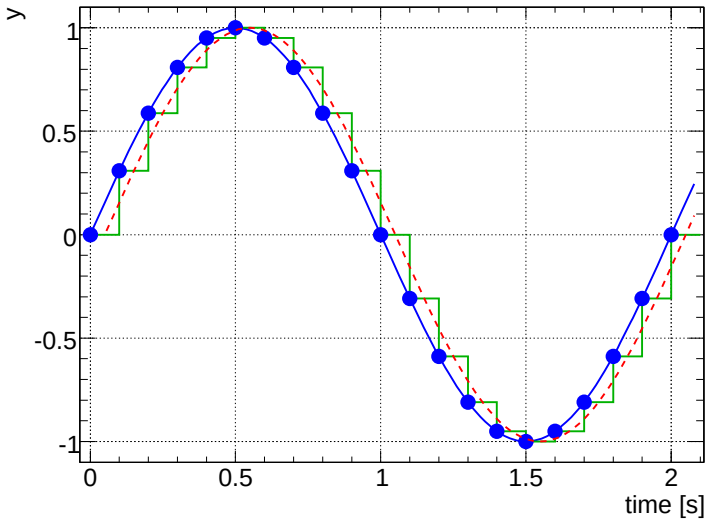


Figure 3.8: Sampling of a continuous system: The continuous system (solid blue), the sampling points, discrete ZOH function and delayed approximation (dotted red) are shown. The delay of half the sampling frequency of the approximated function is visible.

The associated delay λ is half of the sampling frequency:

$$\lambda = \frac{T_s}{2} \tag{3.74}$$

. In the Laplace space this lag can be written in the following general form

$$f(s) = e^{-\lambda s} f_0(s) \tag{3.75}$$

with $f(s)$ the delayed and $f_0(s)$ the original function. It is obvious that causality forbids the inversion of the exponential delay term.

Padé approximation:

$$e^{-\lambda s} \approx \left(\frac{-s + 2k/\lambda}{s + 2k/\lambda} \right)^k \tag{3.76}$$

with k determining the accuracy of the approximation.

3.5.1 Quantisation

Another effect intrinsic to the sampled system and the use of analogue-to-digital and digital-to-analogue converters (ADC/DAC) is the limited and quantised range of the measurement. If the measured signal exceeds the specified maximum range the signal saturates at the ADC's maximum value and if the change to be measured is smaller than the quantisation, the measured signal appears to be unchanged. The quantisation and its range is binary and given in number of bits n where one bit is usually used to distinguish between positive and negative signals. Thus the relative measurement accuracy with respect to the maximum value that ultimately gives a lower bound for the feedback precision is given by:

$$\left. \frac{\Delta x}{x} \right|_{min} \approx 2^{1-n} \tag{3.77}$$

The number of available ADC bits is technically limited by the required sampling frequency and can be as high as 24 bits for slow audio processing (kHz range) and as low as 6 bits for sampling in the GHz range. The LHC beam position monitor, for example, uses a 10 bit ADC at a sampling frequency of about 40 MHz which results in a relative discretisation of about $2 \cdot 10^{-3}$. In order to improve the performance

in the presence of quantisation, often, statistical noise signals with an amplitude of half the discretisation is added, referred to as *dither*, in order to force the last significant bit to oscillate. Using this dither technique, it is possible to increase the precision for very slow signals through statistical averaging. The dither is visible as the white noise floor in the Fourier spectrum of the measured signal, as constant amplitude spectrum for all frequencies.

3.5.2 Aliasing Effect

The maximum sinusoidal signal with frequency f that can be unambiguously resolved with a sampled system is given by the *Nyquist frequency*

$$f_{NQ} = \frac{1}{2} f_s, \tag{3.78}$$

with $f_s = \frac{1}{T_s}$ the sampling frequency. However, in case f is larger than the Nyquist frequency an ambiguity occurs in the reconstruction, referred to as *aliasing*, illustrated in Figure 3.9. For signals with frequencies larger the Nyquist frequency, at least two signals with different frequencies can be fitted through the sampled points. Another representation for the aliasing effect is through the frequency domain. In a

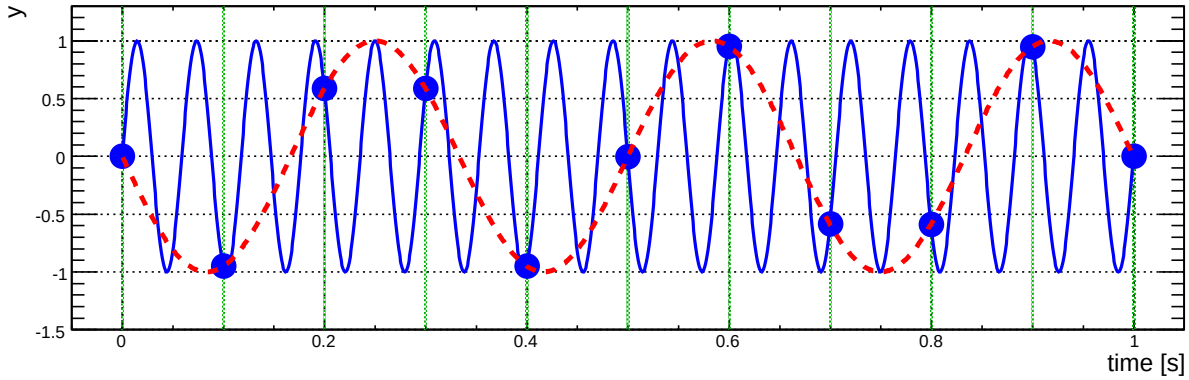


Figure 3.9: Aliasing effect: the initial high frequency oscillation (solid blue line), the sampling (dots) and the erroneous reconstruction (dashed red line) are shown.

sampled system, the signal's Fourier spectra repeats itself around multiples of the sampling frequency f_s with N an integer, as illustrated in Figure 3.10. The frequency range $[-f_{NQ}, +f_{NQ}]$ around zero is referred to as *base band*. In case the bandwidth of the frequency to be sampled is larger than the Nyquist

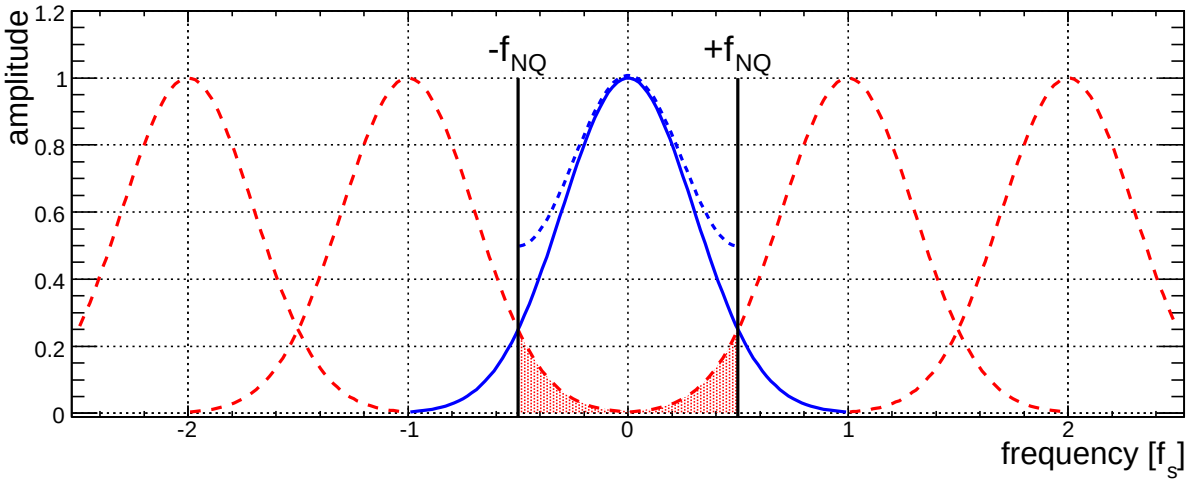


Figure 3.10: Aliasing effect in frequency domain. The base band signal (solid blue line), the side bands (dashed red lines), spectral leakage from the upper and lower sideband into the base band (filled area) and resulting Fourier spectra (blue dotted line) are shown. The larger the bandwidth of the signal the larger the possible leakage into the base band.

frequency, the outermost sides slopes of the Fourier spectra in the side band may extend into the base

band causing the aliasing effect, also referred to as *spectral leakage*. The apparent aliasing frequency f_a is given by:

$$f_a = f - n \cdot f_s \tag{3.79}$$

with $n \in \mathbb{Z}$ chosen so that f_a lies within the base band:

$$-f_{NQ} < f_a < +f_{NQ} \tag{3.80}$$

The aliasing effect may have an impact on feedback performance: In an analogue signal high frequency perturbation signals are usually filtered out through the low-pass characteristic of most processes and loop designs. In a sampled system, these high frequency signals may, due to the aliasing effect, leak into the base band at frequencies where the feedback has significant gain to compensate these spurious signals.

There are two common techniques to reduce the effect of aliasing: Either the initial signal is low-pass filtered in the analogue domain prior to the ADC conversion or the signal is sampled at a higher frequency. The first method effectively reduces the signal's side slopes that may otherwise leak into the base band whereas the second method provides more bandwidth in the base band in order to fit a signal with a potentially broad bandwidth. Digital filters (intrinsically) cannot remove the aliasing effect and the latter method just shifts the aliasing problem for feedbacks to higher frequencies.

Chapter 4

Beam Perturbation Sources

The stability of the beams is an important parameter for safe and reliable machine operation of the LHC. The performance of the LHC Beam Cleaning System, for example, critically depends on orbit stability, with tolerances down to $\approx \frac{\sigma}{3}$ (σ= r.m.s. beam size) corresponding to $\approx 300\,\mu$m at 450 GeV and $\approx 70\,\mu$m at 7 TeV. This chapter discusses the sources for orbit and energy perturbations. The stability and tolerance requirements and comparison to the expected perturbations is further discussed in Section 4.5.

The sources for orbit and energy perturbation can be grouped into three classes:

1. Environmental sources such as ground motion (see section 4.1), tides, temperature and pressure changes (section 4.3), cultural noise, and other effects. These effects are propagated through the magnets (mainly the quadrupoles) and their girders[1] to the beams.

2. Machine-inherent sources such as decay and snap-back of the superconducting magnets, eddy currents on the vacuum chamber and ramp-induced effects as discussed in Section 4.4.1, flow of cooling liquids, vibrations of the ventilation system, and changes of the final focus optics (section 5.2.3). In case of the beta-squeeze, the largest orbit perturbations can reach up to 30 mm, if uncompensated.

3. Machine element failures particularly orbit correction dipole magnets (CODs) and beam separation elements as discussed in Section 5.2.3

4.1 Periodic and Random Ground Motion

This section presents an analysis of periodic and random ground-motion-induced orbit drifts and presents two models describing the propagation and amplification of ground motion on the SPS, LEP, and LHC orbits, qualitatively and quantitatively[2]. Long-term ground settlement effects such as those described and analysed for LEP [61] based on long-term alignment data are not within the scope of this analysis. Two models that describe the propagation and amplification of correlated and uncorrelated ground motion on the SPS, LEP and LHC orbits are presented.

The analysis shows that the largest ground motion contribution to the orbit drift on the time scale ranging from seconds to a month is due to random ground motion and is barely influenced by correlated ground motion. Upper limits of the expected drifts at the LHC, SPS, and the TI8 and CNGS[3] transfer lines due to ground motion are given. The predictions are supported by beam measurements at LEP in 1999 and ground and beam motion based measurements performed in 2004 in the SPS and TI8.

[1] *Girders* are the supporting structure that are used to carry the weight and align the magnets.
[2] The analysis has also been published in [60] as part of my doctoral studies
[3] CNGS: CERN Neutrino to Gran-Sasso, experimental setup for the detection of neutrino oscillation [62].

4.1.1 The Accelerator Tunnels

The SPS and LHC/LEP tunnels have a circumference of about 6.9 km and 26.7 km, and an average depth of about 50 m and 100 m, respectively. Both tunnels are embedded in the *Molasse*, a soft tertiary sandstone on top of a hard rock basin found in the region. The *Molasse* mainly consists of clay and limestone eroded from the surrounding Jura and the Alps and is covered by the *Moraine*, a loose and permeable more recent quaternary erosion from the Jura. An important feature of this geological formation is the different propagation speeds and refraction indices for ground waves in the Moraine and Molasse. Seismic faults and wells in the region of the accelerators are considered to be inactive and are neglected on the time scale ranging from a few hours to a month [63, 64].

4.1.2 Ground Motion Model

The main influence of ground motion on particle beams is through the displacement of quadrupoles and girders. A quadrupole misaligned by ϵ introduces a dipole kick δ proportional to its focusing strength k and length l, $\delta = kl \cdot \epsilon$. This dipole kick creates a perturbation $\Delta x(s)$ of the closed orbit that can be written as:

$$\Delta x(s) = \frac{\sqrt{\beta(s)\beta_j}}{2\sin(\pi Q)} \cdot \cos(\Delta\mu - \pi Q) \cdot \delta \tag{4.1}$$

Respectively, for a transfer line:

$$\Delta x(s) = \sqrt{\beta(s)\beta_j} \cdot \sin(\Delta\mu) \cdot \delta \tag{4.2}$$

It is important to note, that due to the single passage nature, the beam position is only influenced by magnets upstream to the observation point. The betatron oscillations depend on the values of the betatron-function at the location of the kick (β_j) and the observation point ($\beta(s)$); Q is the machine tune and $\Delta\mu$, the positive phase advance from the location of the dipole kick to the observation point. In general, the orbit is sampled at m beam position monitors (BPM). The displacement at monitor i due to n quadrupoles labelled by j is

$$\Delta x_i = \sum_{j=0}^{n} \frac{\sqrt{\beta_i\beta_j}}{2\sin(\pi Q)} \cdot \cos(\Delta\mu_{ij} - \pi Q) \cdot k_j l_j \cdot \epsilon_j \tag{4.3}$$

respectively for transfer lines:

$$\Delta x_i = \sum_{j=0}^{n} \begin{cases} \sqrt{\beta_i\beta_j} \cdot \sin(\Delta\mu) \cdot k_j l_j \cdot \epsilon_j & : \quad \Delta\mu > 0 \\ 0 & : \quad \Delta\mu \leq 0 \end{cases} \tag{4.4}$$

Both equations can be written as

$$\Delta x_i = \sum_{j=0}^{n} R_{ij} \cdot \epsilon_j \tag{4.5}$$

where R_{ij} is the element of the $m \times n$ orbit response matrix and depends on the machine optics. In order to be less dependent on the value of the beta-function at the specific BPM and to give results in terms of beam size, it is useful to normalise Equation 4.5 by $\sqrt{\beta_i}$

$$\Delta x_i/\sqrt{\beta_i} = \sum_{j=0}^{n} (R_{ij}/\sqrt{\beta_i}) \cdot \epsilon_j \tag{4.6}$$

$$\Delta \tilde{x}_i = \sum_{j=0}^{n} \tilde{R}_{ij} \cdot \epsilon_j \tag{4.7}$$

The average effect of misaligned quadrupoles (average quadrupole movement $\bar{\epsilon}$) onto the normalised ($\tilde{\sigma}$) and un-normalised orbit r.m.s. σ can be approximated by factors κ and $\tilde{\kappa}$, respectively:

$$\begin{aligned} \sigma &= \sqrt{\tfrac{1}{m}\textstyle\sum_{i=0}^{m} \Delta x_i^2} = \kappa \cdot \bar{\epsilon} \\ \tilde{\sigma} &= \sqrt{\tfrac{1}{m}\textstyle\sum_{i=0}^{m} \Delta x_i^2/\beta_i} = \tilde{\kappa} \cdot \bar{\epsilon} \end{aligned} \tag{4.8}$$

More generally, the cumulative effect of the ground movement on the orbit strongly depends on the frequency and spatial correlation of the movements. The relation between the power spectrum distribution (PSD) or r.m.s. movement of the ground motion (S_{gm}) and the orbit (S_{orbit}), assuming a rigid girder response, can be approximated using the factor $\kappa(f)$ to:

$$S_{orbit}(f) = \kappa^2(f) \cdot S_{gm}(f) \tag{4.9}$$

and

$$\tilde{S}_{orbit}(f) = \tilde{\kappa}^2(f) \cdot S_{gm}(f) \tag{4.10}$$

To simplify the analysis, the movement can be decomposed into random (fully un-correlated) and coherent (fully correlated) ground motion.

Random Ground Motion

For the application in the SPS and LEP/LHC model, one can assume that the accelerator tunnels are entirely embedded in a homogeneous ground. While this is certainly true for the SPS, the assumption holds for the LEP/LHC tunnel as well, since only about 3 km of the 27 km circumference is located in hard rock regions of the Jura mountains.

For random misalignment, the optical amplification factor κ of Equation 4.9 is independent of the exciting frequency. For a regular FODO lattice with N cells, κ can analytically be approximated to (see Annex):

$$\kappa = \frac{kl\beta_{eff}}{4\sin(\pi Q)} \cdot \sqrt{N} \tag{4.11}$$

Here kl is the average integrated quadrupole strength, Q the tune and β_{eff} the effective beta function, $\beta_{eff} = \sqrt{\beta_{QF}^2 + \beta_{QD}^2}$, with $\beta_{QF(D)}$ the betatron function at the horizontal focusing (defocusing) quadrupoles. This approximation is useful for qualitative purposes and regular FODO lattices. However, it underestimates non-regular lattices with insertions having large quadrupole strengths and large values of beta functions. The choice of fractional tune critically influences the size of the optical amplification due to the factor $\sin(\pi Q)$ in the denominator of Equation 4.11. For comparison, the analytical amplification factors of the SPS and LHC arc FODO lattice (Equation 4.11) are given in the following table.

	horizontal		vertical	
	κ	$\tilde{\kappa}$	κ	$\tilde{\kappa}$
SPS	31	3.0	23	2.2
LEP	36	3.1	48.2	2.4
LHC	26	1.9	20	1.5

Table 4.1: Approximate propagation (scaling) factors for random motion as given by Equation 4.11 for the SPS, LEP, and LHC arc FODO cells.

To take all details of the optics (such as final focus, varying phase advances, etc.) into account, κ is evaluated numerically using Equations 4.3, 4.4 and 4.9. For each sample, all quadrupoles are misaligned. The computed ratio between quadrupole shift and resulting orbit r.m.s. is used to evaluate the individual amplification. The average and r.m.s spread of κ are evaluated from 10^5 samples for a given optics and plane. Since ground motion affects all elements, a corrective BPM shift was applied. In case a BPM is installed next to a quadrupole, it is assumed that the BPM is rigidly connected to the corresponding quadrupole. Otherwise, the BPM is shifted independently. This approximation is valid for LHC and LEP as nearly all BPMs are directly mounted on the quadrupoles. At the SPS, the BPMs are not mounted on the same girder as the quadrupoles, but are very close (≈ 1 m) to the quadrupoles, such that the 'rigid connection' approximation may still be valid. The simulations are performed for the SPS (with LHC beam tunes, $Q_H = 26.13$ and $Q_V = 26.18$), LEP (1999 collision optics, $Q_H = 98.28$ and $Q_V = 92.20$) and LHC (version 6.5, injection and collision optics, $Q_H = 64.28$ and $Q_V = 59.31$). The simulations for the TI8 and CNGS transfer lines are performed similarly, with the exception of the modified beam transfer function. These κ values correspond to the orbit drift at the last TI8 TED (mobile dump) and the CNGS target, respectively. Table 2 summarises the results.
The spread of κ given in Table 2 is an intrinsical uncertainty due to the sensitivity of the response to the exact "seed" of the movement. The actual individual amplification is expected to be within these

	horizontal		vertical	
	κ	$\tilde{\kappa}$	κ	$\tilde{\kappa}$
SPS	40.6 ± 19.3	3.1 ± 1.4	30.1 ± 13.3	4.1 ± 2.0
LEP coll	34.3 ± 13.7	5.0 ± 2.0	119.5 ± 65.7	11.2 ± 6.1
LEP coll*	34.3 ± 13.7	5.0 ± 2.0	58.6 ± 25.9	5.6 ± 2.4
LHC inj	30.5 ± 11.5	3.2 ± 0.8	29.6 ± 9.0	3.0 ± 0.7
LHC coll	63.3 ± 32.5	5.2 ± 2.0	62.1 ± 25.5	4.9 ± 1.9
TI8 TED	19.5 ± 14.6	1.8 ± 1.3	10.6 ± 8.0	1.6 ± 1.2
CNGS target	4.2 ± 3.2	1.3 ± 1.0	4.9 ± 3.6	1.1 ± 0.8

Table 4.2: Simulated orbit amplification factor κ for randomly misaligned quadrupoles. The LHC collision optics (version 6.5) includes the final focus to $\beta^* = 0.55\,\mathrm{m}$ in IR1 (ATLAS) and IR5 (CMS). The LEP '*' values do not include the systematic shifts due to the low-beta quadrupoles (QS0) around the four interaction regions.

limits with a probability of about 68%. The amplification analysis shows that LEP was more sensitive to vertical random ground motion than the LHC. The large difference of the κ factors between the LHC and LEP collision optics is due to the very strong vertical focusing around the interaction region in LEP. In LEP, the insertion quadrupoles contribute to about 50% of the total κ value. Removing these quadrupoles ("QS0") yields comparable numbers for LEP and LHC.

One can use a Brownian process to model time dependence of the random motion of each individual quadrupole as well as to describe the physical properties of the measured data. We ignore, while discussing the propagation of random ground motion only, any coupling between planes and spatial correlation. In this simple model, each quadrupole movement is described by Stoke's differential equation for a viscous medium. The equation describes the dependence of the position x of a damped (α being the damping factor) mass m under influence of an external randomly varying force $F(t)$:

$$m\ddot{x} + \alpha\dot{x} = F(t) \tag{4.12}$$

A more complete model would include a spring-like retracting force ($+kx$) acting on very large excursions, constant and coupling forces. Though an analytical solution of equation 4.12 and a direct prediction of the random movement is not possible, one can derive the following important statistical properties of this type of motion, as shown in the Annexe:

1. For long time-scales, the mean random movement $\mu = \langle x \rangle$ vanishes:

$$\lim_{t\to\infty} \mu = 0 \tag{4.13}$$

2. For large time scales the root-mean-square $\sigma = \sqrt{\langle x \rangle^2 - \mu^2}$ of Brownian motion increases proportionally to the square root of time as derived in the Appendix (A being the diffusion constant):

$$\lim_{t\to\infty} \sigma_{rms} = A \cdot \sqrt{t} \tag{4.14}$$

3. The power spectrum (density) shows a typical 'brown'-noise $\frac{1}{f^2}$ dependence on the frequency.

Coherent Ground Motion

There is a broad frequency spectrum of coherent ground motion sources that affect the orbits:

1. Natural sources such as ground settlement, tectonic motion, tides, earthquakes, microseismic noise, seasonal temperature changes, tunnel temperature and atmospheric pressure changes, change of water levels (lake or ground), etc.

2. Cultural noise, such as noise due to human activity on the surface, for instance railroad traffic, transformer vibrations, civil construction, etc.

3. In situ tunnel equipment: cooling water and cryogenic coolant flow, ventilation, motors, mechanical vibrations of magnets due to time-varying fields, etc.

Only coherent ground waves with source locations outside the tunnel perimeter are considered for the propagation model. These can, in the far field, be described by plane waves. Ground settlement, potential fault activity and tunnel equipment-specific vibrations are neglected.

Rayleigh waves are the most important ground waves that dominate the coherent background. Found only near the surface, they are created through interference of longitudinal "pressure" and transverse-"shear" waves (further referred to as P- and S-waves) from deeper regions while travelling to the surface. Rayleigh waves have longitudinal and vertical particle oscillation with respect to the propagation direction, and the particles describe retrograde ellipses as shown in Figure 4.1.

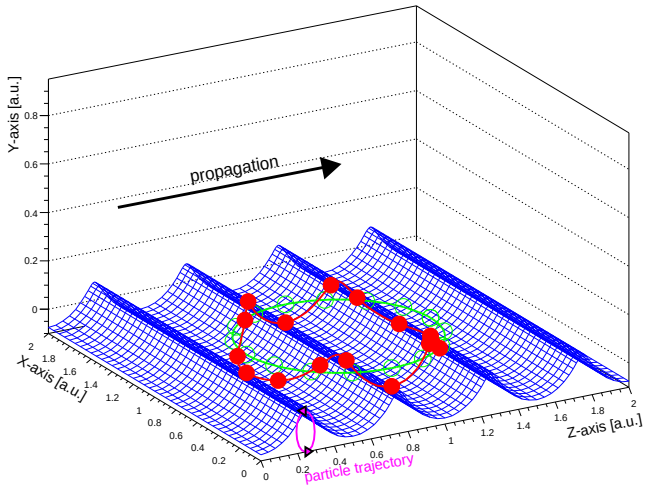


Figure 4.1: Example of an accelerator that is unperturbed (green) or perturbed by a Rayleigh wave (red). The wave consists of transverse and longitudinal oscillations. The particles describe a retrograde ellipse with respect to the propagation direction.

P-, S- and Rayleigh wave propagation velocities differ, fulfilling the relation $c_p > c_s > c_{Rayleigh}$. To first order, the density of the ground increases with depth. Since the propagation velocity of sound is a function of density, Rayleigh waves have a dispersion relation depending on depth, and consequently lose their coherence with depth (dissipation) as well. The accelerator depth is almost constant and hence the dispersive effect is ignored in this analysis. The analytical solution for P-, S- and surface Rayleigh waves for a point source (wavelength λ) show that the damping D at a distance r from the motion source ($r_0 > \lambda$) can be decomposed into three parts, geometric, dissipative (Q_d, site-specific (unit-free) quality factor ranging from 10-25 for near surface to several hundreds for hard rock) and depth-dependent part (h: depth from the surface) as shown in the following equation:

$$\mathrm{D_R} = \sqrt{\frac{r_0}{r}} \cdot e^{-\frac{\pi(r-r_0)}{Q_d \lambda}} \cdot e^{-\frac{h}{\lambda}} \tag{4.15}$$

$$\mathrm{D_{P/S}} = \frac{r_0}{r} \cdot e^{-\frac{\pi(r-r_0)}{Q_d \lambda}} \tag{4.16}$$

Though their wavefront arrives later, Rayleigh waves carry most of the energy over long distances and prevail over P- and S-waves because of the reduced geometric damping (Rayleigh waves $\sim r^{-0.5}$ vs. $\sim r^{-1}$ for P- and S-waves). The depth dependent damping of Rayleigh waves favours deeper underground tunnels as quiet locations. For example, assuming a wavelength of $\lambda = 50\,\mathrm{m}$, an increase of the tunnel depth from 50 m to 100 m has about twice the damping effect than doubling the distance to the source. In addition, ground waves die out faster as the wavelength decreases. Consequently, the coherence (or correlation) length, which is the maximum distance of two points oscillating coherently decreases rapidly with frequency and distance, as shown by measurements in the LEP and TT2 tunnels [65]. These measurements have been reproduced in the Appendix, since this document is difficult to obtain.

The geological configuration around Geneva has an additional effect on the propagation of coherent waves. Molasse and Moraine have different propagation velocities that correspond to different refraction indices for these types of waves. Similar to light optics, coherent waves created on the surface due to human-induced 'cultural noise', for example, will be partially reflected, partially transmitted and

refracted at the boundary layer between Moraine and Molasse, causing an additional reduction of those amplitudes inside the tunnel. Under certain conditions of wavelength and incidence angle with respect to the boundary layer, these waves may even be totally reflected without penetrating the Molasse layer where most of the LEP tunnel is situated. As will be shown later, the largest contribution of coherent ground motion in the LHC tunnel is less due to the cultural activity on the surface, but rather due to long range effects such as the 'ocean hum'.

The amplification factor $\kappa(f)$ of Equation 4.9 was evaluated for coherent waves from a simulation of the accelerator motion. Although there is no simple approximation to describe the wave motion in the near field, one can, for the far field and constant depth, approximate Rayleigh waves by a superimposition of plane P- and S-waves with the same propagation velocity for both planes, as shown in Figure 4.2.

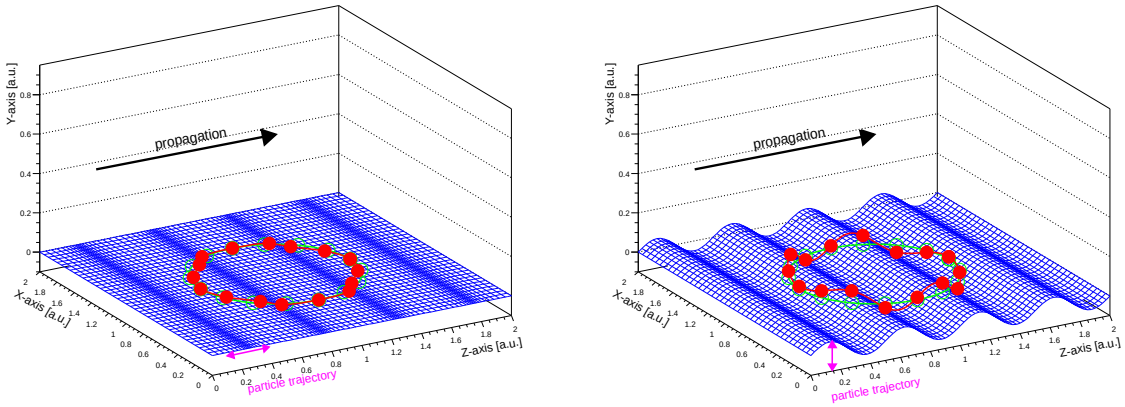


Figure 4.2: Perturbed (red) and unperturbed accelerator (green). Left: Example of a horizontal P-wave oscillation, when the particles oscillate longitudinally with respect to the wave propagation. Right: S-wave oscillation, the particles oscillate transversely with respect to the wave propagation.

This calculation extends earlier results described in [66] that were done for the vertical plane (S-waves) using only an early LHC optics and one main wave incident direction.

The proposed propagation model does not include the dissipative and dispersive nature of the ground that may cause a reduction of the coherence after some distance. Further, it is assumed that the drifts are sufficiently slow that the orbit moves adiabatically.

For the simulation, the quadrupoles are misaligned according to the wave equations and the machine's geodesy data. For the BPMs, the same corrective shifts are applied as for the analysis of random ground motion. The beam response $\kappa(f)$ is defined by the ratio between orbit r.m.s. and ground wave amplitude. The response factors are averaged over different wave phases. The error band in the figures corresponds to the r.m.s. spread over the possible range of wave phases. The wave incidence was varied as a second parameter. The 'zero-degree' angle corresponds to a wave propagation direction parallel to the line going through IP5 (CMS) and IP1 (ATLAS) which is close to the north-south axis.

The results are computed and given as a function of ground motion wavelengths. However, in order to relate the wavelengths to ground motion frequencies, an approximated Rayleigh wave propagation velocity of 2000 m/s is used, which is typical for the *Molasse* at the LHC depth. The actual propagation velocity may differ. The results for the SPS are shown in Figures 4.3 and 4.4 and the results for the LHC injection and collision optics are shown in Figures 4.5 and 4.6. The error bands correspond to r.m.s. spread. One can separate the spectra into three regions:

1. Low frequency region: the amplification κ vanishes since the wavelength exceeds the accelerator diameter and the whole machine is lifted (S-wave) or shifted (P-wave) coherently.

2. Betatron resonance region: Resonant amplification occurs once the wavelength λ (frequency f) drops below the first harmonic of the betatron-wavelength λ_{beta} and the frequency f_{beta}, respectively,

$$\begin{aligned} \lambda_{beta} &\approx \frac{C}{Q}, \\ f_{beta} &\approx \frac{v}{\lambda_{beta}}, \end{aligned} \tag{4.17}$$

C is the accelerator circumference and v the propagation velocity of the ground wave. For the LHC,

the betatron-wave length is about 415 m and 450 m for the horizontal and vertical plane. In the SPS, the betatron-wavelength is about 264 m for both planes.

3. Moire-like pattern regime: Once the wavelength exceeds the minimum distance between quadrupoles l_{min}, the regular harmonic resonances become less important. The amplification is determined by the absolute phase relations between the wave front and individual element location and the local optics. For example, in the case of LHC collision optics, the effect of the insertion quadrupoles becomes increasingly visible once the ground wavelength is about the same distance as between the triplets around the experimental insertion regions. Further, it is visible that the averaged response functions for very high frequency are about the same as the amplification factor for random ground motion, as given in Table 2.

Apart from the 8- and 6-fold symmetry, due to the long straight sections of the LHC and SPS, the LHC injection and SPS amplification spectra do not show any significant angle dependence. The LHC collision optics has a small dependence on the wave propagation direction, as seen in Figure 4.7. The vertical amplification factor is larger for S-wave propagation along the IP3-IP6 axis above about 30 Hz, corresponding to the resonant mode of the inner triplets in IP1 and IP5. The horizontal amplification factor dependence on angles is rather small but is visible above 30 Hz. The dependence on the propagation direction is mainly visible above l_{min}, where its effect is basically negligible as the wave is likely to completely lose either its coherence or its energy over a fraction of the accelerator dimension. This is due to the strong dissipative effect of the ground for high frequencies (Equation 4.15), and confirmed by coherence length measurement performed in the TT2 tunnel [65] (see Appendix F.3 for reprint of figures).

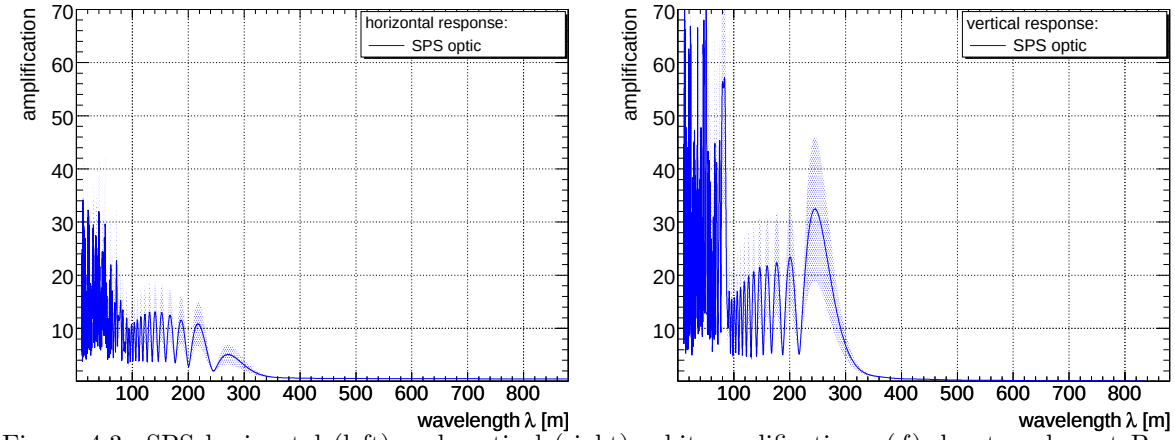


Figure 4.3: SPS horizontal (left) and vertical (right) orbit amplification $\kappa(f)$ due to coherent P- and S-wave oscillations. The amplification is shown as a function of wavelength. The first resonance once the wavelength drops below the betatron-wavelength of about 264 m is visible.

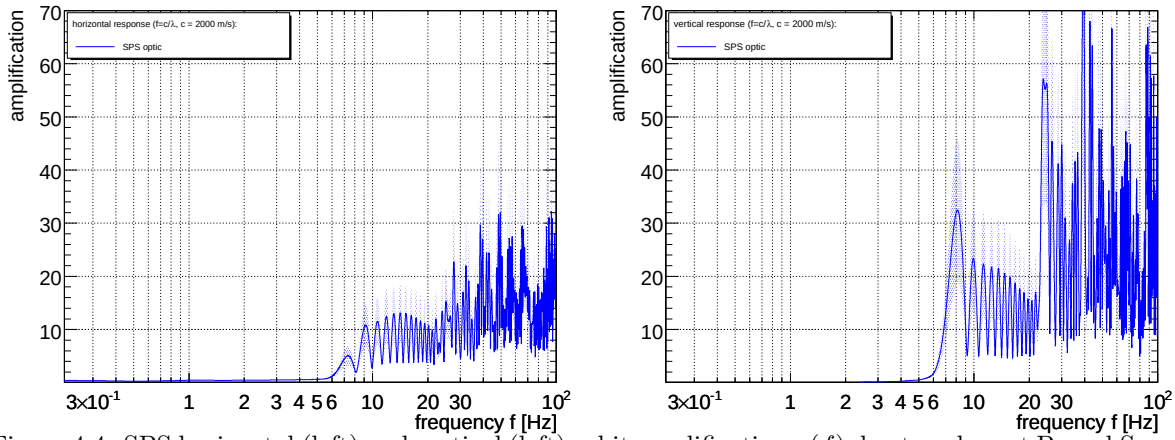


Figure 4.4: SPS horizontal (left) and vertical (left) orbit amplification $\kappa(f)$ due to coherent P- and S-wave oscillations. The amplification is shown as a function of frequency for the horizontal (left) and vertical (right) plane assuming a ground wave propagation velocity in the *Molasse* of $c = 2000$ m/s.

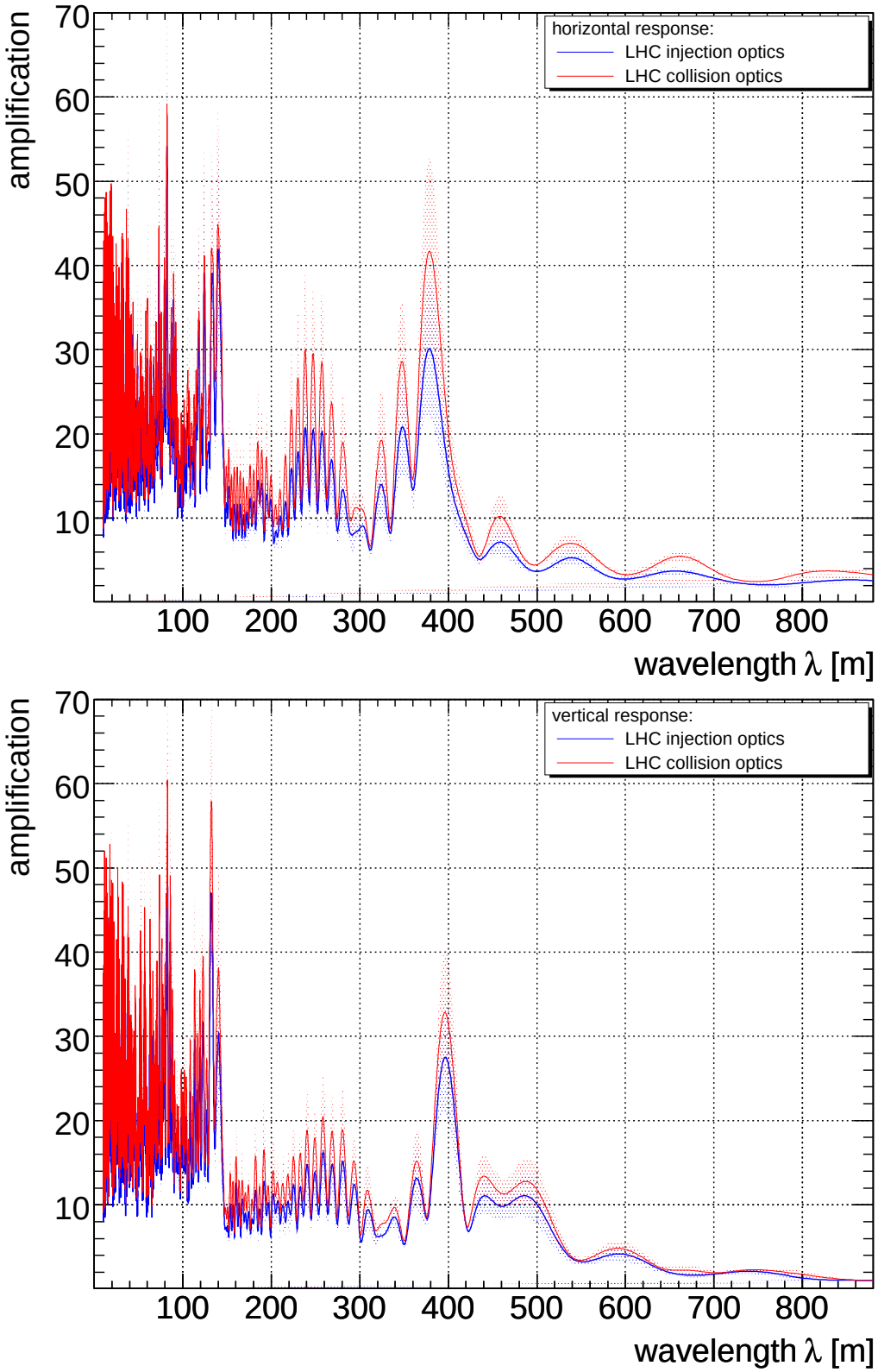


Figure 4.5: LHC horizontal (left) and vertical (right) orbit amplification $\kappa(f)$ due to coherent P- and S-wave oscillations. The first resonance once the wavelength drops below the betatron-wavelength of about 415 m (horizontal) and 450 m (vertical) is visible.

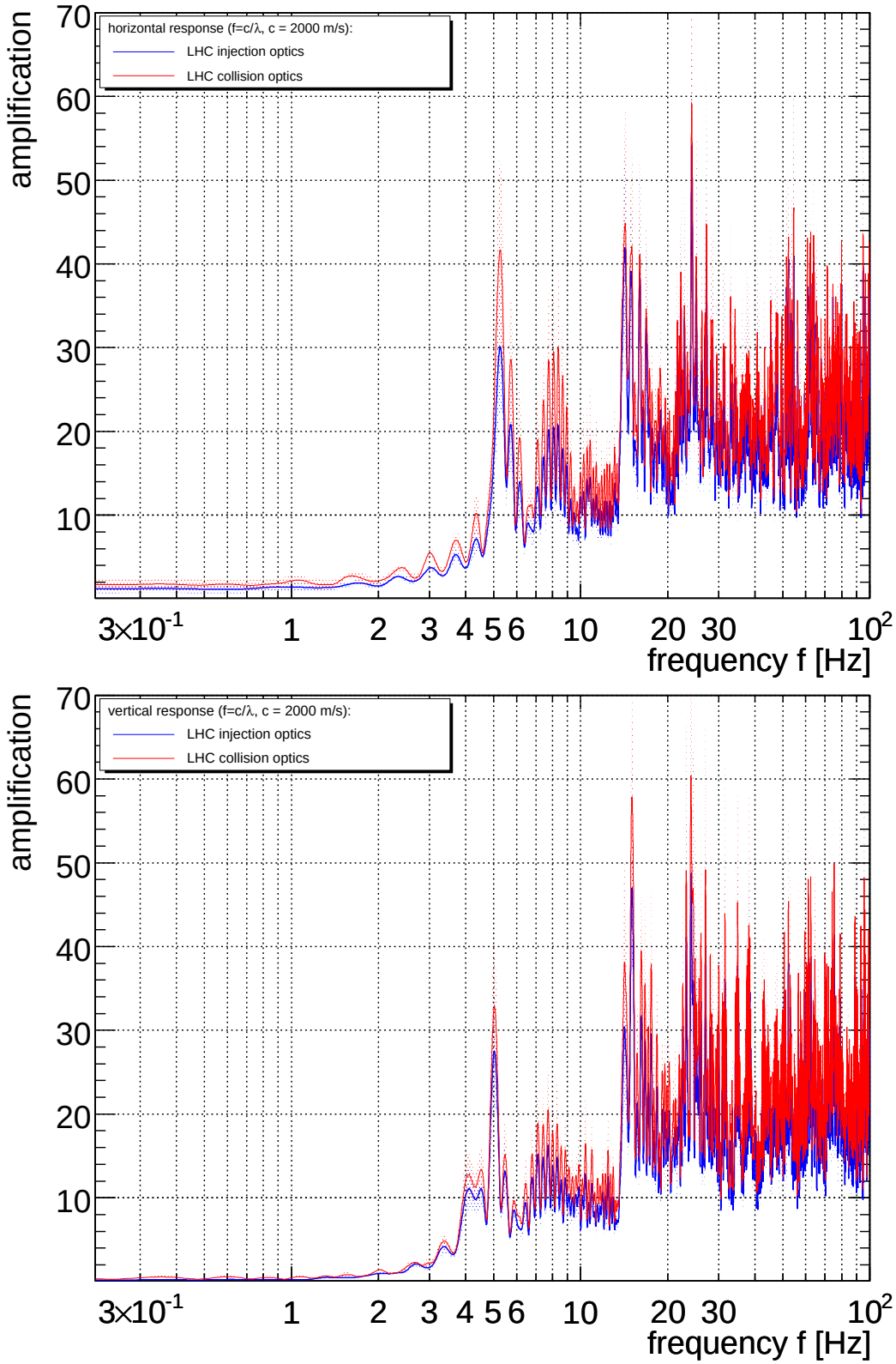


Figure 4.6: LHC horizontal (top) and vertical (bottom) orbit amplification $\kappa(f)$ due to coherent P- and S- wave oscillations. The amplification is shown as a function frequency assuming a ground wave propagation velocity in the *Molasse* of $c = 2000\,\mathrm{m/s}$.

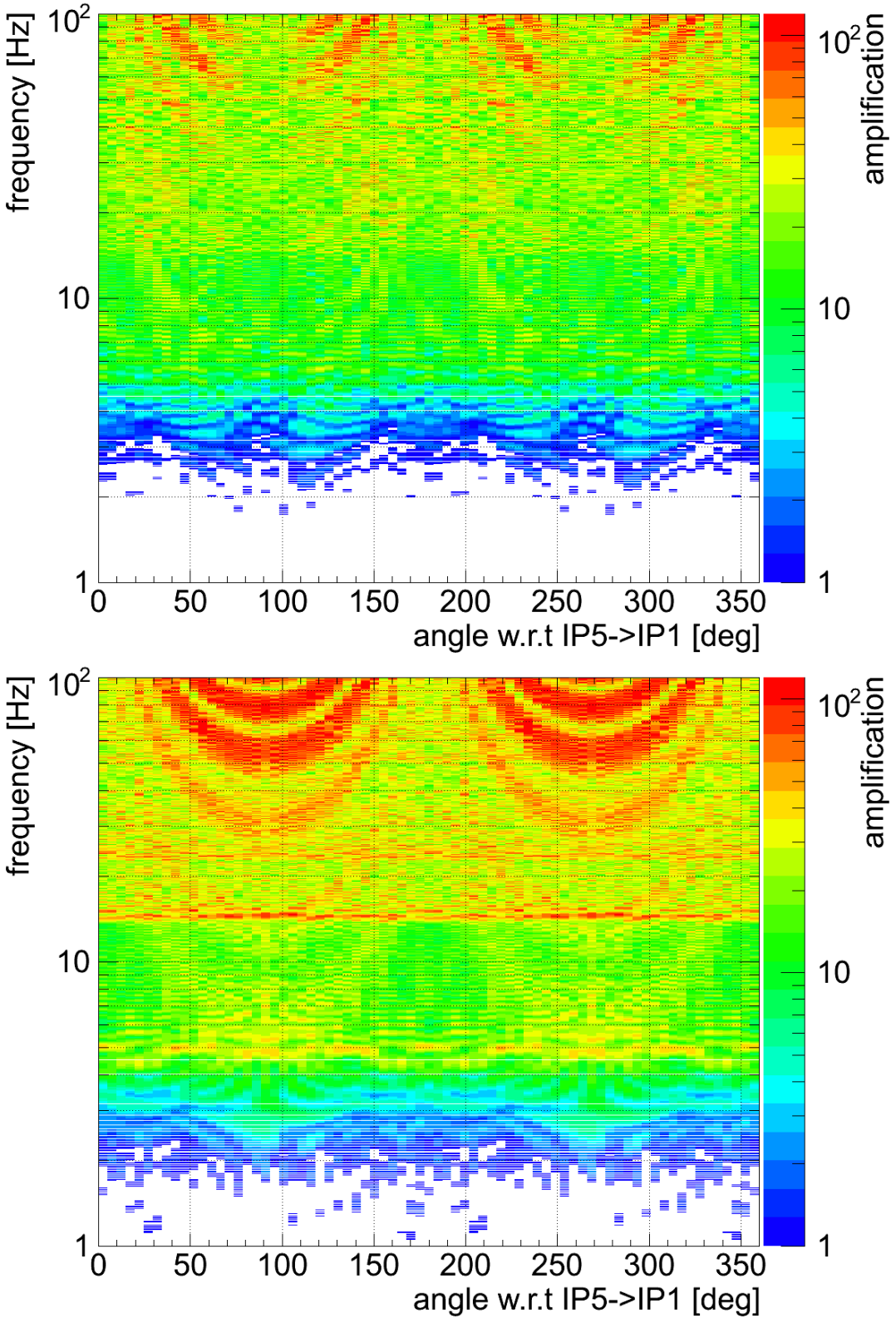


Figure 4.7: LHC amplification factor dependence on the propagation direction for the horizontal (top) and vertical plane (bottom). The 'zero-degree' angle corresponds to a wave propagation direction parallel to the axis going through IP5 (CMS) and IP1 (ATLAS).

4.1.3 Comparison with Experimental Data

Seismic measurements

In 2004, long-term seismic SPS tunnel vibrations and SPS beam drifts were studied. The seismic measurements were performed in the SPS tunnel with a *Guralp CMG-40T* type geophone[4]. The device measures the transverse, longitudinal, and vertical velocity of a damped test mass at a rate of 200 Hz. The sensitivity decreases for very high and very low frequencies, resulting in an underestimation of amplitudes. After calibration, the device is sufficiently linear inside the frequency band from 0.03 to 50 Hz. The probe was positioned such that the reference system corresponds to the accelerator system where 'longitudinal' refers to the axis along the tunnel, the 'tranverse' axis points outwards, and the 'vertical' axis points upwards.

The data were taken between Tuesday, 9^{th} March 2004 and Monday, 14^{th} March 2004 on the ground at the SPS quadrupole QF.522. During the data collection period, installation work was performed in the vicinity (≈ 100 m). Data were acquired for a duration of 6 minutes every half an hour. The sensor sensitivity vanishes for periods larger than 30 s. In order that the Fourier transform is not the limiting factor of the resolution and sensitivity, the analysis window was chosen to be 60 s long. The individual spectra of one acquisition are averaged to reduce noise. We use the Fast Fourier algorithm described and defined in [67]. The velocity spectra ($P_v(f)$) are converted to amplitude spectra ($P_x(f)$) using the following relation:

$$P_x(f) = \frac{1}{(2\pi f)^2} P_v(f) \tag{4.18}$$

The general Fourier transform only specifies that the forward-backward transform of a signal has to yield the same signal but does not explicitly specify the normalisation of the spectra. Hence, there are various FFT spectra definitions. We use the following spectra normalisation in the presented figures: a coherent sinusoidal signal with a constant amplitude A corresponds to a power A^2 in the power spectra and a power density $A^2/\partial f$ in the power density spectra ($1/\partial f$ being the length of the Fourier window). Power spectra (squared Fourier spectra) are better suited for coherent signals, whereas power spectrum density representation is more appropriate when dealing with random signals. For reference, data are given in both forms. For purely random signals, the long-term drift due to frequencies $\geq f$ is obtained from the discrete *integrated r.m.s.* definition

$$I(f) = \sqrt{\sum_{f}^{f_{max}} \mathrm{PSD}(f)\Delta f} \tag{4.19}$$

with $\mathrm{PSD}(f) = P_x(f)/\Delta f$ being the power spectra density and Δf the frequency bin width. This definition is not valid for a superposition of coherent signals. It is important to note that geophones measure the combined effect of random and coherent ground motion. Direct propagation and conclusion from the measured ground motion spectrum to the beam, without knowing the apportionment between the random (correlated) and coherent (un-correlated) part of the spectrum, is generally not possible.

In order to identify the type, coherence, and potential period of a signal, the auto-correlation function that is defined as:

$$A(t) = \int_{-\infty}^{+\infty} x(\tau) \cdot x(t-\tau)\, d\tau \tag{4.20}$$

Figure 4.8 shows the averaged spectra for the longitudinal, horizontal, and vertical planes acquired during quiet periods (weekend). The $\frac{1}{f^2}$ dependence that is typical for random motion and drifts and the peak around 0.1 Hz due to ocean hum is visible. Longitudinal, horizontal and vertical power spectra have similar magnitudes. This is compatible with Rayleigh waves. The contribution due to cultural noise vanishes during lunch times, indicating that the noise is largely due to civil construction activity in the tunnel and less due to surface activities.

A comparison of typical vertical SPS and LEP/LHC ground motion spectra is shown in Figure 4.9. Both tunnels are very quiet and are barely influenced by cultural noise. The spectra are essentially the same. Since both accelerators are embedded in the same ground, it is possible to use SPS orbit data to predict orbit drifts at the LHC, provided the respective factors arising from the beam optics are understood.

For the LHC, random ground motion dominates over coherent ground motion. Though the latter may have a stronger amplification of up to a factor of $\kappa(f) \approx 60$ for frequencies above 3 Hz, they contribute

[4] Device and measurements courtesy S. Redaelli.

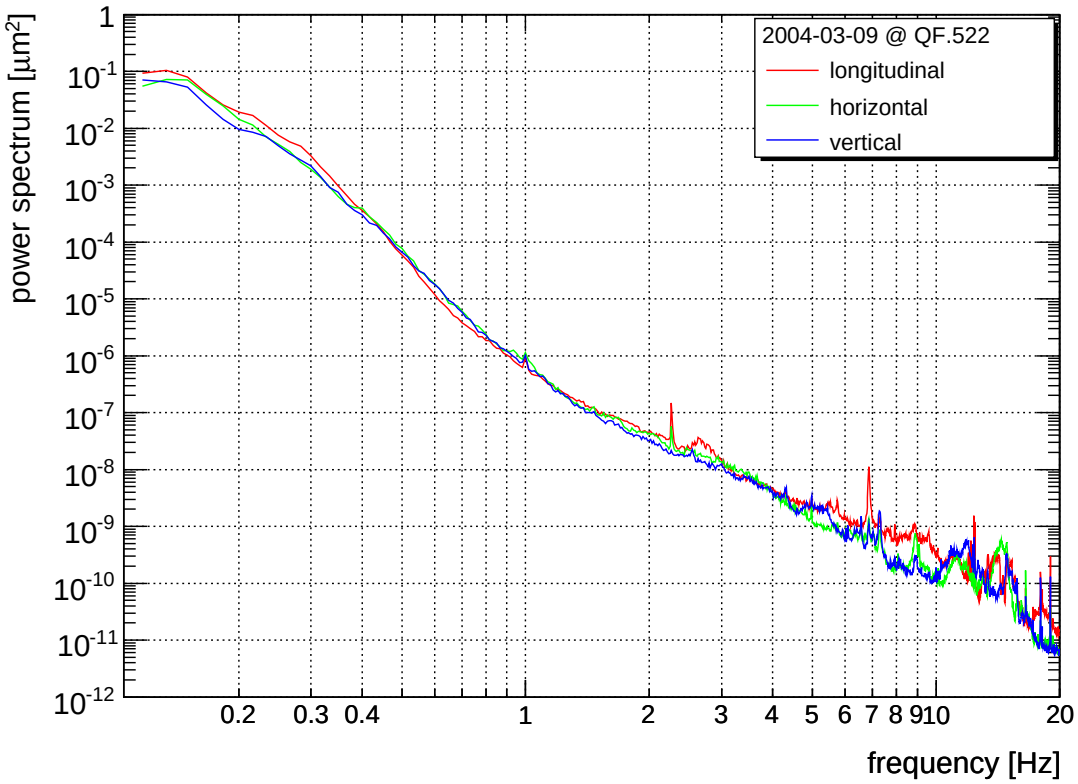


Figure 4.8: Averaged ground motion power spectra in the SPS taken at QF.522. The spectra for the different planes are to first order identical, but show slightly stronger amplitudes for the longitudinal and horizontal planes, compatible with the elliptical motion of Rayleigh waves that predict the asymmetry.

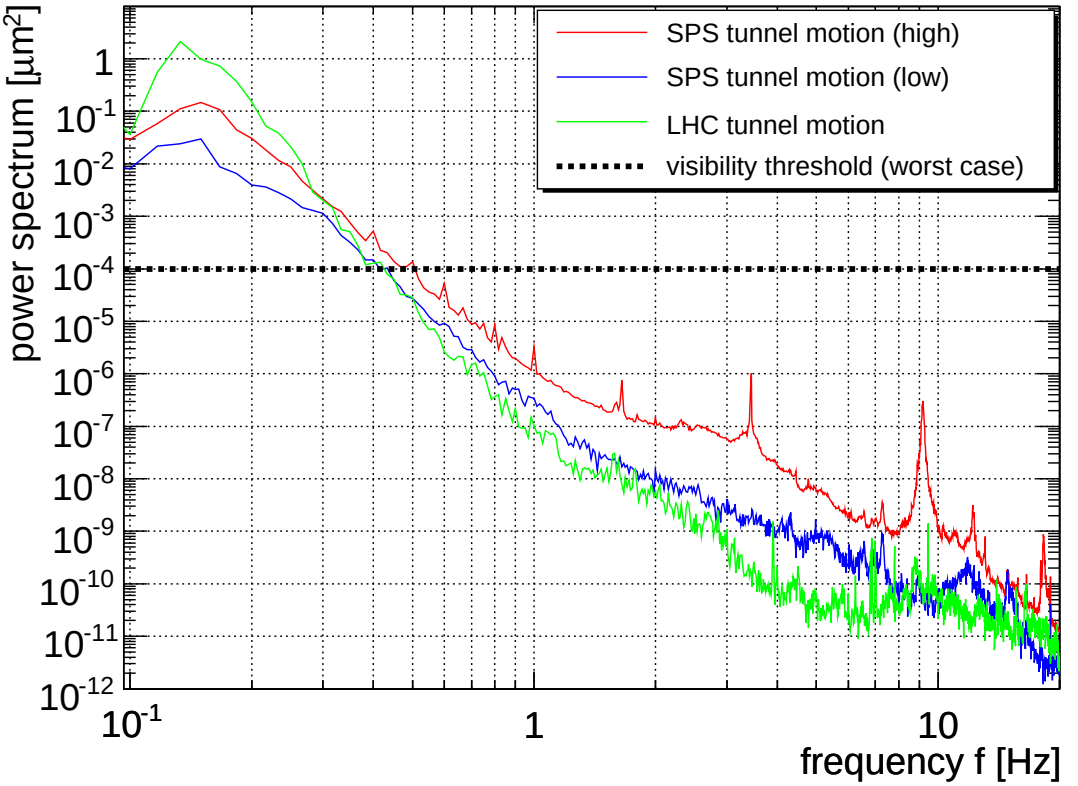


Figure 4.9: Averaged ground motion power spectra in the SPS and LHC tunnel. The 'high' SPS spectrum was recorded during a period of ongoing installation work. The data was taken using a sensitive geophone. The visibility threshold corresponds to the ground-motion level having a 1 μm effect detectable by the LHC beam position monitor, assuming a worst-case constant propagation factor $\kappa = 100$.

less because the power spectra decreases rapidly above this frequency. From Figure 4.9, and assuming $\kappa \approx 100$, it is clear that ground motion above 1 Hz should not pose a problem at the LHC, whether the movement is coherent or not, assuming that the girder response does not show significant resonances.

Typical power spectra (densities) are shown in Figure 4.13, and the corresponding integrated r.m.s. are shown in Figure 4.14. The auto-correlation spectra as a function of acquisition time are shown in Figure 4.11. A repetitive pattern with a period of about 7 s is visible. Cultural noise, such as installation work in the tunnel, manifests itself mainly in the 1 to 10 Hz frequency band, as visible in Figure 4.12. The integrated r.m.s oscillation above 1 Hz is modulated between about 200 and 800 nm, see Figure 4.10.

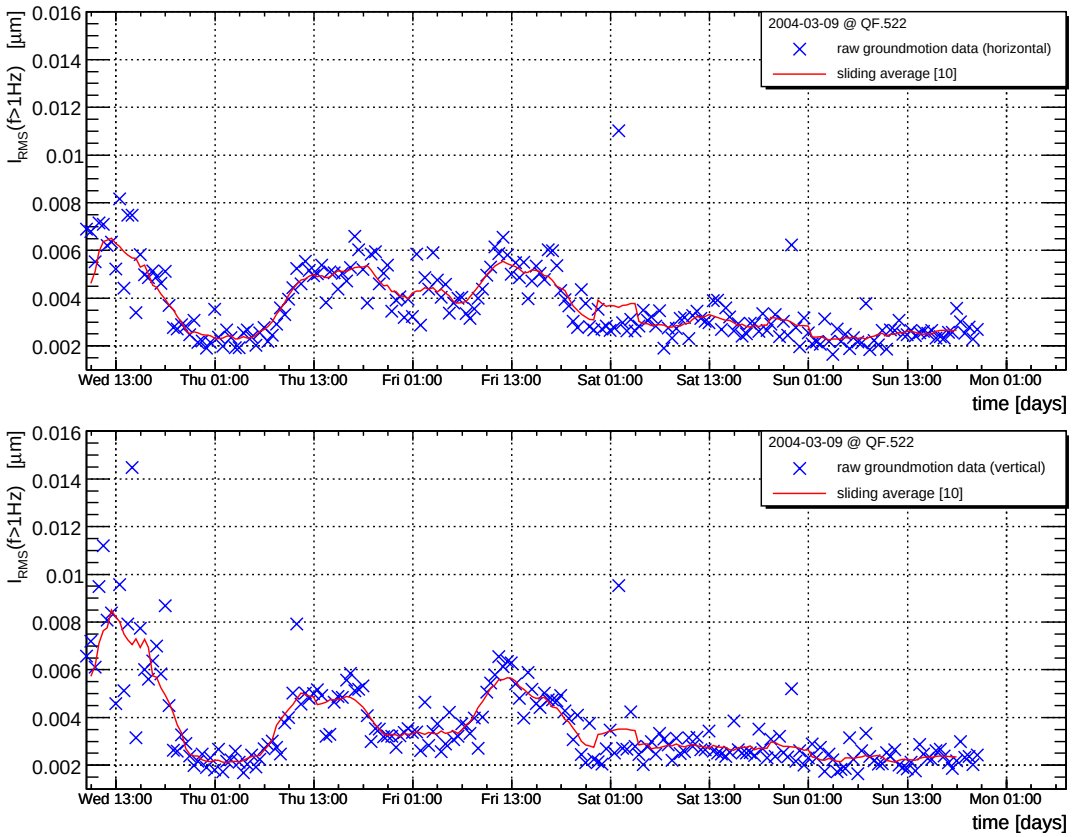


Figure 4.10: Horizontal (top) and vertical (bottom) integrated r.m.s. for frequencies above 1 Hz. The day-night cycle is visible during the week is due to installation work in the tunnel. The amplitude of the cultural noise varies between about 200 and 800 nm.

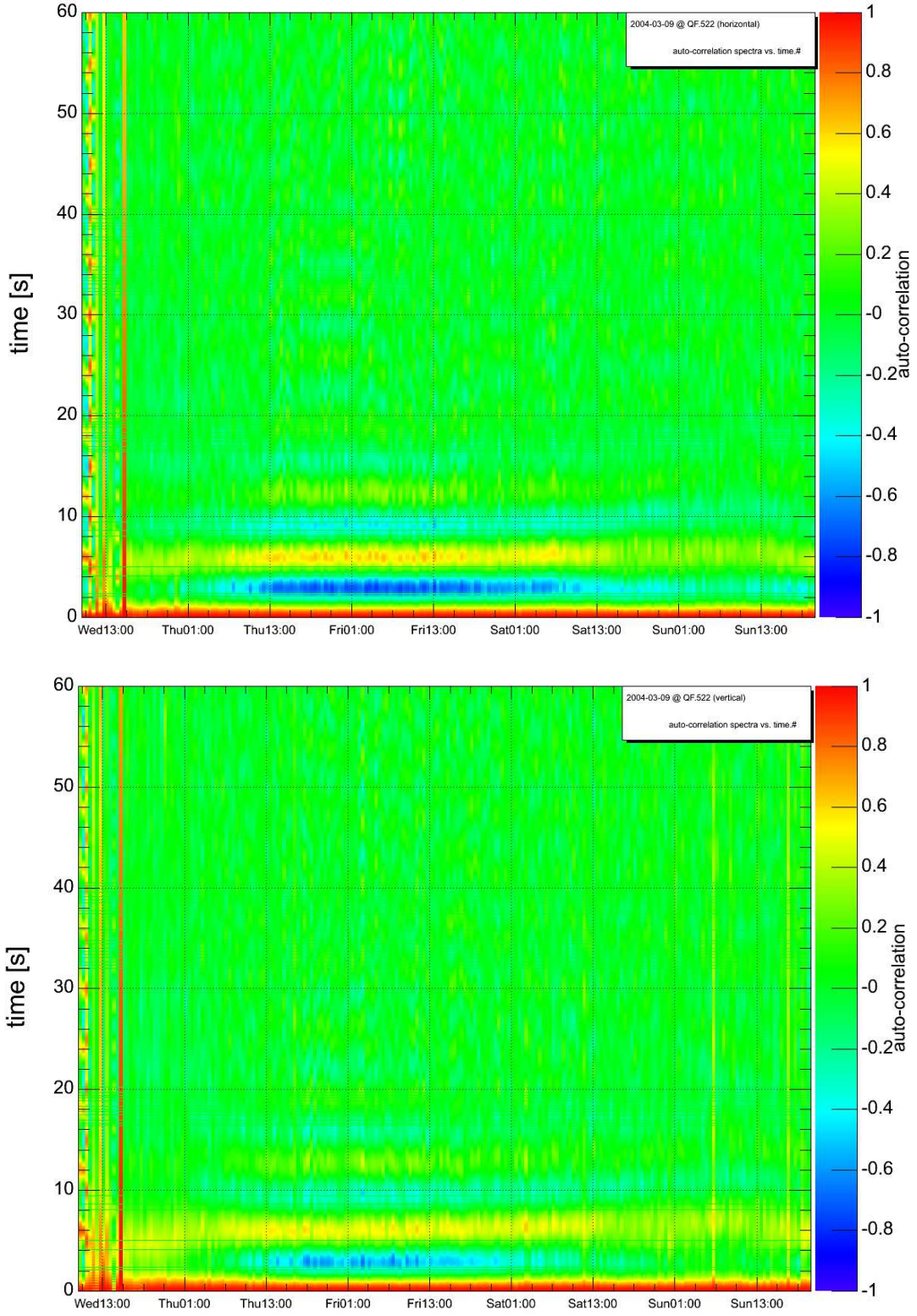


Figure 4.11: Horizontal (left) and vertical (right) auto-correlation spectrum. The coherent signal with a period of about 7 s corresponds to the ocean hum. The decrease of amplitude for long time periods is an artefact of the autocorrelation function due to the finite acquisition window.

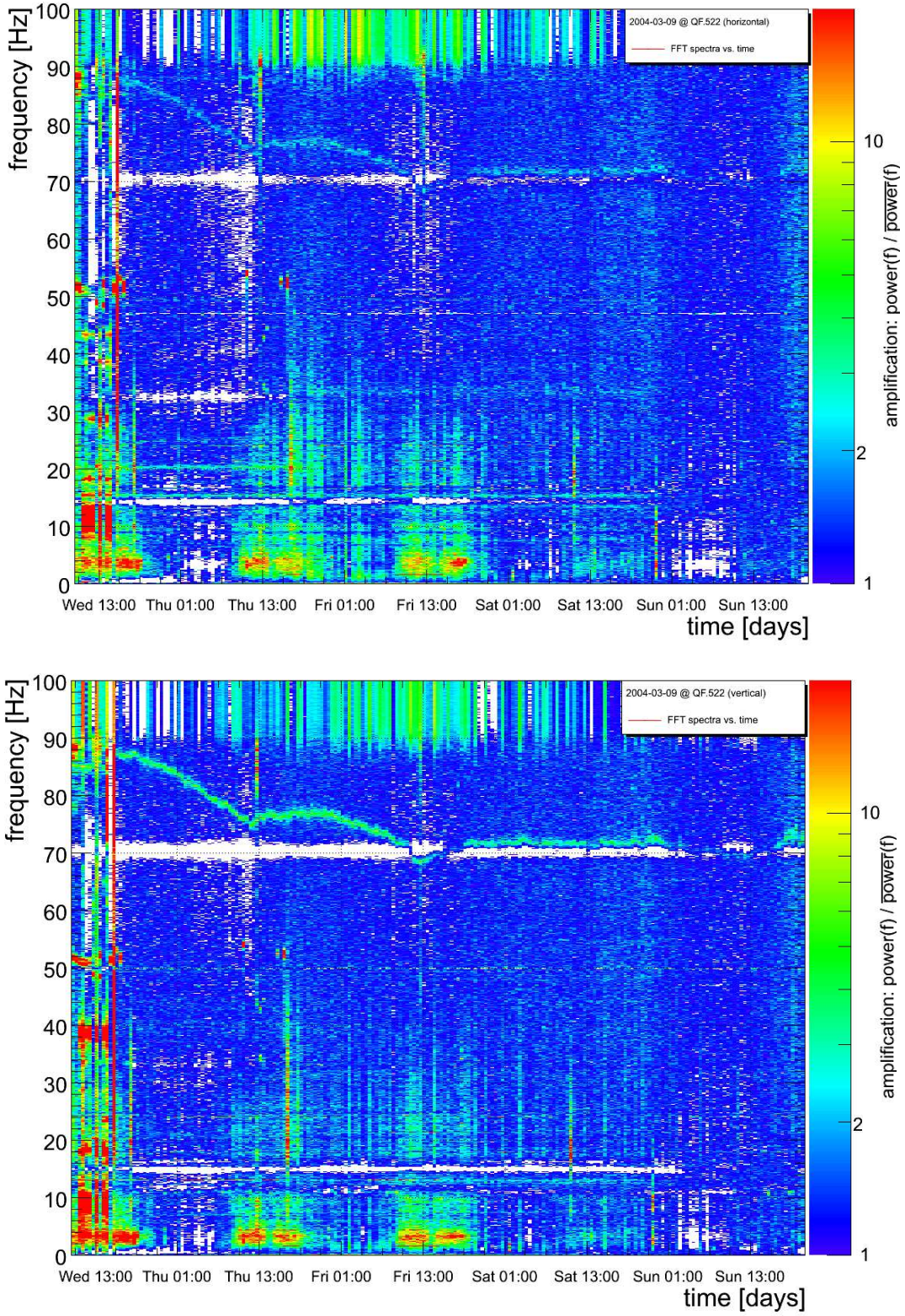


Figure 4.12: Horizontal (top) and vertical (bottom) ground motion spectrum as a function of acquisition time. The spectra are normalised to the averaged low activity (weekend) spectra. Some cultural noise is visible in a frequency band between 1 and 10 Hz. The white spots corresponds to frequency contributions that are less than those from the spectra used for normalisation. The day-night cycles during the work week and the reduced noise level during the weekend are visible. The regular gaps during days with activity in the tunnel correspond to the lunch break.

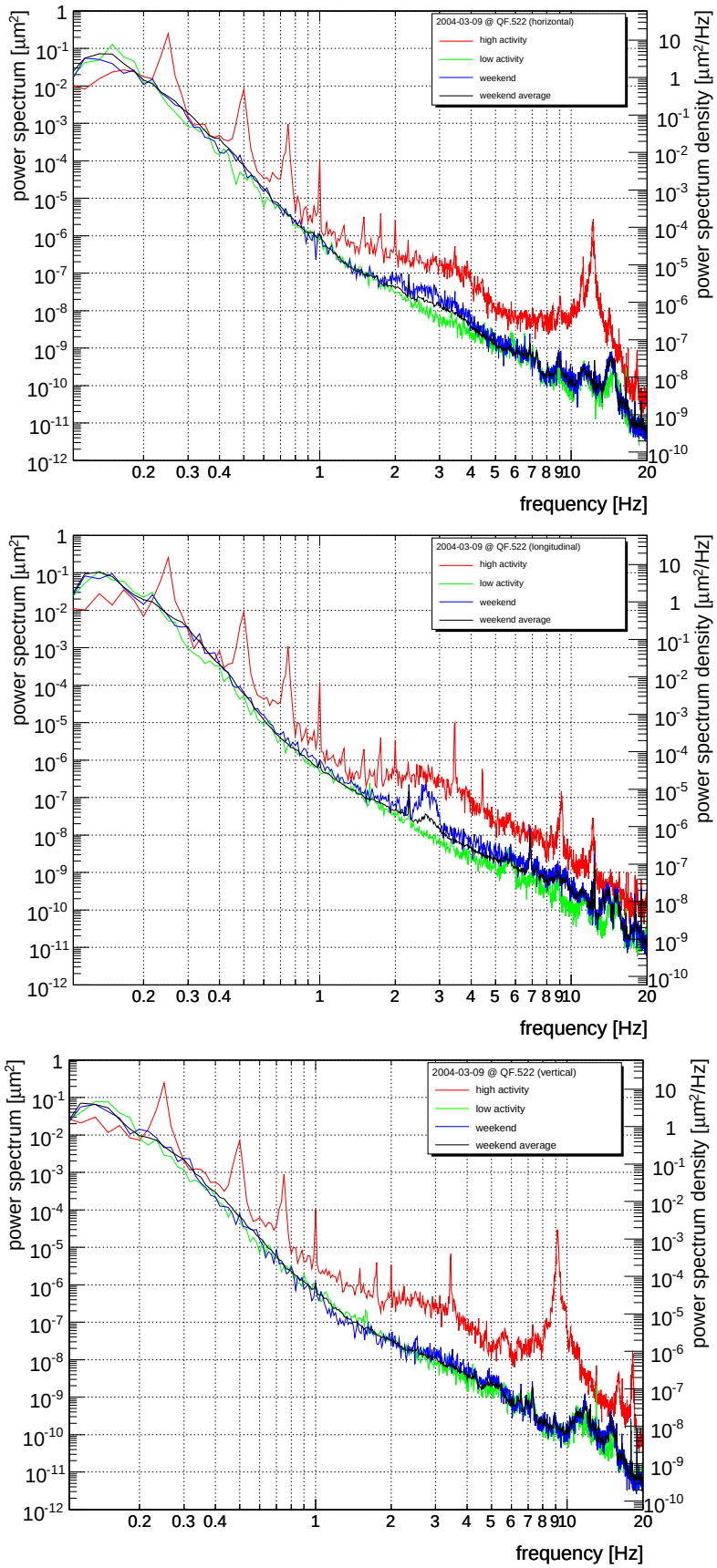


Figure 4.13: Horizontal (top), longitudinal (middle) and vertical (bottom) ground motion spectrum (densities) measured in the SPS tunnel. The 'high activity' spectra correspond to measurements performed during periods of equipment installation in the vicinity of the geophone ($\approx 100\,\mathrm{m}$).

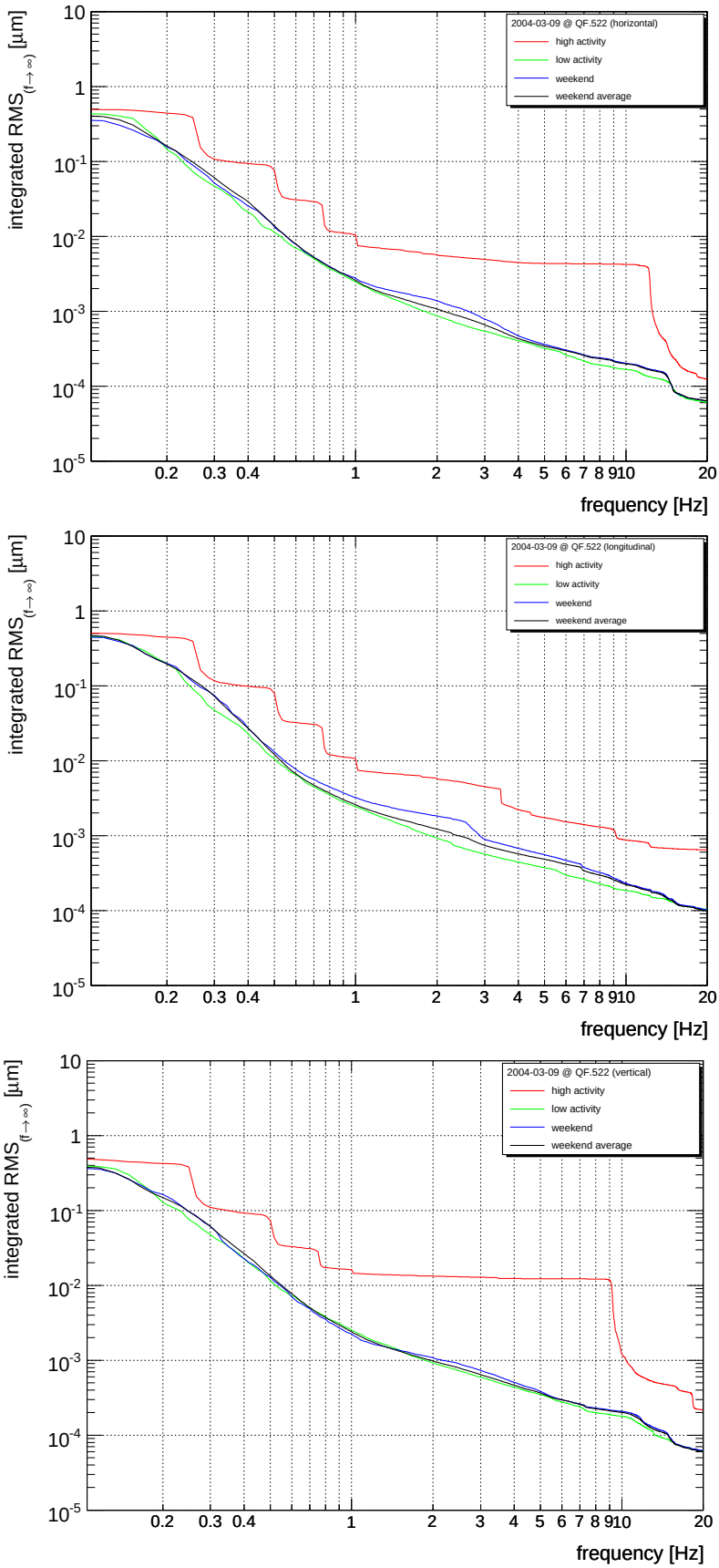


Figure 4.14: Horizontal (top), longitudinal (middle) and vertical (bottom) integrated ground motion spectrum. In normal conditions, the integrated r.m.s. is less than about 2 nm for frequencies above 1 Hz.

SPS Orbit Measurements

In 2004, long-term orbit stability measurements were performed with a 270 GeV coasting[5] beam in the SPS. Figure 4.15 shows an example of the vertical beam motion power spectra of a 270 GeV and 26 GeV coasting beam in the SPS that was sampled at a monitor with LHC readout electronics ($\beta \approx 100$ m). The BPM electronics is based on a bunch-by-bunch wide-band-time-normaliser principle as described in [68] and is designed to be insensitive to a wide range of temperature and bunch intensity changes. The residual bunch intensity dependence is less than 1% with respect to the BPM half aperture (80 mm). The residual white noise floor of about 2 μm r.m.s for frequencies above 0.1 Hz is visible, indicating the BPM electronics noise.

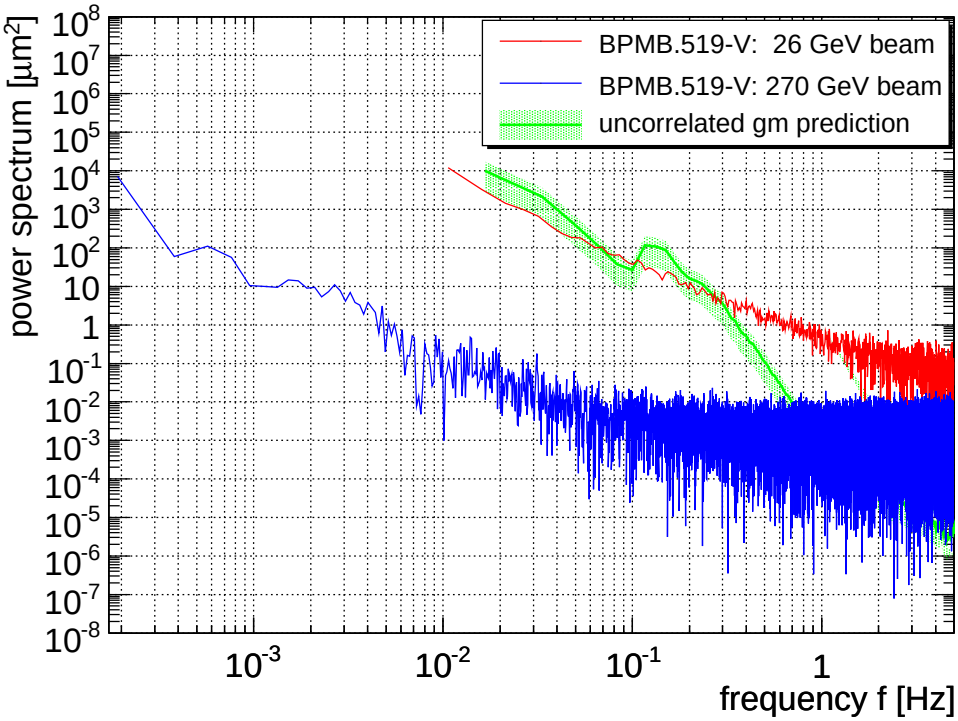


Figure 4.15: Power spectra of orbit movement at 26 GeV and 270 GeV in the SPS. The white-noise floor of the BPM for high frequencies is visible. The 26 GeV coast might be dominated by slow drifts of the magnetic fields rather than by ground motion. The predicted power spectrum for a worst-case (fully uncorrelated) propagation of the tunnel motion on the beam is shown (green band). In comparison with the actual 270 GeV coasting beam, it is clear that the peak below 1 Hz due to the ocean hum is, to a large extent, correlated.

The orbit movements of the 270 GeV beam are much smaller than that of the 26 GeV beam, which indicates that the earlier measurements in 2003 may have been dominated by machine-inherent effects such as drifts of magnetic fields rather than by ground motion. A prediction ($\kappa = 40.6 \pm 19.3$) for the spectrum due to uncorrelated tunnel motion is also shown for comparison.

This measurement confirms that in the range of 0.01-0.7 Hz, the tunnel ground motion is highly coherent. The measured quadrupole girder response is about one and does not show damping for this frequency range, which would explain the missing signal. This coherent wave is also visible with a period of about 7 s in the autocorrelation spectra, shown in Figure 4.11. The repetitive pattern further indicates the oscillatory nature of the signal and excludes that the peak is due to, for example, slow ramp-like drift or Brownian (random) motion. This measurement is in agreement with seismological measurements performed elsewhere [63, 64, 69]. The measurements described in reference [63, 64] identify and locate the cause of the hum around 0.1 Hz to be due to storms on the northern oceans during the northern hemisphere winter and southern oceans during the southern hemisphere winter. From the SPS diameter of 2 km, one can estimate the coherence length of this type and frequency of ground movement to be at least 2 km.

Using the SPS κ amplification factors, the orbit drift can be converted to r.m.s. ground motion drift. In case of the SPS, the white noise floor of the BPMs limits the analysis of orbit drifts to frequencies

[5] Beam that is stored but not accelerated is referred to as *coasting* beam.

below 0.1 Hz or drift times above 10 s, respectively. Random ground motion is a statistical process that, in first order, is described by Brownian motion (see Equation 4.14).

As described above by Equation 4.14, the *r.m.s.* orbit drift, a property of Brownian motion, is proportional to the square root of time. This dependence has been modelled before as an empirical observation by the so called *ATL law* that describes ground motion variance growth σ^2 as a function of a site-specific drift parameter A_{site}, the time T that is elapsed since the initial condition ($\sigma^2 = 0$), and coherence length L of this type of motion as an additional parameter of this motion [70, 71]:

$$\sigma^2 = A_{\text{site}} \cdot TL \tag{4.21}$$

However, the direct measurement and fit of the coherence length L is inaccessible using orbits of circular machines due to the filter mechanism of the lattice described above. Using the vertical κ factors for the SPS and fitting a square root dependence on time to the observed orbit drifts (Equation 4.14), the following SPS ground motion estimate can be obtained:

$$\sigma = A \cdot \frac{1}{\kappa}\sqrt{t} \tag{4.22}$$

$$A_{\text{vert}}^{\text{SPS}} = 1.2 \cdot 10^{-2} \frac{\mu\text{m}}{\sqrt{\text{s}}} \tag{4.23}$$

LEP Orbit Measurements

Over the years, hundreds of thousands of orbits were recorded while LEP was colliding beams for its experiments. The orbit data were analysed to reconstruct the orbit drifts that were compensated by the LEP slow orbit feedback. The data were analysed in the following way, on a fill-by-fill basis:

- The first orbit recorded with stable colliding beams was used as reference.
- The orbit drift for subsequent orbits was reconstructed by subtracting the orbit from the reference and by de-convoluting any correction of the orbit that took place up to that time, with respect to the start of the fill.
- For the horizontal plane, the momentum offset with respect to the nominal momentum was estimated using the dispersion at each monitor. The effect of the estimated momentum offset was subtracted from the orbit. This correction removes the effect of earth tides and of RF frequency changes that were applied to optimise the luminosity.
- For the vertical plane, the difference orbit was corrected for any contribution due to the eight vertical low-beta quadrupoles ("QS0") using a MICADO correction. This correction is applied to remove the systematic contribution of the moving quadrupoles that dominated the drifts due to their strength and their abnormal movements. See [72].
- The BPM readings were finally normalised by $\sqrt{\beta}$ before the raw r.m.s drift was evaluated.

This procedure was reproduced for all fills and the data of all fills finally averaged. The data of different years are consistent. In this analysis, the 1999 data taken at ≈ 100 GeV are used. Approximately $3 \cdot 10^4$ orbits were used for that year, on an average. Figure 4.16 shows the averaged orbit r.m.s. normalised to the monitor beta function of 100 m. Figure 4.17 shows the development in time of the relative spread. In the horizontal and vertical plane, the relative spread seems to reach a steady state of about 30% and 40% respectively, after about 2000 seconds, which is about the same spread as expected from the predicted κ value for LEP (Table 4.1.2). The settling may be due to analysis systematics (correction of QS0 effects and unfolding) and a large number of vertical corrections that were always applied during the first hour of each fill to optimise the luminosity. For the horizontal plane, the drifts are described well by a $\sqrt{t}$ dependence on time, but for the vertical plane the data follow $\sqrt{t}$ only after ≈ 2000 s. For further analysis, 2000 s is chosen as the minimum time for the analysis of the LEP data.

We assume that the residual beam movement is rather due to the remaining residual ground motion. Using a constant optical amplification factor κ for the SPS and LHC, respectively, one can give an approximation for the average ground motion-induced quadrupole shift as shown in Figure 4.18.

Since the influence of other (unknown) effects cannot be fully excluded, the following estimates should be considered to be an upper limit for ground motion. From LEP $\tilde{\kappa}$ amplification factors, the orbit drifts can be converted to r.m.s. ground motion drifts, as seen in Figure 4.18. A fit of these data to $A \cdot \sqrt{t}$

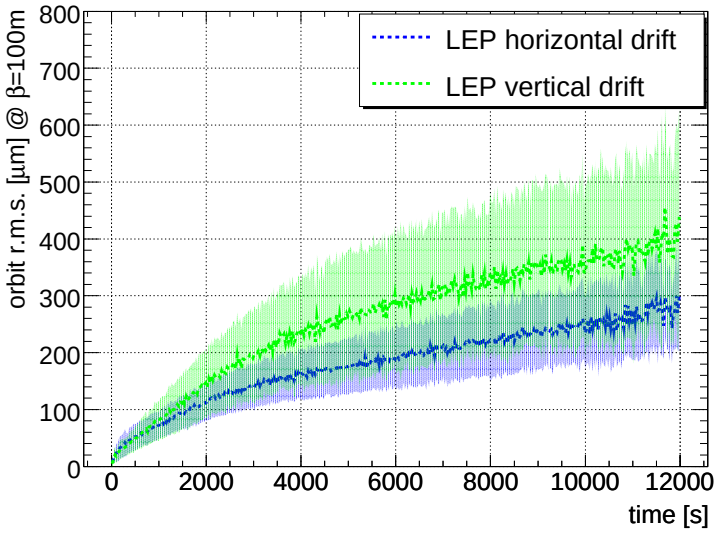


Figure 4.16: Average vertical and horizontal LEP beam motion drift during operation at 100 GeV. The $\sim \sqrt{t}$ growth with time t is visible. Data courtesy J. Wenninger.

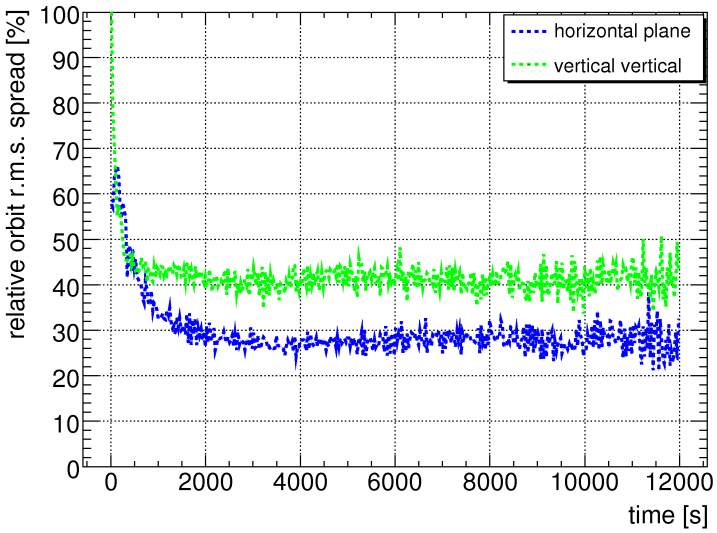


Figure 4.17: Relative LEP beam motion spread

dependence, an analogue to Equation 4.22 yields the following upper limits for the horizontal and vertical ground motion coefficients:

$$A_{\mathrm{hor}}^{\mathrm{LEP}} \approx 5.3 \cdot 10^{-2} \frac{\mu\mathrm{m}}{\sqrt{\mathrm{s}}} \tag{4.24}$$

$$A_{\mathrm{vert}}^{\mathrm{LEP}} \approx 6.0 \cdot 10^{-2} \frac{\mu\mathrm{m}}{\sqrt{\mathrm{s}}} \tag{4.25}$$

The horizontal and vertical parameter spread of about 30 and 40% respectively and is in good agreement with the prediction for the κ spread shown in Table 2. This may indicate that the remaining drift is due to random ground motion only. We use the LEP values for further analysis, since the LEP data are based on much higher statistics than the SPS (based only on a few cycles).

Long-term drifts at LEP

It is important to note that the above estimate is only valid for periods up to a month and does not include long-term systematic drift effects such as fault movements and ground water levels. Long-term LEP accelerator alignment studies described in [61] show that the main part of quadrupole movement over several years seems to be linear. The analysis estimates the drift parameter P to be around 5.5 $10^{-6}\mu$m/s:

$$\langle\sigma\rangle = 5.5 \cdot 10^{-6} \frac{\mu\mathrm{m}}{\mathrm{s}} \cdot t \tag{4.26}$$

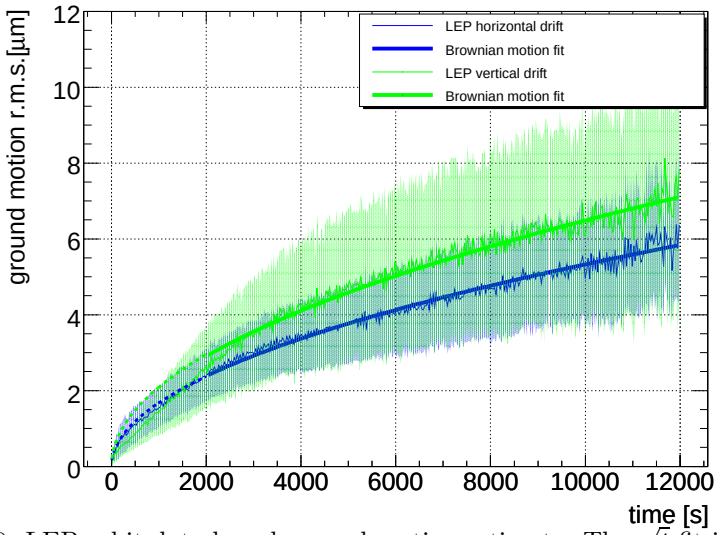


Figure 4.18: LEP orbit data based ground motion estimate. The $\sqrt{t}$ fit is indicated.

In the case of systematic linear drifts of the quadrupoles but with individually different drift velocities per quadrupole, the residual orbit would depend only on the initial seed of individual quadrupole drift velocities and would scale linearly with time. On the time scale of a few hours, the systematic contribution is much smaller than the one due to random ground motion. However, due to the different dependence on time, the systematic component largely exceeds the random motion contribution for time scales above a month to years.

In conclusion, in the time scale of a few hours to a month, random ground motion determines the orbit stability, but on larger time scales, it is exceeded by the systematic ground motion contribution. For completeness, the beam motion predictions include the random and systematic ground motion drift approximations.

Estimate for ground motion induced orbit drift at LHC, SPS and in the transfer lines

As an important application, the above ground motion estimates can be used to predict future beam motion at the LHC, the SPS and the transfer lines TI8 and CNGS. Figure 4.19 and 4.20 show the predicted LHC orbit movement using the LEP and SPS ground motion estimates and the κ factors for the LHC.
The linear prediction is given for comparison. It is important to note that since the actual beam motion depends on the initial seed of the random quadrupole motion, the spread of the prediction is relatively large ($\approx 30 - 40\%$). Further, the crossing between the random and linear ground motion prediction is visible. This crossing can be used to estimate the time frame where the random ground motion approximation is applicable.

Alignment at the collimator is an important issue during LHC operation. The limit at which the orbit exceeds the required $\approx 0.3\sigma$ (σ: beam size r.m.s.) orbit stability at the collimator jaws is indicated in the plot. A critical beam drift due to random ground motion for the LHC injection optics is only reached after about 10 hours. This relaxes the requirement on fast re-steering during injection from the ground motion point of view. However, during injection, other effects such as the decay of the persistent currents will dominate orbit drifts.

The stability of the injection is another important parameter for machine operation and protection. Since the injection plateau is only corrected at every injection, one can derive from the estimates that from fill-to-fill the injection orbit will drift in the order of 0.3-0.4 σ due to random ground motion (assuming a fill every about 15 hours).

Similar estimates can be given for the SPS using the corresponding κ factors. The results are shown in Figure 4.21. After one month, the orbit drift reaches about 1 mm which is of the same magnitude but slightly overestimates the drifts observed during SPS operation (about 500-700 μm per month). Predictions using vertical drift constants obtained with the SPS coasting beam (Equation 4.23) seem to agree better with operational experiences.

Figure 4.22 shows the expected drifts in the CNGS and TI8 transfer lines. The computations are done for the location of the CNGS target and the last TI8 beam dump (TED.87765) of the transfer line. In the case of CNGS, the maximum acceptable drift of 500 μm is determined by the target size. Comparing

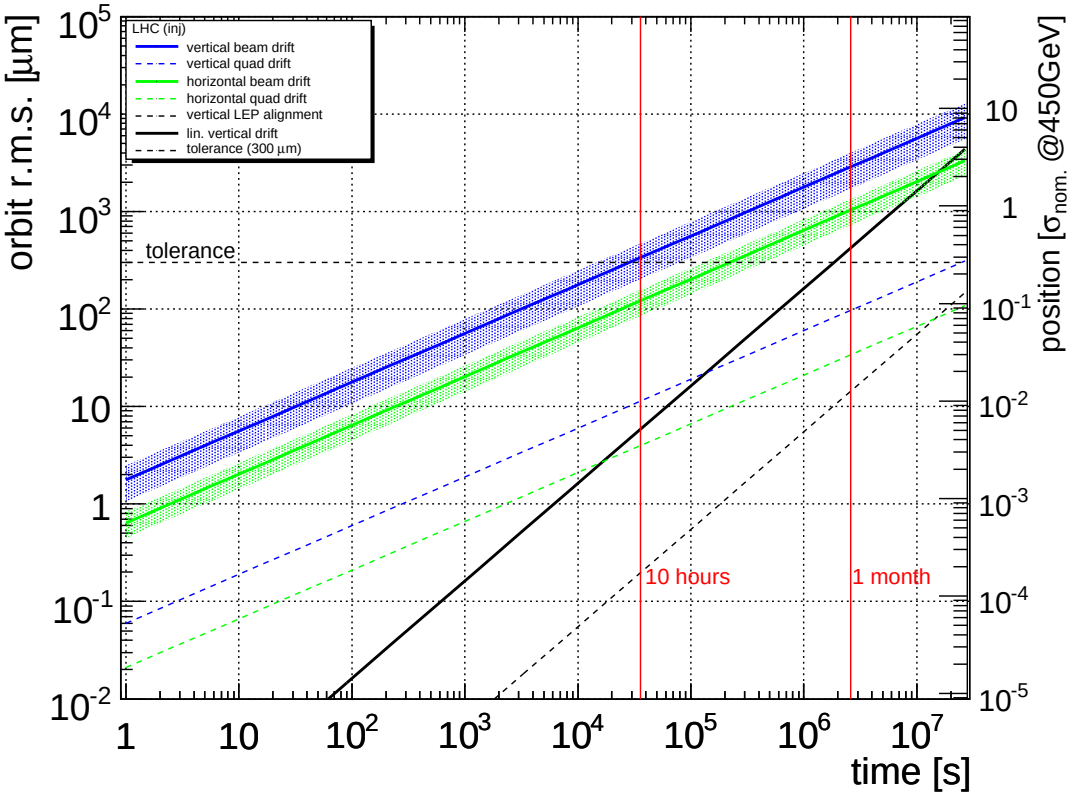


Figure 4.19: Expected LHC r.m.s. orbit drift due to random ground motion for the injection optics. The drift is given in units of nominal beam size σ (3.75 μm rad emittance and for 450 GeV) on the right scale. The spread corresponds to the spread of parameter κ.

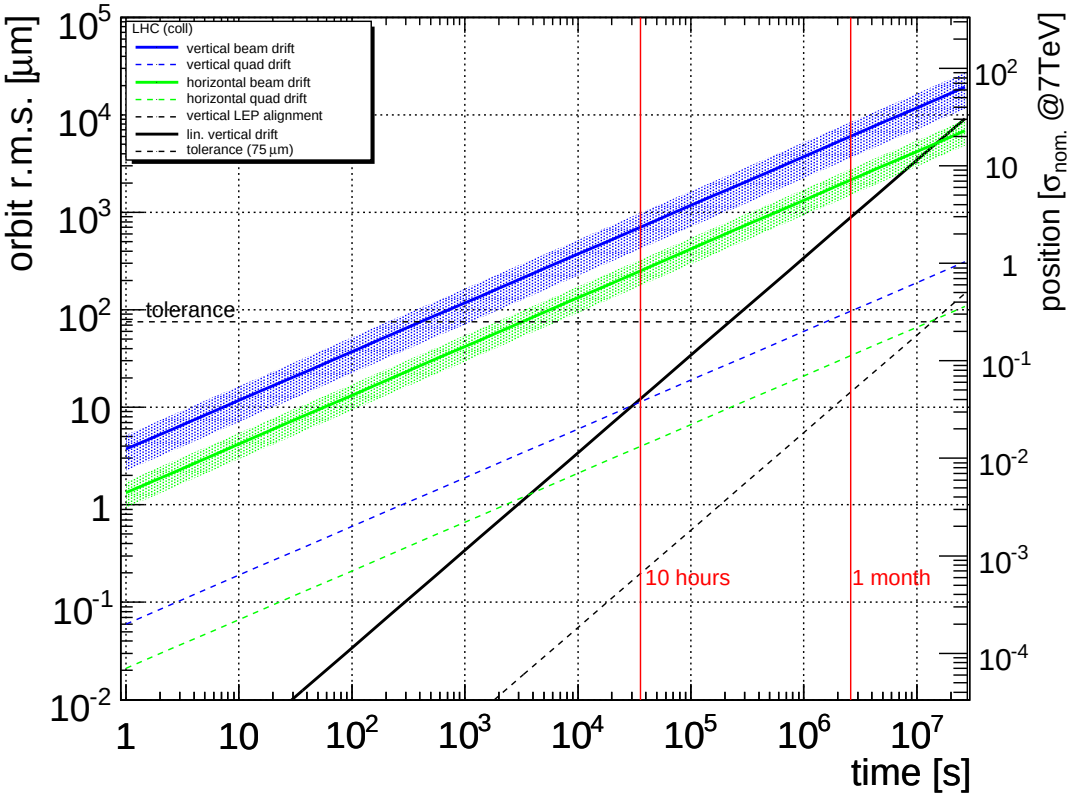


Figure 4.20: Expected LHC r.m.s. orbit drift due to random ground motion for the collision optics. The drift is given in units of nominal beam size σ (3.75 μm rad emittance and for 7 TeV) on the right scale. The spread corresponds to the spread of parameter κ.

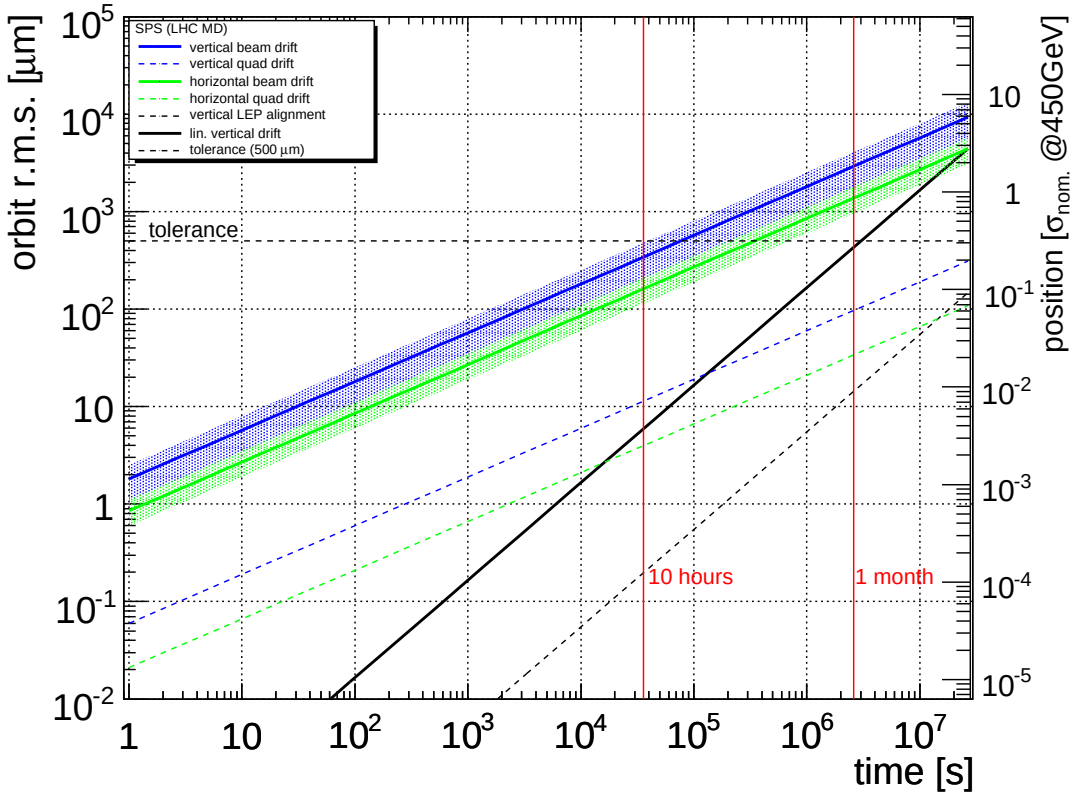


Figure 4.21: Expected SPS r.m.s. orbit drift due to random ground motion. The corresponding drift in units of nominal beam size σ (3.75 μm rad emittance and for 450 GeV) is given on the right scale. The range corresponds to the parameter κ spread of the model. The tolerance corresponds to the maximum permissible orbit shift at the extraction point to the LHC.

the expected drift at the target over one month with the required stability, it is clear that the drift at the target is dominated rather by the stability of the SPS orbit, the extraction kicker and the septum magnet than by the magnet misalignment due to random ground motion. Similar numbers can be obtained for the TI8 transfer line that was commissioned with beam in 2004. For TI8 the tolerance of about 200 μm corresponds to the largest allowed injection oscillation to ensure good emittance preservation. Measurements described in [73] give an upper limit for the transfer line stability of about $\sigma/8$ and $\sigma/14$, and for a period of about 6 hours and 1 hour, respectively. These are compatible with the ground motion prediction. Measurements described in [73] assume that the residual beam jitter is due to the power converter ripple of the transfer line magnets, in particular the extraction septum magnet. The plotted tolerance corresponds to the required upper limit of $\sigma/5$.

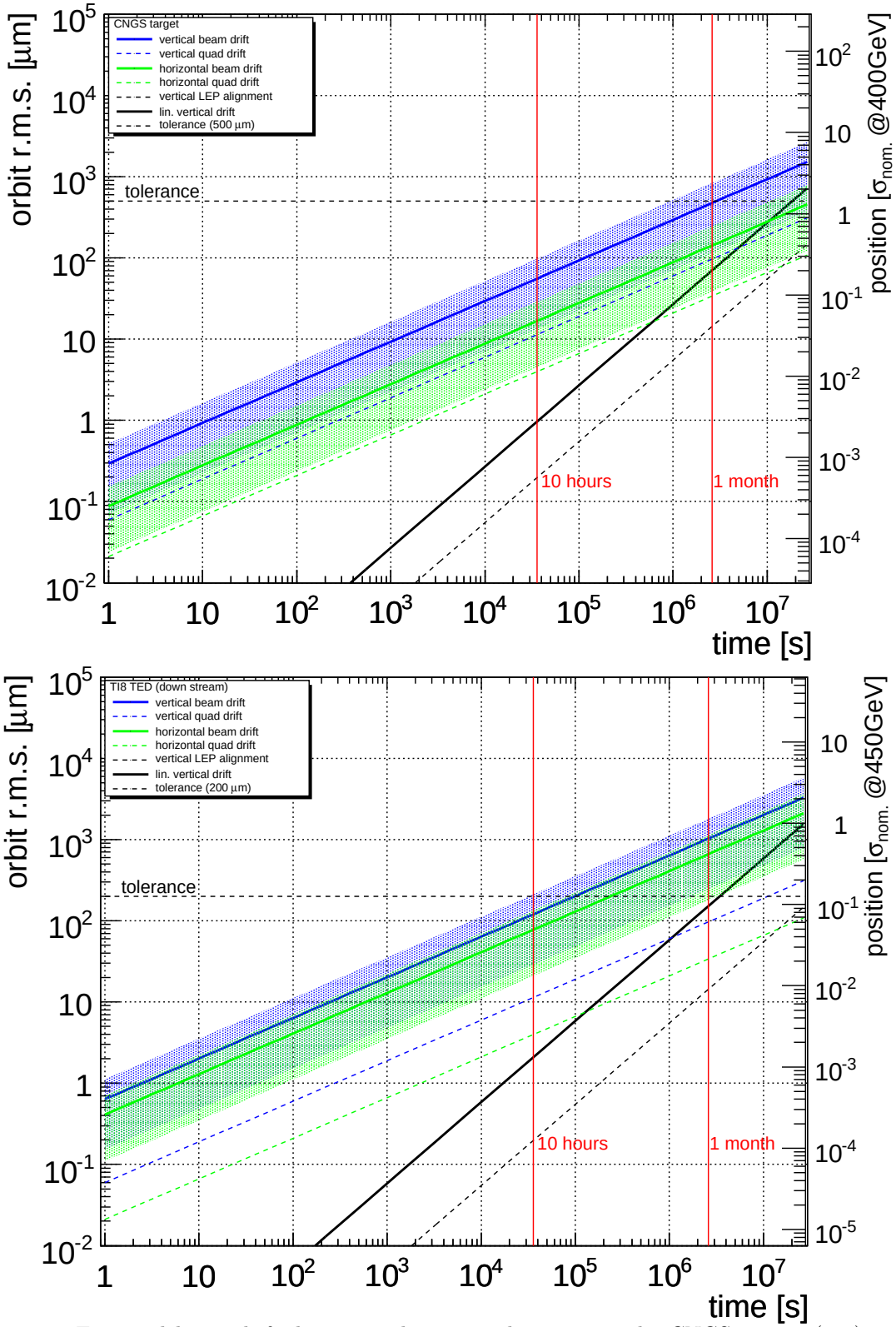


Figure 4.22: Expected beam drift due to random ground motion at the CNGS Target (top) and TI8 transfer line TED (bottom). The effect of the orbit drift in the SPS at the extraction kicker is not included.

4.1.4 Conclusions

Ground motion affects the beam orbit during all operational phases. The two models that were presented describe the propagation of coherent and random ground motion on the beam orbit both qualitatively and quantitatively. The ground motion analysis is based on geophone and beam motion measurements in the SPS and LEP. The measured spectra are very quiet and are essentially the same for the SPS and LHC tunnels. The ground motion spectra are barely influenced by so-called 'cultural noise' that originate at the surface.

These measurements confirm that ground motion contribution can be neglected for the LHC for frequencies above 1 Hz, independent of whether the ground motion is correlated or not, assuming that the girder is sufficiently stiff and that its response does not significantly enhance the coherent part of the ground motion spectrum. The actual vibration level of the LHC superconducting magnet, not discussed in this paper, remains to be investigated.

The comparison of beam motion and geophone measurements reveal that the large peak in the ground motion spectrum between 0.01 and 1 Hz is mostly correlated and is of an oscillatory nature. Since the amplification for correlated ground motion on the beam vanishes in that frequency range for SPS, LEP, and LHC, the peak does not measurably influence the beam. In fact, for frequencies below 1 Hz, orbit drifts are dominated by random ground motion.

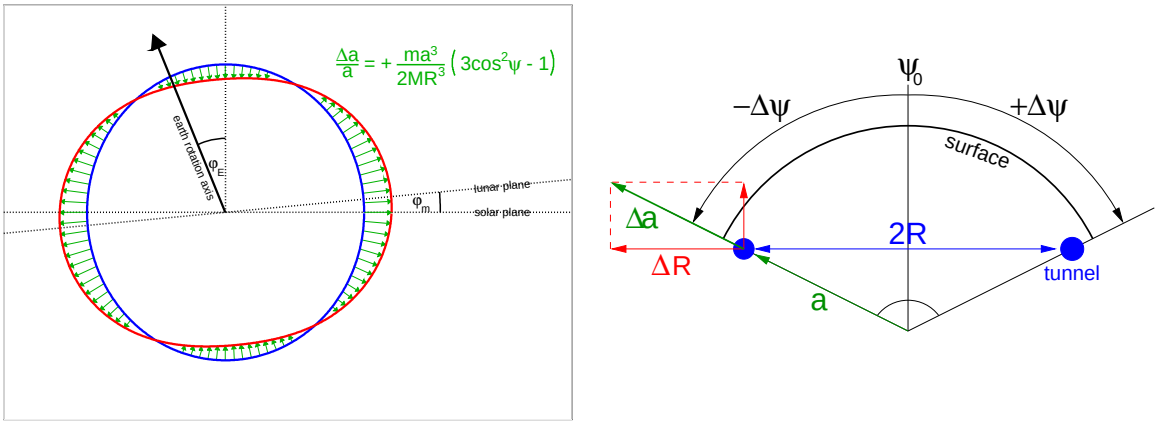
The drift constants and optical propagation factors (lattice amplification) can be used to predict upper limits for orbit drifts in the LHC, the SPS and the CNGS and TI8 transfer lines. The LHC orbit drift due to random ground motion alone is expected to reach the critical ($0.3\,\sigma$ tolerance for the collimation system only after about 10 hours for the injection optics and after about 0.5 hours for the collision optics under the pessimistic assumption that the orbit feedback is not running. The LHC injection orbit is estimated to drift about 0.3-0.4 σ from fill-to-fill, assuming a fill about every 15 hours. This may slightly reduce the available aperture during the first injection into an empty machine. The drift estimates for the CNGS and TI8 transfer lines are in good agreement with experimental data obtained during the TI8 commissioning and show compared to the SPS extraction stability that ground motion-induced orbit drifts in the transfer lines are much smaller than drifts due to the SPS extraction stability.

It is important to note that these predictions only give an upper limit of ground motion-induced orbit drifts. The actual orbit drift may be dominated by the machine environment and operation. The ground motion-induced orbit drift may, over the long-term, be dominated by systematic drifts due to specifics of the LHC tunnel described in [61].

4.2 Lunar and Solar Tides

It is well known ([74, 75]) that the gravitational forces of the moon and sun affect the operation of large accelerators. Both cause terrestial tides that induce a quadrupolar deformation of the earth crust which eventually changes the physical circumference of the machine. The lunar periodicity of the perturbation driven by the moon is 24 hours, 48 minutes, and one lunar month lasts about 27.32 days (27 days, 7 hours, 43 minutes, 11.6 seconds). The effect of the solar tidal force is about 45% of that of the moon and has (by definition) a periodicity of 24 hours. The LHC centre is located at a latitude of about 46.23^o north and a longitude of about 6.05^o east.

The gravitational force is the main mechanism that creates two bulges, one towards the moon/sun and one pointing away, causing a local lift Δa of the ground and change of the gravity Δg as illustrated in Figure 4.23(a).



(a) tidal deformation of earth surface

(b) horizontal strain changing tunnel radius

Figure 4.23: Tidal deformation of Earth's surface (left) and tunnel radius (right). The inclination of Earth's rotational axis is $\varphi_E = 23.5^o$ with respect to the solar axis (vertical). The lunar plane is tilted by $\varphi_E = 5.1^o$ with respect to the solar plane. The illustrations are not to scale. The order of magnitude of the deformation is usually very small and around $\frac{\Delta a}{a} \approx 10^{-8}$.

As shown in [76], the relative change of gravity $\Delta g/g$ and relative elevation of the earth $\Delta a/a$ can be analytically derived to

$$\frac{\Delta g}{g} = -\frac{m_{S/M} a^3}{M R_{S/M}^3} \left(3\cos^2(\psi) - 1\right) \tag{4.27}$$

$$\frac{\Delta a}{a} = +\frac{m_{S/M} a^3}{2M R_{S/M}^3} \left(3\cos^2(\psi) - 1\right) \tag{4.28}$$

with $\frac{m_{S/M}}{M}$ the mass ratio and $R_{S/M}^3$ the distance between the given celestial body and the earth. The indices indicate parameters for sun ('S') and moon ('M'), respectively. The radius to the point of interest, in our case the LHC tunnel, is denoted by a, which is typically about 6400 km. The angle ψ is defined by the angle between the point of interest on the Earth's surface and the centre of the moon and sun, respectively. The angle ψ_0 depends on the relative position of earth, moon and sun and can be precisely predicted to within a few decades.

The total tidal effect is a superimposition of the lunar and solar bulges. The effect is strongest if earth, moon, and sun are aligned. One refers to *spring tides* if the sun and moon are on the same side of the earth causing a maximum tide, or *neap tides* if the sun and moon are opposite with respect to the earth resulting in a minimum tide.

Though the qualitative physics is well described by equations 4.27 and 4.28, the actual prediction is limited by the precision of the parameters entering these equations. The moon has an approximate elliptical orbit with a perigee of about 363300 km, a mean radius of about 384400 km and an apogee of about 405500 km. The lunar, solar and earth's rotational plane have different inclinations with respect to each other, making the prediction of the moon-sun constellation more complex. As a result, the earth describes a nutation around its rotational axis.

The relative change of the accelerator circumference $\Delta C/C$ is proportional to the relative lift $\Delta a/a$ in the centre of the accelerator

$$\frac{\Delta C}{C} = \lambda_T \frac{\Delta g}{g} = -2\lambda_T \frac{\Delta a}{a} \tag{4.29}$$

with λ_T being a constant. As illustrated in Figure 4.23(b), $\Delta\psi$ is the dihedral angle between the outermost tunnel centres and the centre of the earth. The externally driven radius change ΔR is given by the local lift Δa in the centre of the accelerator. The local lift Δa depends on the latitude, distance of the accelerator a from the centre of the earth and the angle $\Delta\psi$. The external force is partially compensated by the factor ϵ which is related to the elastic property of the Molasse and the concrete of the tunnel walls. Hence, the actual effect on the radius change is reduced by a factor of $(1-\epsilon)$. The resulting radius change ΔR is thus given by:

$$\begin{aligned} \Delta R &\approx (1-\epsilon)\cdot \Delta a(\psi_0 \pm \Delta\psi)\cdot \sin(\Delta\psi) \\ &= (1-\epsilon)\cdot \frac{R}{2a}\cdot \Delta a(\psi_0 \pm \Delta\psi) \end{aligned} \tag{4.30}$$

In first order, the relative change in circumference is equivalent to the relative change of the accelerator's average physical radius.

$$\frac{\Delta C}{C} \approx \frac{\Delta R}{R} = +\frac{(1-\epsilon)}{2}\cdot \frac{\Delta a(\psi_0 \pm \Delta\psi)}{R} \tag{4.31}$$

Comparison of equations 4.29 and 4.31 yields the following approximate for the constant λ_T:

$$\lambda_T \approx -\frac{(1-\epsilon)}{4}\cdot \frac{3\cos^2(\psi_0 \pm \Delta\psi) - 1}{3\cos^2(\psi_0) - 1} \tag{4.32}$$

It is visible that the factor λ_T in this simplified model has to be between -0.25 and 0. The parameter ϵ depends on the elasticity of the Molasse and other specifics and is difficult to estimate analytically. The parameter λ_T was determined from LEP energy calibration data described in [75]. Since the LHC will be housed in the same tunnel, one can assume that the experimental value for λ_T is similar and about:

$$\lambda_T = -0.15 \pm 0.01 \tag{4.33}$$

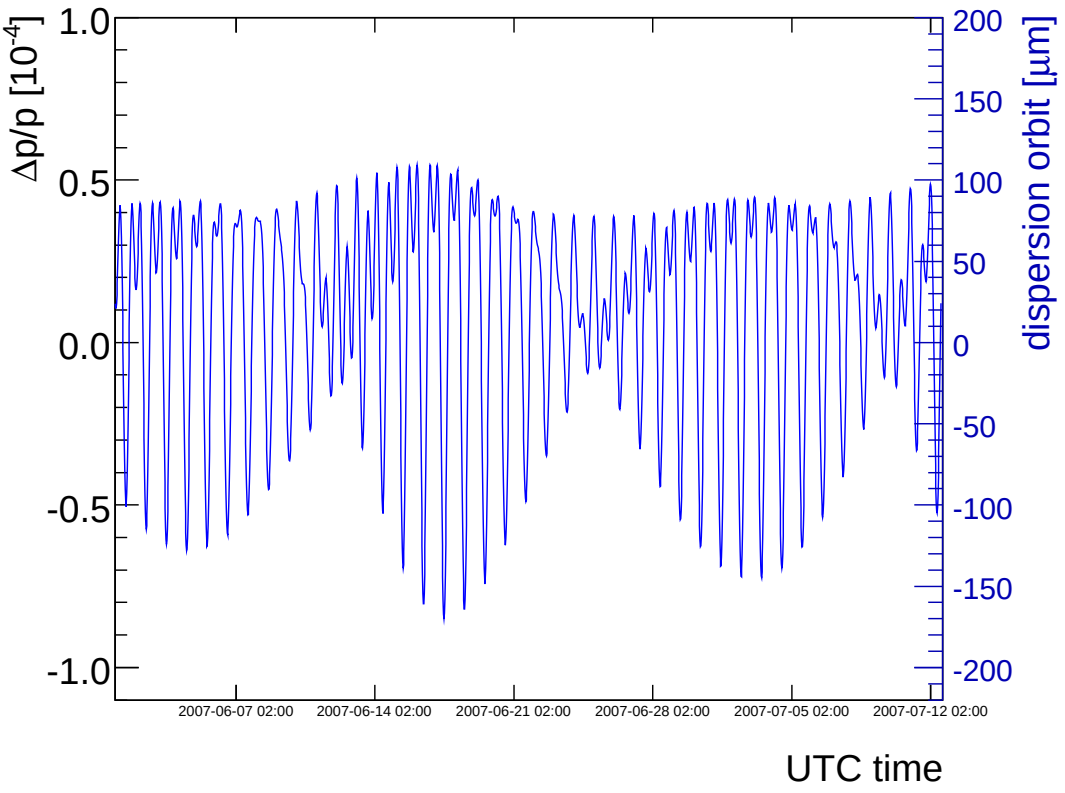
The RF frequency is commonly kept constant during collisions. Thus, the primary effect of the circumference change is a shift in energy of the particles which causes the particle to move on a dispersion orbit as described by equation 2.23. While the vertical dispersion $D_y(s)$ is negligible in the LHC, the horizontal dispersion function $D_x(s)$ ranges in the arcs between about 1 and 2 m as shown in Figure 2.4. Thus the effect is primarily dominant in the horizontal plane at locations with large values of the dispersion function. Figures 4.24(a) and 4.24(b) show the estimated LHC momentum change and corresponding dispersion orbit due to the lunar and solar terrestial tides. The changes are shown for the time scale of a month and a year, respectively. It is visible that the peak-to-peak $\Delta p/p$ oscillation amplitude is about $1.5\cdot 10^{-4}$ which corresponds to an dispersion orbit of about 150 μm. Due to the lunar and solar periodicity of about 24 hours with two bulges per period, the maximum change occurs within about 6 hours. Figures 4.24(a) and 4.24(b) implicitly assume a dispersion function of 2 m. However, the largest dispersion can reach values of 2.3 m as, for example, inside the cleaning insertion IR3. The energy change and resulting orbit drifts are very slow and can be easily corrected.

However, due to the different nature of the driving mechanism, this effect has to be differently compensated than the betatron oscillation created by the periodic and random ground motion described in the previous section. The control algorithms for orbit and energy are further discussed in detail in Chapter 6.

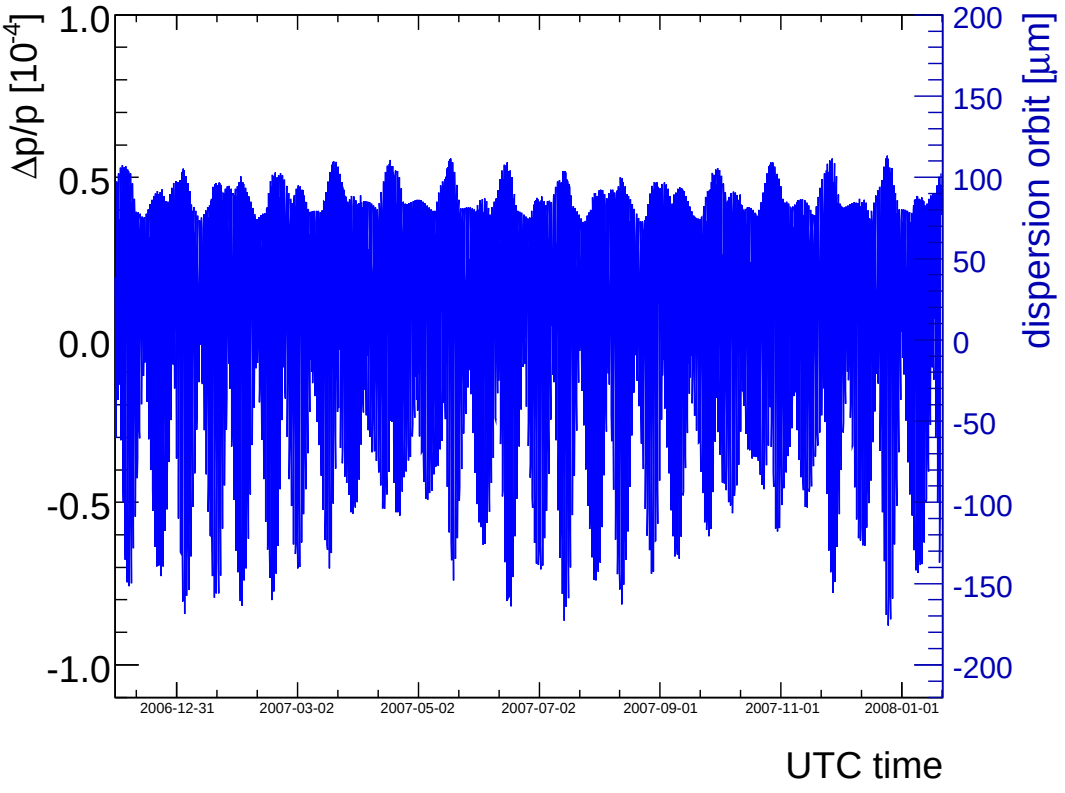
4.3 Thermal Expansion of Magnet Girders and Tunnel Walls

The change of tunnel air and wall temperature as another environmental perturbation source may lead to thermal expansion of magnet girders and the accelerator tunnel itself, affecting medium to long-term stability of the beam orbit and momentum. While the expansion of girders cause orbit perturbations essentially through feed-down dipole momenta of the shifted quadrupoles, the thermal expansion of the tunnel itself may change the circumference of the machine and thus may affect the beam momentum.

The thermal effects play an important role in machine operation and are widely recognised in so-called 'light sources', which are electron synchrotron accelerators with the main purpose of creating high intensity synchrotron light used for spectroscopy, physics, chemistry and other applications. The



(a) tide prediction for June 2007



(b) tide prediction for 2007

Figure 4.24: Tidal oscillation on the scale of a month and year. The deviation of the LHC momentum and dispersion orbit is shown. The basic period is about 12 hours. The peak-to-peak oscillation amplitude is about $1.5 \cdot 10^{-4}$ units in terms of $\Delta p/p$ and is equivalent to an orbit oscillation amplitude of about 150 μm at a location with local dispersion of $D_x \approx 2$ m. Tidal prediction algorithm courtesy of J. Wenninger.

third generation of these light sources even applies, for example, a so-called 'top-up' operation: frequent injection are used to maintain a constant beam current and thus heat load and temperature inside the machine in order to achieve sub-micron stability [77].

4.3.1 Girders

Thermal movements of the quadrupoles, in particular those used for the final-focus, were already observed to have an influence on the LEP beam orbits as described in [72]. There, the orbit perturbation effect was particularly enhanced by the design of the final focus quadrupoles. As for the ground motion induced quadrupole shifts, the effect of thermal quadrupole misalignment onto the orbits is enhanced by the orbit lattice response function estimated above for ground motion.

The same effect may have an impact on the beam orbits in the LHC. The effect is, as for LEP, most pronounced for the final focus magnets with an integrated quadrupoles strength kl_{mag} between about $0.048-0.055\,\mathrm{m}^{-1}$ and a local betatron function of about $800-4000\,\mathrm{m}$. For comparison, the arc quadrupole magnets have an integrated strength of $|kl_{mag}| \approx 0.027 - 0.028\,\mathrm{m}^{-1}$. The maximum betatron function of the FODO lattice is about 180 m. Using equation 4.3, for a constant girder movement an individual triplet quadrupole has a 4 to 10 time larger feed-down effect on the orbit than typical arc quadrupole.

By design, the focus phase advance between the quadrupoles surrounding the IP is approximately π. As a consequence of this phase advance, the effect of a vertically shifted quadrupole on the left of the IP onto the global orbit is very similar to the effect of a vertical shifted quadrupole on the right of the IP and can only be resolved using beam position monitors close to these magnets. Due to low BPM sampling around the LEP IPs, the precise beam alignment relied on the tracking of the mechanical alignment of the magnets as described in [72].

At the LHC, the left-right ambiguity of insertion quadrupole movement can in principle be resolved with two warm monitors between the Q1 triplet quadrupoles ('BPMSW.1L5x' and 'BPMSW.1R5x' with 'x' being the IR number). Using these monitors, it is in principle possible to minimise the systematic feed-down dipole kick of the Q1 quadrupole and to preserve the beam crossing angle inside the IP.

Anticipating the use of feedbacks, the automated LHC orbit feedback control will continuously compensate for these slow thermal-induced quadrupole feed-down drifts. As discussed later, this compensation is limited by beam position measurement noise and the chosen locality of the correction. Consequently the correction will maintain a constant beam position with respect to the quadrupole centre. A-priori, this type of compensation seems to be favourable from the point of view of keeping aperture constraints: the machine aperture is in first order defined by the vacuum chambers. Since these are connected to the quadrupoles to which coordinate system the orbit is fixated, the systematic shifts due to thermal expansion are transparent with respect to keeping the vacuum aperture constraints.

However, since the beam position monitors are rigidly mounted to the quadrupoles, the automated stabilisation will driven by the thermal expansion effectively move the beam orbit position with respect to an external reference point that is not subjected to thermal drifts. This systematic reference shift is in good approximation identical to the thermal expansion Δx of the girders that is described by

$$\Delta x = x_0 \alpha \Delta T \tag{4.34}$$

with x_0 the initial height of the girder, ΔT the temperature change and α the thermal expansion coefficient, which for steel, varies between $\alpha = 10 - 17.3 \cdot 10^{-6} K^{-1}$ (BS: 970 and DIN 18800 standard). The absolute expected shift clearly depends on the temperature variation over the tunnel length and its stability. The design of the LHC ventilation system is based on the assumption that the air is blown into the tunnel at the even IPs at a temperature of $18 \pm 1\,^{\mathrm{O}}\mathrm{C}$ and released at the odd IPs at an expected design temperature of $23 \pm 6\,^{\mathrm{O}}\mathrm{C}$ as described in [78]. Since the LHC is still under construction, it is rather difficult to accurately estimate the actual spatial and temporal temperature variation inside the tunnel. On the short term, one may expect temperature variation due to the rather static load heating of the electronic racks and dynamic loads due to power converters, warm magnets and their interconnections that, depending on operation and length of the physics run, may slowly heat up the tunnel on the timescale of a few hours. Though the largest part of the heat is led away through water cooling systems a given amount will be dissipated into the air and heat up neighbouring elements and among other devices, girders.

Table 4.3 summarises the expected shifts of the local quadrupole based reference system per degree of temperature change. It is visible that the reference system shifts are relatively small. Even under the assumption of a maximum relative temperature swing of $\Delta T \pm 5\,^{\mathrm{o}}\mathrm{C}$, the relative misalignment error with respect to an external fixed reference point is less than approximately $160\,\mu$m, which is small and takes place on a timescale from hours, days, to a month. In any case, the temperature gradient will likely be less

	Girder Height x_0 [mm]	**Relative Shift** $\Delta x/\Delta T$ $[\mu m/^o C]$
Cold Quadrupoles and Main Bends	340 ± 20	$3.4 \pm 0.2 \ldots 5.8 \pm 0.3$
Warm Quadrupoles in IR3/IR7	≈ 950	$9.5 \ldots 16$
Collimators	≈ 950	$9.5 \ldots 16$

Table 4.3: Shift of beam position and quadrupole reference system with respect to an external static fix-point due to thermal expansion of magnet girders. The cryostats of the main quadrupoles are supported by the same girders as the main dipoles [79].

than $\pm 1\,^{\mathrm{O}}$C over a few hundred meters and consequently the relative alignment between adjacent magnets should not pose a problem for most beam alignment requirements that require orbit stabilities in the order of 0.1 to 1 mm. For comparison, modern light sources that require an orbit stability of 1 μm stabilise the tunnel temperature, cooling water and electronics to $\pm 0.1\,\mathrm{C}^{\mathrm{o}}$ ([77, 80]). Further, these accelerator use the Invar (FeNi36) alloy for their BPM girder that has a minimal thermal expansion coefficient (see [81]) and hydrostatic levelling systems in order to quantify the thermal expansion of girders ([82]).

A possible exception may be given for the LHC Cleaning System and TOTEM experiment that request a beam stability in the order of a few ten micrometers or less, as discussed in Section 6. The betatron collimation system in IR7, may be affected by temperature induced drifts due to the possibly large temperature gradient in that region. Due to cleaning and inevitably larger particle losses, the air in that region may potentially become radioactive. Thus sections 4 and 5 left of IR7 will be hermetically sealed for nominal operation and the regular tunnel air flow is bypassed through that section. A schematic overview of this section is shown in Figure 4.25. Due to this separation, the temperature inside this

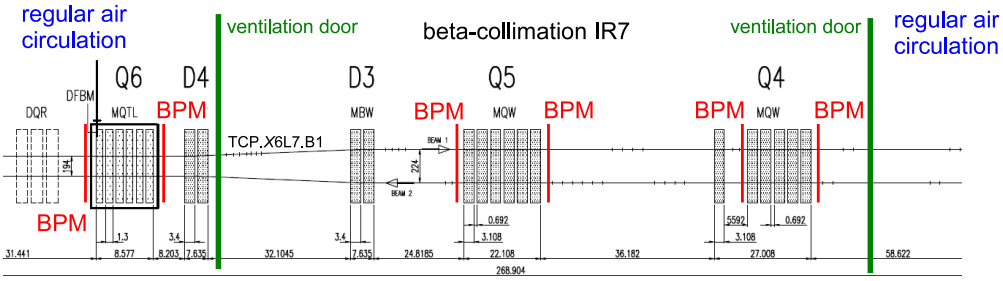


Figure 4.25: Schematic overview of the beta-cleaning insertion in IR7. The quadrupoles in this region are equipped with two BPMs each. The two ventilation doors separate, for nominal operation, the primary collimator 'TCP.X6L7.B1', and the Q4 and Q5 quadrupoles from the regular air circulation, which may lead to increased air temperature in this section.

section is expected to be higher than in the neighbouring tunnel segments. Present forecasts assume that the thermal heating power of about 50 kW may heat up air temperature inside the section, up to $30 - 35\,^{\mathrm{O}}$C compared to a temperature of about $23 - 29\,^{\mathrm{O}}$C in the neighbouring cells [83]. The actual spatial and temporal temperature perturbations could be evaluated during early LHC commissioning.

4.3.2 Tunnel and Machine Circumference

Thermal tunnel temperature imbalances may also affect tunnel and machine circumference. These elongations, as for the tides, lead to a change of beam momentum as observed in [80, 84, 85].

The concrete in the tunnel wall usually has a thermal expansion coefficient of $\alpha_{concr.} \approx 10 \cdot 10^{-6} K^{-1}$ (EC2 standard) which is similar to steel. Inserting the LHC tunnel circumference ($C = 266658.8832$ m) in equation 4.34

$$\Delta C = C \cdot 10^{-5} \cdot \Delta T \tag{4.35}$$

would lead to a theoretic circumference change of up to about 0.26 m per $^o C$. Obviously, these large expansions would inevitably lead to cracks and deformations of the tunnel walls. Thus these potential elongations are mitigated through expansion joints every 39.5 m between adjacent tunnel floor segments. The length of these segments were initially matched to the average distance between quadrupoles in LEP [86].

Main Dipole Magnets								
	static components		*dynamic components*					
	geometric		persistent		eddy		decay	
	μ	σ	μ	σ	μ	σ	μ	σ
b_1		10.0	-8.63	0.49	-1.5/-5/0	0/1.1/1.0	0.8	0.74
a_1		10.0	0.0	0.40	0	0/0.13/0	x	x
b_2	1.2	0.63	-0.003	0.31	0.01	0.06	-0.03	0.11
a_2	-0.37	1.21	0.0	0.77	0.05	0.19	0.00	0.24
b_3	-3.78	1.3	-11.0	0.29	-0.04	0.18	1.73	0.44
a_3	-0.01	0.33	0.0	0.20	0.00	0.02	0.03	0.07

Main Quadrupole Magnets								
	static components		*dynamic components*					
	geometric		persistent		eddy		decay	
	μ	σ	μ	σ	μ	σ	μ	σ
b_2	0.0	10.0	-5.6	0.5	-16.82	2.7		

Table 4.4: Excerpt of error Table '0605' based on [87–89]. The relative multipole field errors are given in units of 10^{-4} with respect to the main field component and a reference radius of 17 mm. The field errors are decomposed into systematic μ and random σ components.

As a first order approximation, assuming a steady surface temperature T_0, the underground temperature T in an idealised homogenous medium increases with depth. In a dynamic case, the perturbation of the surface temperature by $\pm\Delta T$ in time will, similar to a driven oscillator, lead to a variation of the underground temperature with the same periodicity. The phase lag and amplitude of this oscillation depends on the heat flux, thermal conductivity of the Molasse and the depth inside the ground. As shown in [85], the lag increases and the temperature modulation decreases with depth. In contrast to most other accelerator tunnels, the LEP/LHC tunnel is embedded deep underground with an average depth of about 100 m. Hence it is unlikely that the tunnel temperature variations are driven by surface temperature oscillations propagated through the Molasse.

However, similar to the discussion of the thermal girder expansion, the stability of the air ventilation system may have a long-term influence on the tunnel wall and thus tunnel segment temperatures, and may potentially lead to circumference changes visible as a modulation of the beam momentum.

In any case, these perturbations are expected to be small and may be negligible in comparison with tides and other effects, both with respect to their magnitude and their time scale. Accurate estimates of the magnitude of this effect are difficult to obtain due to the complex interaction and heat exchange processes and is like to be accessibly only through measurements with the beam.

4.4 Machine-inherent Sources

An important contribution to beam perturbations at injection energy is inherent to accelerator operation and the use of superconducting magnets: the magnetic field imperfections and resulting influences on the beam are most pronounced for the main dipole and quadrupole magnets when they are set to their injection fields. The imperfections of these magnets can be split into a static and dynamic component:

The static components are caused by imperfect coil and collar geometry, beam screen imperfections, stress and deformation during the magnet transport, cooling down of the magnet and other sources.

The dynamic effects are essentially driven by the decay and snapback of the multipole components described in Section 1.3.2, eddy currents inside the magnets and vacuum chamber while ramping the magnets, and decay of the persistent currents during the ramp and ripple of the power converters.

4.4.1 Multipole Field Errors

Table 4.4 summarises the expected main dipole and quadrupole multipole field errors that are derived from measurements and simulations. The errors are given in relative units of 10^{-4} with respect to the main field component of the magnet. As defined in equation 1.6: b_n denotes the n^{th} regular and a_n the n^{th} skew component of the magnet field error. Three different scenarios are given for eddy current contribution of the b_1 and a_1 components. The a_1 component is suppressed by design of the dipole magnets. However, roll misalignments of the dipole magnets may cause a feed-down of main dipole

component to the a_1 component. Since the propagation effect is similar for the horizontal and vertical plane, further analysis assumes the same magnitude of field errors for both planes.

The eddy currents inside the magnet and beam screen created by a varying main dipole field have opposite signs with respect to the initial field change, as expressed by the Maxwell equations, also known as LENZ' LAW. Hence the b_1 eddy current contributions are expected to partially shield the main field component. In order to comply with this general physical observation, in contrast to the original error table, the expected b_1 eddy current contributions have been inverted to reflect the above physical properties. For further simulation results, a systematic $\mu(b_1) = 5$ units and vanishing systematic $\mu(a_1) = 0$ eddy field error is assumed. The random eddy current components are due to magnet-to-magnet variations of the coil and beam screen geometry.

Similar to ground motion propagation discussed in Section 4.1, the effect of the main dipole and quadrupole imperfections onto the beam can be decoupled with respect to space and time. Since the absolute field errors usually decrease for successively higher multipole field components, it is sufficient to take into account only the low orders, dipoles and quadrupoles, for the perturbation treatment of orbit and energy.

For example: in contrast to the quadrupoles where an off-centre beam results in an effective dipole kick that has a significant impact on the orbit, the feed-down of the next higher order – the lattice sextupoles – is negligible. For illustration: an arc quadrupole ($k_1 l_{mag} \approx 0.028 \frac{1}{m}$) with an off-centre beam at half of the arc half-aperture $\Delta x = 11$ mm yields the following effective dipole kick

$$\delta_{MQ} = k_1 l_{mag} \cdot \Delta x \approx 308\,\mu\text{rad} \tag{4.36}$$

due to feed-down effect whereas a lattice sextupole ($k_2 l_{mag} \approx 0.062 \frac{1}{m^2}$) with an off-centre beam at the same amplitude creates only a dipole kick of

$$\delta_{MS} = k_2 l_{mag} \cdot \Delta x^2 \approx 7.5\,\mu\text{rad} \tag{4.37}$$

which is about 40 times smaller than the effect of a lattice quadrupole.

For small perturbations, it is possible to approximate the propagation by a linear scaling law that can be written in a general form as:

$$\Delta U = \lambda_\mu \mu \tag{4.38}$$
$$\Delta U = \lambda_\sigma \sigma \tag{4.39}$$

where U is the specific beam parameter (e.g. orbit 'x' or momentum shift $\frac{\Delta p}{p}$) and λ a linear propagation factor. Similar to the κ factor introduced for ground motion in equation 4.9, the factor λ depends on the specific field error component and on whether it is systematically (μ) or randomly (σ) distributed over the magnets. Since these effects are treated linearly, the analysis is largely independent from the assumed field errors and can easily be rescaled to other choices and assumptions on magnetic field errors.

Impact on beam parameters

The final effect of a b_1 field imperfection onto the beam depends strongly on whether this error is systematic ($\mu(b_1)$) or randomly distributed ($\sigma(b_1)$) over the main dipoles:

- For particle velocities close to the speed of light, a systematic $\mu(b_1)$ component corresponds to a shift of the reference bending field of all dipole magnets that defines the beam momentum. While the beam is captured by an RF system, the RF frequency fixes the circumference C and thus the bending radius ϱ of the main dipole magnets. In first order the equilibrium orbit stays constant. However, since the momentum p is defined by the bending radius ($p \sim \varrho B$), the systematic change of the main dipole field changes the beam momentum. Using the multipolar field expansion definition 1.6 it is visible that:
$$\frac{\Delta p}{p} = 10^{-4} \cdot \mu(b_1) \tag{4.40}$$

- A random $\sigma(b_1)$ component corresponds to random kicks distributed over the dipoles. The average effect on the global orbit can be estimated through a Monte-Carlo simulation similar to ground motion . The simulation is based on the LHC injection optics (V 6.5). For each seed, each main dipole is perturbed randomly by 1 unit r.m.s. kick and the cumulative effect of each dipole onto the orbit $x(s)$ is evaluated using the dipole response given in equation 2.71. The residual orbit r.m.s. is plotted for each seed in Figure 4.26. The propagation depends on the individual seed and

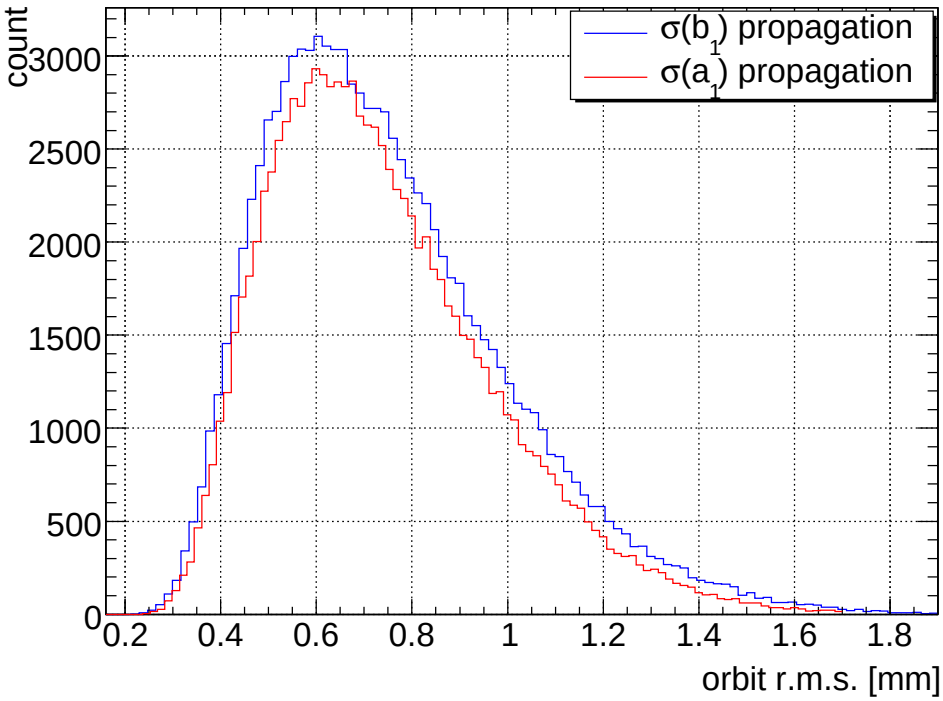


Figure 4.26: Simulated horizontal and vertical orbit r.m.s. distribution based on random b_1 and a_1 field errors of one unit r.m.s. per dipole. The average orbit r.m.s. is about 0.7 mm with a r.m.s. spread of about ±30%. The Monte Carlo simulations are based on 10^5 orbits (seeds).

distribution of dipole kicks across the main bending magnets. Similar to the case of the quadrupole response the orbit r.m.s. distribution is Non-Gaussian although the underlying error distributions are Gaussian. The shape depends on the beta-function at the dipole magnets and beam position monitors. From Figure 4.26 one can find:

$$\Delta x\ [mm] = (0.68 \pm 0.23)\ [mm/unit] \cdot \sigma(b_1)\ [units] \quad (4.41)$$
$$\Delta y\ [mm] = (0.67 \pm 0.21)\ [mm/unit] \cdot \sigma(a_1)\ [units] \quad (4.42)$$

with Δx and Δy being the orbit r.m.s. due to a random $\sigma(b_1)$ and $\sigma(a_1)$ perturbation of the main dipole. The prediction, similar to the quadrupole lattice response, has a relative spread of about 30 %.

Similar Monte-Carlo simulations were performed for tune, chromaticity and coupling. The results are summarised in the following equations:

$$\Delta x[mm] \approx (0.68 \pm 0.23)[mm/units] \cdot \sigma(b_1)[units] \quad (4.43)$$
$$\Delta y[mm] \approx (0.67 \pm 0.21)[mm/units] \cdot \sigma(a_1)[units] \quad (4.44)$$
$$\frac{\Delta p}{p} = 10^4 \cdot \mu(b_1)[units] \quad (4.45)$$

minimal or no dependence on: $\mu(a_1), \sigma(b_1), \sigma(a_1)$

$$\Delta Q_{x/y} \approx 8 \cdot 10^{-3} \cdot \mu(b_2) \quad (4.46)$$
$$\Delta Q'_x \approx +44 \cdot \mu(b_3) \quad (4.47)$$
$$\Delta Q'_y \approx -39 \cdot \mu(b_3) \quad (4.48)$$

Eddy Currents

As described in Section 2.3, the magnets are ramped synchronously with RF frequency to accelerate the protons from 450 GeV to 7 TeV in order to maintain a constant machine circumference. The time required for acceleration is usually determined by the maximum ramping speed of the main dipoles. As shown in equation 2.10, the induced voltage due to the changing dipole field is less than about 3 V and has a negligible effect on the beam momentum with respect to the total acceleration. However this voltage

is sufficiently large to drive eddy currents in the superconducting filaments and beam screens and may become large due to the small resistance and create higher order field momenta that may perturb the beam.

The eddy currents are most pronounced for the main dipole and quadrupole magnets due to their large magnetic flux change during the ramp. Table 4.4 summarises the cumulative effect of eddy currents in the superconducting cables, magnet yoke and beam screens. The multipole momenta $c_n^{eddy}|_{inj}$ are given for the maximum ramping rate of 10 A/s and are normalised to the injection energy as a reference. The resulting energy independent dipole kick scales with $\frac{1}{p} \sim \frac{1}{B} \sim \frac{1}{I}$. The eddy current field imperfections can thus be written as:

$$c_n^{eddy}(t) = c_n^{eddy}|_{inj} \cdot \frac{1}{I(t)} \frac{\partial I}{\partial t} \tag{4.49}$$

Direct transit from the injection plateau ($I = const$) to the nominal current ramp ($\Delta I/\Delta t = 10\,\mathrm{A/s}$) would yield a large transient due to the discontinuity of the derivation at the start of the ramp. In order to compensate for this effect, the baseline LHC main dipole ramp has been designed to minimise this transient as described in [90]. The main dipole ramp function is composed of a parabolic, exponential and linear function before the nominal current is reached and connected through a parabolic function prior to reaching its nominal current. The proposed ramp and the residual eddy current multipole field component is shown in Figure 4.27. The optimised main dipole ramp requires about 26.3 minutes, which

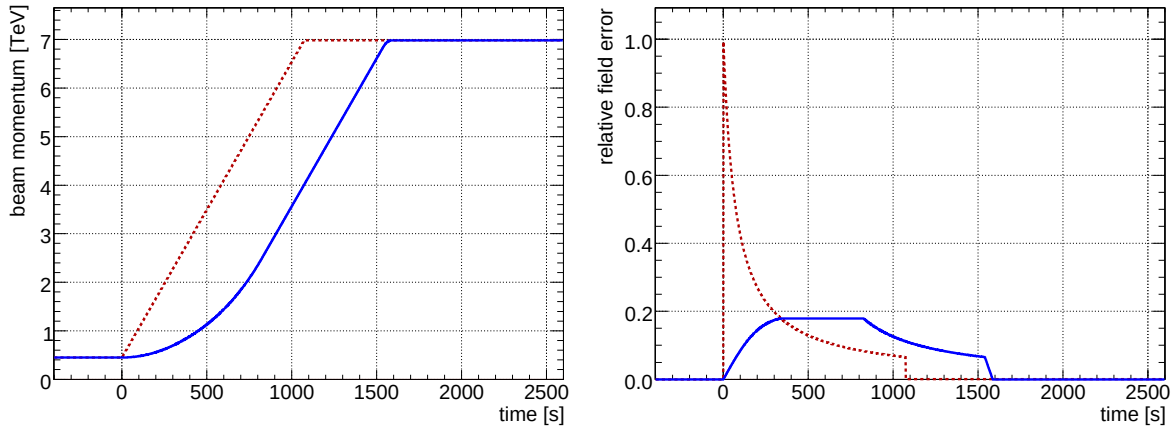


Figure 4.27: Main dipole current ramp (left) and rate of change (right). The optimised baseline (blue line) and a linear (dotted red line) ramp are shown. Though the optimised ramp requires 26.3 minutes, which is about 8.3 minutes more compared to the linear ramp, it reduces the amplitude of transients by a factor of about 5 [90].

is about 8.3 minutes longer than the faster linear ramp but reduces the peak transient due to the eddy currents by a factor of about 5.

Persistent Currents

The persistent currents are most pronounced during injection while only a small filament cross section is penetrated by the main dipole current. According to [91], the magnitude of the persistent currents $C_n^{pers.}$ depends on the actual magnet current I, magnet temperatures T and other magnet and material specific properties. The persistent current contribution C_n can be written as

$$C_n^{pers.}(t) = C_n^{pers.}|_{inj} \left(\frac{I(t)}{I_{inj}}\right)^{2-p_n} \left(\frac{I_c - I(t)}{I_c - I_{inj}}\right)^{q_n} \left(\frac{T_c^{2.7} - T^{2.7}}{T_c^{2.7} - T_{meas}^{2.7}}\right)^2 \tag{4.50}$$

with $C_n^{pers.}|_{inj}$ the n-th normal or skew multipole field component measured at injection setting, $I_{inj} = 739.29\,\mathrm{A}$ the injection current, I_c the critical current and T_c the critical temperature of NbTi. The fit parameter $p_n = +1.11$ and $q_n = +1.11$ are specific for the individual magnet. The field component decreases for increasing current and the time scale of the persistent current field reduction is determined by the ramp speed of the main dipole and quadrupole magnets.

Decay and Snapback

The *decay* and *snap-back* phenomena[6] are unique to accelerators using superconducting magnets. While resting on the injection plateau, the magnet's field components slowly decay and *snap-back* to their original at the start of the ramp. The approximative exponential decay occurs with a time constant of about $\tau \approx 1000\,\mathrm{s}$. It can be further decomposed into a fast and slow component. As described in [91, 92], the decay can be written as

$$b_n^{decay}(t)(t) \sim \delta_n \left[d \cdot \left(1 - e^{-\frac{t-t_{inj}}{\tau}} \right) + (1-d) \cdot \left(1 - e^{-\frac{t-t_{inj}}{9\tau}} \right) \right] \tag{4.51}$$

δ_n is a complex function that depends on the duration of magnet operation at nominal top energy, on injection porch duration, actual flat top current and on several other fitting parameters. Table 4.5 gives a qualitative example of the measured decay values for the main dipole's Δb_3 component. The variation covers an order of magnitude from $\Delta b_3 \approx 0.25$ to $\Delta b_3 \approx 3.31$. The *snap-back*, the reverse effect of the

	MB b_3 decay [units]
Nominal Decay	3.31
Beam abort at 4kA	0.64
Beam abort after 1 minute	0.25
30 minutes pre-injection stop	1.59
15 minutes injection	1.03
Magnet quench	0.25

Table 4.5: Variation of the main dipole's b_3 field component depending on different operational scenarios [88].

decay occurs at the start of the ramp can be approximated by an exponential decay

$$b_n^{snap}(t)(t) = \Delta b_n^{decay} \cdot e^{-\frac{I(t)-I_{inj}}{\Delta I}} \tag{4.52}$$

with $I(t)$ being the current of the ramping magnet. It has been observed and described in [92] that ΔI is proportional to the maximum decay b_n^{decay} prior to the start of the ramp, which is favourable with respect to the time scales of the required correction. A larger decay causes larger perturbations but also decreases the snap-back induced parameter drift velocities. Based on measurement, the analysis described in [92] gives the following estimate:

$$\Delta I \approx 6\,\mathrm{A} \cdot b_n^{decay} \tag{4.53}$$

Using equations 4.43 and 4.52 and multipole predictions found in Table 4.4, it is possible to simulate and predict the expected orbit perturbation due to the random snap-back of the main dipole's b_1 component only. The result is shown in Figure 4.28. The maximum orbit drift is about $550\,\mu$m r.m.s. The predicted maximum average orbit drift velocity is less than $15\,\mu$m/s r.m.s. during the snap-back.

Combined Results

Figure 4.29 and 4.30 show the combined effect of the snap-back, decay of persistent currents, and eddy currents onto the beam orbit and energy. The superimposition and prediction of the cumulative effects depend on whether the magnitude and sign of the snap-back is correlated to the magnitude and sign of the persistent current decay. The predictions show both random and systematic superimposition of these effects. Though the ramp design foresees a continuously differentiable main dipole current, it is further visible that the drift velocities change fast at about 800 and 1500 seconds due to theoretical discontinuities in the second derivative of the dipole current during the ramp. The drift velocities remain well below $20\,\mu$m/s, which provides an estimate for the required tracking speed or feedback bandwidth, which will be discussed later.

The given orbit predictions estimate the drift of the average r.m.s. equilibrium orbit; according to the statistical definition of the r.m.s., the drift may be three times larger at a particular monitor with a probability of 5%.

[6] See Section 1.3.2 or [20] for details.

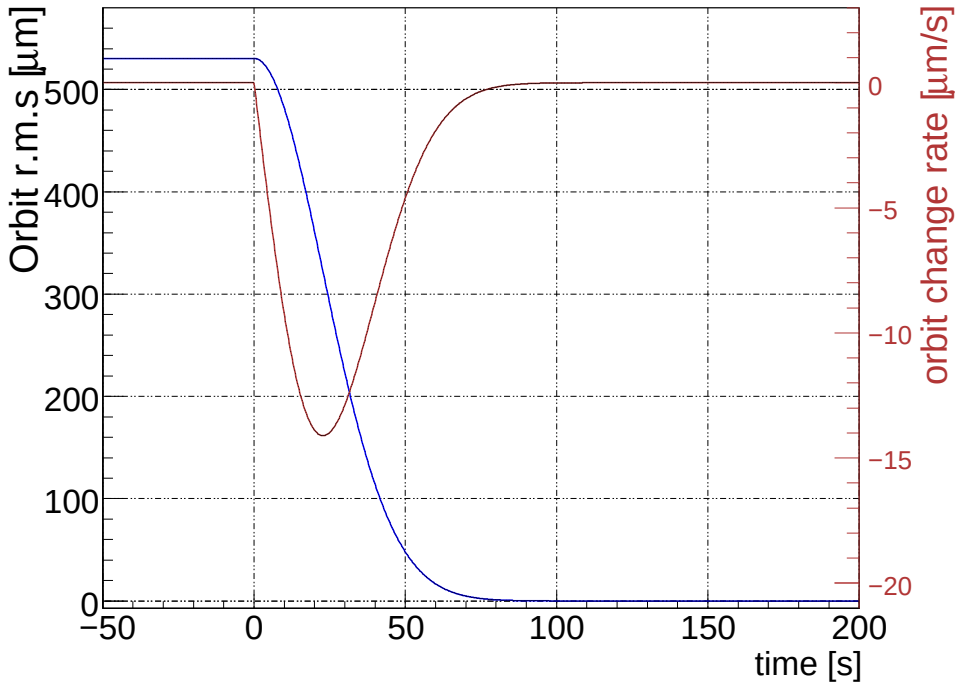


Figure 4.28: Time dependence of the orbit r.m.s. change during the snap-back at the start of the ramp. The predictions are based on the main dipole field errors given in Table 4.4 and propagation constants in equations 4.43 and 4.44.

4.4.2 Change of Final Focus

The optic is usually kept constant during injection and during the ramp to 7 TeV. Once the nominal momentum is reached, the beam sizes σ^* is decreased by reducing the betatron function β^* at the location of the IP, in order to optimise the luminosity. This change of optic is usually performed using the final focus triplet and other surrounding quadrupole magnets in the experimental insertions. The change of the final focus is also referred to as *squeeze*. Apart from the experimental insertions, the rest of the accelerator optic is usually kept constant during the squeeze.

During injection, the insertion optic is similar to the one found in the arc. Figure 4.31 shows the relative change of quadrupole strengths, betatron and phase advance due to the optic squeeze around IP5 (CMS) with respect to the injection optics. To provide a large change of the betatron function at the IP, the largest strength variations are performed at the quadrupoles in the dispersion suppressors, whereas the largest betatron and phase advance changes occur in the vicinity of the IP. This is related to the fact that, as shown for example in [17], the betatron function around the IP can be written as

$$\beta(s) = \beta^* + \frac{a^2}{\beta^*} \tag{4.54}$$

with a the distance from the IP. The betatron function increases quadratically with the distance to the IP, which links the large values of the betatron function inside the final focusing quadrupole magnets to the required space of the detectors and the closest possible location the final focus quadrupoles can be positioned. Usually, the final focus quadrupoles are placed as close as possible to the IP. In case of LEP the final focus quadrupoles where located inside the detectors [72]. Figure 4.32 shows the insertion optic of IR5 before and after the squeeze.

The effect of the final focus on the orbit is given by equation 2.66, that describes the feed-down of dipole momentum due to a change of quadrupole strength Δk in the presence of an off-centre beam Δx. In order to minimise the effect, the beam should ideally be steered through the centre of each quadrupole. However due to magnet alignment imperfection and other effects described above there is always a residual orbit perturbation within the quadrupoles.

A simulation of the squeeze has been performed assuming a random quadrupole misalignment of about 0.5 mm r.m.s. The largest squeeze-induced orbit drifts was found to reach amplitudes of up to 30 mm, if un-compensated, which would clearly exceed the available aperture in the machine and thus needs to be controlled. The average amplification factor between the r.m.s. quadrupole misalignment Δx_{quad} and

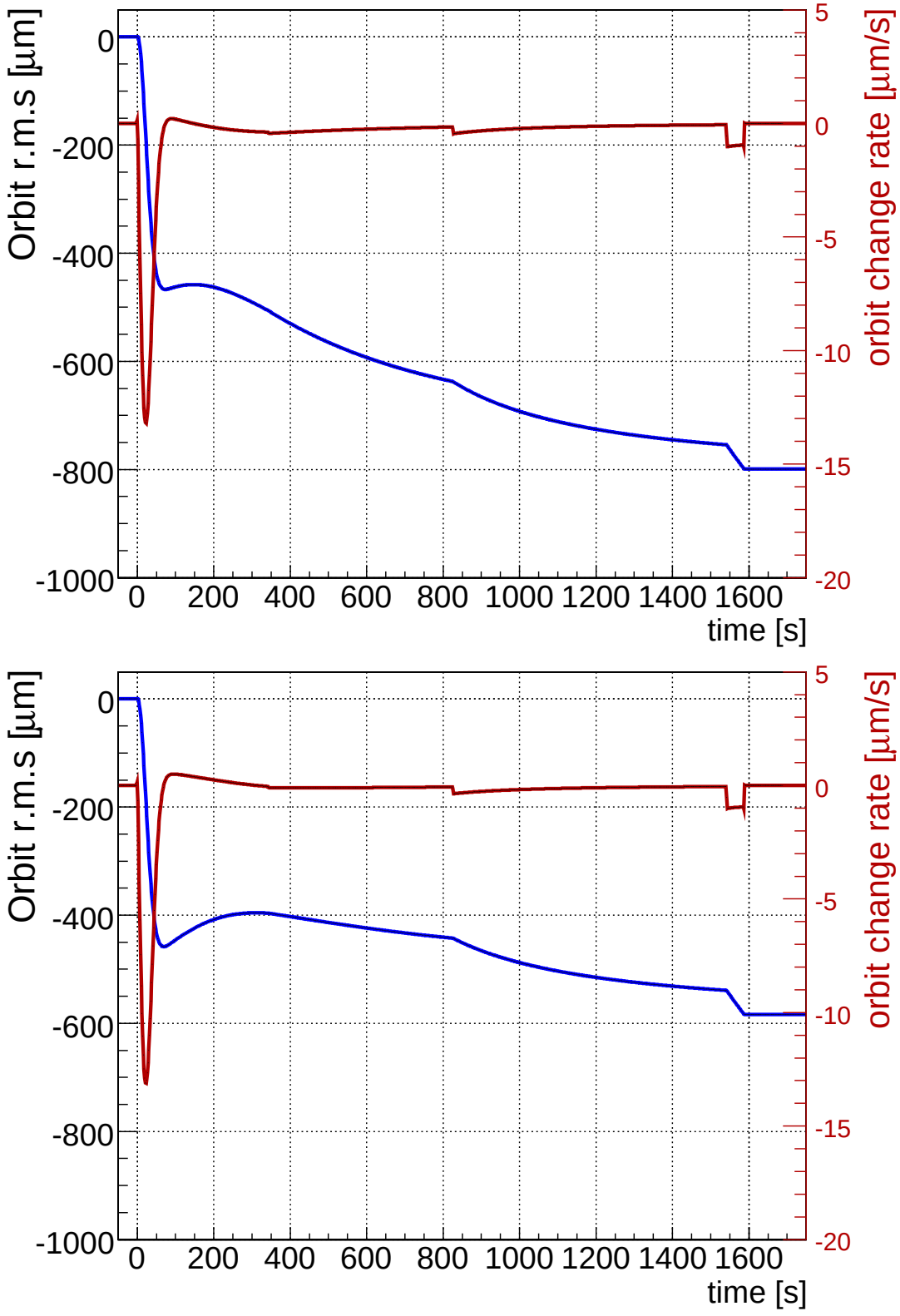


Figure 4.29: Expected total orbit perturbations during the ramp. The predictions are shown assuming correlated (top) and un-correlated (bottom) random b_1 contributions of the snap-back and persistent currents. The expected orbit drift velocity is less than 15 μm/s and maximum at the start of the ramp. The fast velocity changes at about 800 and 1500 seconds are caused by "theoretical" discontinuities in the second derivative of the dipole current during the ramp. In these figures the sign of the r.m.s. change has been deliberately changed to indicate snap-back to the original orbit before the start of the decay.

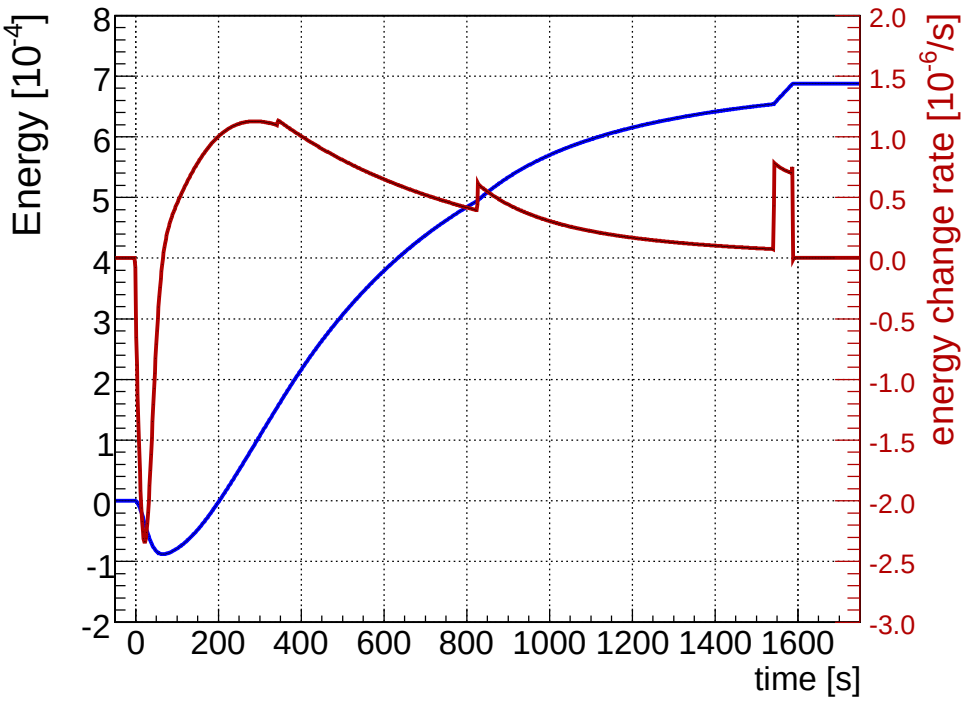


Figure 4.30: Expected total energy perturbation during the ramp. The expected maximum energy transient is about $\frac{\Delta p}{p} \approx 8 \cdot 10^{-4}$ with drift velocities less than $\left|\frac{d\Delta p/p}{dt}\right| < 2.5 \cdot 10^{-6}/s$. The fast velocity changes at about 800 and 1500 seconds are caused by discontinuities in the second derivative of the dipole current during the ramp.

squeeze induced orbit shift $\Delta x_{FF-orbit}$ can similarly to the ground motion be written as

$$\Delta x_{FF-orbit} = \kappa_{squeeze} \cdot \Delta x_{quad} \quad \text{with:} \tag{4.55}$$

$$\kappa_{squeeze} \approx 60 \tag{4.56}$$

The actual transient depends strongly on the chosen orbit correction algorithm and parameters and is further discussed in Chapter 6. The time scale of the final focus change is variable and depends strongly on the operation of the machine. The present baseline foresees a duration in the order of 20 minutes, which would yield a maximum orbit drift velocity of about

$$\left|\frac{\Delta x}{\Delta t}\right|_m ax \approx 25\,\mu/\mathrm{s} \tag{4.57}$$

which would make the squeeze one of the fastest expected orbit perturbations during machine operation. Though large parts of this drift will be compensated using the orbit correction dipole magnets next to the insertion quadrupoles, the required corrector current change may potentially exceed the ramping limit of the power supplies, discussed in Section 5.2. It is foreseen to increase the duration of the squeeze in order to slow down these squeeze induced orbit shifts in case of problems.

In a steady-state machine with not-moving quadupoles, above effect could in principle be fully compensated by feed-forwarding the changes from one cycle to the next, since the quadrupole field changes are known and the quadrupole misalignment is static. In a real machine, as shown in Section 4.1, the r.m.s. quadrupole drifts are expected to be within about $5 - 10\,\mu$m after 10 hours. Using this drift estimate and equation 4.55 yields an expected random orbit perturbation component during the squeeeze between about $300 - 600\,\mu$m that can only be compensated by beam-based feedbacks.

4.4.3 Flow of Cryogenic Coolant

Another potential source of machine-inherent orbit perturbations may be due to the flow of cooling liquids such as superfluid helium used for the superconducting LHC dipole magnets or cooling water that is used for the warm magnets. Cryogenics-induced orbit oscillation have been observed at the Relativistic Heavy Ion Collider (RHIC) at the Brookhaven Laboratory [93] where the cryogenics causes coherent

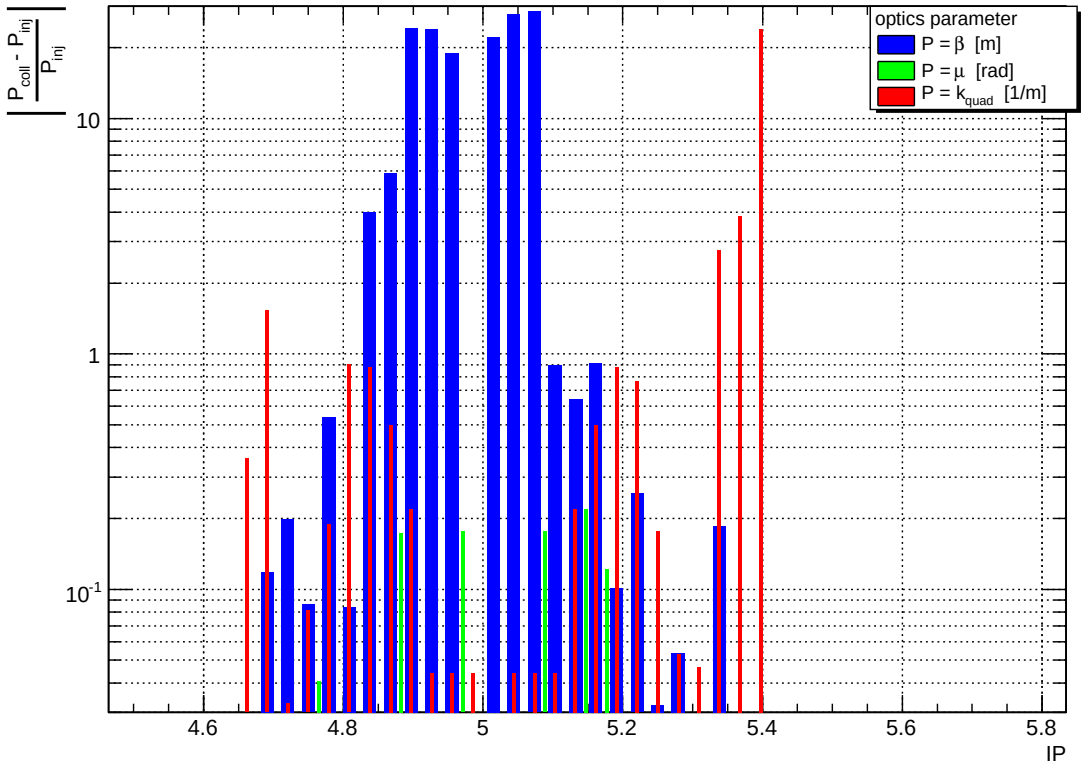


Figure 4.31: Relative optic change due to squeeze to $\beta^* = 0.55\,\mathrm{m}$. It is visible that the largest betatron function changes around the IP are mainly due to quadrupole strength changes in the outer dispersion suppressor quadrupoles.

orbit oscillations of up to $300\,\mu\mathrm{m}$ at a relatively low frequency between 10 and 11 Hz. RHIC has less strong requirements on the orbit and thus this effect is considered to be less critical.

The effect of cryogenics-induced orbit oscillation is most pronounced for the quadrupoles, in particular for those that are part of the final focus around the IPs. The flow of the cooling liquids may induce mechanical vibrations that dynamically misalign the magnetic centre of the quadrupole and thus create feed-down dipole kicks onto the beam, similar to ground-motion and thermal-induced perturbations, as described in equation 2.65. These vibrations are enhanced if their frequency matches one of the structure girder eigenfrequencies. The effect onto the beam is further amplified according to the accelerator lattice response presented in Section 4.1.

A possible forecast of this effect is limited by the poor knowledge of the actual quadrupole vibrations. An experimental verification without beam is difficult due to the limited space inside the magnet and the cryogenic temperatures since, for accurate results, the geophones or other similar devices should be operated within the magnet bore.

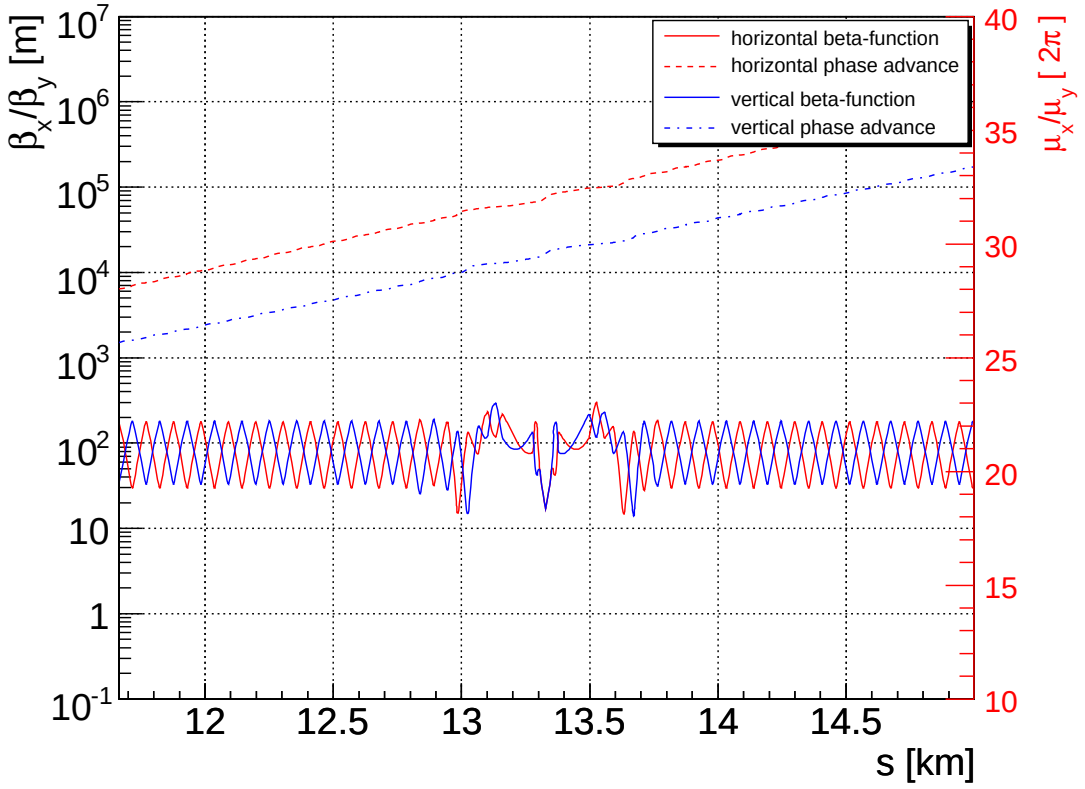
In the hypothetical case that there will be similar cryogenics-induced oscillations in the LHC, these effects could be minimised either by:

- changing the driving frequency through modifying the coolant flow velocity inside the cryogenic system,
- an active damping system or increase of the bandwidth of the presented systems.

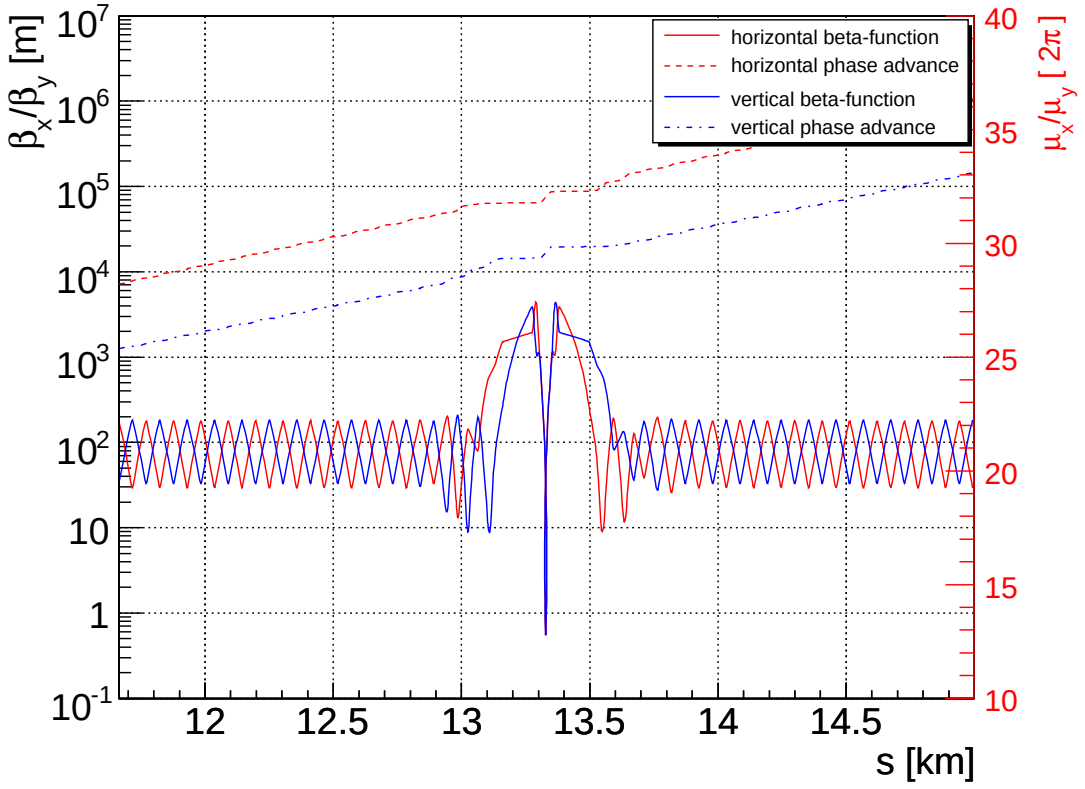
For the time being, this effect is rather speculative. If such an effect arises at the LHC, it will be measured. Depending on the size of the effect it may be necessary to re-evaluate the necessity of further control.

4.4.4 Orbit Corrector Hysteresis

During operation, the LHC corrector magnets will perform multiple field changes: their states will not be precisely known at the end of each run, due to magnet hysteresis. Earlier analysis described in [94, 95] estimated that these effects may have a significant impact on fill-to-fill stability at injection, reproducibility of settings, and operation of feedbacks, assuming the maximum possible width of the



(a) IR5 (CMS) injection optics



(b) IR5 (CMS) collision optics

Figure 4.32: Experimental insertion optic in IR5 occupied by the CMS detector. The horizontal and vertical beta-function and phase advance is plotted as a function of position s in the ring. In case of collision optics, the enlargement of the beta function at the triplets is visible. The horizontal (red) and vertical (blue) betatron function are superimposed and nearly identical close to the IR. IR1 has a similar optic.

hysteresis loop. Only a few magnets are expected to run at full current. Thus it was proposed to perform a detailed measurement campaign using more likely (small) settings and current changes in order to assess the effect of the hysteresis under more realistic beam steering conditions and to estimate the effect of pre-cycling the corrector magnets on fill-to-fill stability and on reproducibility of injection settings.

MCBH(V) Magnets

The 1060 orbit corrector dipoles of the LHC can be grouped into 8 families. This analysis focuses on the stability and hysteresis of the 'MCBH(V)' type COD family, which represent, with 752 out of 1060, the majority of all corrector dipole magnets. The results should qualitatively apply for the other magnet types, since (except of the 'MCBX' and 'MCBW' types) they have a similar design, parameter, and location in the machine. Table 4.6 summarises the parameter of all COD families. Each 'MCBH(V)'

Magnet Type	**B**	$\mathbf{L_{mag}}$	$\mathbf{BL_{mag}}$	$\mathbf{I_{nom}}$	$\mathbf{\lvert\Delta I/\Delta t\rvert_{max}}$	**N**
	[T]	**[m]**	**[Tm]**	**[A]**	**[A/s]**	
MCBH(V) @1.9K	2.93	0.647	1.90	55	0.5	752
MCBCH(V) @1.9K	3.11	0.904	2.81	100	1.0	156
MCBCH(V) @4.5K	2.33	0.904	2.11	80		
MCBYH(V) @1.9K	3.00	0.899	2.70	88	1.0	88
MCBYH(V) @4.5K	2.50	0.899	2.25	72		
MCBXH	3.35	0.45	1.51	550	5.0	48
MCBXV	3.26	0.48	1.56	550		
MCBWH(V)	1.1	1.7	1.87	500	5.0	16

Table 4.6: Available LHC corrector types. BL_{mag} is the integrated dipole strength, I_{nom} the nominal current, $|\Delta I/\Delta t|_{max}$ the maximum current ramping rate given by the power converter and N the number of magnets installed in the LHC. A complete parameter list can be found in [3, 96].

magnet has a maximum integrated dipole field strength of $BL|_{max} = 1.896\,\mathrm{Tm}$ that corresponds to maximum possible deflection of:

$$\begin{array}{rcrll} \delta_{max} & = & 1260\,\mu\mathrm{rad} & @450\,\mathrm{GeV} & \text{and} \\ \delta_{max} & = & 81\,\mu\mathrm{rad} & @7\,\mathrm{TeV} & \end{array} \tag{4.58}$$

Using the LHC arc injection lattice and $\beta_{max} \approx 180\,\mathrm{m}$ and equation 2.71, each COD can create a maximum beam orbit excursion of about $\delta x_{max} \approx 144\,\mathrm{mm}$ and $\delta x_{max} \approx 9\,\mathrm{mm}$ at 450 GeV and 7 TeV, respectively. Since the aperture of the beam screen is about 44 mm, each corrector magnet is capable of deflecting the beam outside the vacuum chamber at injection. The analysis focuses on beam stability at the injection energy (450 GeV) as the beam is more sensitive to field errors and power converter ripples at low energies. It is noteworthy that the superconducting dipole corrector magnets are made of strands in contrast to the main quadrupole and dipoles where the strands are further grouped to cables. Thus the dipole correctors do not show excessive decay or snapback effects. However, since their strands are made out of superconducting NbTi filaments, they are still subject to filament magnetisation and filament-related hysteresis phenomenon.

In 2005, the hysteresis properties of one MCB orbit magnet was measured at 1.9 K (courtesy W. Venturini, [97]). The measurement campaign was designed to clarify:

- the reproducibility and deviation of the hysteresis after a predefined cycle, e.g. cycling through saturation or using a 'De-Gauss'-cycle
- to check whether there is a minimum required current change in order to change the magnetic field/deflection of the CODs.

The purpose of the first measurement was to provide an estimate of the expected contribution of the MCB CODs to fill-to-fill injection stability and reproducibility of settings due to the hysteresis, whereas the second determines the expected orbit correction convergence, as a possible dead-band or quantisation effect might eventually limit the correction schemes of the feedback loop.

Reproducibility after Pre-cycling

There are two classes of proposed current pre-cycles that are sketched in Figure 4.33.

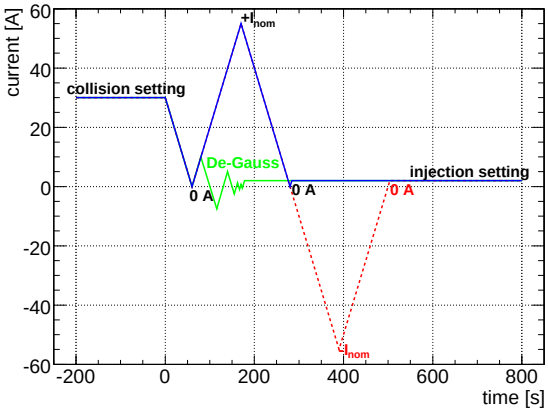


Figure 4.33: Schematic COD pre-cycles: cycle through positive saturation only (blue curve), through positive and negative saturation (red curve), or using a De-Gauss cycle with reducing amplitude (green curve).

1. Cycling through saturation of the magnets (either through maximum '$+I_{nom}$' and/or minimal nominal current '$-I_{nom}$') ensuring that the magnetic history of the persistent current is erased; maximum remanent field is expected.

2. A De-Gauss cycle that applies an oscillating current with decreasing amplitude to the magnet. The initial current amplitude has to be larger than the maximum expected remanent field. This cycle not only ensures that the magnetic history of the persistent current is erased, but also that the remanent field converges to zero.

In order to simplify controls, the currents are first set to zero before and after the pre-cycle prior to the new injection setting, for both pre-cycle types. The required time for both pre-cycle types is, for all orbit corrector magnets, in the order of 5- 10 minutes. During LHC operation, the pre-cycles could be performed in the shadow of the main dipole magnet ramp down.

Reproducibility Measurement

A-priori, both pre-cycle options are applicable. However, as it is easier to be set-up, the first pre-cycle option was chosen to test the reproducibility and the magnets were cycled through positive saturation only ('$I_0 = 0\,\mathrm{A} \leftrightarrow I_{nom} = +55\,\mathrm{A}$). After each cycle, the reproducibility of the remanent field at zero current was measured. Figure 4.34 shows the measurement results. Successive measurements show that

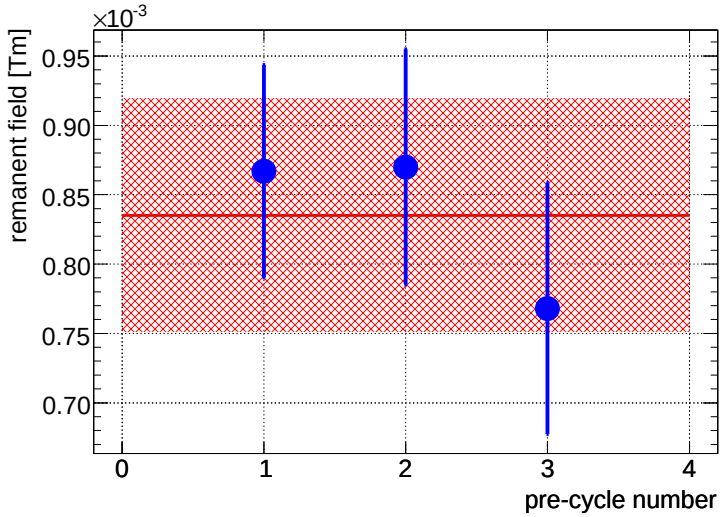


Figure 4.34: Field reproducibility at 0 A after pre-cycling the magnet through saturation. The plotted error bars (area) correspond to $1\,\sigma$ r.m.s.

the reproducibility was about $2.5 \cdot 10^{-5}$ Tm, as measured at 1, 10 and 50 A. The reproducibility is smaller than the cycle-to-cycle variation and thus the measured spread can be contributed to the latter effect.

Averaging over the measured cycles gives the following estimate for the reproducibility of the remanent integrated dipole corrector field and corresponding kicks, respectively:

$$\begin{aligned} BL_{mag} &\approx \mu \pm \sigma \\ &\approx (8.4 \pm 0.8) \cdot 10^{-4}\,\mathrm{Tm} \\ \text{resp.} \qquad \delta_{\mathrm{cod}} &\approx (560 \pm 53)\,\mathrm{nrad} \end{aligned} \tag{4.59}$$

It is important to note that these numbers are based on a low statistic of only three cycles, which strictly speaking, correspond to a statistical confidence of less than one σ. The analysis revealed that the measurement may have been influenced by power converter stability. It is known that converter stability around zero current is an issue. Further, these tests were performed using a ± 600 A power converter that has, at the nominal MCB current of 55 A, a worse stability than the ± 60 A converter foreseen for the MCB type magnets in the LHC. In case more detailed measurements of these magnets are requested, the operational working point after cycling should be different from zero current and if possible a nominal ± 60 A power converter should be used.

Hence we believe that these numbers represent rather worst-case estimates. However, they are still a good estimate for fill-to-fill reproducibility of better than 10^{-4} Tm.

Implications for Fill-to-Fill Reproducibility

The remanent field given in equation 4.59 can be broken down into a systematic

$$< \Delta\delta_{cod} >= 560\,\mathrm{nrad} \tag{4.60}$$

and random

$$\sigma(\delta_{cod}) = 53\,\mathrm{nrad\ r.m.s.} \tag{4.61}$$

component that affect the beam in two different ways.

The pre-cycle was chosen to go always to positive saturation only. Hence the systematic kick $\Delta\delta_{cod}$ has the same sign for all correctors. In case of horizontal correctors, this increases the total integrated dipole field and, as a result, the energy of the LHC as well. The maximum expected momentum shift $\Delta p/p$ due to the horizontal MCB hysteresis, which can be evaluated using equation 2.78, is about $2 \cdot 10^{-5}$. Compared to the expected energy shifts of $1.5 \cdot 10^{-4}$ caused by the b_1 decay of the main dipole field and lunar and solar tides, this contribution is rather small. This shift is reproducible from fill-to-fill and has, in principle, to be corrected only once. If required, a 'De-Gauss' pre-cycle would minimise this contribution.

The $\sigma(\delta_{cod})$ component of about 53 nrad r.m.s. ($\sigma(\delta_{cod})/\delta_{max} \approx 4 \cdot 10^{-5}$) around the systematic part of the hysteresis causes a random orbit perturbation $\Delta x(s)$, described in equation 2.71

$$\Delta x(s) = \frac{\sqrt{\beta_i \beta(s)}}{2\sin(\pi Q)} \cos(\mu(s) - \pi Q) \cdot \sigma(\delta_{cod})$$

around the ring that contributes to the total random fill-to-fill variation. Using either an analytical approach and applying an incoherent sum of the orbit corrector response (equation 2.71) or a numerical evaluation of the orbit response matrix due to random dipole kicks leads to the following estimate of the propagation factor between the random COD deflection $\sigma(\delta_{cod})$ and resulting orbit r.m.s. σ_H and σ_V, respectively:

$$\sigma_H \approx (966 \pm 245)\,[\mathrm{m/rad}] \cdot \sigma(\delta_{cod}) \tag{4.62}$$

$$\sigma_V \approx (1004 \pm 275)\,[\mathrm{m/rad}] \cdot \sigma(\delta_{cod}) \tag{4.63}$$

The simulated estimates are based on the LHC injection optics (LHC 6.5) and a seed of about 10^4 different orbits. The spread of the prediction reflects the distribution of different beta function and phase advance combinations that are sampled by the different seeds. See section 4.1 for details.

The expected orbit r.m.s during injection due to hysteresis can be estimated to about 50 μm r.m.s. (0.05σ with σ being the beam size), using equations 4.61 and 4.63. This orbit excursion, solely due to the MCB hysteresis, is very small compared to the available aperture of about 11 mm and expected ground motion contribution (see Section 4.1; $0.3 - 0.5\sigma$). As shown in Section 5.1, it is barely detectable with a LHC BPM shot-by-shot resolution of about $50 - 100\,\mu$m for a single nominal LHC bunch.

Small Hysteresis Loops

The orbit perturbations on the injection plateau, which are corrected by the orbit feedback, are expected to be in the order of about 0.5 mm, as described in Section 4.1. Assuming an arc dipole corrector at $\beta = 180$ m, this amplitude corresponds to an average current modulation around the initial corrector working point of about 0.2 A at 450 GeV, which is small compared to the nominal current (55 A). Thus, 1 A and 10 A are likely working points of the CODs, assuming a static random misalignment of the quadrupole magnets of 0.5 mm r.m.s. during injection and collision, respectively.

The small hysteresis loop of 0.2 A was measured around 1 and 10 A as shown in Figure 4.35. The

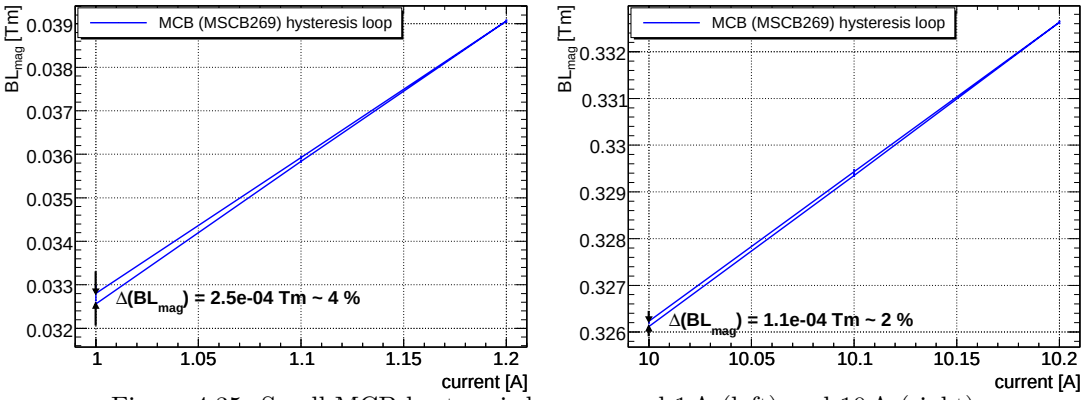


Figure 4.35: Small MCB hysteresis loop around 1 A (left) and 10 A (right).

width of the small hysteresis loops are $2.4 \cdot 10^{-4}$ Tm and $1.1 \cdot 10^{-4}$ Tm at 1 A and 10 A, which correspond to deflections of about $\Delta\delta_{cod} = 167$ nrad and $\Delta\delta_{cod} = 73$ nrad, respectively. This additional deflection due to the hysteresis can be translated into a relative scaling error ϵ_{scale} of about 4 % with respect to the maximum deflection. In a feed-forward-only environment, this scale error would directly translate into a 4 % error with respect to the given reference orbit. It is important to note that, though the requested field change is less due to hysteresis, neither quantisation nor a dead-band effect has been observed. This shows that even a small current change yields an immediate change of the field, and hence deflection of the magnet. The effect of the hysteresis can thus be cast into a time and history dependent multiplicative process error that has to be taken as a uncertainty into account while designing the feedback loop.

4.4.5 Power Converter Noise

The lower bound for orbit stability is given by the stability of power converters in particular those used for the orbit corrector and main dipole magnets. In 2005, a $\pm 8\,\mathrm{V} / \pm 60\,\mathrm{A}$ power converter that is foreseen for the LHC MCBH(V) magnets was tested with a cryogenic MCB load attached [98]. The stability of the power converter was measured to $\frac{\Delta I}{I}\big|_{MCB} = 5 \cdot 10^{-6}$ with respect to its nominal current $I_{nom} = 55$ A and the maximum peak-to-peak excursion observed was less than $\frac{\Delta I}{I}\big|_{MCB} = 19 \cdot 10^{-6}$. The current stability corresponds to an r.m.s. and peak-to-peak deflection of each MCB magnet in the LHC of

$$\sigma(\delta_{cod}) = 6.3\,\text{nrad r.m.s.} \tag{4.64}$$

$$\sigma(\delta_{cod})|_{pp} = 25.2\,\text{nrad} \tag{4.65}$$

at a beam momentum of 450 GeV. Using equations 4.64, 4.65 and 4.63, the noise floor due to the COD power converter can hence be estimated to

$$\Delta x \approx (6 \pm 2)\,\mu\text{m r.m.s.} \tag{4.66}$$

$$\Delta x|_{pp} \approx (24 \pm 26)\,\mu \text{ peak-to-peak} \tag{4.67}$$

corresponding to a stability of about 0.01 and $0.04\,\sigma$ (σ being the maximum arc beam size), respectively.

As visible in Table 2.1, the magnetic deflection reduces with increasing beam momentum and at 7 TeV, is a factor of about 15.5 smaller. The noise floor given by the orbit corrector dipole power converter is thus for collision energy about:

$$\Delta x \approx (0.4 \pm 0.1)\,\mu\text{mr.m.s} \tag{4.68}$$

$$\Delta x|_{pp} \approx (1.5 \pm 1.7)\,\mu\text{peak} - \text{to} - \text{peak} \tag{4.69}$$

4.5 Beam Stability Requirements

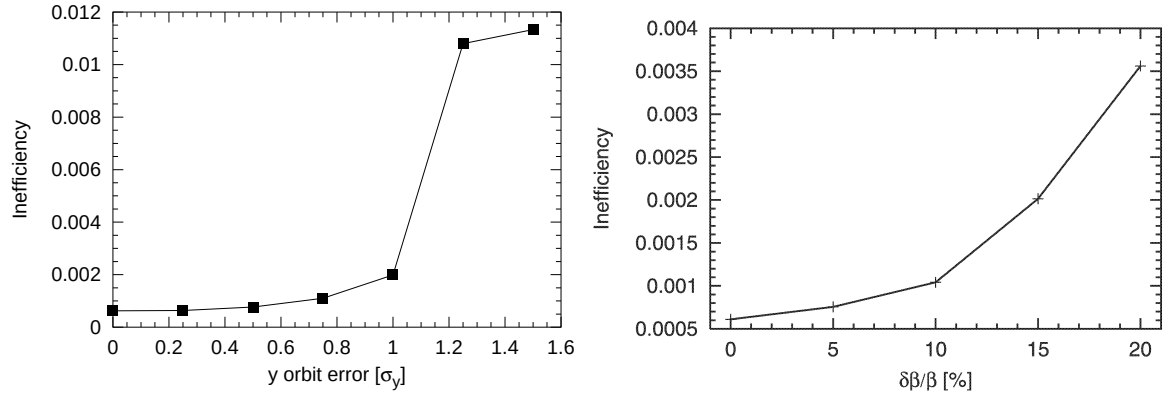
The need for an automated orbit and energy correction is mainly driven by requirements derived from collimation and machine protection. In addition to these, there are a number of additional requirements based on the reduction of feed-down effects, operational efficiency, electron cloud and optimisation of measurement procedures such as the transverse damper and the TOTEM experiment.

4.5.1 Collimation

The performance of the LHC Cleaning System critically depends on the orbit. The cleaning inefficiency η is defined as the ratio between the number of protons impacting the primary collimator and the number of protons escaping the cleaning system and getting lost in the cold aperture that requires protection. As analysed in [29, 99], the maximum allowed cleaning inefficiency is determined by the minimum quench limit R_q of the superconducting magnets, the total number of stored protons N_{max}, the average dilution length L_{dil} and the minimum acceptable lifetime τ_{min}

$$\eta = \frac{\tau_{min} \cdot R_q \cdot L_{dil}}{N_{max}} \tag{4.70}$$

Inserting the expected nominal values for $R_q \approx 7.6 \cdot 10^6$ protons/s, $N_{max} \approx 3 \cdot 10^{14}$, $L_{dil} = 50\,\text{m}$ and $\tau_{min} := 10\,\text{min.}$ while running at 7 TeV, the cleaning inefficiency has to be smaller than $\eta = 10^{-3}$ (see [99, 100] for details). To meet nominal requirements, the LHC Cleaning System consists of a two-stage collimation system. Figure 4.36 shows its cleaning inefficiency with respect to a peak-to-peak orbit error and relative beta-beat between the primary to secondary collimator. The secondary collimator is retracted by $1\,\sigma$ ($\sigma \approx 160\,\mu$m being the r.m.s. beam size at the collimator). In a worst case estimate,



(a) peak-to-peak orbit (b) beta-beat

Figure 4.36: Collimation inefficiency vs. peak-to-peak orbit error (left) and peak-to-peak beta-beat (right) [100]. A fast increase of the cleaning inefficiency is visible as the orbit error approaches $1\,\sigma$, with σ being the beam size at the collimator.

the errors of orbit and beta-beat add up linearly. Thus, as visible in Figure 4.36, the total orbit error between primary and secondary collimators should be less than $0.6\,\sigma$ and the beta-beat less than 5%, in order to be able to provide the required cleaning inefficiency. The total stability margin of about $0.6\,\sigma$ is shared between different systematics such as jaw positioning precision, jaw surface flatness, beta-beat and orbit at the jaws, that are summarised in Table 4.7. As a trade-off between collimation and orbit

	Tolerances
mechanical tolerances	$40 - 60\,\mu$m
collimator setup	$50\,\mu$m
orbit	$50 - 30\,\mu$m
beta-beat	$5\% \longrightarrow \approx 20\,\mu$m
Total (0.6σ)	$160\,\mu$m

Table 4.7: Tolerance margin with respect to collimator at $\beta = 150\,\text{m}$ [101, 102].

feedback design, the orbit has an assigned budget of about $0.3\,\sigma$, corresponding to a peak-to-peak orbit error of about $30-40\,\mu$m between the primary and secondary collimator. The orbit error given in Figure 4.36 is measured between primary to secondary collimators. At a single collimator the required stability scales thus down by a factor of two and is given by:

$$\Delta x \overset{req.}{\leqslant} 0.15\,\sigma \approx 15-25\,\mu\text{m} \tag{4.71}$$

During 'Stage I', it is expected to accelerate only up to 43 nominal bunches with a bunch intensity of $5\cdot10^{12}$ protons and total intensity per beam N_{max} of about $5\cdot10^{12}$ protons. Comparing the reduced total intensity with equation 4.70 and assuming operation at 7 TeV, the maximum acceptable cleaning inefficiency is more relaxed:

$$\eta \lesssim 0.05 \tag{4.72}$$

Comparing the required inefficiency with Figure 4.36, a peak-to-peak orbit stability of about $1\,\sigma$ seem to be sufficient for Stage I ([103]) which relates to the required orbit stability at the location of the collimator of:

$$\Delta x \overset{req.}{\leqslant} 140\,\mu\text{m} \tag{4.73}$$

4.5.2 Machine Protection

In addition to the stability of the orbit at the collimators, to ensure proper function of the Cleaning System and protection devices, the orbit in the arc has to be controlled to a level which guarantees that the protection devices and collimators always define the aperture. For instance: At 450 GeV, the injection protection absorber TDI, found in IR2 and IR8 for Beam 1 and 2, respectively, are positioned at $7\,\sigma$ in order to protect against injection kicker failures [16]. Presently, the estimated arc aperture during injection is about $7.5\,\sigma$ [104]. A possible bump of $1\,\sigma$ in the arc, created either deliberately or naturally through ground motion drift or persistent current effect, would reduce the distance of the beam with respect to the aperture. In case of an injection kicker failure, the TDI would pass a beam with an oscillation amplitude up to about $6\,\sigma$, which in addition with a bump inside the arc, would lead into a direct particle loss into the cryogenic aperture. The criticality of this failure mode lies in the combination of a slow orbit error that moves the beam close to the aperture and a fast machine failure due to, for example, the injection, extraction, other kicker magnets or failure of other beam steering elements inside the machine.

Thus, the distinction between global and local orbit requirements is less evident, and as a consequence, the global orbit has to be steered to about the same level as inside the protection and collimation regions. The present estimate of available aperture indicates that the orbit should be steered at the level of $0.2-0.3\,\sigma$ at injection.

A first order scheme that minimises these types of failures is to use a two staged injection scheme: as a general rule, a single bunch with a low bunch intensity is always injected first into the LHC in order to test the correctness of the injection kicker settings and machine optics. The further injection of high intensity beams is only permitted once the correctness of machine optics and settings has been verified. This first beam is commonly referred to as *pilot* bunch and has a intensity of about $5\cdot10^9$ protons per bunch. This rather low intensity was chosen in order to minimise magnet quenches and prevent damage to machine equipment at 450 GeV. However, at 7 TeV, even a single pilot bunch could be sufficient to cause damage to the machine if lost accidentally [24].

4.5.3 Performance

Beside collimation and machine protection, there are other requirements to the orbit stability that are less critical with respect to machine safety but important for machine optimisation and operation:

- The transverse damper is a feedback system installed in IR4, foreseen to stabilise injection oscillations and other (coupled) beam instabilities that may occur during regular operation [105,106]. The transverse damper relies on two beam position monitors and a fast kicker magnet that is used to excite and damp the beam. In order to exploit the maximum resolution of the oscillation measurement, it is favourable to have a stabile orbit near the centre of the feedback BPMs. The required stability is in the order of $200\,\mu$m at the location of the BPMs.
- The same applies to the beam position monitors that are installed close to the extraction and protection devices in IR6. These monitors detect fast orbit movements and emit an interlock signal once the amplitude exceeds about 4 mm, in order to ensure a clean extraction. A orbit stability at the BPMs of about $200\,\mu$m is required.

- Stability of collision points: To maximise the luminosity for the LHC experiments, the stability of the two LHC beams at the interaction point is increasingly important. In circular particle-antiparticle accelerators such as Tevatron or LEP, for example, a shifted quadrupole has the same effect on both beams due to the CP symmetry of the involved particle beams. However, as an intrinsic property of accelerators that collide particles of the same species, external quadrupole shifts and resulting feed-down dipole momenta always separate the two beams at the IP. For the reduction of the luminosity (equation 1.8) due to non perfect beam overlap (for a round beam and same beam sizes for Beam 1 and 2 at the IP) is given by

$$\mathscr{L} \approx \mathscr{L}_0 \cdot e^{-\left[\frac{(\overline{x}_2-\overline{x}_1)^2}{(2\sigma_x)^2}+\frac{(\overline{y}_2-\overline{y}_1)^2}{(2\sigma_y)^2}\right]} \tag{4.74}$$

  with σ_x the horizontal and σ_y the vertical beam size, and $\overline{x}_i$ the horizontal and $\overline{y}_i$ the vertical beam position at the IP of Beam 1 and 2, respectively. To stabilise the relative luminosity to better than for example 1%, the beam separation must not exceed $0.2\,\sigma^*$, which is $3.2\,\mu$m with $\sigma^* \approx \sigma_x \approx \sigma_y \approx 16\,\mu$m. For comparison, a separation of the two beams by $1\,\sigma^*$ corresponds to a luminosity reduction of about 20 %. The first BPMs ('BPMSW.1RN', with N the IP number) are located at a distance of about 20 m around the IP. Since there are no further optical elements between the IP and these monitors, the required position and measurement stability translates one-to-one to the beam stability at the beam position monitors. The final luminosity optimisation will be performed using luminosity monitors.

  In addition, for nominal operation, the two LHC beams collide with a crossing angle θ causing a further reduction of luminosity, as described in equation 1.9. As given in [107], it is requested that these crossing angles are stabilised within $7\,\mu$rad for regular LHC operation and $0.2\,\mu$rad while providing collisions for the TOTEM experiment. This crossing angle stability translates to beam stability at the BPMSW monitors of about $140\,\mu$m for nominal operation.

  The requested stabilities are quite demanding, especially with respect to the expected LHC BPM design resolution of about $5\,\mu$m [108]. It is important to note that, besides the crossing angle, the actual luminosity optimisation will be based on a feedback using luminosity monitors that are within the TAN absorber blocks, installed around the IPs, as described in [107]. In order to simplify and to optimise its operation, it is however favourable to pre-stabilise the orbit. The actual required level of stabilisation has to be established during LHC operation.

- A limit to the nominal LHC performance may be given by the *electron cloud* effect. This effect arises due to the emission of photo electrons emitted from the beam screen that protects against synchrotron radiation loss and related heating of the superconducting aperture. The effect of the electron cloud is further enhanced by the 25-ns bunch structure of the LHC beam that leads to electron multi-pacting. The effect depends on the material properties of the screen, such as the secondary emission yield, and beam parameters such as bunch intensity and bunch spacing [109]. The maximum possible bunch intensity is essentially limited by the available cooling power of the beam screen that can lead away the heat losses created by the electron cloud effect. As described in [109], besides the geometry and alloy choices for the beam screen, keeping the high intensity beam at the same location inside the vacuum chamber may reduce the effect of electron cloud. It has been observed that the electrons of the clouds condition the chamber surface and effectively reduces the secondary emission yield of the surface. In order to exploit this effect, it is foreseen to perform a so-called *scrubbing run*.

  In order to ensure the efficiency of the scrubbing, the beam must remain to the cleaned surface. A stability of better than $1\,\sigma$ seems to be sufficient.

Feed-Down of Higher Multipole Magnets

A good control of the global orbit also helps to minimise the dynamic feed-down of higher order magnets such as shifted quadrupoles, inducing a dipole kick. One important feed-down may arise due to the vertical orbit offset inside the lattice sextupoles that, similar to tilted quadrupoles, cause transverse coupling. Compared to other multipole field expansions given in Table 2.1, it is visible after inserting $y = y_0 + \Delta y$ into the sextupole definition that the shifted sextupole generates a skew quadrupole with strength

$$\kappa_{feed-down}(s) = m(s)\Delta y \tag{4.75}$$

and horizontal dipole kick that depends on the square of the vertical orbit with respect to the magnetic sextupole centre:

$$\delta_{x_{feed-down}}(s) = \frac{1}{2} m(s) \Delta y^2 \tag{4.76}$$

In the presence of coupling between planes, the transverse betatron oscillations are described by the two eigenmodes-modes frequencies Q_1 and Q_2 that are given by (see [33, 110] for details):

$$\begin{cases} Q_1 &= Q_{x,0} - \frac{1}{2}\Delta + \frac{1}{2}\sqrt{\Delta^2 + |C^-|^2} \\ Q_2 &= Q_{y,0} - \frac{1}{2}\Delta + \frac{1}{2}\sqrt{\Delta^2 + |C^-|^2} \end{cases} \tag{4.77}$$

with Δ the unperturbed horizontal-vertical tune split (LHC: $\Delta = 0.003$ for injection and $\Delta = 0.001$ for collision optics), and $Q_{x,0}$ and $Q_{y,0}$ the unperturbed horizontal and vertical tune, respectively. The global coupling parameter C^- can be computed through the following equation ([110]):

$$C^- = |C^-| e^{i\xi} = \frac{1}{2\pi} \oint \sqrt{\beta_x \beta_y} \kappa(s) e^{i(\mu_x - \mu_y - 2\pi\Delta \cdot \frac{s}{C})} ds \tag{4.78}$$

with $\kappa(s)$ the skew quadrupole momentum, defined in Table 2.1, at the location s in the ring. To obtain a rough decoupling of the planes, $|C^-|$ should be smaller than the unperturbed tune split Δ. However, in practice and in particular for feedbacks that rely on fully decoupled planes, it is convenient if the coupling is at least a factor five to ten smaller than the unperturbed tune split Δ.

Using equations 4.75 and 4.78, one can estimate the induced coupling contribution due to vertical orbit misalignments in the lattice sextupoles. For this estimate, the quadrupoles were randomly misaligned, the corresponding orbit evaluated, and the coupling evaluated using equation 4.75 based on the beam position and sextupole strength of the individual magnet. The distribution of the coupling parameter $|C^-|$ is based on a seed of 10^6 orbits that was normalised to 1 mm r.m.s., shown in Figure 4.37. With

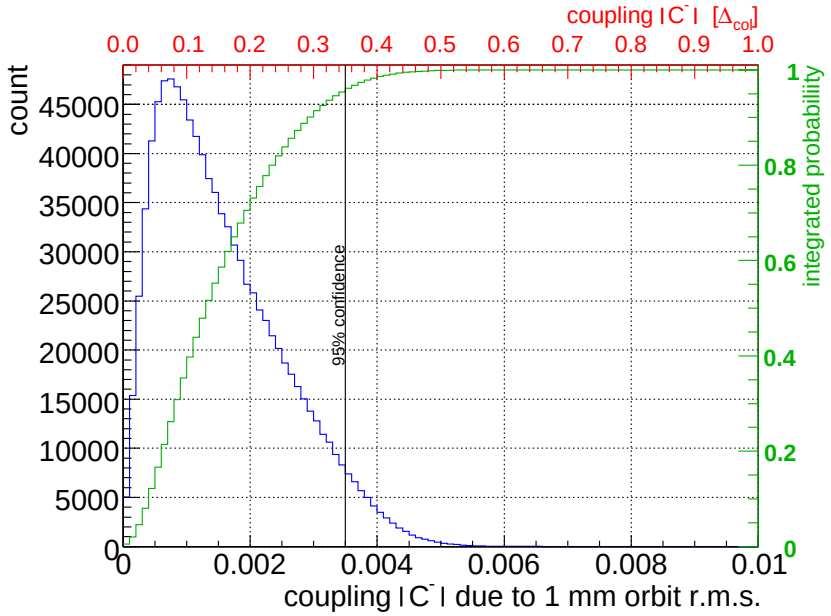


Figure 4.37: Global coupling $|C^-|$ due to skew quadrupole feed-down of the lattice sextupoles and vertical orbit r.m.s. of 1 mm. The simulation is based on 10^6 seeds.

95% confidence, $|C^-|$ is below 0.0035 and 0.35 for the collision tune split $\Delta_{col} = 0.01$. In order to meet the requirement that the coupling is less than a tenth of the tune split, the orbit r.m.s. should thus be stabilised better than about 300 μm r.m.s. during collision and better than about 860 μm r.m.s. during injection, taking the three times larger tune split of $\Delta_{inj} = 0.03$ during injection into account.

Luminosity and Total p-p Cross-section Measurements

A special stability requirement arises from the TOTEM Luminosity measurement, and the determination of the total proton-proton cross-section. As described in [12], TOTEM intends to measure the total

proton-proton cross-section at 7 TeV using the optical theorem and study diffractive physics at very low transverse momenta in the order of about $t \approx 10^{-3}\,\mathrm{GeV}^2$. These momenta correspond to an angle of about 10^{-5} rad with respect to the beam axis. The total proton-proton cross-section can be used to normalise and calibrate other relative luminosity measurements described in [107, 111].

In order to measure these small transverse momenta, TOTEM intends to use the parallel-to-point focusing method, illustrated in Figure 4.38. In this scheme, the scattering angle at the IP is translated

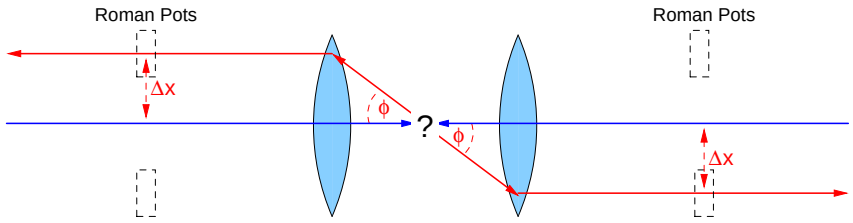


Figure 4.38: Schematic parallel to point focusing used for the TOTEM experiment.

to a beam position shift at the Roman Pots. The Roman Pots are special vacuum tanks that can move close to the beam, and contain silicon detectors that track the beam position of the scattered particle in the beam halo. A similar setup will be used by ATLAS. The parallel-to-point focusing condition can be written using the matrix formalism as

$$\begin{pmatrix} u(s) \\ u'(s) \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} u^* \\ u'^* \end{pmatrix} \tag{4.79}$$

with u^* the position, u'^* the scattering angle at the IP, $u(s)$ the position and $u'(s)$ the angle of the scattered particle at the location of the Roman Pot and T_ij the element of the beam transfer matrix given in equation 2.43. In order to translate the scattering angle into a position, the following condition has to be met:

$$u'^* = \frac{u(s) - T_{11}u^*}{T_{12}} \tag{4.80}$$

Comparing with equation 2.43 it is visible that:

$$T_{11} = \sqrt{\frac{\beta_u(s)}{\beta^*}}\cos(\Delta\mu) \tag{4.81}$$

$$T_{12} = \sqrt{\beta_u(s)\beta^*}\sin(\Delta\mu) \tag{4.82}$$

$$\tag{4.83}$$

In order to maximise the sensitivity, the strength of the final focus triplets are matched in order to satisfy:

1. that the phase advance between IP and the Roman Pot is matched to: $\Delta\mu = \frac{\pi}{2}$
2. the beta at the IP should by maximised.

A special optics has been developed to generate a $\beta^* = 1533$ m at the IP. As visible in equation 4.54, the large β^* does not significantly change the betatron function at the location of the triplets and is, besides the slight phase advance change, very similar to the regular LHC collision optics with respect to orbit control. As seen in equation 4.80, the systematic errors of the scattering angle u'^* measurement are given by the beam stability $u(s)$ at the Roman Pot and the stability of the crossing angle, which due to larger bunch spacing during TOTEM will be zero degree (*head on*). In order to meet their design goals, TOTEM requires ([12]) an orbit stability of $5 - 10\,\mu$m or better at the location of the Roman Pots and, as mentioned above, a stability of the crossing angle of $0.2\,\mu$rad, which translates to a stability of about $2 - 5\,\mu$m at the location of the warm BPMs located at a distance of 20 m closest to the IP.

4.5.4 Energy Stabilisation Requirements

To minimise RF capture losses of the injected beam, the energy offset between the SPS and the LHC due to b_1 decay and tides should be minimised using the horizontal arc correctors in the LHC. A priori, the control of energy is not urgently required for low intensity beams and early LHC operation. However, it may help to keep capture losses below an acceptable limit and to minimise possible population of

the longitudinal 'abort gap' that is required for the extraction kicker in order to reach its nominal deflection. Since it would simplify the setup of nominal beam after commissioning the capture of pilot bunches, control of energy should be performed at an early stage. Once the control loop is implemented, maintenance of nominal stability of $\frac{\Delta p}{p} < 10^{-4}$ is desired [112]. In addition to the minimisation of capture losses, stabilisation of the energy during injection would reduce and minimise the expected perturbations of the tune due to the momentum tracking error. As visible in equations 2.26 and 2.27, a momentum shift in the order of about 10^{-4} would yield a tune shift of about $\Delta Q \approx -0.014$. For comparison: for nominal operation, a tune stability better than 0.003 and 0.001 during injection and collision, respectively, is required [48, 49]. A minimisation would allow reduced gains for the tune feedback and possibly reduce its criticality.

Tides, thermal expansion of the tunnel walls, and other effects may change the physical circumference of the machine and move the orbit with respect to the magnets' centre. As presented above, the expected drifts are less than about 100 μm. Comparing with the dispersion function of about 0.4 m and required orbit stability of about 20 μm at the location of the primary betatron cleaning collimators, the corresponding energy change has to be controlled at the level of about $\frac{\Delta p}{p} \approx 10^{-5}$, as visible in equation 2.23. One may choose to compensate for this off-centre orbit using radial steering which adjusts the frequency of the RF systems, to re-centre the orbit and optimise the machine aperture.

4.6 Summary

There are various sources that may perturb the orbit and energy of the two LHC beams. These effects usually range from rather slow environmental effects such ground motion and tides to faster machine-inherent effect such as the final focus induced orbit change. The perturbation sources, their timescale and the machine phase in which they occur are summarised in Table 4.8. Ground motion affects the

Perturbation Source	**Orbit r.m.s.**	$\mathbf{\lvert \Delta x/\Delta t \rvert_{max}}$	$\mathbf{\Delta p/p}$	**Phase**
	$[\mu m]$	$[\mu m/s]$	$[10^{-4}]$	
Random Ground Motion	$(200-300)$	< 0.01	$8 \cdot 10^{-3}$	all
Tides (max/min)	$+100/-170$	< 0.01	$+0.5/-0.9$	all
Thermal Girder Expansion	$(9.5 \ldots 16)/^{O}C$	$< 10^{-3}/^{O}C$	-	all
Cryostat vibration	unknown	-	-	all
Decay	530	< 0.5		injection
Snapback	530	< 15		start ramp
Eddy currents	129	< 0.3	-1	ramp
Persistent currents	340	< 0.2	-9	ramp
Ramp total	600-700	< 15	8	ramp
β^* squeeze[1]	< 30 mm	< 25	-	squeeze
COD power supply ripple	6	noise	-	injection
	0.4	noise	-	collision
COD hysteresis	50	static	0.2	first injection

Table 4.8: Summary of expected orbit and energy perturbations. As described in the text, the value of the orbit perturbations as a spread of about 30 %. [1] Based on a 0.5 mm r.m.s. quadrupole misalignment and no orbit correction.

beam orbit during all operational phases and the presented propagation models are both qualitatively and quantitatively supported by geophone and beam motion measurements in the SPS and LEP. These confirm that ground motion contribution to the orbit can be neglected in the LHC for frequencies above 1 Hz. This conclusion is independent of whether the ground motion is correlated or not, and assumes that the girders supporting the magnets are sufficiently stiff. The actual vibration level of the LHC superconducting magnets remains to be investigated. For frequencies below 1 Hz, orbit drifts are dominated by random ground motion and are expected to reach collimation system tolerance after about 10 hours for the injection optics and after less than 0.5 hours for the collision optics, under the pessimistic assumption that the orbit feedback is not running. The injection orbit is estimated to drift about 0.3-0.4 σ from fill-to-fill, assuming a fill about every 15 hours. This may slightly reduce the available aperture during the first injection into an empty machine after a nominal physics cycle.

Lunar and solar tides are known to periodically change machine circumference. While the RF system maintains the same path length of the beam, the mismatch between the physical machine and beam circumference results in a momentum mismatch that can reach up to $\Delta p/p \approx (+5/-8) \cdot 10^{-5}$ depending

on the alignment of the earth, moon and sun. The corresponding dispersion orbit at a location with a 2 m dispersion function is about $(+110/-170)\,\mu$m. The peak-to-peak change is slow and occurs within 6 hours. In contrast to quadrupole-induced dipole errors, it is more efficient to compensate this effect through the LHC RF frequency than through the horizontal orbit correctors. However, during injection, a constant RF frequency is mandatory. Thus the momentum of the LHC will be adjusted through the horizontal orbit corrector magnets.

Temperature changes may affect beam orbit and its energy through thermal expansion of the supporting magnet girders and tunnel walls, respectively. Due to the small temperature gradient between neighbouring cells, the relative orbit perturbation is expected to be small in most parts of the machine. An exception may be in the collimation section in IR7 that is foreseen to be hermetically sealed from the rest of the tunnel. There, the systematic BPM and quadrupole reference may change up to $9.5-16\,\mu$m per ${}^{O}C$ temperature difference between the collimation section with respect to the neighbouring cells outside the sealed region. This mainly affects the long-term stability and fill-to-fill reproducibility of the orbit inside the collimation region. However, present expected temperature gradient estimates are vague, may possibly be overestimated, and only be established once LHC operation with beam has started.

Machine-inherent sources are primarily driven by multipole field errors of the main bending magnets, the change of final focus in the experimental insertions, possibly the flow of cryogenic coolant in the magnets, and other effects. The decay and snap-back phenomenon is driven by current loops in the superconducting alloy that builds up while the magnets are kept at low field during injection and restored to the pre-decay value as soon as the magnet is ramped. Compared to the snap-back, the decay is rather slow with a time constant in the order of 1000 s. Assuming nominal magnet operation, the magnetic field imperfections induced by orbit drifts can reach up to $600-700\,\mu$mr.m.s. during the ramp with maximum drift velocities less than about $15\,\mu$/s, which is within the compensation capabilities of the dipole corrector magnets. The largest and fastest dynamic orbit perturbations may be given by the final focus optics squeeze to a $\beta^* = 0.55$ m and may reach up to 30 mm, if uncompensated and assuming a worst-case r.m.s. quadrupole misalignment of 0.5 mm in the experimental insertions. The corresponding drift velocity is less than about $25\,\mu$m/s and – if not further reduced by other means – estimates the required effective feedback bandwidth to be in the order of 2 Hz. Assuming a perfect feed-forward between fill-to-fill, the residual random orbit perturbation that would have to be compensated by feedbacks during squeeze is about $300-600\,\mu$m.

The hysteresis of the superconducting orbit corrector dipoles can be considered as static on the targeted feedback time scales and mainly affects the fill-to-fill injection orbit and momentum stability. The expected injection r.m.s. orbit uncertainty is about $50\,\mu$m and the systematic momentum shift of $\Delta p/p \approx 2 \cdot 10^{-5}$. Both effects are mitigated and minimised through feedback operation with the initial pilot or low-intensity beam prior to the injection of the first high intensity beam. The minimal orbit stability during injection as a result of the orbit corrector power converter ripple is about $5\,\mu$m. The effect of the dipole corrector power converter ripple can be neglected while operating at a particle momentum of 7 TeV.

The orbit stability requirements on the two LHC beams can be grouped into: LHC cleaning system, machine protection, active beam instrumentation, and issues related to the accelerator performance, summarised in Table 4.9.

As a novelty and in contrast to previous and present hadron accelerators, the LHC has constraints on the orbit stability that are similar to those found at modern light sources [77, 113]. The required stability ranges from $0.5\,\sigma$, with σ being the beam size, which is about 1 mm during injection and about $300\,\mu$m during collision at a maximum $\beta \approx 180$ m in the arc, down to about $5\,\mu$m at the location of the Roman Pots in IR5 and IR1 during TOTEM operation. Most of the required stabilities are easily exceeded without beam control by the presented orbit perturbation sources. The most tight tolerances are found in the insertion regions. A distinction between global and local orbit stability requirements is less obvious. Essentially, for nominal operation, excluding TOTEM operation, a global stabilisation of the orbit at the level of better than about $0.3\,\sigma$ is required, corresponding to a global orbit stability of about $30-90\,\mu$m r.m.s. in the LHC.

System	Tolerance	Region
LHC Cleaning System:		
Orbit alignment at the collimator jaws	$\sim 0.15\,\sigma$	IR3, IR6
Machine Protection & Absorber:		
TCDQ (protection against asynchronous beam dumps)	$< 0.5\,\sigma$	IR6
Injection collimator & absorber	$< 0.3\,\sigma$	IR2, IR8
Tertiary collimator for collisions	$< 0.2\,\sigma$	IR1, IR5
Injection arc aperture w.r.t. protection devices:	$0.3 - 0.5\,\sigma$	global
Active systems:		
Transverse damper	$\sim 200\,\mu$m	IR4
Beam interlock BPM	$\sim 200\,\mu$m	IR6
Injection Beam Momentum for RF capture	$\frac{\Delta p}{p} < 10^{-4}$	global
Performance:		
Stability of collision points	minimise drifts	IR1,2,5,8
TOTEM/Atlas Roman Pots	$\sim 5 - 10\,\mu$m	IR1, IR5
Reduction of higher multipole feed-down effects	$300\,\mu m r.m.s.$	global
Maintain beam on cleaned surface (e-cloud)	$1\,\sigma$	global

Table 4.9: LHC orbit and energy stabilisation requirements: The magnitude of requirements are similar; a distinction between local and global requirements is less evident in the LHC. The for the collimation system given stability requirement is given with respect to the individual collimator.

Chapter 5

Feedback Sub-System Analysis

In order to provide a sufficient stability of key beam parameters such as orbit, energy, tune, coupling and other observables in the presence of machine uncertainties, semi-automated accelerator operation and feedback control systems rely on beam based measurements of these observables. The quality of the orbit control is determined by the accuracy, resolution and noise of the involved beam position monitors, orbit corrector magnets and their individual power supplies. This chapter provides an analysis of the subsystems involved in the orbit and energy feedback: beam position monitors, orbit corrector magnets, controls and communication infrastructure.

5.1 Beam Position Monitors

This section gives a short summary of the electromagnetic pickups principles that are used to measure the average beam position and bunch oscillations in the LHC. Since the position is the most fundamental parameter, these pick-ups are commonly referred to as Beam Position Monitors (BPM). An introduction and overview of beam instrumentation, in particular BPMs, can be found in [114–117].

5.1.1 Position Measurement Principle

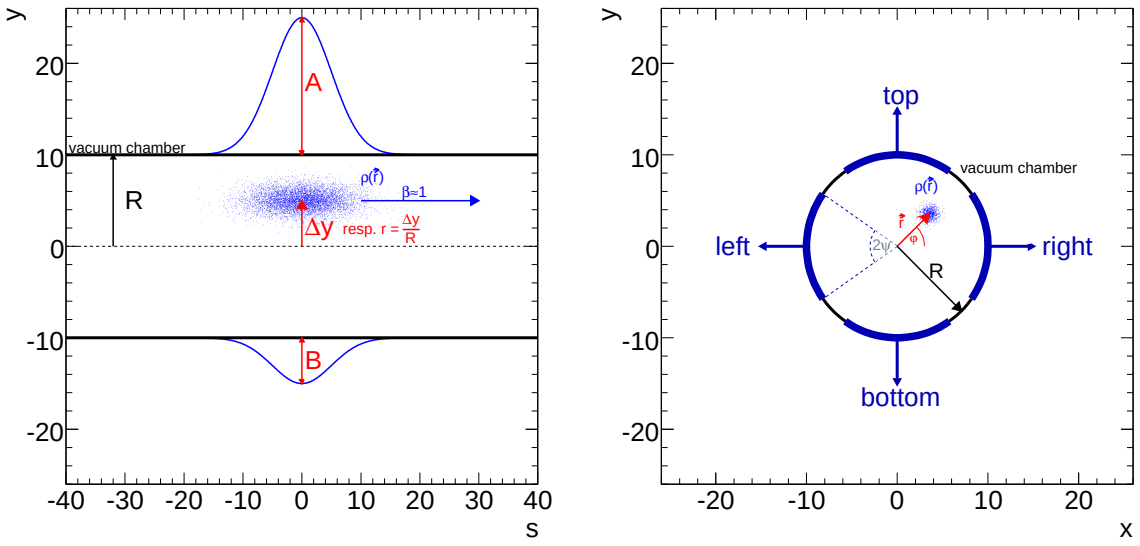
The apparent electromagnetic field distribution changes when observed from a frame of reference that is boosted with a velocity $\vec{\beta} = (\beta_1, \beta_2, \beta_3)^T = \vec{v}/c$ with respect to the initial frame of reference. As shown in Section 2.1.1, the three-dimensional equipotential surface is squeezed into a disc in the transverse plane. Thus the longitudinal image charge density that is travelling along on the vacuum chamber is nearly identical to the longitudinal bunch profile as illustrated in Figure 5.1(a). The distribution of the image charge on the vacuum chamber boundaries usually depends on the beam position and its transverse and longitudinal distribution. For an off-centre beam, the image charge density on a wall segment at an angle θ and vacuum chamber radius R can be written as

$$J_{image}(r, \phi, R, \theta) = \frac{J_{bunch}(r, \phi)}{2\pi R} \frac{1 - \left(\frac{r}{R}\right)^2}{1 + \left(\frac{r}{R}\right)^2 - 2\frac{r}{R}\cos(\theta - \phi)} \tag{5.1}$$

with (r, ϕ) the beam displacement with respect to the centre as illustrated in Figure 5.1(b) and $J_{bunch}(r, \phi)$ the charge distribution. The current travelling through a given wall segment, as indicated in Figure 5.1(b), is given by the integral

$$I_{segment} = \int_{-\psi}^{+\psi} J_{image}(r, \phi, R, \theta) \mathrm{d}\theta \tag{5.2}$$

with ψ the opening angle of the selected wall segment, (r, ϕ) the coordinate of the charge and (R, θ) the point of observation located on the vacuum chamber wall with radius R. As shown in [117], equation 5.1



(a) longitudinal pickup cross section

(b) transverse pickup cross section

Figure 5.1: Electromagnetic pickup principle: The longitudinal (left) and transverse (right) cross section are shown. For relativistic particles the longitudinal bunch structure to first order is projected onto the image charge current on the vacuum chamber. The buttons are isolated from the vacuum chamber.

can be expanded into:

$$J_{image}(r,\phi,R,\theta) = \frac{J_{bunch}(r,\phi)}{2\pi R}\left[1+\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^2\cos n(\theta-\phi)\right] \tag{5.3}$$

Inserting this expansion into equation 5.2 and integrating over the 'left' and 'right' button area indicated in Figure 5.1 leads to the following currents flowing through the 'left' (I_L), 'right' (I_R) segment respectively

$$I_L = \frac{I_w}{2\pi}\left[2\psi + 2\frac{x_0}{R}\sin\psi + \frac{x_0^2-y_0^2}{R^2}\sin(2\psi) + \text{h.o.}\right] \tag{5.4}$$

$$I_R = \frac{I_w}{2\pi}\left[2\psi - 2\frac{x_0}{R}\sin\psi + \frac{x_0^2-y_0^2}{R^2}\sin(2\psi) + \text{h.o.}\right] \tag{5.5}$$

with I_w is the longitudinal bunch current density, x_0 the horizontal and y_0 the vertical bunch position in Cartesian coordinates with respect to the BPM centre. The corresponding terms for the 'top' (I_T) and 'bottom' (I_B) segments are similar, with exception of the swapped x and y coordinates.

Early beam position monitor designs were based on segmented vacuum chambers that directly measured the image current with many ohmic resistances and sensitive voltmeters. However, these types of monitors have shown to be prone to noise due to parasitic currents of power supplies, vacuum pumps and other sources conducted by the vacuum chamber in addition to the beam-induced image current. Thus, it is common nowadays to use capacitive monitors that have isolated chamber segments. The induced voltage V on these monitors is given by

$$V = \frac{Q}{C} \tag{5.6}$$

with Q the charge on the electrode and C its capacitance, which usually depends on its shape and length. One distinguishes two cases depending on whether the bunch is longer or shorter than the electrode. In the first case, the electrodes become fully charged during the bunch passage. The bunch current is differentiated with the electrode capacitance C as differentiation constant. The button beam position monitors are the most common type of this monitor class. For the second case, where the bunch is much shorter than the electrode, the signal corresponds to the longitudinal bunch profile. The second case is applied in strip-line beam position monitors that are isolated conductors parallel to the beam. With strip-line monitors, it is possible to retrieve the beam signal either at the up-stream or down-stream end of the strip-line. The output signal depends on the type of termination that affects the reflection of signals. If the termination is matched to the strip-line impedance only the down-stream port show signals. In the LHC, for example, this discrimination is used by the monitors in the experimental insertions, where the

two beams share a common vacuum chamber in order to distinguish between beam 1 and counter-rotating beam 2.

Independent of their time structure, it is possible to define the sum- ($\sum$) and difference (Δ) signal as

$$\Delta := (I_L - I_R) \cdot Z_0 \qquad \text{and} \qquad \sum := (I_L + I_R) \cdot Z_0 \tag{5.7}$$

with I_L and I_R the current at the left and right, respectively, and Z_0 the impedance of the electrode. Inserting equations 5.4 and 5.5 in above definition yields:

$$\Delta = \frac{2 i_w Z_0}{\pi R} (\sin \psi) z + \dots \tag{5.8}$$

with z either the horizontal or vertical coordinate, depending on the context. The Δ signal is proportional to the position z and the bunch current (intensity) i_w inside the BPM. The 'sum' signal $\sum$ is independent of the beam position and thus usually used to normalise the difference signal in order to make it insensitive to the actual bunch current. The normalised position z can thus be written as

$$z_{norm} = \frac{\Delta}{\sum} = \frac{I_L - I_R}{I_L + I_R} \tag{5.9}$$

The geometric sensitivity S of a beam position monitor is defined as:

$$S := \frac{\sum}{i_w z} = \frac{2 Z_0}{\pi R} (\sin \psi) + \dots \tag{5.10}$$

To enhance the sensitivity, half-aperture R should be small and the button opening angles should be large, up to $\psi = \pi/2$. In order to compare the performance of BPMs with different apertures it is common to normalise the position with respect to the BPM half-aperture R:

$$z_{norm} = \frac{z}{R} \tag{5.11}$$

5.1.2 LHC Beam Position Monitors

In the LHC nearly every quadrupole is equipped with a BPM, resulting in 1056 beam position monitors dedicated to the control of orbit and circumference. Up to 18 BPMs are grouped in VME front-end crates that contain the BPM acquisition electronics, a timing card and a Power-PC computer. There are 8 BPM front-ends per IR, yielding a total of 64 BPM front-ends in the LHC.

The majority (≈ 900) of LHC BPMs are located in regions where the beams are travelling in separated chambers. Such BPMs are made of buttons inserted into the vacuum chamber as shown in Figure 5.2. The BPM half-aperture is about 22 mm. A button opening angle of close to $\psi = \pi/2$ would maximise the geometric sensitivity ($\sin(\psi) \approx 1$) and also further minimises the coupling between horizontal and vertical beam position ($\sin(2\psi) \approx 0$). However in order to be able to measure in both planes, the vacuum chamber is segmented into four buttons with an half-opening angle of $\psi = \pi/4$. Most LHC BPMs are connected either directly or indirectly through the orbit dipoles to the quadrupoles and are located inside the same cryostat.

The sections where the two LHC beams share the same vacuum chamber are equipped with strip-line BPMs in order to distinguish between the clock-wise and counter-clock-wise rotating beam. The strip-lines have a length of 120 mm, which is much shorter than the LHC bunch length of about 300-600 mm. Hence, the strip-line BPM is similar to the button BPM response and the same acquisition electronic can be used. If 'U' is denoting the up-stream and 'D' the down-stream port of the strip-line BPM with respect to Beam 1, then the acquisition electronic measuring Beam 1 is connected to 'U' and the acquisition electronic measuring Beam 2 to 'D', respectively.

5.1.3 Wide Band Time Normaliser Principle

During the different modes of operation, the stored beam charge extends from $\approx 5 \cdot 10^9$ proton charges ($e_p \approx 1.6 \cdot 10^{-19} As$), for a single pilot bunch used for commissioning and check-out, up to about $5 \cdot 10^{13}$ charges spread over 2808 bunches. The corresponding beam currents range from about 9 μA for a single circulating pilot bunch up to 0.53 A for a nominal LHC beam [108, 118].

Classic BPM electronics often average over a given number of turns and all bunches inside the accelerator, and use the ratio between the difference Δ and sum $\sum$ signal, as shown in equation 5.9, in order

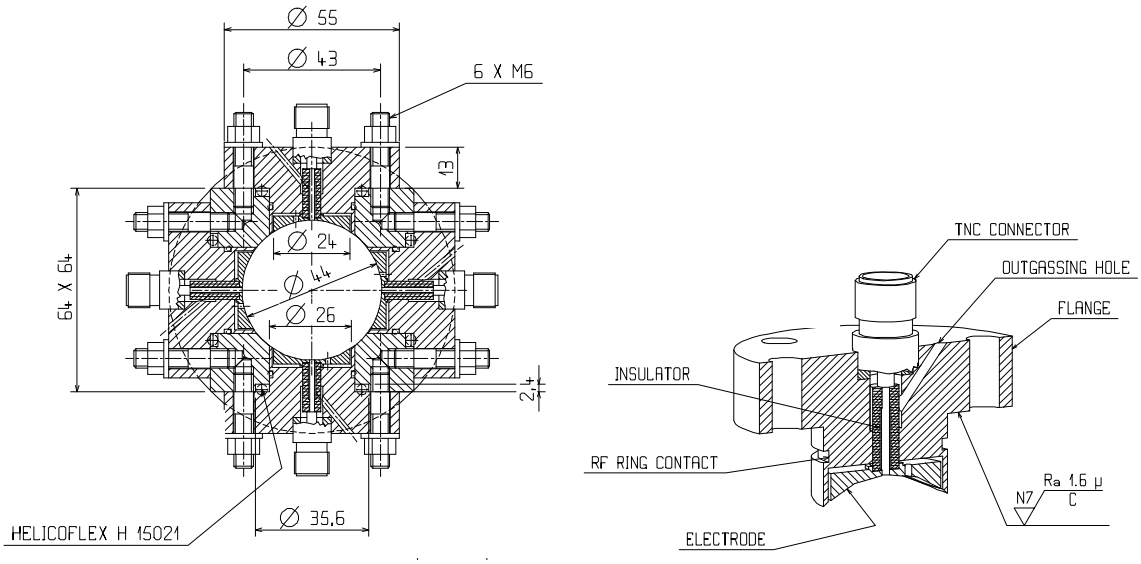


Figure 5.2: Mechanical LHC BPM cross section (left) and BPM button detail (right). It is visible that the button size is optimised in order to almost completely cover the chamber circumference.

Precision goal	**Coarse** (pilot bunch)	**High** (other beams)
Scale error	-	$\pm 4\%$
Offset	$\pm 750\,\mu$m	$\pm 100\,\mu$m (relative offset $< \pm 30\,\mu$m in IR's)
Non-linearity	-	$\pm 200\,\mu$m over ± 4 mm
Resolution	$200\,\mu$m r.m.s.	$\pm 50\,\mu$m r.m.s. (traj.), $5\,\mu$m r.m.s. (orbit)

Table 5.1: Specified BPM accuracy [118].

to estimate the beam position. In the LHC, a similar scheme would require a dynamic input range of up to four orders of magnitude for the BPM acquisition electronics, corresponding to about 80 dB, which is quite demanding with respect to the required precision. The specified accuracy is summarised in table 5.1.

In order to meet the precision requirements, the bunch-by-bunch resolution, and to be more flexible and independent with respect to the actual number and bunch filling pattern, the LHC BPM electronic measures the position of each individual bunch on a per turn basis using a wide-band-time-normaliser circuit (WBTN). The WBTN principle is described in [68]. For flexibility and robustness with respect to radiation, the BPM signal processing is split into two parts: a robust analogue WBTN circuit connected to and located close to the BPM, and a digital processing logic located at the surface. Both parts are connected through optical fibre. The digital processing logic is controlled by a front-end computer [119].

The analogue front-end circuit converts the individual bunch position into a time-difference signal that is encoded into two laser pulses and transmitted through optical fibre to the receiver card. The receiver card measures the time delay using an integrator circuit and inverts the WBTN process using third order polynomial to a normalised beam position. The block diagram of the analogue circuit is shown in Figure 5.3. The WBTN principle is applied to both BPM planes individually. The differentiated left and right (or top and bottom) signals, respectively, are referred to as $A(t)$ and $B(t)$. For the time-normalisation, signal B is delayed by τ_1 and added to the signal of button A and vice versa, yielding the normalised signals A' and B':

$$\begin{aligned} A' &= A(t) + B(t+\tau_1) \\ B' &= A(t+\tau_1) + B(t) \end{aligned} \tag{5.12}$$

The time-normalised signal $\Delta\tau$ is given by the time difference of the individual zero-crossings t_A and t_B of A' and B', respectively.

$$\Delta\tau = t_B - t_A \qquad \text{with} \qquad A'(t_A) = 0 \wedge B'(t_B) = 0 \tag{5.13}$$

For causality reasons, $\Delta\tau$ has always to be positive, whether the beam has a positive or a negative positions inside the BPM. Thus the output signal of B is delayed by an additional lag τ_2 prior to the

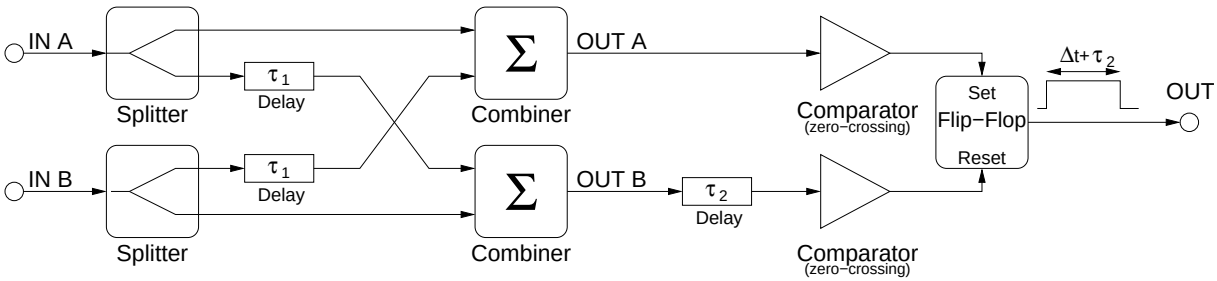


Figure 5.3: Schematic Wide-Band-Time-Normaliser (WBTN) block diagram that is used by the LHC BPM system. The output signal of B is delayed by an additional lag τ_2 in order to account for positive as well as negative positions (w.r.t. the BPM centre) prior to detection of the zero-crossing time difference of both signals.

detection of the zero-crossing. The additional delay is subtracted after digitisation inside the receiver card. The corresponding WBTN signals are shown in Figure 5.4. The position can be obtained from the

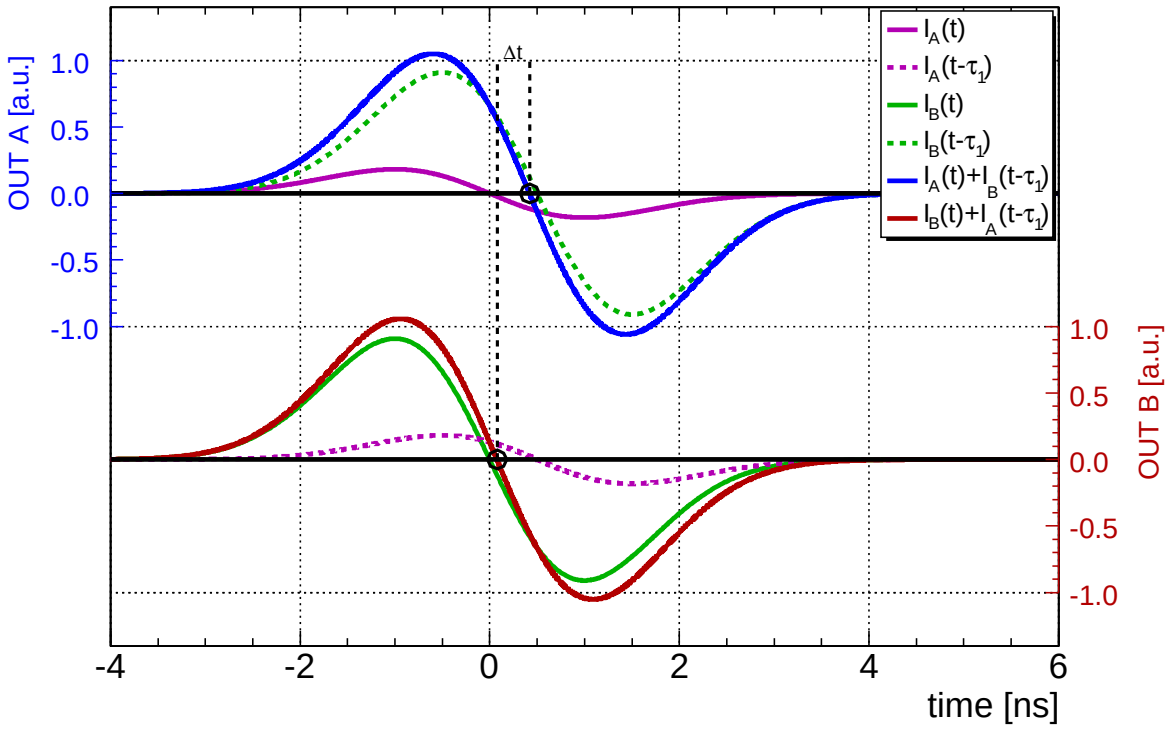


Figure 5.4: Schematic Wide-Band-Time-Normaliser signals.

time-difference signal $\Delta\tau$ through a third order polynomial transfer function. The polynomial coefficients depend on the geometry (half-aperture) of the BPM and must be determined by calibration for each individual circuit.

Proof of Equivalence: In order to study the BPM dependence on certain beam parameters and to relate the WBTN principle to a traditional position measurement using the ratio between the BPM difference Δ and sum $\sum$ signal, the following approximative model can be used for beam position close to the BPM centre. The static bias can be neglected, since it is the same for both buttons. Further ignoring the higher order terms, the maximum signal amplitude on the buttons is proportional to the position z_{norm} normalised to the half-aperture. The time structure of the button signal is given by the longitudinal bunch structure $\varrho_b(t)$ which, in first order, can be approximated by the following cosine distribution

$$\varrho_b(t) \cong q_b \cdot \frac{2}{B}\left[\cos(\pi\frac{t}{B})\right]^2 \tag{5.14}$$

with B the finite longitudinal bunch length and q_b bunch charge. Since in the LHC the bunches are much longer than the BPM buttons, the button signals are proportional to the derivation of the longitudinal bunch profile and can (using equations 5.4 and 5.5) be written as

$$\begin{aligned} A(t) &\cong \frac{\psi+\tilde{z}}{\pi}\cdot C\cdot\frac{\partial}{\partial t}\varrho_b(t) \\ &\cong -4(\psi+\tilde{z})\cdot\frac{q_bC}{B^2}\cos(\pi\frac{t}{B})\sin(\pi\frac{t}{B}) \qquad \text{and} \qquad (5.15) \\ B(t) &\cong -4(\psi-\tilde{z})\cdot\frac{q_bC}{B^2}\cos(\pi\frac{t}{B})\sin(\pi\frac{t}{B}) \qquad (5.16) \end{aligned}$$

with C the button capacitance, $\tilde{z}$ the normalised beam position. The expression for top-button button pair is similar. Since we are interested in zero-crossing detection, equations 5.15 and 5.16 can be expanded around zero and approximated to:

$$\begin{aligned} A(t) &\cong -4\pi(\psi+\tilde{z})\cdot\frac{q_bC}{B^3}\cdot t = -A_0\cdot t \qquad (5.17) \\ B(t) &\cong -4\pi(\psi-\tilde{z})\cdot\frac{q_bC}{B^3}\cdot t = -B_0\cdot t \qquad (5.18) \end{aligned}$$

Using equations 5.17, 5.18, the normalisation relation given in equation 5.12 and perfect zero-crossing detection yields

$$\begin{aligned} A' = A(t)+B(t+\tau_1) &\approx -A_0\cdot t - B_0\cdot(t+\tau_A)\stackrel{!}{=}0 \implies t_A = -\frac{B_0}{A_0+B_0}\cdot\tau_1 \\ B' = A(t+\tau_1)+B(t) &\approx -A_0\cdot(t+\tau_B) - B_0\cdot t\stackrel{!}{=}0 \implies t_B = -\frac{A_0}{A_0+B_0}\cdot\tau_1 \end{aligned} \qquad (5.19)$$

Inserting into the time difference definition (equation 5.13)

$$\Delta\tau = t_B - t_A = \tau_1\frac{A_0-B_0}{A_0+B_0}\sim z_{norm} \qquad (5.20)$$

and comparing equation 5.9 in this first order model establishes the equivalence between the time normalisation and sum-over-difference measurement principle.

5.1.4 Beam Position Monitor Errors, Failures and their Detection

Measurement errors and equipment failures are part of real-life accelerator operation. The reliability and availability of the LHC BPMs, for example, have an important impact on beam operation and machine performance, not only since they are used to measure beam properties but also they are used to drive the orbit feedback system and thus the stabilisation of the two LHC beams. One can distinguish between BPM 'errors' and 'failures': In case of an 'error', the beam position is acquired but with an inconsistency between measured and true beam position. A detected BPM error may either lead to a correction, for example through calibration and follow-up adjustment of, for instance, the BPM scale and offset, or may cause the transition to the 'fault' or 'failure' state. A 'failure' refers to a BPM error exceeding specified limits (see Table 5.1) or to the unavailability of the measurement. The criticality of BPM failures and errors depends on a number of factors, in particular their total number, spatial distribution and the correction strategy that is discussed in Section 6.1.4.

Possible sources for systematic BPM errors are the WBTN dependence on bunch length and intensity, discussed in the next section, analogue noise of the WBTN or digitisation noise of the receiver card ADC.

Systematic Bunch Length and Intensity Dependence

The strength of the WBTN circuit is that it can be run self-triggered on the first zero-crossing detection and is thus robust with respect to the synchronisation of bunch arrival, position acquisition trigger and RF frequency. For perfect zero-crossing detection that is identical for both WBTN channels, the circuit is independent of the bunch length and bunch intensity and thus would require only one gain setting for all beam types. This is a clear advantage compared to other BPM acquisition systems that require multiple gain settings and complex calibration procedures.

Unfortunately, reflections and other type of noise on the buttons and strip-lines are inevitable. The reflections correspond typically to 1% of the maximum bunch signal. Compared to the expected bunch

intensities, a nominal bunch produces a reflection signal that has an amplitude similar to a single pilot bunch and would thus produce a *ghost bunch* acquisition.

In order to minimise spurious triggers on reflections and noise, the WBTN uses a comparator threshold level S_c which is set above the expected noise level for the zero-crossing detection which may slightly differ for the two WBTN channels. It was decided to implement two modes of BPM operation: a 'HIGH' sensitivity' comparator setting for low intensity beams up to about $5 \cdot 10^{10}$ protons per bunch, such as the pilot and ion bunches; and a 'LOW' sensitivity setting for beams with larger bunch intensities. The schematic comparator zero-crossing detection is illustrated in Figure 5.5. The difference between the

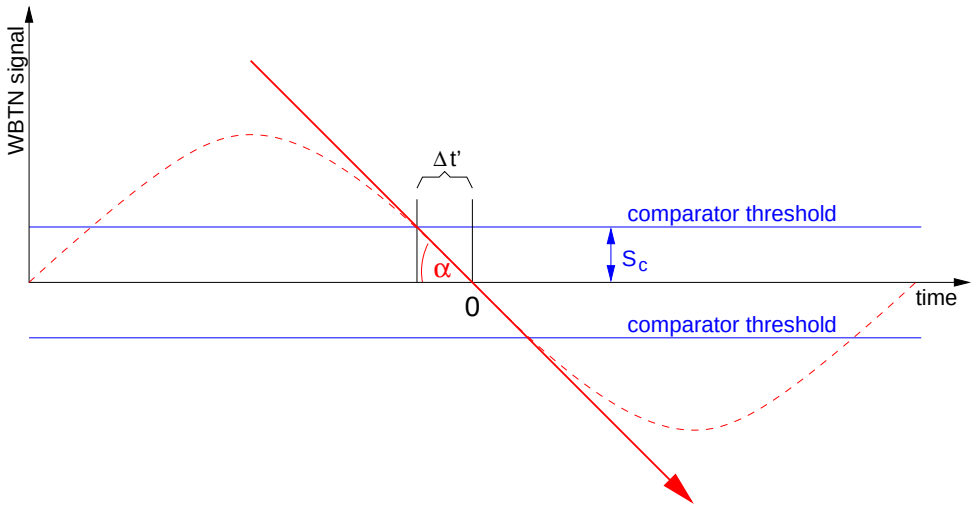


Figure 5.5: Schematic zero-crossing detection of the wide-band-time-normaliser signal using a comparator with threshold S_c. The trigger time error $\Delta t'$ between the detected and actual zero-crossing event is indicated. The slope α depends on bunch length and intensity. The systematic trigger error $\Delta t'$ increases with decreasing intensity and decreasing bunch length.

measured and the actual zero-crossing $\Delta t'$ depends on the threshold level S_c and on the gradient of the button signal around the zero-crossing. As visible in equation 5.17 and 5.18, the gradient depends on total bunch charge and bunch length. If the threshold levels differ between the buttons, the triggering error $\Delta t'$ is thus given by

$$\tan(\alpha) \sim \frac{I_0}{B^3} \sim \frac{S_c}{\Delta t'} \qquad \Longleftrightarrow \qquad \Delta t' \sim S_c \cdot \frac{\sigma_b^3}{I_0} \tag{5.21}$$

with the r.m.s. bunch width $\sigma_b \approx \frac{B}{4}$. Using equation 5.20, the systematic position measurement error Δx_{error} is proportional to:

$$\Delta x_{error} \sim S_C \cdot \frac{B^3}{q_b} \tag{5.22}$$

The sensitivity to the bunch length and intensity was considered during the WBTN design stage: for that reason two comparator thresholds and a filter were introduced. The filter is installed at the input stage of the WBTN circuit and has a bandwidth of about 167 MHz [68]. Qualitatively, the filter "increases" the bunch length of short bunches, as seen by the WBTN, to a r.m.s. bunch length of $\sigma_f \approx 6$ ns which corresponds to the filter bandwidth. It leaves the signal of larger bunches unaffected. Using a Gaussian bunch approximation, the effective bunch length that is seen by the WBTN circuit is given by:

$$\sigma_{eff} = \sqrt{\sigma_b^2 + \sigma_f^2} \tag{5.23}$$

with $\sigma_b \approx 4 \cdot B$. For very short bunches, the time dependence of the button signal converges to the signal given for bunches with a length σ_f.

The sensitivity to bunch intensity and bunch length were measured using six BPMs and LHC beams in the SPS. The BPMs are equipped with the same acquisition electronics and pre-processing chain that is foreseen for the LHC. The SPS scraper was used to test the systematic error on the bunch intensity. The scraper consists of a block of copper that can be moved into the path of part or the whole beam. Particles that encounter the block are scattered and eventually lost at the aperture. Figure 5.6 shows the measured

vertical orbit sampled at a LHC prototype monitor and the average bunch intensity as a function of the SPS cycle time. The beam consisted of 24 bunches at a momentum of 26 GeV. The measured position shift at 14 s coincides with the moving in of the scraper device. The position shift of about 100 μm corresponds to about 0.2% of the BPM half-aperture (41.5 mm). Though this measurement established

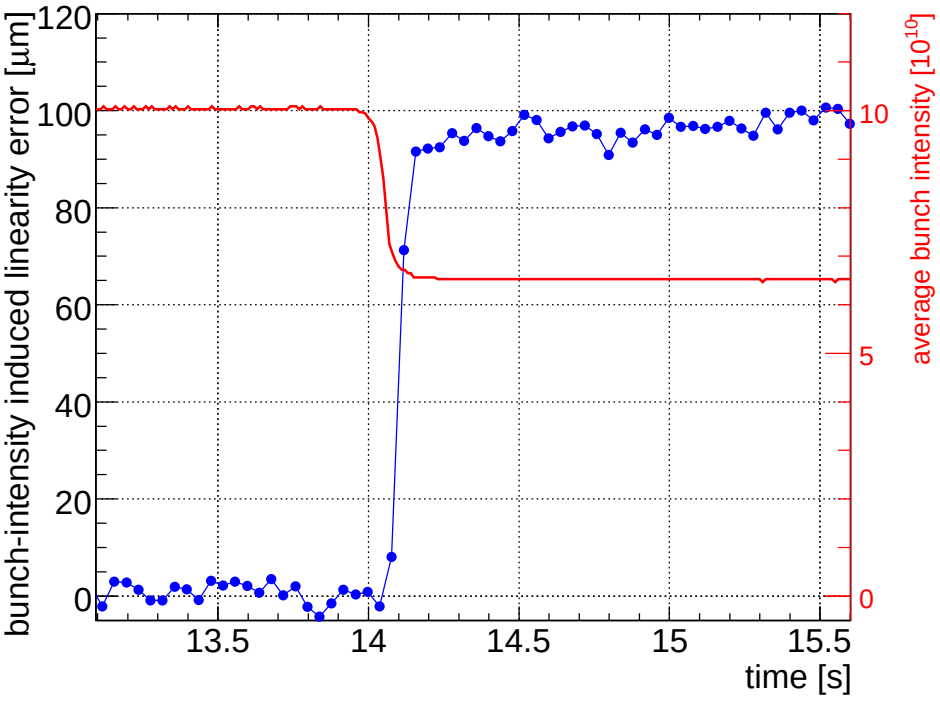


Figure 5.6: WBTN sensitivity to varying bunch intensity. The measurement shift at about 14 s coincides with the scraper movement that reduced the average bunch intensity.

the qualitative dependence on the bunch intensity, the bunch intensity change using the scraper showed to be too coarse and further resulted in a large non-uniform spread of the 24 individual bunch intensities. Figure 5.7 shows the WBTN dependence on the bunch intensity. The measured data points are based on test-bench measurements using generated signals that are similar to beam signals and that were injected into the WBTN circuit (measurement data courtesy R. Jones). An important transition error occurs when the bunch intensity drops below the selected sensitivity range. The button signal may then be smaller than the fixed comparator threshold and thus may no longer generate acquisition triggers. Near the transition, or for beams with large bunch-to-bunch intensity spreads, this may lead to missing or erratic acquisition triggers. This transition is expected to occur for nominal operation once the bunch intensities drop from the nominal value of $1.15 \cdot 10^{11}$ below about $5 \cdot 10^{10}$ protons per bunch. To limit the effect of erratic or missing triggers, the comparator level is changed from 'LOW' to 'HIGH' for all BPMs before the threshold is reached. Related to the change of threshold is an associated systematic shift of the measured position. However, since the threshold change is known in advance, this error can be "anticipated". It is possible to determine the shift if the orbit is sufficiently stable using the measured orbit before and after the threshold change. To minimise cross-talk, the orbit feedback should be paused for the calibration and while switching thresholds.

The WBTN dependence on the bunch length was also measured with beam in the SPS. The bunch length was adiabatically modulated using the total RF voltage between 0.5 and 7 MV. As given by equation 2.63, the change of bunch length is, in first order, proportional to the square-root of the RF voltage. Figure 5.8 shows the measured orbit error as a function of time. The measurement error is given by the difference between the closed orbit with RF voltage modulation and the average orbit prior and after the RF modulation. The measured BPM signal follows the RF voltage modulation. Compared to the other voltage excitation, the RF modulation of 7 MV indicates that the excitation was too fast and that the measurement did not reach a steady state. This measurement was thus not taken into account for the analysis. Figure 5.9 shows the measurement error as a function of bunch length change. The non-linearity is most dominant for short bunch lengths. The signs of these non-linearities are fairly independent of the position of the beam inside the BPM. It was however observed that the sign varies from BPM to BPM. The origin of these sign differences are not fully understood, but might be due to BPM cabling

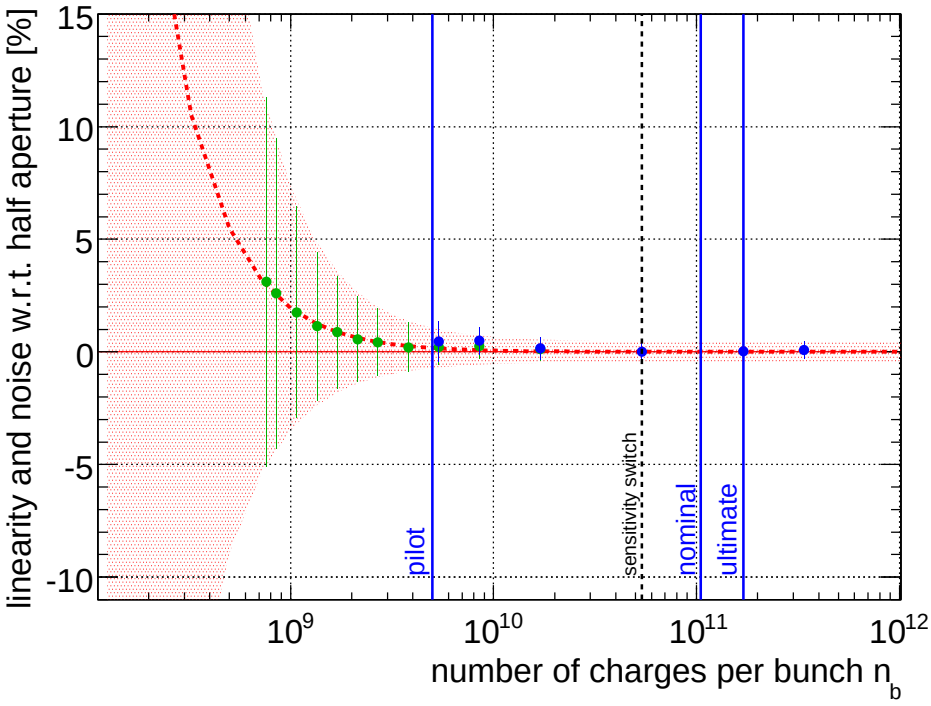


Figure 5.7: Linearity errors with respect to BPM half-aperture plotted versus the bunch intensity. The model prediction (dotted line) and measurements (points) are indicated. The filled area corresponds to the measurement noise. The free model parameter S_c was fit against the measured data points. The systematic shifts (dots) were measured through averaging the random r.m.s. noise of the individual measurement at a given position. Measurement data courtesy R. Jones.

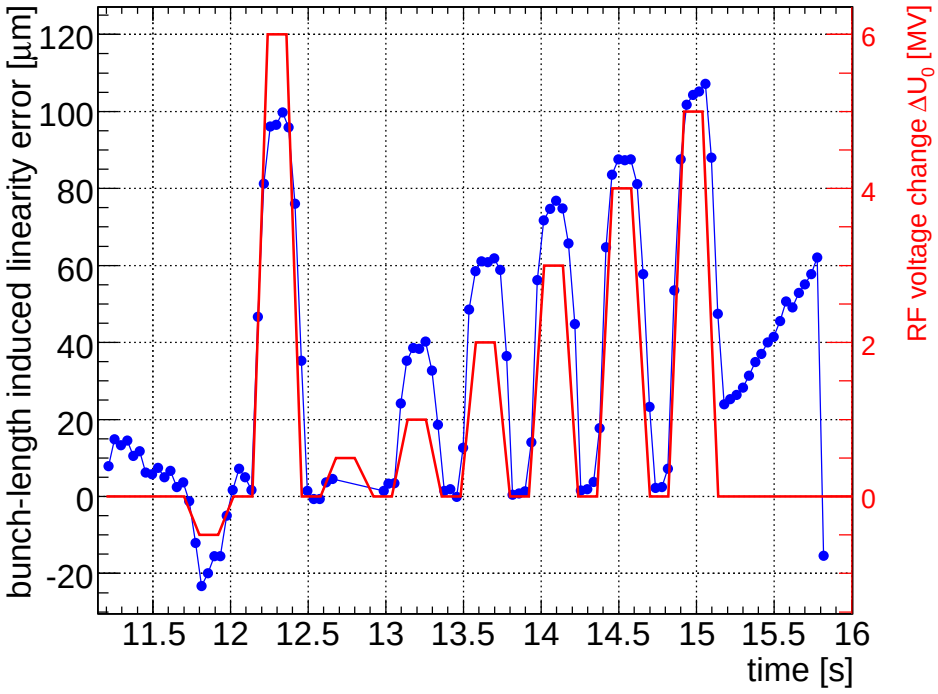


Figure 5.8: WBTN sensitivity to varying bunch length. The bunch length was modulated using the total RF voltage of the SPS cavities.

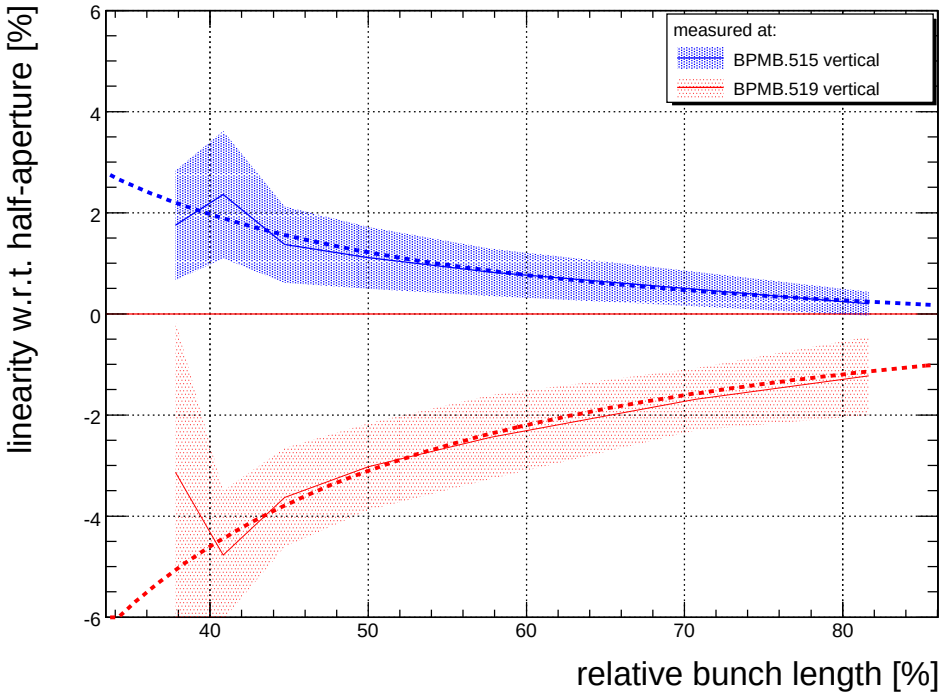


Figure 5.9: Linearity error normalised to the BPM half-aperture versus relative bunch length change. The simple WBTN model fit is superimposed.

errors that were in the end corrected through calibration. For very short bunch lengths, the button signal is determined by the filter bandwidth, see equation 5.23. At 26 GeV in the SPS, LHC bunches have an approximate length of 4 ns which decreases to about 2 ns during acceleration to 450 GeV. Thus the change of bunch length has an impact on the system in the SPS. In the LHC, bunch lengths are expected to be in the order of 1-2 ns, which is much smaller than the filter bandwidth of 6 ns. Hence, for the LHC, it is expected that bunch length induced non-linearities on the position measurement will be less compared to the BPM bunch intensity dependence.

The results agree to first order with the simplified model. A fit to the simulation and measurement data, including larger non-linearities of the WBTN signal for small bunch intensities, shows the following slightly modified systematic BPM measurement error:

$$\Delta x_{error} \sim \frac{\sigma_{eff}^3}{n_b^{1.5}} \tag{5.24}$$

For comparison, a WBTN non-linearity of 1% corresponds to a systematic orbit measurement error of about 220 μm for the LHC BPM half-aperture of 22 mm. As visible in Figure 5.7, the bunch intensity dependence may become important when the sensitivity is changed from 'LOW' to 'HIGH' and for very low intensities of a few 10^9 protons per bunch.

The bunch intensity dependence of 1% with respect to the BPM half-aperture is small compared to the few specialised high-precision BPMs used at LEP [120]. However, it will ultimately define the minimum orbit stability at the collimator jaws due to the systematic position measurement errors. As given in equation 4.71, the available nominal orbit stability budget is about 25 μm corresponding to about 0.1% with respect to the BPM half-aperture. If the collimation system is for example commissioned with nominal bunches ($1.15 \cdot 10^{11}$ protons per bunch), the available position stability margin would be reached, since the orbit feedback intrinsically steers on these systematic measurement errors, as soon as the bunch intensities drop below about $5 \cdot 10^9$ protons per bunch, as visible in Figure 5.7.

A further compensation of this intensity dependence would require additional processing in the WBTN receiver card and knowledge on the individual bunch intensities. For the time being this compensation is not foreseen [118].

Beam Position Measurement Noise

The average orbit, which is the equilibrium beam position, can be obtained through the average position of all (N) individual bunches that are circulating over a given number of turns N_{turns} in the machine.

The closed orbit z_{co} and its resolution σz_{co} thus given by

$$z_{co} = \frac{1}{N \cdot N_{turns}} \sum_{i}^{N} \sum_{j}^{N_{turns}} \hat{z}_{ij} \tag{5.25}$$

$$\sigma z_{co} = \frac{1}{\sqrt{N \cdot N_{turns}}} \sum_{i}^{N} \sum_{j}^{N_{turns}} \sigma(\hat{z}_{ij}) \tag{5.26}$$

with $\hat{z}_{ij}$ the measured position of the i-th bunch and turn j and $\sigma(\hat{z}_{ij})$ the shot-by-shot noise. As shown in Chapter 3 (equation 3.35), for large gains and bandwidths, the closed loop stability is determined by the noise of the exciter and measurement device.

In 2004, long-term measurements were performed with a feedback prototype using six LHC-style BPMs and a 270 GeV coasting beam in the SPS [121]. The LHC prototype feedback system showed good overall performance and could maintain an orbit stability of 2 μm r.m.s (corresponds to about $\frac{2}{1000}\sigma$, with σ the r.m.s. beam size) over the length of one coast as shown in Figure 5.10. The measured beam

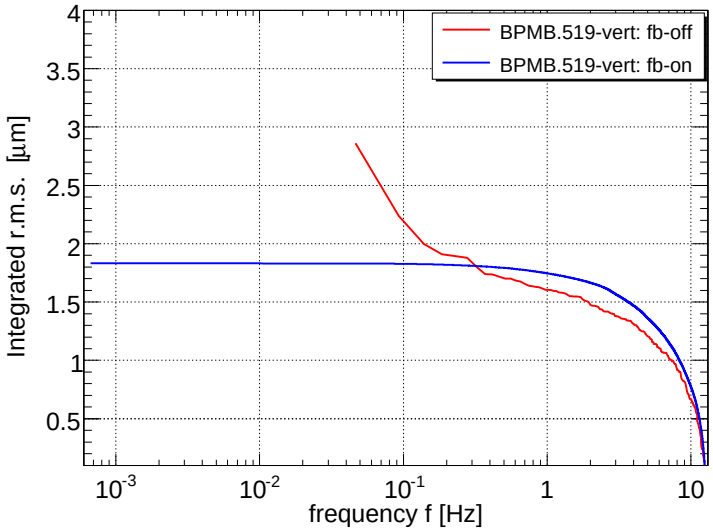


Figure 5.10: Integrated r.m.s. orbit stability with feedback 'on' and 'off' for a 270 GeV beam in the SPS. The stability is limited by the residual 1.8 μm BPM measurement noise. The excitation of the orbit to the BPM noise level for frequencies above the effective orbit feedback bandwidth of 0.3 Hz is visible.

consisted of 16 nominal bunches with about $1.05 \cdot 11$ protons each. The feedback sampling frequency of 25 Hz was used and the effective bandwidth was set to about 0.3 Hz, compared to a maximum possible bandwidth of about 14 Hz given by the SPS orbit corrector magnets. Since the effective bandwidth of the LHC correctors is higher than the tested effective bandwidth, the above results should, to first order, apply to the LHC.

Figure 5.10 (and Figure 4.15) indicates that the SPS prototype system is essentially limited by the residual noise of the WBTN circuit of the BPM system. Each closed orbit measurement is obtained by averaging all bunch positions over 255 turns. This default value corresponds to 20 ms in the LHC. It was chosen to suppress potential 50 Hz noise of the BPMs. From the measurements shown in Figure 5.10 and equation 5.26, the single-turn measurement noise is thus less than 115 μm. The LHC-style BPM has a half-aperture of about 42 mm. For the LHC BPM half-aperture of 22 mm, the resolution scales down to

$$\sigma\hat{z}_{LHC} \approx 61\,\mu\text{m} \tag{5.27}$$

which is only slightly above the specified LHC performance shown in Table 5.1. The test and first order compliance with the specification is considered as an important milestone and proof that the foreseen base-line with respect to BPM and use in feedback are feasible.

BPM Calibration

Systematic BPM scale and offset deviations are inevitable in real accelerator operation. The difference between design trajectory and the local BPM reference system is given by the systematic BPM electronic

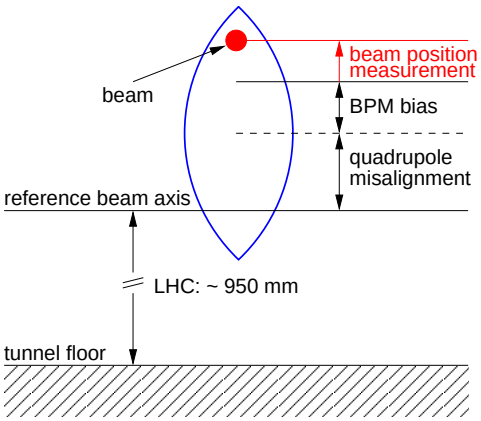


Figure 5.11: Local reference system. Not to scale.

offset (bias) and quadrupole misalignment as illustrated in Figure 5.11: Systematic errors may arise, for example, due to errors of the BPM alignment and survey fiducials [122], systematic uncertainties of the mechanical BPM half-aperture and electric uncertainties such as the presented WBTN dependence on the bunch length and bunch intensity. The uncertainties are in the range of 0.2-0.5 mm r.m.s. The dominant electric uncertainties arise from systematic production errors on the effective button capacitance C and uncertainties of the delay-lines that are required by the WBTN to generate the delays τ_1 and τ_2. As shown in equation 5.21, the button capacitance C affects the angle α of the zero-crossing detection.

A signal generator, prior to the WBTN generating beam a-like signals for the WBTN circuit, can be used to remove or minimise the electronic errors. All LHC WBTN cards are equipped with such a calibration unit. For the calibration, the signal generator creates three reference signals that correspond to a minimum, maximum and zero position reading. The measured deviation is fitted using a third order polynomial. It is foreseen to monitor and, if necessary, re-calibrate the BPMs in situ.

However, such electronics calibrations do not account or compensate for uncertainties of the BPM half-aperture and misalignments of the BPM bodies. Most accelerators apply so-called *beam based* alignments techniques, described in [123–125], that take advantage of knowledge of the lattice or field calibrations of the magnets. The numerous error sources can be grouped into offset and scale errors:

$$x_{true} = \text{'offset'} + \text{'scale'} \cdot \mathrm{x_{meas.}} \tag{5.28}$$

Quadrupole Modulation Quadrupole or 'k-modulation' is a well established beam-based alignment procedures. It is used to measure the offset of the BPM with respect to the quadrupole centre, shown in Figure 5.11. This method exploits the quadrupole feed-down of a beam passing off-centre through the magnet, creating a dipole deflection that is proportional to the off-centre beam position Δz and quadrupole strength k, see equation 2.65. For the k-modulation, the individual quadrupole strength is changed dynamically by a small amount Δk which leads to a change of the quadrupole induced dipole kick and to a small closed orbit oscillation. This oscillation vanishes if the orbit is going through the magnetic centre of the quadrupole. Thus, while modulating the quadrupole strength, the beam position is measured and changed at the BPM of interest, using a closed orbit bump described in the Appendix, until the modulation signal on the orbit vanishes. The corresponding measured orbit position corresponds to the relative offset of the BPM reference system with respect to the quadrupole centre and may be incorporated into the individual BPM calibration.

The advantage of this method is that it is independent of the quadrupole strength or other calibrations that may have large errors. The method is only limited by the resolution of the position monitors, the intrinsic stability of the orbit without k-modulation. The method however requires individually powered quadrupoles.

At the LHC the arc lattice quadrupoles are powered by two circuits per octant whereas most quadrupoles inside the insertions are powered individually. This beam-based method could be applied at the LHC to maximise the aperture and reduce the feed-down effects of the insertion quadrupole magnets – in particular the inner triplet magnets and quadrupoles in the dispersion suppressors in order to reduce the squeeze induced orbit drifts presented in Section 4.4.2.

Cross-Calibration The scale errors may be determined through a cross-calibration of the BPMs against other monitors or orbit corrector dipole magnets. The dipole response that is measured by a

BPM can be decomposed into:

$$\Delta z_i = A_i \cdot R_{ij} \cdot B_j \cdot \delta_j \tag{5.29}$$

with Δz_i the measured beam position at the i-th monitor and δ_j the dipole deflection at the j-th closed orbit dipole. The terms A_i and B_j are calibration factors of the BPM and COD and R_{ij} is the element of the orbit response matrix as given by equation 2.71 to:

$$R_{ij} = \frac{\sqrt{\beta_i \beta_j}}{2 \sin \pi Q} \cos(\mu_{ij} - \pi Q)$$

Assuming a perfect and linear machine optic (R_{ij} is constant), the calibration can be established by individually exciting a sufficient number of dipole correctors and by measuring their response at the BPMs. Assuming m BPMs and n correctors, this measurement yields $m \times n$ equations for $m+n$ unknown calibration factors A_i and B_j which can be written in a least square form and solved using standard linear algebra techniques such as the Singular-Eigenvalue-Decomposition (SVD) described in Section 6.1.3.

A more rigorous and more precise approach includes also a possible momentum deviation, the strengths and misalignments of the quadrupoles and other higher order multipoles that may affect R_{ij}. Optic errors and their correction are beyond the scope of this thesis. A full calibration and optics measurement and correction based on the SPS and its transfer-line TT40 can, for example, be found in [126]. A review of beam-based alignment techniques can be found in [32].

5.1.5 Detection of Single BPM Failures

A BPM error can be classified as 'failure' if the measurement error exceeds, for instance, the targets given in Table 5.1, is un-correlated to the real beam position, or the expected measurement is unavailable. Most of the observed BPM failures arise due to flaws of the WBTN circuit, out-of-bound beam parameter or failures in the BPM front-end computers processing and sending the data. Figure 5.12 shows an example for an orbit acquisition of a 270 GeV beam in the SPS with large BPM failure rate.

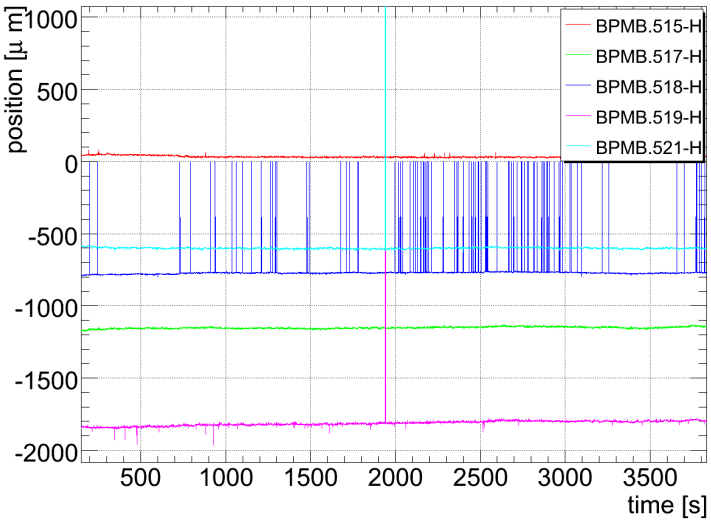


Figure 5.12: Example for erroneous raw orbit acquisition at the LHC BPM acquisition system in the SPS. Spikes ranging from a few μm up to several millimetre are visible. Zero orbit acquisitions correspond to acquisition failures that have been detected by the front-end computer.

The observed temporal BPM failures can be classified into:

- transient or single acquisition failures, also referred to as 'spikes' or 'outliers',
- cascading failures or bursts, which are a consequence of prior failures or errors,
- systematic failure or total unavailability where the beam position cannot be acquired at all.

Most of the observed acquisition failures, as for example illustrated in Figure 5.12, are transient errors. The spikes range from a few μm above the r.m.s. orbit measurement noise level to several mm clearly exceeding the physical BPM aperture. Apart from sensitivity-related gain changes and fast bunch length

changes, no step or systematic changes have been observed with the SPS system. Cascading failures have been observed when the average bunch intensity approaches the minimum intensity of the specified intensity range. These types of failures appear as spikes or single acquisition unavailability with an increasing Mean-Time-Between-Failure (MTBF) rate until the BPM totally fails, assuming the intensity threshold has not been changed.

One source of cascading failures and spikes related to the WBTN can be explained as a result of large intensity spread between bunches, caused by electron cloud [109] and other bunch-to-bunch instabilities: while the first bunches in the bunch train have a relatively high life-time, the life-time of the successive and last bunches is usually poorer as a result of the electron cloud and other effects. As described before, the orbit acquisition consists of an average over all bunches. During operation, bunches at the end of the train may drop, with their intensities below the WBTN comparator threshold, yielding none or erratic triggers, while the first bunches remain above this threshold and yield correct triggers. A change of threshold setting to 'HIGH' would yield better triggering for the low intensity bunches but also cause spurious and erratic triggers on reflections of the high intensity bunches. Thus, this problem can only be solved by minimising the intensity spread between bunches.

Apart from failures that are detected and marked by the BPM front-ends themselves, the most obvious failures can be detected if the measured position exceeds the physical BPM aperture. Further, as shown in Section 4.6, the fastest expected orbit drifts are below about $\frac{\Delta z|_{max}}{\Delta t} \leq 25\,\mu$m/s, hence a measured orbit change faster than this rate may be used to indicate either an erroneous spike or, if occurring consecutively, a step in the response. The orbit is sampled with a period T_s and $z(k)$ is the orbit at the (actual) sampling interval k. The average (windowed) orbit $\bar{z}(k)$ and $\bar{z}_\mu$ are given by:

$$\bar{z}(k) = \frac{1}{n_{avg}} \sum_{i=0}^{n_{avg}} z(k-i) \tag{5.30}$$

$$\bar{z}_\mu(k) = \frac{1}{n_{rms}} \sum_{i=0}^{n_{rms}} z(k-i) \tag{5.31}$$

The two average estimates differ in the number of samples (n_{avg} and n_{rms}). The orbit r.m.s. measured over the last n_{rms} samples is given as

$$\sigma z_{rms}(k) = \frac{1}{n_{rms}} \sum_{i=0}^{n_{rms}} (z(k-i) - \bar{z}_\mu)^2 \tag{5.32}$$

Equations 5.30 and 5.31 essentially differ in the number of samples n_{avg} and n_{rms}, respectively, the orbit is averaged over. The following cuts may be applied to filter erroneous position acquisitions and failures:

1. BPM position outside physical BPM aperture:

$$z(k) > \text{BPM half-aperture} \tag{5.33}$$

2. Transient errors and spikes:

$$|z(k) - \bar{z}(k-1)| > \begin{cases} n_\sigma \cdot \frac{\Delta z|_{max}}{\Delta t} \cdot T_s & \text{if } \frac{\Delta z|_{max}}{\Delta t} \cdot T_s > \sigma z_{rms}(k-1) \\ n_\sigma \cdot \sigma z_{rms}(k-1) & \text{otherwise} \end{cases} \tag{5.34}$$

3. Systematic and step failures:

   The number of spikes classified by equations 5.33 and 5.34 are counted by N_{spikes} while contributions older than n samples are removed from the sliding sum. The presence of a systematic BPM offset error may be indicated if the sliding sum

$$N_{spike} > 2 \tag{5.35}$$

As discussed in Chapter 4, the largest expected orbit drift velocities are below about $25\,\mu$m/s. Assuming a sampling of 10 Hz, the maximum expected drift over one sample $\frac{\Delta z|_{max}}{\Delta t} \cdot T_s$ is about $2.5\,\mu$m which is slightly above the expected measurement r.m.s. shown in Figure 5.10. However, while sampling with 25 Hz or even 50 Hz, these drifts per sampling may even below the expected minimum random r.m.s. BPM noise. In order that the cut is more robust with respect to BPM measurement noise, the branching condition was introduced in equation 5.34 that sets the rejection threshold always above the estimated

orbit measurement noise. Tests show that the cut parameter n_σ between 4-8 seem to be appropriate. An intrinsic feature of the second cut is that it automatically adapts itself to the present beam conditions: a large orbit r.m.s. or unstable beam leads to more relaxed rejection conditions, while a stable beam with very small measurement noise may enable the detection of spikes in the range of a few micrometres. This may be helpful with respect to the changing beam conditions during injection, ramp, squeeze and 'physics'. If required, the minimum possible rejection threshold given on the right side of equation 5.34 could be fixed to an arbitrary level. The spikes are further mitigated with respect to beam steering by the feedback controller that acts on the average of several position acquisitions.

The third cut, given by equation 5.35, detects systematic steps of the BPM measurement, which should a priori be treated different. For spikes, the missing position acquisition may replaced by an interpolation of the last valid measurement (zero-order-holder) or a predefined reference value. Large measured steps may indicate either a large systematic measurement error or failure of the WBTN. A simple count of number of spikes N_{spike} that occurred during the last n samples seem to be sufficient to distinguish between spikes and steps in the response. Further, this cut filters BPM failures due to cascading failures or bursts.

However, the third cut needs a further refinement procedure, as its systematic use would, with a given probability, lead to systematic disabling of all available BPMs in the machine. Thus in addition to the third cut, in case a BPM has been disabled, the filter continues evaluating above cuts and re-enable the BPM with the new average beam position as reference value if the following cut applies:

- if $t_{step} > t_{wait}$ and $N_{spike} < 5$ re-enable BPM if temporarily disabled by filter
- if $t_{step} > t_{wait}$ and $N_{spike} \geq 5$ mark BPM as faulty and remove it from the correction scheme

with t_{step} the time since the BPM measured a spurious step change and t_{wait} a user specified grace period during which the BPM position acquisition may settle. The given grace period t_{wait} should be larger than the filter lengths given by $n_{rms} \cdot T_s$ and $n_{avg} \cdot T_s$ that are required for estimating the averaged position and r.m.s. measurement noise.

The BPM filter function is illustrated in Figure 5.13 by means of simulated and measured orbit data. The orbit is sampled at 25 Hz. The following rejection thresholds were used for the BPM filter:

$$t_{wait} = 10\,\mathrm{s} \tag{5.36}$$
$$n_\sigma = 8 \tag{5.37}$$
$$n_{rms} = 20 \tag{5.38}$$
$$n_{avg} = 4 \tag{5.39}$$

The BPM spike and step rejection show satisfying results and are independent of the actual r.m.s. orbit noise. The filter favours a fast sampling: For a sampling of 25 Hz for example, the expected orbit drifts per sample are below the expected orbit measurement noise. Thus any position reading exceeding the r.m.s. measurement noise by a factor of four can directly map to a BPM failure. Further, a reduced sampling period reduces the latency until a possible step failure is detected and filtered out.

More advanced algorithms may include interpolation of the position deviation at a given point using neighbouring BPMs and the knowledge of the beam transfer function in between. The deviation of interpolated to measured position may be used as an indication for a BPM failure. Such a scheme is for example implemented at the APS light source [127]. These schemes proved to be reliable due to the high BPM sampling of the spacial betatron oscillation and precisely known optics between BPM.

For the LHC this scheme is not considered due to the larger BPM spacing and uncertainties of the beam propagation between adjacent cells. It is impossible to fully disentangle these deflections between those due to additional dipole kicks onto the beam and dipole compensations of quadrupole feed-downs driven by ground motion. As a future improvement one may interpolate and use the relative orbit change of the neighbouring BPMs. In this case the ground motion induced orbit drifts may be neglected with respect to the spikes on the measurement. In any case, the systematic on bunch length and intensity can neither be detected nor compensated with this scheme since these effects affect all BPMs coherently.

5.1.6 BPM front-end failures

The errors and failures discussed in the previous sections mostly affect the BPMs on an individual basis, or concerning the systematic WBTN dependencies on bunch length and bunch intensity, coherently for all BPMs installed in the LHC.

The failure of individual BPM front-ends is another important source of failures. In the LHC, up to 18 BPMs are grouped and controlled through one VME crate containing the BPM acquisition electronics

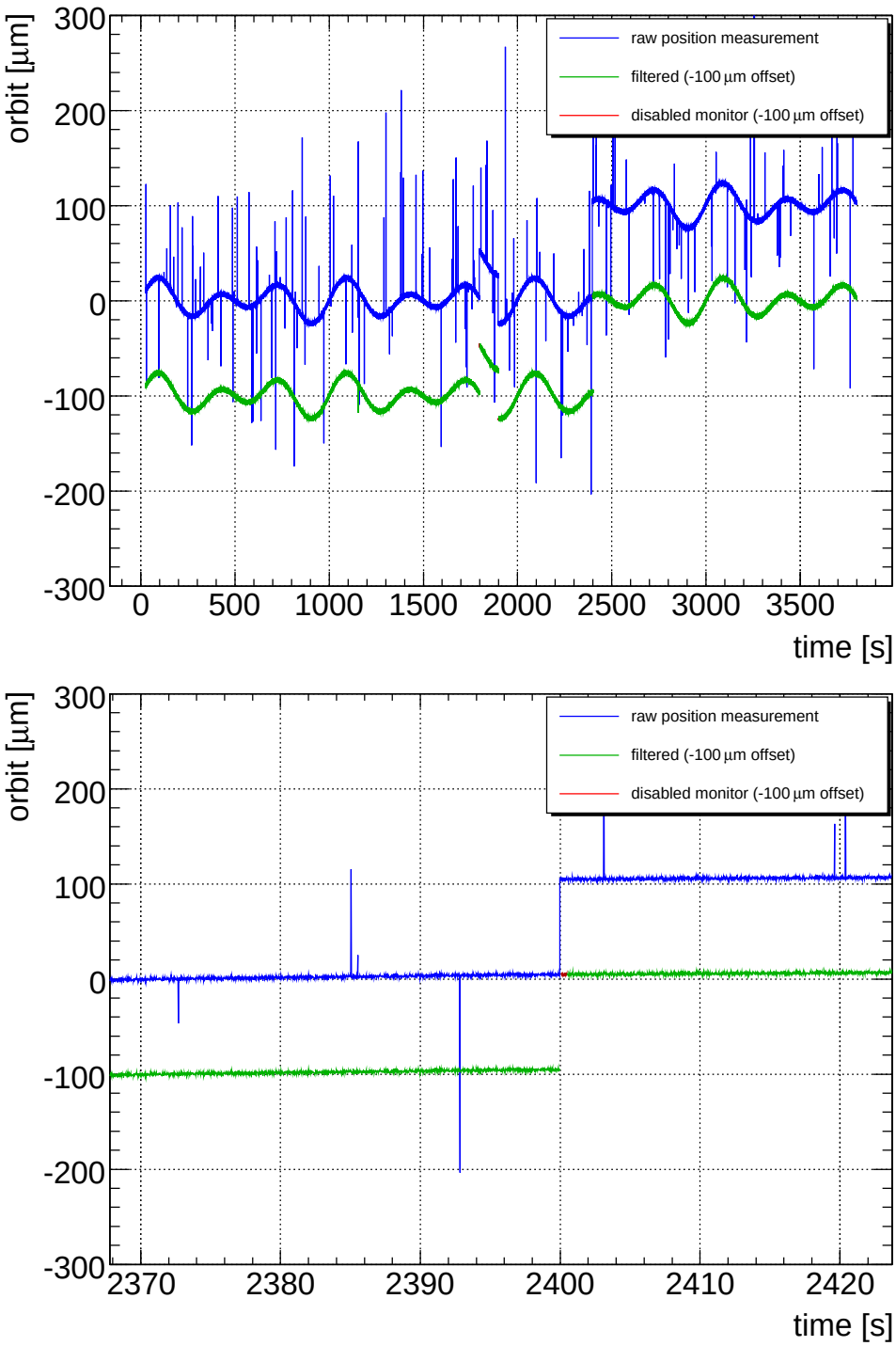


Figure 5.13: Example of BPM failure detection using simulated (top) and measured erroneous orbit data taken during an SPS test (Date: 2004-10-05, SPS Cycle 35369). Though the actual LHC BPM system did not have any step failures, a simulated step failure was added at about 2400 seconds in the mock and measured data. The full (top) and zoom (bottom) is shown. The filtered data points were shifted by 100 μm for better visibility.

(WBTN receiver cards), digital preprocessing cards and a VME Power-PC front-end computer. There are 8 front-end crates per IP and 64 crates in total responsible for acquiring, preprocessing and sending the BPM position information to the central feedback controller or data concentrator, respectively.

The spatial sensitivity to single BPM failures is in principle mitigated by the large number of BPMs that sample the orbit with an approximate phase-advance of 45^O in the LHC. However, in a simple configuration where the BPMs are grouped consecutively according to their position in the ring to BPM frontend crates, a single front-end failure would cause the loss of up to 18 BPMs concentrated within a few cells, leading to a complete loss of beam position. Possible sources for front-end failures are:

- failure of the front-end hardware, in particular power supplies,
- front-end operating system (OS),
- errors in the front-end software pre-processing the position data,
- erroneous data transmission due to delays, data congestion on the front-end network card or on the network infrastructure that is effectively treated as data loss if exceeding a given threshold,
- failures in the data transmission protocol,
- "hammering" or improper use of the software functionality etc.

For comparison: The HP Proliant DL 585 server (see [128] for details), used for some of the orbit feedback tests, is optimised for high availability and has an MTBF of about 60000 hours, accounting for the server hardware only. The VME power supplies used at CERN have an MTBF between 45000 and 85000 hours [129]. This corresponds to an MTBF of one out of 64 BPM-front ends due to hardware failure of about 30-55 days. However, frequent failures may be given by front-end operating systems and software reading out the acquisition electronic, pre-processing and sending the data. Usually these estimates are very vague due to the complexity of dependencies, libraries and coding style involved.

In order to provide additional passive robustness and to reduce the impact of a single BPM front-end failure, the LHC BPMs for both rings will be mixed together and distributed over these front-end computer. The distribution scheme, illustrated in Table 5.2, distributes the BPMs of one half-sector over four crates '*A*','*B*','*C*' and '*D*' starting from the first BPM close to the IR to the last monitor in the arc centre (34R respectively 33L): If a single front-end crate fails, the BPM information of every second

BPM Nr. :	1	2	3	4	5	6	7	...	16	17	18	...	30	31	32	33	(34)
Beam 1 :	A	B	A	B	A	B	A	...	B	A	B	...	B	A	B	A	B
Beam 2 :	C	D	C	D	C	D	C	...	C	D	C	...	D	C	D	C	D

Table 5.2: BPM organisation scheme: The present scheme is optimised against a loss of consecutive BPMs in case of a single front-end computer failure.

cell can be acquired and the beam is sampled at about every second quadrupole, which is sufficient for detection and correction of arbitrary betatron oscillation. The arrangement does not interleave hardware of the two beams.

BPMs installed at consecutive quadrupoles, usually alternate also between large and small values of the betatron function. For this scheme, a single BPM front-end failure may cause either a loss of all low-β, which is acceptable from the sensitivity point of view, or the loss of all high-β BPMs and thus reduce the sensitivity to betatron oscillation, as visible in equation 2.71. A more advanced scheme would include the shuffling of consecutive monitors of Beam 1 and 2 over different crates. Table 5.3 illustrates the scheme: This scheme minimises the worst case effect of losing all high β BPMs and is further less

BPM Nr. :	1	2	3	4	5	6	7	...	16	17	18	...	30	31	32	33	(34)
Beam 1 :	A	B	C	D	A	B	C	...	B	C	D	...	C	D	A	B	C
Beam 2 :	B	C	D	A	B	C	D	...	C	D	A	...	D	A	B	C	D

Table 5.3: BPM organisation scheme: In addition to the scheme presented in Table 5.2, this scheme also minimises the loss of all high β monitors in case of a single front-end computer failure.

sensitive to noise in case of a front-end failure. However, it is also more complex and may possibly lead to mapping errors during LHC installation. For the time being it was decided to use the scheme shown in Table 5.2 and to keep the advanced scheme (Table 5.3) as a future option in case the MTBF and impact of BPM front-end failures must be reduced.

5.2 LHC Orbit Correction System

The LHC orbit correction system consists of about 530 horizontal and 530 vertical correction dipole magnets for both beams. In order to maximise the response on the orbit (see equation 2.71), most of the horizontal correctors are located at the horizontally focusing quadrupoles and vice versa. A few quadrupoles, namely the triplets in the experimental insertions, are equipped with correctors that can deflect the beam in both planes. Each magnet is individually powered and controlled by a digital Function Generator Controller (FGCs). Up to 30 of these FGC cards are connected over a WorldFIP bus to a gateway that can be accessed through the network. About 50 Gateways will be used to control the ≈ 1200 closed orbit dipoles in the LHC, thus presenting the majority of power converter and independently powered magnet types. Most closed orbit dipoles are superconducting except for a few (16) normal conducting magnets (MCBWH/V) in the cleaning insertions in IR3 and IR7. The magnet and power converter parameters are summarised in Table 5.4.

Magnet Type	**B**	$\mathbf{BL_{mag}}$	**L**	**R**	$\mathbf{I_{nom}}$	$\mathbf{U_{max}}$	$\mathbf{\lvert\Delta I/\Delta t\rvert_{max}}$	$\mathbf{N_{LHC}}$
	[T]	**[Tm]**	[H]	$[m\Omega]$	**[A]**	**[V]**	**[A/s]**	
MCBH(V) @1.9K	2.93	1.90	6.02	90	55	± 8	0.5	752
MCBCH(V) @1.9K	3.11	2.81	2.84	12-200	100	± 40	1.0	156
MCBCH(V) @4.5K	2.33	2.11			80			
MCBYH(V) @1.9K	3.00	2.70	5.26	13-40	88	± 10	1.0	88
MCBYH(V) @4.5K	2.50	2.25			72			
MCBXH	3.35	1.51	0.287		550	± 10	5.0	48
MCBXV	3.26	1.56	0.175		550			
MCBWH(V)	1.1	1.87	0.02		500	± 40	5.0	16

Table 5.4: Parameters of the LHC orbit corrector magnets and power converters: B is the maximum field, BL_{mag} the maximum integrated dipole field, L the magnet inductance, R is the magnet and cable resistance, I_{nom} is the nominal current of the magnets and U_{max} the maximum voltage delivered by the PC. N_{LHC} is the number per type of magnets installed. A complete parameter list can be found in [3,96]. Resistances according to LHC cabling database (courtesy M. Zerlauth).

The majority of the LHC CODs are either MCBH(V) or MCBCH(V) type magnets. According to equation 2.34, the deflection (in *rad*) a dipole corrector magnet bends the beam is given by:

$$\delta_{kick} = c \cdot 10^{-9} \frac{BL_{mag}[Tm]}{p[GeV/c]} \cdot \frac{I}{I_{nom}} \tag{5.40}$$

with c being the speed of light. The corresponding maximum kick strengths at 0.45 TeV/c and 7 TeV/c, respectively are given by:

$$\delta_{max} = \begin{cases} 1.26\,\text{mrad} & \text{for } p = 0.450\,\text{TeV/c} \\ 80.76\,\mu\text{rad} & \text{for } p = 7\,\text{TeV/c} \end{cases} \tag{5.41}$$

Using the closed orbit dipole response given in equation 2.71 and the LHC arc lattice, the maximum response of a single close orbit dipole kick δ observed at a BPM, both at locations with $\beta = 180$ m, is given by

$$\Delta z|_{max} \approx 100 - 110\delta \tag{5.42}$$

with $\Delta z|_{max}$ the maximum horizontal or vertical orbit shift, respectively. The maximum achievable amplitudes may thus reach about 140 mm during injection and about 8 mm during collision. It is visible that during injection already a single corrector dipole magnet is sufficient to deflect the beam into the edge of the vacuum chamber. Thus, one may opt for machine protection to survey orbit correctors and their settings and to capture erroneous settings before they lead to a direct beam loss into the cryogenic magnets.

5.2.1 Maximum Current Ramp Limit

In first order, the dynamic properties of the orbit corrector circuits are given by the cable resistance R and inductance L of the magnet, and maximum available voltage U_{max} that can be applied by the power converter. The simplified circuit is illustrated in Figure 5.14. In case of a power supply failure or equivalent open circuit, the induced voltage is proportional to the current change rate and inductance. Due to the high inductance L of the magnets and high currents these voltages may reach large values

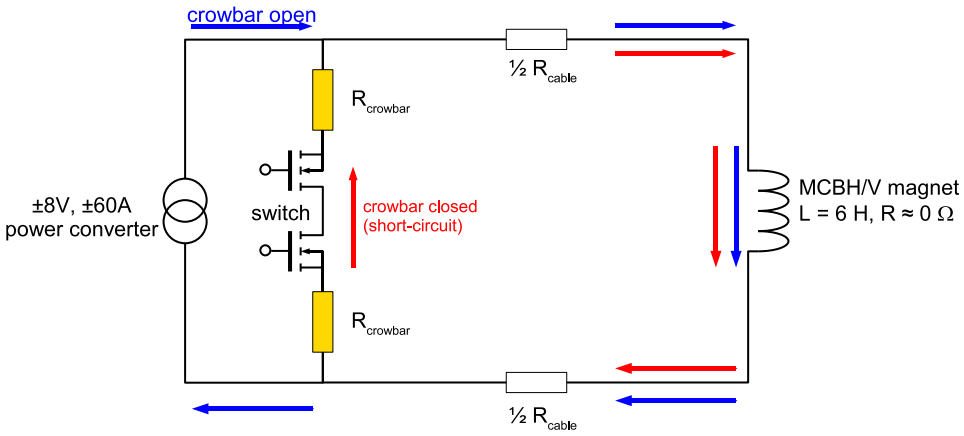


Figure 5.14: Schematic orbit corrector circuit.

leading to violent electric arcs and possible damage to the equipment. Thus, for additional protection, each circuit contains a so-called *crow-bar* that shorts the circuits in case of a power supply failure. The different failure types are further discussed in Section 5.2.3.

The differential equation describing the current $I(t)$ in the circuit, is given by the power converter's variable excitation voltage $U(t)$, the circuit resistance and inductance L of the magnet:

$$U(t) = R \cdot I(t) - L\frac{dI(t)}{dt} \tag{5.43}$$

Comparing with equation 3.4, it is visible that equation 5.43 describes a first order system with a natural response time:

$$\tau = \frac{L}{R} \tag{5.44}$$

The natural response time of all circuits is shown in Figure 5.15. It is visible that the natural time of most

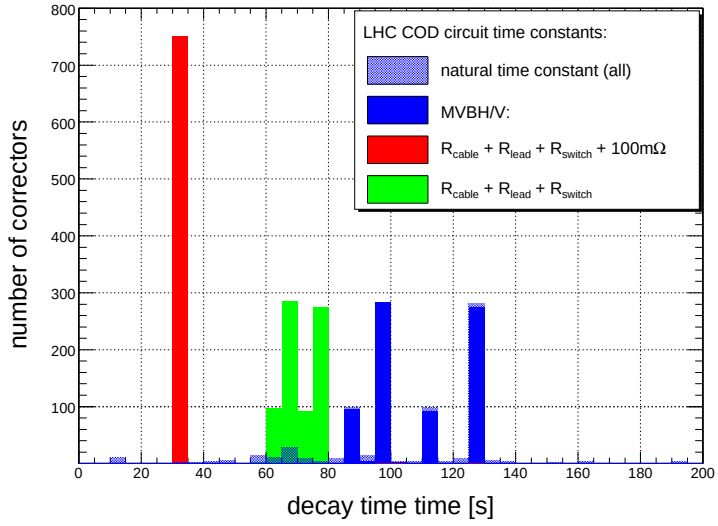


Figure 5.15: Natural time constant distribution of the LHC orbit corrector circuits with and without crow-bar circuit. Data on cable and magnet resistances is taken from LHC cabling database (Courtesy M.Zerlauth).

magnets (MCBH/V) is accumulated around 100 and 130 seconds. In the LHC the converters powering the dipole magnets of one complete FODO cell are located underneath the quadrupole cryostats. The difference in resistance is due to the different cable length of the dipoles installed directly above with respect to those installed close to the next quadrupole. The cable length differs by about 100 m (twice the half-FODO length). The time constants below 20 s correspond to the 16 normal conducting magnets.

In case of a failure and short circuit through the crow-bar, the time constants drop for all magnets to 35 s with and 60-80 s without addition crow-bar resistance $R_{crow} = 100\,\Omega$.

In order to improve and reduce the circuit time constant, the LHC power converter itself contains a digital feedback loop running at 1 kHz. The loop overdrives the actuator (voltage) signal $U(t)$ to decrease the response time of the electrical circuit, as described in [130]. The voltage is stabilised with an internal loop running at 1 kHz. Due to radiation limitations on the WorldFIP bus that connects the gateways to the FGCs the maximum rate at which the current can be changed is presently limited to 50 Hz. Accordingly, the maximum possible feedback sampling frequency for the orbit and other feedback control systems is thus essentially limited to 50 Hz. It is important to note that the feedback bandwidth is usually less than the feedback sampling frequency.

The maximum excitation signal $U(t)$ is limited by the maximum voltage (5.4) the power converter can provide

$$|U(t)| < U_{max} \tag{5.45}$$

and the maximum current is given by $I_{max} = \frac{U_{max}}{R}$. For small changes around a constant set current I_0 of the power converter feedback loop, the maximum current change is given by

$$\left|\frac{\Delta I}{\Delta t}\right|_{max} \approx \frac{U_{max} - RI_0}{L} \tag{5.46}$$

The maximum change depends on the working point I_0. Using the values given in Table 5.4, the maximum current change per time interval, depending on the working point ($I_0 = \pm I_{nom}$), is:

$$0.51\,\frac{\mathrm{A}}{\mathrm{s}} < \left|\frac{\Delta I}{\Delta t}\right|_{max} < 2.15\,\frac{\mathrm{A}}{\mathrm{s}} \tag{5.47}$$

To obtain a power converter response that is independent of the working point, the maximum current change is limited by the converter controller design to:

$$\left|\frac{\Delta I}{\Delta t}\right|_{wp} := 0.5\,\frac{\mathrm{A}}{\mathrm{s}} \tag{5.48}$$

The current rate limit is a tradeoff between predictability (with respect to ramping) and tracking accuracy which is particularly critical for the main dipoles. The drawback is a reduced maximum rate of current change (speed of the corrector). But the rate limit allows to precisely predict the response of the power converter.

5.2.2 Closed Orbit Dipole (COD) Frequency Response

The Bode response can be estimated using a Fourier decomposition of the requested current change. A current oscillation with an amplitude of I_a and frequency f involves current changes $\frac{dI}{dt}$:

$$I(t) = I_a \sin(2\pi f t) \tag{5.49}$$

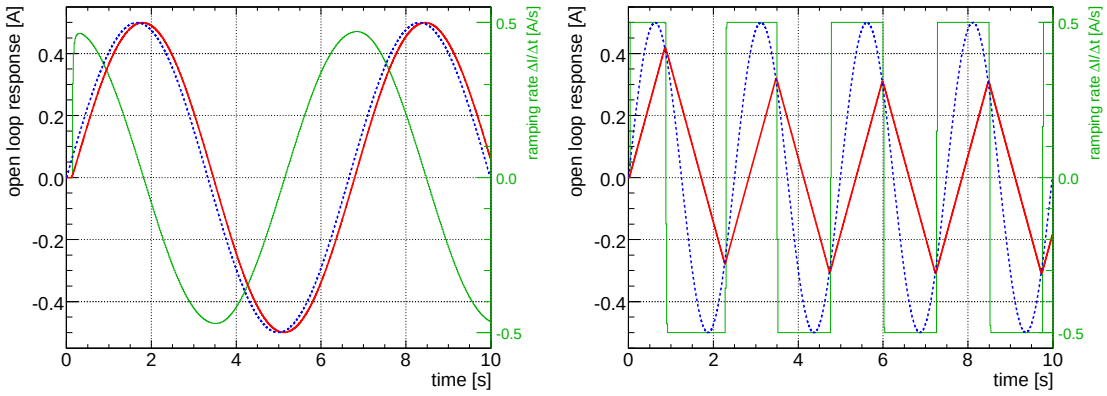
$$\frac{dI}{dt} = 2\pi f I_a \cos(2\pi f t) \overset{!}{<} \left|\frac{\Delta I}{\Delta t}\right|_{max} \tag{5.50}$$

It follows that the power converter can precisely produce the sinusoidal response up to a frequency

$$f_{max} = \frac{1}{2\pi I_a}\left|\frac{\Delta I}{\Delta t}\right|_{max} \tag{5.51}$$

before the current rate reaches the maximum slew rate, gets clipped and the converter starts to produce a distorted approximative saw-tooth response. The linear and distorted clipped open loop power converter response is illustrated in Figure 5.16. The first response shows shows an excitation with 0.15 Hz excitation frequency which is within and a second response with 0.4 Hz excitation frequency which is above the linear regime given by $f_{max} \approx 0.16$ Hz. Within the linear regime, the magnitude of the frequency response equals one (perfect tracking) and shows only a lag of 10 ms due to the 50 Hz power converter sampling [1]. For higher frequencies the power converter cannot track the sinusoidal reference but produces a triangular

[1] It is implicitly assumed that the excitation is phased in with the power converter's 50 Hz clock and the absence of other additional control delays



(a) $f = 0.15\ Hz, \quad \Delta I = 0.5\ A$ (b) $f = 0.4\ Hz, \quad \Delta I = 0.5\ A$

Figure 5.16: Open loop response to a sinusoidal input signal with requested 0.5 A amplitude. The perfect tracking with only the sampling lag for the low frequency and the distortion due the 0.5 A/s rate limit and higher frequency is visible.

response. The amplitude of the response drops by 10 orders per decade as typical for first order systems. The response magnitude is

$$|M(f)| = \begin{cases} 1 & \text{if } f \leq f_{max} \\ \frac{1}{2\pi I_a} \left|\frac{\Delta I}{\Delta t}\right|_{max} \cdot \frac{1}{f} & \text{if } f > f_{max} \end{cases} \tag{5.52}$$

Beside limiting the amplitude, the rate limiter induces an additional delay $\Delta\tau$ as illustrated in Figure 5.17. It is visible, apart from the 20 ms control delay, that at $t = t_0 + \Delta\tau$ the excitation amplitude and

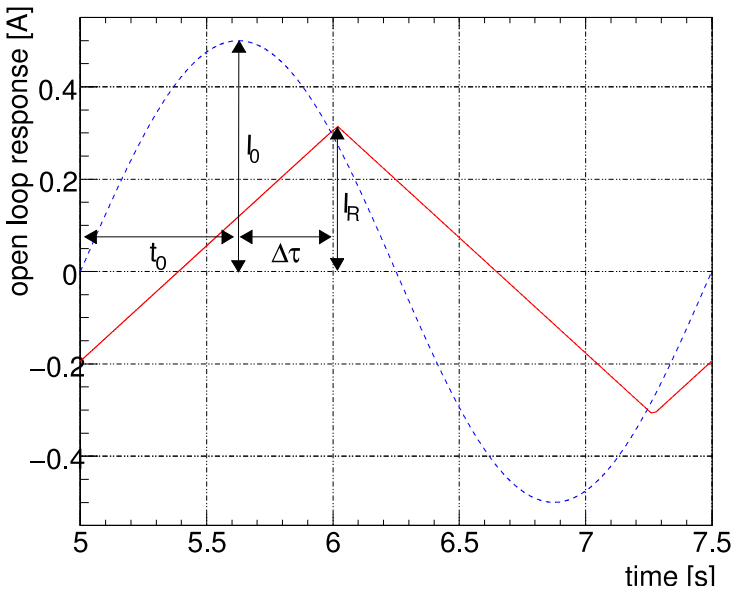


Figure 5.17: Rate limiter response. The additional delay due to the rate limiter is visible.

response I_R are equal. It follows that the signal $I(t)$ can be written as

$$I_R = I_0 \sin\left[\omega(t_0 + \Delta\tau)\right] \tag{5.53}$$

Comparing and substituting the ratio between I_R to I_0 by $|M(f)|$, given in equation 5.52, and replacing $\omega t_0 = n \cdot \frac{\pi}{2}$

$$cos(\varphi) = \frac{[I(t)]_{max}}{I_0} = |M(f)| \tag{5.54}$$

yields the phase advance $\varphi = 2\pi f \cdot \Delta\tau$

$$\varphi = \arccos(|M(f)|) \tag{5.55}$$

The full response $M(f)$, including control T_c and sampling delay T_s, is given by

$$M(f) = |M(f)| \cdot e^{i \cdot \varphi(f)} \tag{5.56}$$

$$|M(f)| = \begin{cases} 1 & \text{if } f \leq f_{max} \\ \frac{1}{2\pi I_a} \left|\frac{\Delta I}{\Delta t}\right|_{max} \cdot \frac{1}{f} & \text{if } f > f_{max} \end{cases} \tag{5.57}$$

$$\varphi(f) = \pi f \cdot (T_s + T_c) + \arccos(|M(f)|) \tag{5.58}$$

with I_a the requested current, f the requested frequency, $\left|\frac{\Delta I}{\Delta t}\right|_{max}$ the maximum current ramp limit and f_{max} the transition frequency, given in equation 5.51, which is the limit between the linear and non-linear regime. The open-loop frequency response for various requested excitation amplitudes I_a is shown in Figure 5.18. The maximum bandwidth and phase margin decrease with increasing current.

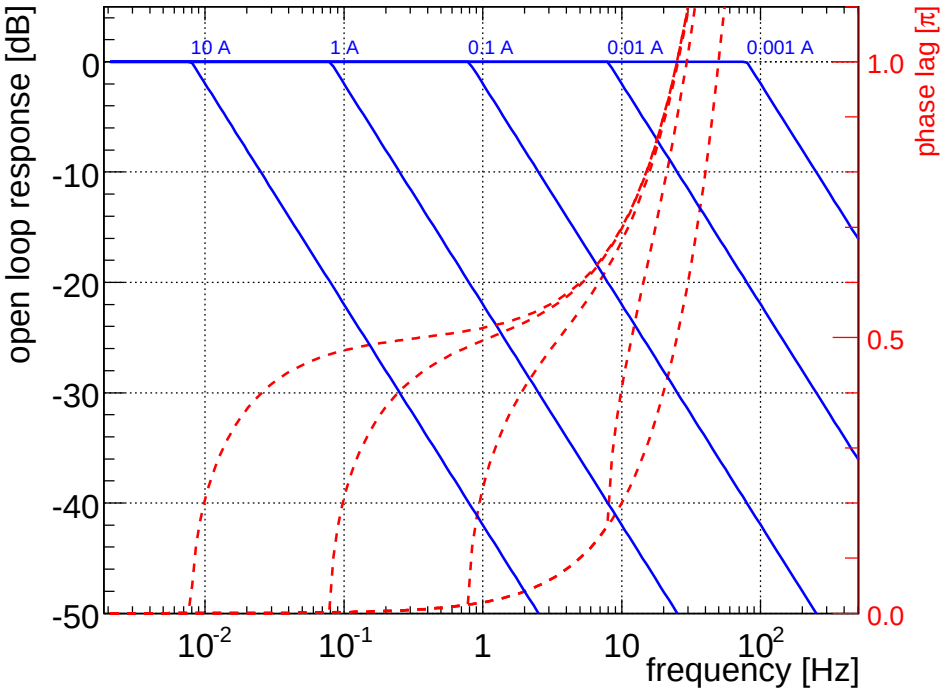


Figure 5.18: Open loop response of the $\pm 60A$ converter that powers the MCBH and MCBV magnets. The non-linear magnitude and phase lag are plotted for various reference currents. A current change of 0.01 A corresponds to an average orbit r.m.s. change of about 15 μm at 7 TeV. The reduced bandwidth and increased phase advance for increasing amplitudes I_a is visible. The phase delay is plotted for a perfect sampling which matches the controller 50 Hz sampling and its phase.

Using equations 5.42, 5.40 and 5.51, it is possible to give a rough estimate of the maximum sinusoidal excitation the corrector can produce at 7 TeV in an open-loop mode. Figure 5.19 shows the maximum dipole and orbit excitation as a function of closed loop bandwidth. The maximum sinusoidal orbit change Δx, tracked perfectly at frequency f and 450 GeV and 7 TeV beam momentum, is given by the relation:

$$\Delta x \cdot f \overset{!}{\lesssim} \begin{cases} 202\ \mu m/s & \text{at } p = 450\,\text{TeV} \\ 13\ \mu m/s & \text{at } p = 7\,\text{TeV} \end{cases} \tag{5.59}$$

In case of a ramp-like drift the maximum orbit drift that can be compensated by a single MCB magnet is given by:

$$\left|\frac{\Delta x}{\Delta t}\right|_{max} \overset{!}{\lesssim} \begin{cases} 1269\ \mu m/s & \text{at } p = 450\,\text{TeV} \\ 81.5\ \mu m/s & \text{at } p = 7\,\text{TeV} \end{cases} \tag{5.60}$$

At 0.1 Hz the corrector is able to produce a sinusoidal excitation on the orbit with maximum amplitude of $\approx 130\,\mu$m at 7 TeV beam momentum which, compared to the expected ground motion and magnet induced amplitudes, shows that the correction system's margin seems to be sufficient to cope with the expected oscillation. However, the correction of the final focus induced orbit drifts is close to the above limit stressing the requirement of good alignment of the insertion quadrupole magnets.

However the available gain and phase margin may shrink due to unavoidable imperfection of the total feedback system due to noise, measurement and feedback sampling delay and actual steering and controller

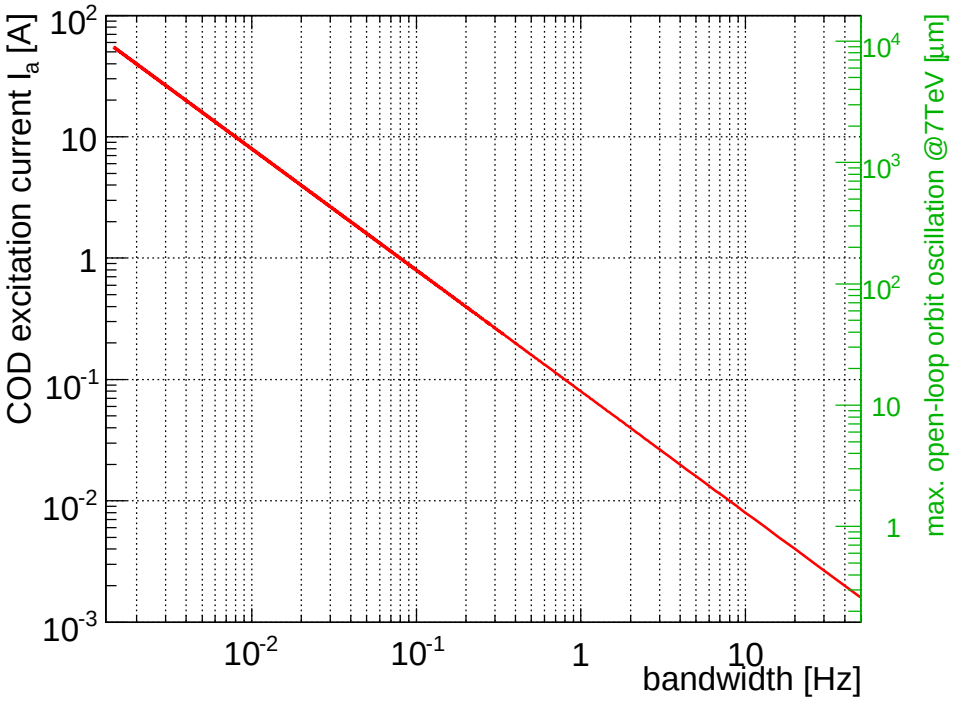


Figure 5.19: Open loop bandwidth of the $\pm 60A$ converter that power the MCBH and MCBV magnets. The maximum sinusoidal current amplitude is plotted as a function of bandwidth. The corresponding maximum orbit excursion prediction in the arc driven by the MCB magnet at a $\beta = 180\,\mathrm{m}$ is given at the right hand scale.

imperfections. The following example illustrates how the actual gain reduces due to the additional phase lag: The MCB orbit correction circuits can (in theory) produce and correct orbit oscillations of up to $1.6\,\mu\mathrm{m}$ at a frequency of 8 Hz. However, as visible in Figure 5.18, the phase margin for the corresponding $I_a = 0.001\,\mathrm{A}$ is already less than about $0.8\,\pi$ at this frequency. If one assumes a closed loop sampling of 10 Hz, the sampling would induce an additional phase advance of about $0.8\,\pi$. This additional phase would fully exploit the phase stability and drive the system under verge of instability.

5.2.3 Converter Failures

In contrast to beam position monitor failures, the consequence of a faulty corrector magnet is more severe because the magnet may be compensating for quadrupole misalignments and may be used to steer the beam. At the LHC, the criticality of corrector failures is given by their large number and by the fact that each failing dipole may independently deflect the beam into the collimator or even vacuum chamber.

Under the assumption that the correctors compensate a quadrupole r.m.s. alignment error of 0.4 mm, the corresponding average corrector deflection is about $\delta \approx 8\,\mu\mathrm{rad}$. This angle corresponds to an average current in the magnet of about 0.4 A during injection and 5.2 A during collision. In case a single corrector dipole fails, the expected average orbit drift (arc, $\beta = 180\,\mathrm{m}$) is about 0.9 mm or, with a probability of about 89% a failing circuit causes an orbit perturbation of more than $100\,\mu\mathrm{m}$. These drifts are significantly above the orbit tolerance, in particular those given by the beam cleaning and protection system. Though these failures are not as fast as a strong normal conducting dipole magnet failure, orbit corrector dipole failures may still cause increased particle loss inside the collimation region and thus lead to a beam dump request due to excessive beam loss.

Two main fast COD failure sources have been identified: quench of the corrector magnet and power converter circuit failures.

The orbit dipoles and other smaller corrector circuits do not have a dedicated energy extraction system. In case of a quench, the magnet becomes resistant and the energy is dissipated inside the magnet. Figure 5.20 shows the decay from a maximum current of 60 A due to quench. The decay is not exponential but can be approximated by a Gaussian distribution. The decay time is about 350 ms and the total current is about zero after one second. The corresponding maximum current decay rate is about $|\Delta I/\Delta t|\,|_{max} \approx 150 A/s$ for an initial current of 60 A. However, MCB quenches are expected to be rare

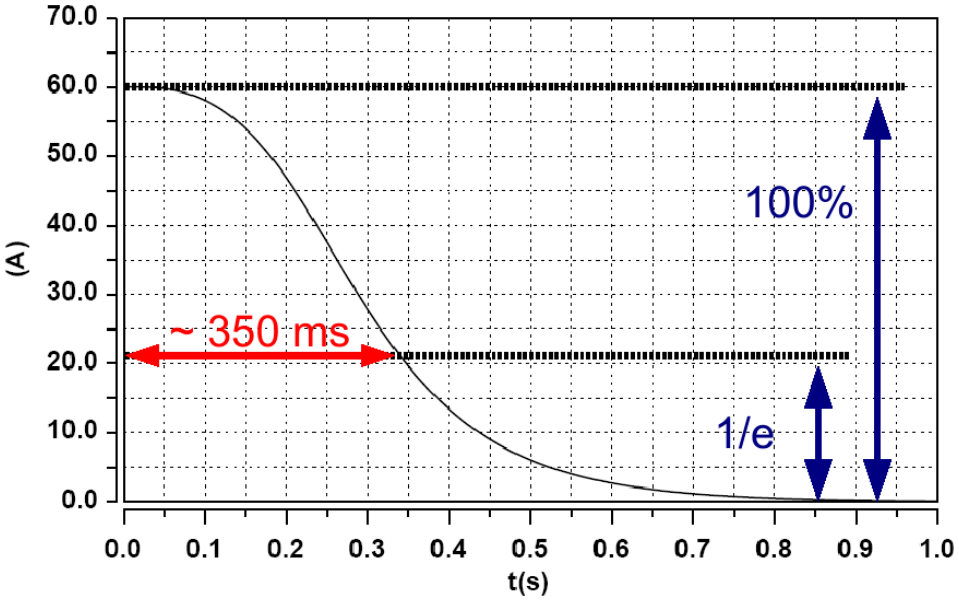


Figure 5.20: Current decay from 60 A in an MCB type orbit corrector after an induced quench. (Courtesy F. Rodriguez Mateos)

and less likely than a quench of the neighbouring quadrupole magnet, which will in any case result in a stop of the machine.

The second important failure source is due to failing power supplies. The 60 A power converter design, foreseen for the orbit corrector dipole magnets, is based on a MTBF of about 100 000 hours which takes into account that the converters are operated below the quadrupoles in a radiation environment [131]. Assuming a cycle length of 10 hours, this MTBF corresponds to a total failure rate of about one every 10 machine cycles for all magnets. The actual MTBF may be verified after a few months of real LHC operation.

The initial converter crow-bar design included two additional 50 Ω resistances that were inserted in order to limit the maximum currents flowing through the crow-bar switch and to deliberately reduce the current decay time in order to enable faster access during power converter tests. The corresponding decay time including the additional 100 Ω resistance corresponds to about 30-35 s which is about half the decay time without the additional resistance (60-80 s). From the machine protection point of view, it is favourable that, in case of a failure, the magnetic decay be as slow as possible to reduce the criticality of the failures. As discussed in Section 6.1.4, a larger decay time also enables compensatory measures. The orbit feedback may minimise and possibly fully compensate the excessive transient due to the power converter failure inside the collimation region. Thus it may keep the beam inside the machine while eventually increasing the beam availability for physics.

After verification, it was shown that the additional resistances are neither required for protection of the crow-bar switch nor for the protection of the magnet. This thesis contributed to the engineering change request that removed these resistances [132][2].

5.3 Technical Network and Computational Infrastructure

The use of computer based simulations and analyses becomes increasingly important and demanding with respect to required computing power. Nowadays, the benefit of the analysis and simulation in high-energy physics is given not only by the correctness of the result but also by the time frame the result is achieved. In particular, sub-fields such as *Lattice QCD*, a numeric evaluation of QCD fields and compound particle in the non-perturbative regime [133, 134], rely on massive parallel computing and hardware optimised simulation routines [135].

Within control systems, the validity of the result, in addition to the correctness of the computation, also depends on the time the result is obtained and applied as correction to the process. Computational

[2] A minor draw-back of the removal is the doubling of the time required until the current inside the circuit has dropped below a safe threshold before a regular power converter intervention can be performed by the repair team.

delays, data transmission and jitter are inevitable in digital systems. Often it is required that the total system delay response λ stays constant within a given margin corresponding to stable sampling and delay zero locations. Thus for describing the physical behaviour of feedback control systems it is also important to study the behaviour of the involved digital devices contributing to the total delay.

In computer science, the term 'real-time' is often used to describe an application responding to a given stimuli within an upper limit of response time, equivalent to stable sampling and delay zeros. Another common use of the term 'real-time' is that a physical process is simulated and updated at the same rate as the real physical process or at least with a constant mapping between simulated and real time. A more formal definition of 'real-time' can be found in [136]. Often, the prefixes "hard" or "soft" are added to indicate or distinguish the benefit of the system in case the specified upper response time boundary is exceeded by the computation: 'hard' refers to a catastrophic outcome if the deadline is not met and 'soft' a reduction of the system quality by whatever measurement scale is applicable.

A common misconception about 'real-time' is that the process has to run "fast" or even "as fast as possible". However, the emphasis lies on the determinism of the maximum (minimum) time of the signal processing independent on its length.

The foreseen delay compensation scheme is further discussed in Section 6.2.2. From the control theory point of view, it is merely required that the delay λ and its distribution are known.

5.3.1 LHC Technical Network

The foreseen LHC feedback systems are large distributed systems that need to exchange information on the scale of the 27 km machine circumference. As a compromise between quality of service, determinism, price and required data transmission bandwidth it was decided to use LHC's Technical Network for communication between the beam position monitor front-ends, central feedback controller and power converter gateways.

The Technical Network having a 2 GBit/s backbone bandwidth is based on store-and-forward switches and routers, respectively, that are distributed corresponding to the general LHC eight-fold symmetry [137]. The core of the network is based on Enterasys' X-Pedition 8600 router, which are redundant composite routers providing store-and-forward capabilities with a bandwidth of up to 64 GBit/s per back-plane and packet filtering based on up to Level 4 information. Further hardware details can be found in [138]. The front-ends are connected to the router through switches (3Com SuperStack 3 Switch 4400 Family) that provide two bandwidths of up to 100 MBit/s and 1 GBit/s, respectively. The schematic LHC Technical Network topology is shown in Figure 5.21. The Gigabit backbone is highly redundant and shows a two-

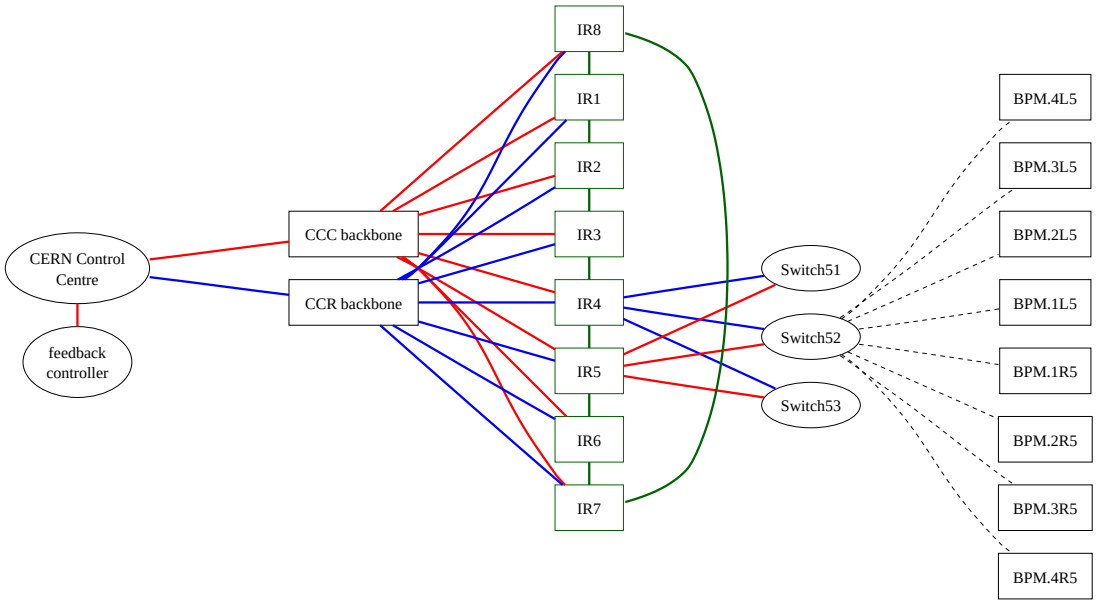


Figure 5.21: Reduced schematic LHC technical network topology including only the Gigabit backbone (bold lines) and its routers (boxes) and exemplarily some computers connected in IR5 through Fast Ethernet connections (dashed lines). Triple and higher redundancies are visible. The transmission delays are essentially determined by the fibre length given in table E.1.

fold, three-fold or higher level of redundancy depending on the packet source and sink location. Each

LHC Point is equipped with a router that is interconnected through Gigabit Ethernet to the neighbouring points. Each of these routers is redundantly connected to two central routers, one located in the CERN Control Centre (CCC, building 874) and the other in CERN's Computing Centre (CCR, building 513). The feedback controller is indirectly connected to the CCC and CCR backbone router. The routers have an internal two-fold redundancy and predicted MTBF of about 200 000 hours [138]. The optical fibre lengths of the Technical Network are given in table E.1

Packet collisions and losses are intrinsically prevented by the store-and-forward router design as long as the actual bandwidth does not exceed the router's buffer capacities. In case of X-Pedition 8600 router, the buffer size is about 3 MB per gigabit port yielding, together with the switching bandwidth of up to 64 GBit/s an approximative maximum latency of 375 μs inside the router before packet loss occurs. Delay and stress test on the real hardware show an average latency of about 12.6 μs for this system. The switches' latency is 3.3 μs for the 3COM switches (Data courtesy M. Zuin).

It is foreseen to use User Datagram Protocol (UDP [139]), a connection-less level 4 communication protocol for transmitting feedback data from the BPM front-ends and to the power converter gateways. UDP, commonly used for real-time applications such as voice and video transmission over the Internet, was chosen in favour of TCP (Transmission Control Protocol [140]) due to its predictability, simplicity and consequently low computational overhead. In particular, in case of a packet loss, TCP implements a packet retransmission that could cause additional latencies which may be larger than the foreseen feedback sampling frequency. Due to the high update rate, the feedback design does not have to rely on the information of a single data packet. Thus it is preferable to rather drop a (possibly) delayed packet and work with the information of the newest measurement than to wait for a resend packet. This policy also circumvents cascading failures of the TCP protocol which may possibly lead to a total failure and re-initialising of the connection, equivalent to a BPM front-end unavailability that can last up to several seconds depending on the protocol implementation.

Presently, the foreseen BPM UDP packet size is about 1.5 kByte per front-end, and 96 kByte in total. The amount of data transmitted per sample from the 64 BPM front-ends is small compared to the maximum bandwidth of the Technical Network. The transmission delay, if the whole amount of data would be transmitted through a single backbone-router to backbone-router connection, would be less than about 400 μs. This is small compared to the minimum feedback sampling time of 20 ms and further mitigated since the packets travel different paths within the network. In reality the delay is about ten times larger due to delays of the routing/switching hardware and travel delays of light in the optic fibres. The transmission time of a 500 byte packet from IP5 to the control room ($\approx$ 20 km) through the technical network has been measured in a laboratory test setup using optical fibre with length corresponding to those found in the LHC. The experimental results are shown in Table 5.5: The total delay can be

Network element:	SW	T	R	T	14 km	R	T	4 km	R	T	SW
Delay(μs):	3.3	41.0	12.6	4.1	70	10.9	4.1	20	12.6	41.0	3.3

Table 5.5: Travel delay for a 500 Byte packet transmitted from IP 5 to PCR. Times are given in μs (courtesy Marianna Zuin).

decomposed into delays due to switches (SW), routers (R), data transmission (T: 100 and 1000 MBits/s respectively) and delays due to transmission through optical fibre. The resulting total transmission delay is 223 μs. About 20% of the delay is due to routing hardware whereas 80% is due to the light propagation in the optical fibres $c_{medium} \approx 200000 \frac{\text{km}}{\text{s}} < c_{vacuum}$

The maximum latency measurement has been verified with the actual Technical Network installation. The measurement consists of a network host located in the CCC sending ICMP packets ([141]) to a specified destination in IR5 that send a reply packet ('ping' mechanism). Earlier measurements, shown in Table 5.5, indicated that the switches have a response latency in the order of 3.3 μs which is much less than the expected $20 - 30\,\mu$s inherent to the use of front-ends and general purpose operating systems. Thus, in order to obtain more accurate results and to account only once for the sending front-end related delays, a switch as part of the network infrastructure was chosen as the destination. A standard HP Proliant DL 380 rack-mountable PC with two Xeon processors (2.8 GHz) and two Gigabit Ethernet cards (Broadcom chip-set) were used as a sender. The computer was running a real-time enhanced standard Linux kernel (version 2.6.15 and Ingo Molnars RT patch [142, 143]). In order to avoid bouncing of the network driver and packet sending process between the two CPUs, the two processes CPU affinity were further set to CPU 1 and 2, respectively. The relative round-trip travel delay given as the difference between send and arrival time of the packet was noted with microsecond resolution. The measurement was repeated with an approximative periodicity of 1 Hz over a duration of about 77 days (2006-01-25 till

2006-04-12) resulting in a total number of 6640117 samples. During the whole measurement period, only 71 packets were lost. Further investigation revealed that this loss was due to a systematic crash of one of the back-bone routers which led to a reset every about 200 days due to an internal buffer overflow error [144]. The problem was known and solved by the router vendor and the router firmware it is planned to update during the next accelerator shutdown. Apart from this single failure, no other packet losses were observed. The packet transmission delay distribution from the CCC to IP5 is shown in Figure 5.22. With three sigma confidence, all packets arrived within 400 μs. The worst-case delays are all within the

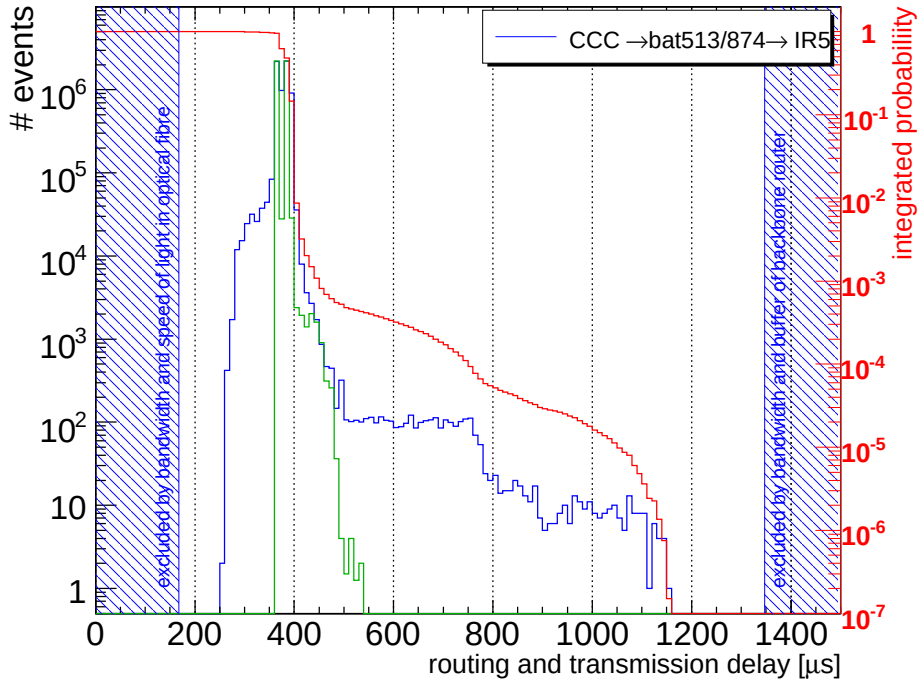


Figure 5.22: Half-round-trip latency distribution of 1 KB packet send from the CCC to IP5 (blue). The maximum and minimum boundaries due to hardware limitations and speed of light are indicated. The distribution for a local loop (green) is superimposed. See text for further details.

expected boundaries given by the minimum travel delay due to the speed of light in the optic fibre and the maximum possible delay in case of a router buffer overflow. The delays are larger than those measured with a dedicated network analyser during the stress tests presented in Figure 5.5. Two local maxima are visible. The actual packet travelling route is not necessarily given by only the shortest distance or delay but also by load balancing the traffic over the two central backbone-routers in the CCC and CCR. The time difference between the two maxima corresponds to the longer cable length going through the CCR with respect going through the CCC backbone-router to IP5.

The additional delays are, to a large extent, due to delays in the front-end hardware and operating system. Figure 5.22 shows superimposed distribution of the latencies of sending the same packets locally within the same front-end while omitting the frontend's network hardware and Technical Network. The internal kernel logging facilities revealed that nearly all the outliers up to about 1.1 ms were due to the network card driver implementation, which is not surprising keeping in mind that these devices where developed optimising the maximum throughput performance and not worst-case latencies. Further, the absence of the larger outliers can be explained by the fact that the local latency measurement was taken in burst-mode: The same amount of packets that were sent over two months through the technical network was sent within 2 hours locally in order to obtain similar statistics. However, this type of sampling does not account for rare cases leading to large delays that occur only once every few hours or days.

Summarising, the maximum travelling delays are well below 1 ms and given by the travelling delay of light in the optical fibres and unpredictabilities in the front-end hardware and operating system sending and receiving the data packets.

Chapter 6

Orbit and Energy Control

As discussed in Section 2, the beam position shift $\Delta z(s)$ with respect to the design trajectory can be decomposed into

$$\Delta z(s) = \Delta z_{co}(s) + \Delta z_D(s) + \Delta z_\beta(s) \tag{6.1}$$

with Δz_{co} the contribution due to closed orbit, Δz_D the dispersion and Δz_β betatron oscillation, described by equations 2.71, 2.23 and 2.17:

$$\begin{aligned}
\Delta z_{co}(s) &= \frac{\sqrt{\beta(s)\beta_i}}{2\sin(\pi Q)} \cdot \cos(\Delta\mu_i(s) - \pi Q) \cdot \delta_i \\
\Delta z_D(s) &= D(s) \cdot \frac{\Delta p}{p} \\
\Delta z_\beta(s) &= \sqrt{\epsilon_z \beta_z(s)} \cdot \sin(\Delta\mu_z(s) + \phi_z)
\end{aligned}$$

where δ_i are the dipole kicks due to orbit correctors, quadrupole feed-downs, main dipole's b_1 imperfections and other effects discussed in Chapter 4. While measuring the closed orbit, which is the equilibrium beam position, the term z_β is usually suppressed through averaging over the individual particle and bunch oscillations. These term may however still have an effect in the orbit measurement through inducing a residual orbit measurement noise Δz_{rms}, for example due to bunch-to-bunch oscillations and bunch instabilities.

Further, the following notation is used: The orbit is sampled at m BPMs with z_i being the individual position reading and can be written as a column vector $\underline{z} = (z_1, \ldots, z_i, \ldots, z_m)^T$. In a similar way, the deflection δ_j can be written as $\underline{\delta} = (\delta_1, \ldots, \delta_j, \ldots, \delta_n)^T$ with n the number of available orbit corrector dipoles. The closed orbit dipole corrector magnet (COD) induced orbit shift $\underline{z}$ can be consequently cast into matrix form

$$\begin{aligned}
\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_m \end{pmatrix} &= (R_{ij}) \cdot \begin{pmatrix} \frac{\Delta p}{p} \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix} \\
\Delta\underline{z} &= \mathbf{R} \cdot \underline{\delta}
\end{aligned} \tag{6.2}$$

with $\mathbf{R} = (R_{ij}) \in \mathbb{R}^{m \times n}$ the orbit response matrix determined by the coefficients in equation 6.1, $\Delta\underline{z} \in \mathbb{R}^m$ a column vector containing the individual BPM measurements and $\underline{\delta} \in \mathbb{R}^n$ a column vector containing the individual orbit perturbation sources. In this form, momentum offset and quadrupole misalignments are treated similar to dipole deflection. However, the individual matrix elements differ for those perturbations.

As visible in equation 6.1, for accurate orbit control, it is not only necessary to compensate the dipolar induced orbit perturbations but also to stabilise the beam momentum and machine circumference, respectively. The stabilisation of the dipolar perturbation relies on the use of orbit correctors that are evenly distributed inside the accelerator. For efficiency, these correctors are often only at focusing quadrupoles that have large values of the betatron function. The LHC insertion regions are an exception: there quadrupole is equipped with at least one 'MCBX' type orbit corrector magnets that can deflect the beam in both planes.

The energy or momentum compensation depends on the machine state: during injection, for example, the RF frequency and phase has to be fixed with respect to the SPS and thus the energy correction usually relies on the horizontal dipole correctors that effectively increase or decrease the integral b_1 field, defining the machine momentum and compensating for the main dipole's b_1 drift. During collision, however, the main contribution to the energy perturbation is given by the tidal ring deformation that effectively changes the machine circumference with respect to the beam trajectory circumference given by the RF system. Due to the limited strength that reduces with energy it is generally less favourable to adjust the beam momentum and centre the beams using the dipole correctors at top energy. But, in contrast to injection, the RF phase and frequency are not constrained and can thus be used to adjust the beam momentum through a coherent change of the two beam's RF frequency.

The orbit and energy control can be decomposed into two steps:

1. Finding an optimal set-point of the orbit which involves establishing the circulating beam, verification of the polarity and rough calibration of BPMs and CODs while keeping the constraints given on to the orbit. This step is usually performed manually and establishes finally a *reference orbit* $\Delta x_{ref}(s)$. The reference orbit is also referred to as *golden orbit*.

2. Stabilising the established reference orbit while minimising small perturbations, described in Chapter 4, around the reference orbit. This stabilisation is performed periodically at a given frequency by the machine operator or orbit feedback using the beam position monitors, dipole correctors and RF frequency.

The control may alternate between the two above steps.

The differential equation describing the beam movement x_i at a given BPM with index i can be written as the superimposition of n dipole deflections δ_j that are coupled through the orbit response matrix R_{ij}:

$$x_i \approx \sum_{j=0}^{n} R_{ij} \cdot f_j(\delta_j, t) \tag{6.3}$$

$$\frac{\Delta p}{p} \approx \left(\frac{1}{\gamma} - \frac{1}{\alpha_p}\right) \cdot \left[\frac{\Delta f}{f} + \sum_{j=0}^{n} \frac{D_j}{C_0^2} f(\delta_j)\right] + \quad \text{h.o.} \tag{6.4}$$

The time dependent response $f_j(\delta_j, t)$ may contain non-linearities and is, due to different cable lengths etc., usually specific for each individual COD. As described in Section 5.2, one non-linearity, for example, is given by the current rate limiter inside the COD power converter. In order to simplify the feedback design it is convenient to decouple equations 6.3 and 6.4 into what is further referred to as *space* and *time domain*.

- The task in the *space domain* is to find a set of steady-state dipole deflection $\underline{\delta}_{ss} = \lim_{t\to\infty} (\delta_1(t), \ldots, \delta_n(t))^T$ that minimises the residual r of

$$r = \left|\underline{z}_{ref} - \underline{z}_{actual}\right| = \left|\mathbf{R} \cdot \underline{\delta}_{ss}\right| < \epsilon \tag{6.5}$$

  with $\mathbf{R} = (R_{ij})$ the orbit response matrix, $\underline{z}_{ref}$ the reference and $\underline{z}_{actual}$ the measured orbit displacement[1]. Using only two correctors the problem is equivalent to finding the minimum of the two dimensional error surface, illustrated in Figure 6.1. Several minimisation methods for the higher dimensional case of more than two correctors are discussed in Section 6.1.

[1] Here, the bars $|\underline{x}|$, usually denoting the 'absolute value' of a complex or real number, are used in the context of matrices in a "relaxed" sense to abreviate the 'p-norm' that has been normalised by the number n of elements in the vector. In a more rigorous mathematical treatment, this 'p-norm' can be written for a vector $\underline{x}$ as

$$|\underline{x}| \equiv \frac{1}{n} \|\underline{x}\|_p = \frac{1}{n} \left(\sum_{i=0}^{n} |x_i|^p\right)^{\frac{1}{p}} \tag{6.6}$$

and for a matrix $\mathbf{A}$ as:

$$|\mathbf{A}| \equiv \|\mathbf{A}\|_p = \max_{\|\underline{x}\|_p = 1} |\mathbf{A}\underline{x}| \tag{6.7}$$

This analysis, unless noted otherwise, uses the special case of the '2-norm'.

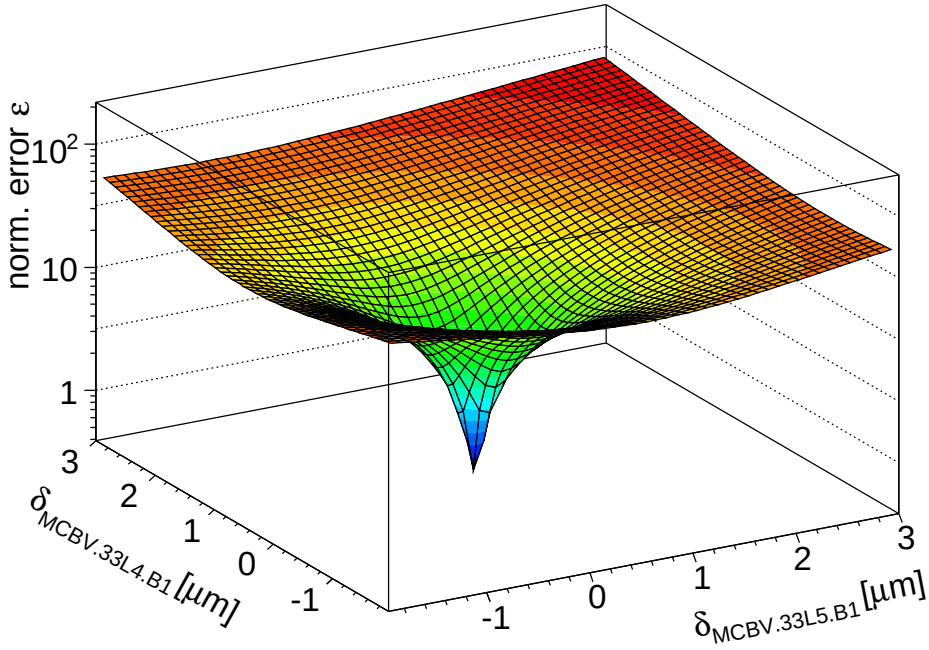


Figure 6.1: Error Surface ε.

- In the *time domain* the task is to find an appropriate control law $u(t)$ that causes a transition of the individual initial to the defined steady-state dipole deflections

$$\lim_{t\to\infty} \sum_{j=0}^{N} u_j(t) * f_j(\delta_j, t) = \underline{\delta}_{ss} \tag{6.8}$$

  while keeping given constraints such as required bandwidth, minimisation of overshoot and robustness with respect to model and measurement errors.

The decoupling makes the choice of orbit correction strategy and the controller adjusting the dipole currents more flexible, particularly in the presence of element failures that require quick adjustments of the feedback controller. The state-space formalism, used for example at SLAC [145], is another approach that solves equations 6.3 and 6.4 as one whole system.

This analysis assumes a sufficiently decoupled machine with the coupling parameter $|C^-|$ about a factor five to ten smaller than the tune split of the machine.

6.1 Space Domain

Orbit correction requires a good knowledge of the beam response to a dipole perturbation which may either rely on the accelerator lattice design or on beam based measurements. A review of common lattice response measurement methods can be found in [32]. Exemplarily, the following methods may be used to determine the corrector-dipole-to-orbit response function:

- Measurement of the column element of R_{ij} using the beam position monitors measuring the difference of the orbit before and after the excitation of each single COD:

  The measurement resolution can be further improved by reducing the jitter through averaging over several orbit acquisitions. The advantage of this method is that it yields a self-consistent response matrix with respect to orbit correction and thus is less prone to possible calibration errors of the monitors, correctors and machine optics. This method is limited by systematic drifts of the orbit and machine optics that occur while averaging the orbits. Further, the measurement time that is required for the excitation and measurement of the response of each individual COD may restrict its use to a few cases due to the large number of CODs (1060) to be tested in the LHC. For example, in the LHC, a reasonable excitation would include displacement of the orbit in the order of 1-2 mm during injection. This excitation corresponds to an approximate COD current of about 0.5 A. In order to reach this current, it would require, neglecting the control system overhead involved, about one second for ramping the magnet up and down plus the time chosen for averaging the two orbits. The averaging period could, for example, be in the order of about 5 s. Due to operational and control system limitations the efficiency may be as low as about 60% yielding to a measurement time of about 15 s per COD excitation and about 4-5 hours total measurement time for all CODs. For the calibration of the BPMs, only a few COD excitations are sufficient and this procedure can be done more quickly.

  In spite of the long total measurement time, this method may be performed in parallel during early LHC commissioning with beam while verifying the polarity of correctors with respect to the beam position monitors. This method does not seem to be suitable for regular verification and re-calibration of the CODs, monitors and machine optics.

- Another common response matrix estimate is based on the turn-by-turn oscillation induced by a single fast dipole kick onto the beam. As shown in [32,146], the phase advance $\Delta\mu$ can be estimated by the phase difference $\Delta\psi$ of the tune peak found in the Fourier spectra between adjacent BPMs. This method requires a minimum number of three consecutive BPMs. The design and measured phase advance between the $i-th$ and $j-th$ monitor is given by $\Delta\mu_{ij}$ and $\Delta\psi_{ij}$. The first order correction to the designed betatron function β at the first BPM is then given by ([146])

$$\beta = \beta_0 \left(\frac{\cot(\Delta\psi_{12}) - \cot(\Delta\psi_{13})}{\cot(\Delta\mu_{12}) - \cot(\Delta\mu_{13})} \right) \tag{6.9}$$

  It is visible that equation 6.9 may become singular if the phase advance of the selected BPMs is a multiple of $\frac{\pi}{2}$. For this reason, this method was less favourable for accelerators such as the SPS or LEP. In the LHC, however, each quadrupole is equipped with a BPM that measures in both planes thus resulting in an average phase advance of BPMs of $\pi/4$. Equation 6.9 can be used iteratively till the required level of convergence for β is reached.

  This method can be performed much faster than methods relying on COD excitation. The natural damping of the betatron turn-by-turn oscillation usually limits the useful data to less than about 1000 turns corresponding to a measurement time of about less than 100 ms. In practise, the time for this measurement is limited by the time required for the read-out of the BPMs and processing - essentially (fast) Fourier analysis - of the turn-by-turn BPM data. The response matrix can consequently be computed using the closed orbit response due to a dipole kick given in equation 2.71 and the reconstructed values of the betatron function and phase advances obtained from equation 6.9.

  The advantage of this method is the relative short measurement time and insensitivity to BPM or COD calibration errors which is favourable for some accelerator studies and operational scenarios where the COD and BPM calibration are known and the optics is suspected to have changed.

- The dispersion function is measured through changing the relative beam momentum $\Delta p/p$ and measuring the relative beam displacement $\Delta \underline{z}_D$ at the location of the BPMs. The momentum is

changed by changing the relative frequency $\Delta f/f$ of the LHC RF cavities as given by equation 2.51. The dispersion function is given by

$$\underline{D} = \frac{\Delta \underline{z}_D}{\Delta p/p} = \frac{\Delta \underline{z}_D}{\eta^{-1} \cdot \Delta f/f} \tag{6.10}$$

The measurement can be performed fairly quickly, and since the RF frequency change can be measured very accurately, the accuracy is essentially limited by the time window used for averaging and systematic drifts that may occur during the measurement. The advantage of this technique it the insensitivity to calibration errors and systematics of the BPMs. The measurement can be performed within less than a minute for both beams and, apart from the required adiabatic change of the RF frequency, is only limited by the control system overhead involved in changing the accelerator parameters.

6.1.1 Closed Orbit Bumps

The choice of orbit correction techniques and algorithm depends on the type of stability requirement and assumption on the perturbation sources that affect the orbit.

For instance, the requirement of local stability at few insertion devices only, leads to the use of so-called *three-* or *four-corrector bumps* that control the orbit at given locations in the machine while leaving the rest of the ring unaffected. Depending on whether only the position or the position and angle at a given location should be controlled, the bump consists of one BPM and three correctors (three-corrector-bump) or two BPMs, measuring the position and angle, and four correctors (four-corrector bumps). The required kick strength can be found in the Annexe B.1. Figure 6.2 shows an exemplary three corrector bump in the LHC arc. The largest kicks are found at the first and last correctors. It is visible that the

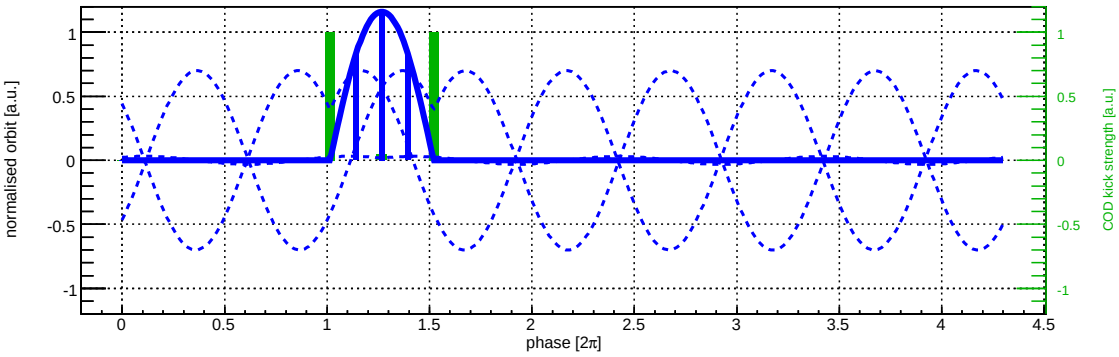


Figure 6.2: Example of a three corrector bump in the LHC arc. The residual orbit is plotted as a function of phase advance. The discrete BPM measurements and COD deflections are shown as bars. It is visible that the superimposition of the individual dipole's closed orbit response vanishes outside the bump region. In the LHC, three corrector bumps in the arc are sampled at three monitors.

superimposition of the individual dipole's closed orbit response vanishes outside the bump region. For a one-to-one steering, the bump is usually controlled by the position of the central BPM between the outermost CODs, for example. For lattices with BPM spacing of $\pi/2$ such as the SPS, this scheme could, in principle, be extended to a global correction scheme by enqueuing several three-corrector bumps next to each other.

In case of the LHC arcs, a three-corrector bump is usually sampled by at least three BPMs which, in order to avoid cross-talk between neighbouring bumps, makes two of the three BPMs superfluous using this scheme. Further, since the COD strength is driven by the position at a single BPM, this scheme is sensitive with respect to BPM failures and errors. Still, the three-corrector bump and its extension is still used for specific cases where the control of the orbit at a single location is required.

6.1.2 MICADO

For global orbit stability, the chosen algorithms commonly differ in the assumption on the locality of the perturbation sources. Depending on whether the errors are due to a few or a large number of distributed sources, the global orbit correction algorithm may use only a few or all available orbit correctors. Two commonly used algorithm are MICADO and SVD: MICADO (Minimisation des CArrés des

DISTORTIONS D'ORBITE) is based on the global orbit optimisation using a few orbit corrector dipoles with rather large deflections and SVD (Singular-Value-Decomposition) using all available correctors with usually small individual deflections. The SVD correction scheme is further discussed below.

As described in [147], MICADO searches for a set of "most effective correctors", which minimise the RMS orbit in all BPMs using a minimum ("most effective") corrector kick by means of the SIMPLEX algorithm [2]. Commonly MICADO exploits Householder transformation in order to reduce the numerical complexity of the minimisation problem

The algorithm consists of two steps that may be repeated iteratively till the required level of tolerance is reached:

1. Find the most effective dipole corrector minimising the residual

$$r_i = |\Delta \underline{z} - \mathbf{R} \cdot \underline{\delta}_i| \tag{6.11}$$

   using one corrector only ($\underline{\delta}_i = (0, \ldots, \delta_i, \ldots, 0)^T$). After this step, only the COD yielding the smallest residual r_i is selected, the other are discarded.

2. Minimise the equation using the SIMPLEX method while varying the last and all previously selected CODs.

The number of correctors (iterations) is a free parameter of the algorithm. Due to the pivot selection in the first step, the algorithm is effective in identifying single perturbation sources such as single large quadrupole shifts. For further analysis, the *attenuation* G of the orbit correction is defined as

$$G_{corr} = \frac{\left|\underline{z}_{initial} - \underline{z}_{final}\right|}{|\underline{z}_{initial}|} \tag{6.12}$$

with $\underline{z}_{initial}$ the initial perturbed orbit, $\underline{z}_{final}$ the corrected orbit and $|\underline{z}|$ the orbit r.m.s. given by the 2-norm of the vector divided the number of elements. After comparison with equation 3.29, it is visible that the attenuation (equation 6.12) corresponds to the *disturbance rejection* of the correction algorithm and is thus also referred to as *gain* of the correction. The residual orbit r.m.s. r is accordingly related to the initial perturbation causing the orbit shift $\Delta \underline{z}$ as:

$$r = G_{corr} \cdot \Delta \underline{z} \tag{6.13}$$

If required, the r.m.s. calculation is performed only for BPM in regions of interest. Figure 6.3 shows an example of the achievable attenuation of the MICADO algorithm correcting the vertical LHC orbit of Beam 1 and 2 using all BPMs. It is visible that already a few orbit correctors can decrease the residual orbit significantly. As further discussed below, the attenuation also depends on optics imperfection and decreases for increasing peak-to-peak beta-beat (optics failures). MICADO is used for initial correction during machine commissioning and situation where a single or a few elements are suspected to malfunction or to be the cause of the initial orbit perturbation.

The draw-back of this algorithm is that it is sensitive to so-called *singularities* of the response matrix. Singularities may occur due to correctors that are located close to each other or with similar phases that cannot be resolved by the BPMs and accordingly produce similar orbit response. MICADO usually produces large corrector strength in these cases that eventually may exceed the physical COD limitations. As a first order compensation, the response matrix is often "conditioned" by removing correctors that may lead to these singularities. As a side-effect of this conditioning, these correctors are not used to minimise local orbit perturbations or to minimise the overall COD r.m.s. deflections. From a control theory point of view, these singularities correspond to unstable poles of the transfer function.

The "conditioning" usually employs elaborate algorithms which require continuous adjustments in the presence of multiple BPM or COD failure. Hence, beside its cubic numerical complexity further discussed below, it is not ideal for fully automated feedback control systems. Nowadays, nearly all continuous orbit correction schemes rely on SVD based approaches [77, 149, 150].

[2] The SIMPLEX algorithm is an improved version of the classic Gaussian elimination [148].

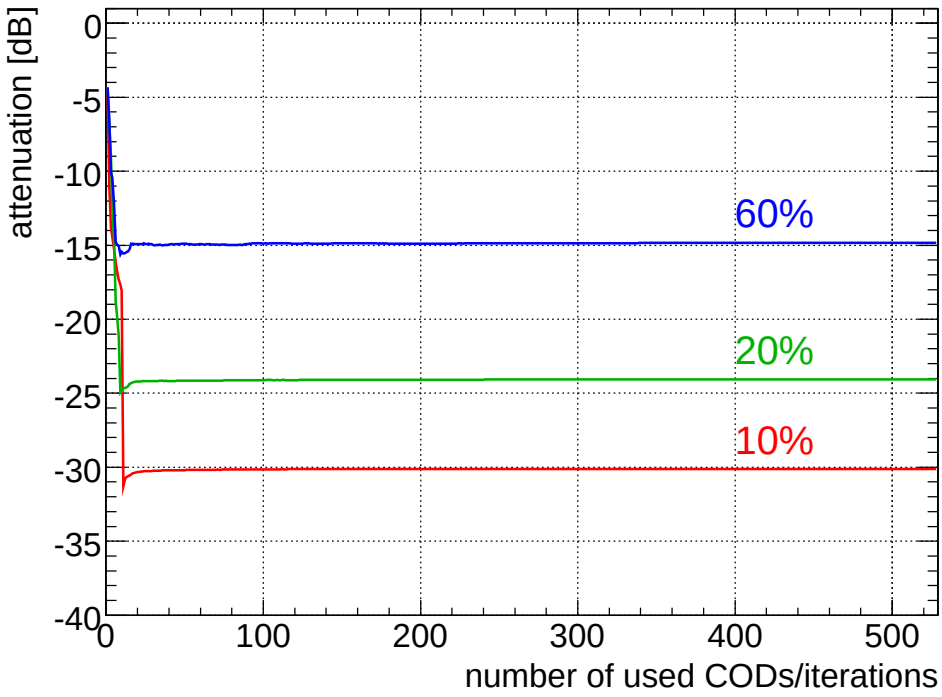


Figure 6.3: Attenuation of the MICADO based vertical orbit correction during injection as a function of used CODs and peak-to-peak beta-beat optics imperfection. The correction is performed for both beams using all available monitors and based on a seed of 100 different orbits per COD iteration.

6.1.3 Singular-Value-Decomposition (SVD)

The Singular-Value-Decomposition (SVD) is a generalisation of the eigenvalue decomposition of matrices [151]. It extends and generalises classic methods such as the historically important Jacobi eigenvalue decomposition [152] to the general case of non-square matrices.

As shown in [151], any matrix $\mathbf{R} \in \mathbb{R}^{m\times n}$ with $m \geq n$ can be decomposed into

$$\mathbf{R} = \mathbf{U} \cdot \lambda \cdot \mathbf{V}^T \tag{6.14}$$

with $\mathbf{U} \in \mathbb{R}^{m\times n}$ being a dense unitary matrix, $\lambda = diag(\lambda_1, \ldots, \lambda_n)$ a diagonal matrix holding the eigenvalues of $\mathbf{R}$ and $\mathbf{V} \in \mathbb{R}^{n\times n}$ an orthogonal matrix containing the eigenvectors of $\mathbf{R}$ in its columns. The decomposition is illustrated in Figure 6.4. Without limiting the universality, it is further assumed

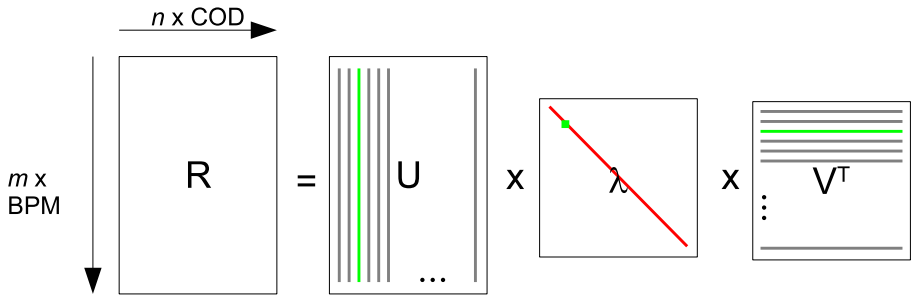


Figure 6.4: Schematic of the Singular Value Decomposition (SVD): An arbitrary real matrix $\mathbf{R}$ with dimension $m \times n$ can be decomposed into a rectangular matrix $\mathbf{U} \in \mathbb{R}^{m\times n}$, diagonal eigenvalue matrix $\lambda = diag(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n\times n}$ and transpose eigenvector matrix $\mathbf{V}^T \in \mathbb{R}^{n\times n}$.

that the eigenvalues λ_i are positive and sorted according their size and the columns and rows of $\mathbf{U}$ and $\mathbf{V}$ swapped and adjusted accordingly. The matrices and eigenvalues are related through the following

eigenvalue relations:

$$\lambda_i \underline{u}_i = \mathbf{R} \cdot \underline{v}_i \quad (6.15)$$
$$\lambda_i \underline{v}_i = \mathbf{R}^T \cdot \underline{u}_i \quad (6.16)$$

and

$$\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} \quad (6.17)$$
$$= \mathbf{V}\mathbf{V}^T \quad (6.18)$$
$$= \mathbf{1} \quad (6.19)$$

with $\mathbf{1}$ the $n \times n$-dimensional unity matrix. Further, as a consequence of the orthogonality of the columns of $\mathbf{U}$, the following holds:

$$\mathbf{U}\mathbf{U}^T = \mathbf{1} \in \mathbb{R}^{m \times m} \quad (6.20)$$

In the context of orbit correction, the columns $\underline{u}_i$ of $\mathbf{U}$ can be interpreted as *orbit eigenvectors*, an orthogonal basis spanning the range of all possible orbits that can be generated with the given number and configuration of closed orbit correctors. A selection of orbit eigenvectors is shown in Annexe C. It is visible that for large eigenvalues (low index) the vectors correspond to solutions with many nodes similar to free betatron oscillations. For small eigenvalues (large index), the orbit eigenvectors become more local and alike to few orbit corrector bumps. This relation is visible in the spacial distribution of the COD kicks linked to a single arc BPM shown in Figure 6.5. A graphical visualisation of the inverse orbit response

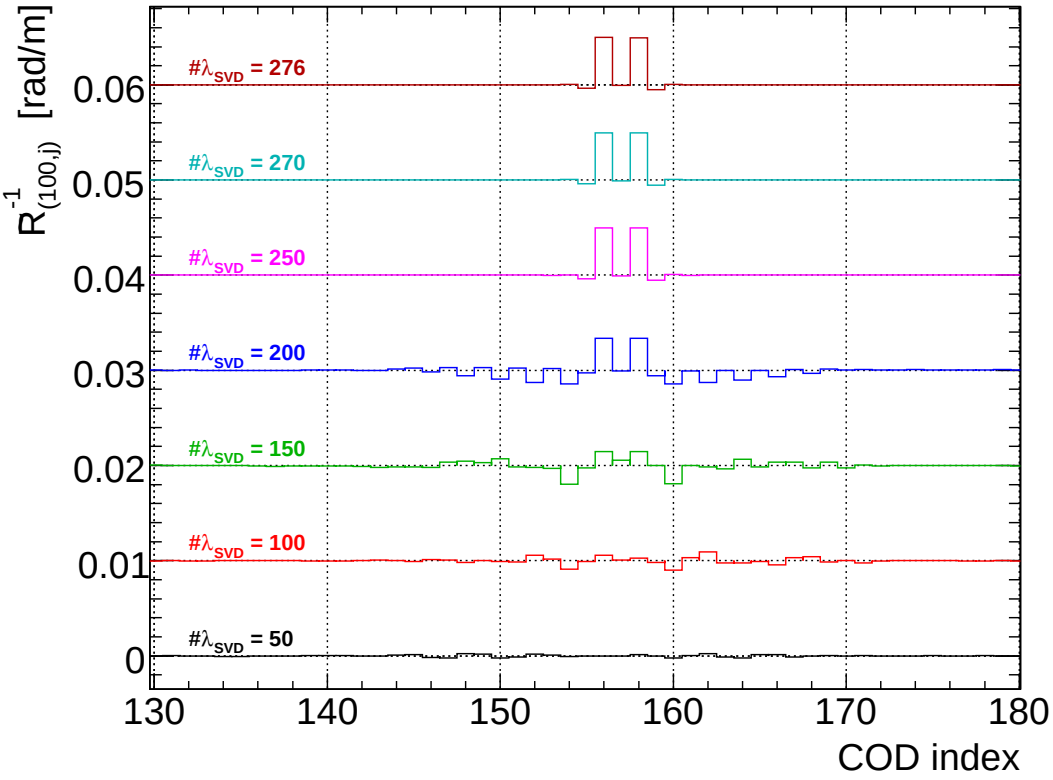
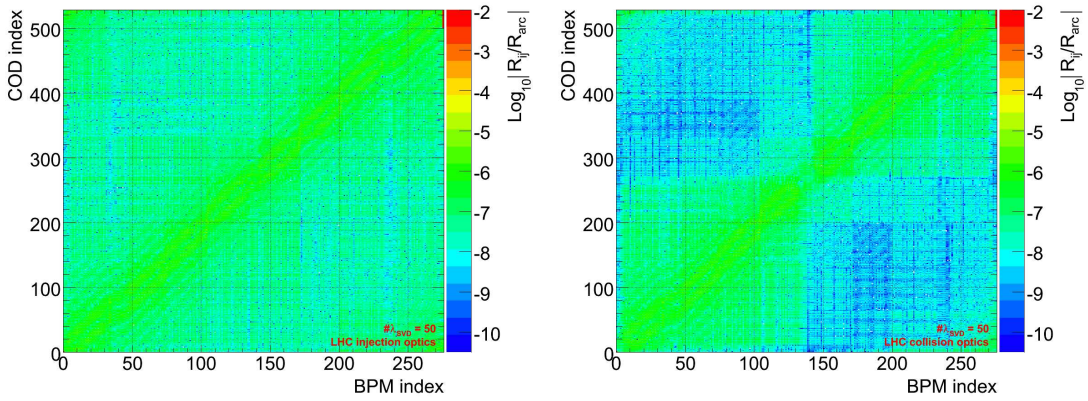


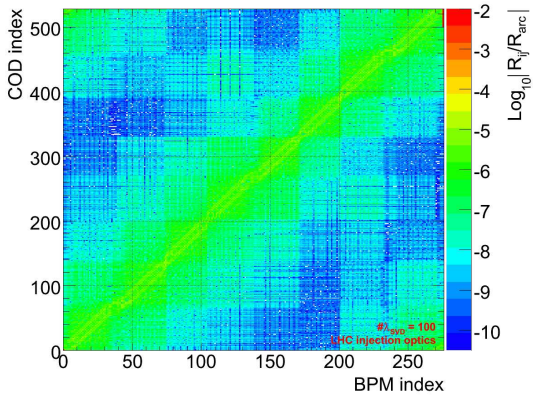
Figure 6.5: COD corrector pattern as a function of COD position and number of used eigenvalues for the inversion. The SVD inversion was performed for beam 1 only. For clarity, the individual COD correction patterns are shifted vertically by a fixed amount for each number of eigenvalues chosen.

matrix is shown in Figure 6.6. For better visibility the matrix is shown for only one plane and beam, and the individual elements were normalised to the response of a typical arc BPM. The logarithm is taken of the final value to reduce the range of orbit responses with respect to BPMs/CODs located in the arc and inside the insertion region. The eight-fold symmetry and insertions of the LHC are visible. While using more eigenvalues for the computation of the pseudo-inverse response matrix, the COD strength compensating for a orbit offset at a given BPM becomes more larger and the correctors used for the bump are found in the local vicinity of the BPM.

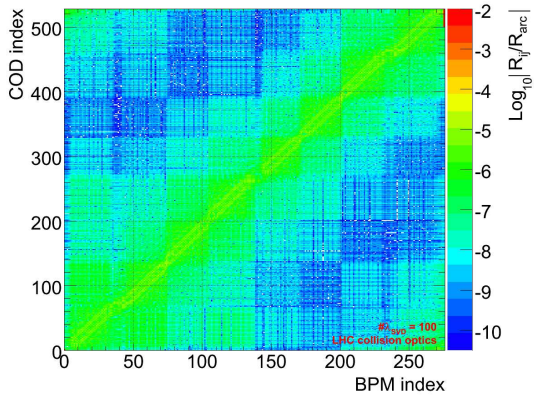


(a) injection optics, $\#\lambda_{SVD} = 50$

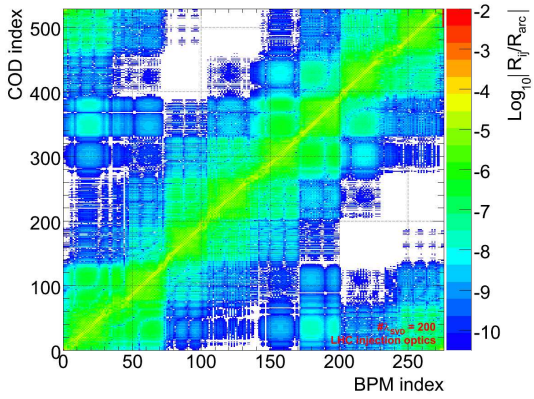
(b) collision optics, $\#\lambda_{SVD} = 50$



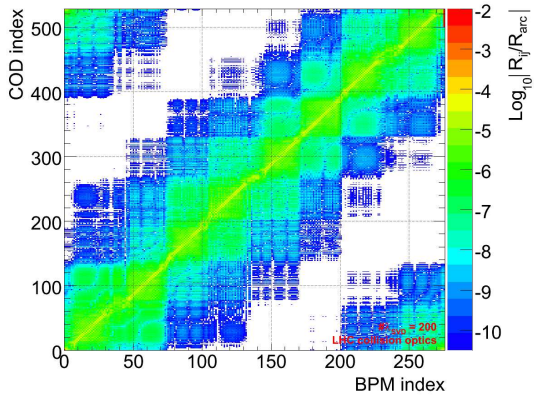
(c) injection optics, $\#\lambda_{SVD} = 100$



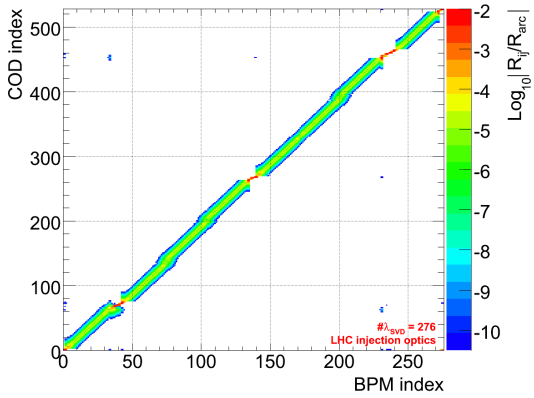
(d) collision optics, $\#\lambda_{SVD} = 100$



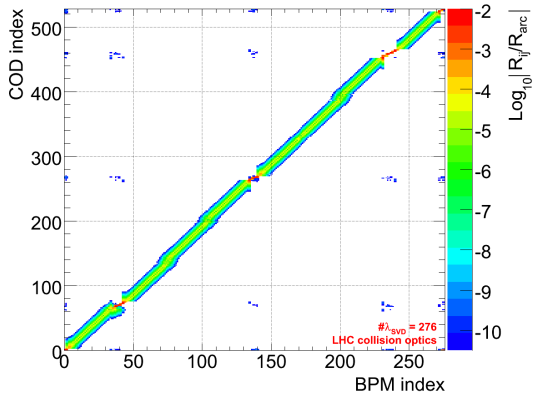
(e) injection optics, $\#\lambda_{SVD} = 200$



(f) collision optics, $\#\lambda_{SVD} = 200$



(g) injection optics, $\#\lambda_{SVD} = 276$



(h) collision optics, $\#\lambda_{SVD} = 276$

Figure 6.6: SVD inverted horizontal injection (left) and collision (right) orbit response matrix for a selected number of used eigenvalues $\#\lambda_{SVD}$.

The strength of the SVD algorithm lies in the property that the decomposition can be used to compute a so-called *pseudo-inverse* response matrix $\tilde{\mathbf{R}}^{-1}$ that describes the required steady-state COD strengths $\underline{\delta}$ as a function of the measured orbit error $\Delta\underline{z}$:

$$\underline{\delta} = \tilde{\mathbf{R}}^{-1} \cdot (\underline{z}_{meas} - \underline{z}_{ref}) \tag{6.21}$$

$$= \tilde{\mathbf{R}}^{-1} \cdot \Delta\underline{z} \tag{6.22}$$

with

$$\tilde{\mathbf{R}}^{-1} = \mathbf{V} \cdot \lambda^{-1} \cdot \mathbf{U}^T \tag{6.23}$$

In the first instance, it was assumed (equation 6.23) that all eigenvalues are non-zero and thus their inverse is simply given by:

$$\lambda_i^{-1} = \frac{1}{\lambda_i} \tag{6.24}$$

As an important property of the SVD algorithm, the presence of singularities can be easily removed by setting the inverse of the eigenvalues that are close or equal to zero yielding the further called '*pseudo*-inverse'. The pseudo-inverse of the eigenvalues can be written as

$$\tilde{\lambda}_i^{-1} := \begin{cases} \frac{1}{\lambda_i} & \text{if} \quad \lambda_i > \lambda_{cut} \\ 0 & \text{if} \quad \lambda_i \leq \lambda_{cut} \end{cases} \tag{6.25}$$

with λ_{cut} an arbitrary cut-parameter. Alternatively, since the eigenvalues are assumed to be sorted according to their values, one can equivalently choose to remove those eigenvalues whose index is larger than the cut index $\#\lambda_{SVD}$.

$$\tilde{\lambda}_i^{-1} := \begin{cases} \frac{1}{\lambda_i} & \text{if} \quad i > \#\lambda_{SVD} \\ 0 & \text{if} \quad i \leq \#\lambda_{SVD} \end{cases} \tag{6.26}$$

The attenuation of the SVD based orbit correction is exemplarily shown in Figure 6.3 as a function of the cut parameter $\#\lambda_{SVD}$ for a vertical global orbit correction of both LHC beams during injection. It

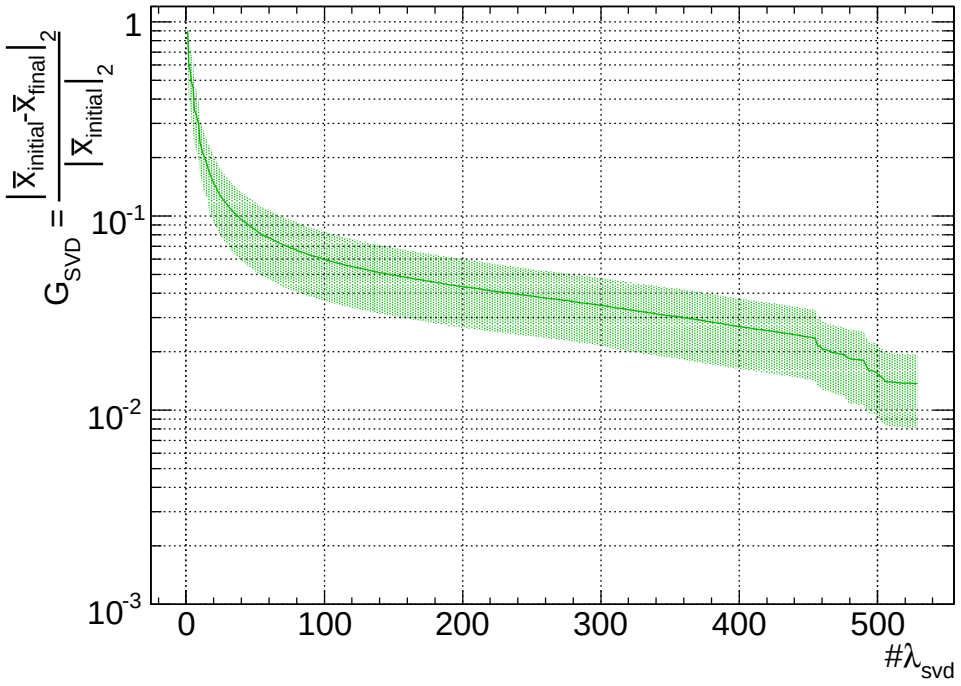


Figure 6.7: Attenuation of the SVD based vertical orbit correction during injection as a function of the cut parameter $\#\lambda_{SVD}$. The correction is performed for both beams using all available monitors and based on a seed of 100 different orbits per COD iteration.

is visible that the residual orbit r.m.s. decreases with an increasing number of used eigenvalues and is always less than one. The SVD orbit correction simulation is based on a seed of 10^5 different orbits. The

resulting r.m.s. width of the correction gain is indicated as an error band. The non-zero correction limit is given by shifts of defocusing quadrupoles that are usually not equipped with dipole correctors and thus can not be fully compensated.

An important feature using SVD follows from the decomposition inverse response matrix: the pseudo-inverse response matrix given in equation 6.22 not only minimises the residual orbit given in equation 6.5 but also the r.m.s. of the required corrector strength for the given orbit. This property is helpful since the cut parameter $\#\lambda_{SVD}$ can also be used to limit the average expected corrector deflection. The corresponding r.m.s. COD deflections in Figure 6.7 are shown as a function of used eigenvalues $\#\lambda_{SVD}$ in Figure 6.8. The required r.m.s. COD deflection increases with the number of used eigenvalues and

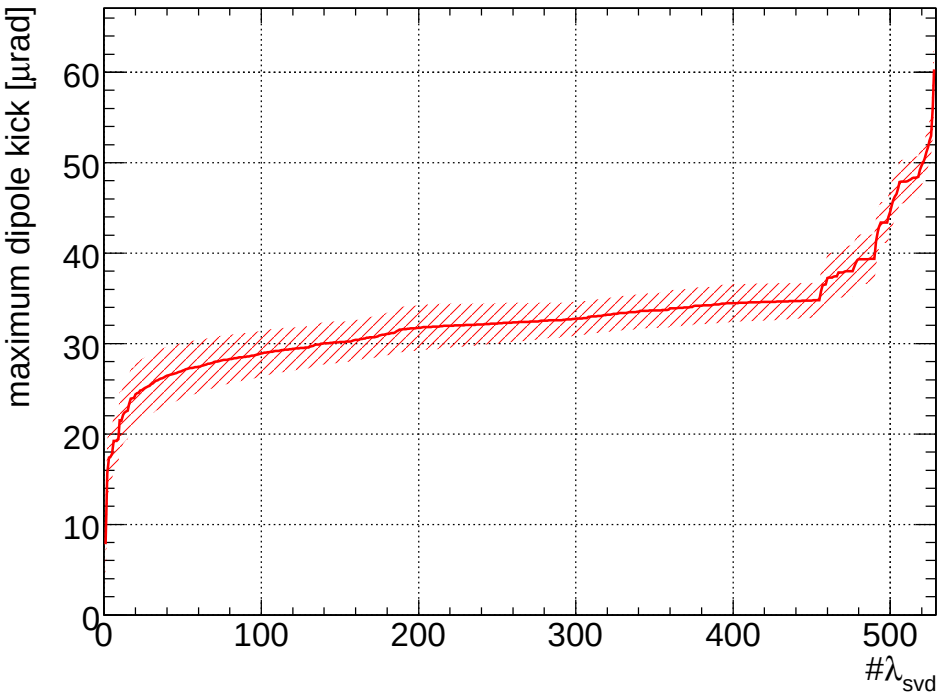


Figure 6.8: R.M.S. COD deflection strength of the SVD based vertical orbit correction during injection as a function of the cut parameter $\#\lambda_{SVD}$. The correction is performed for both beams using all available monitors, based on a seed of 10^5 orbits due to a random quadrupole misalignment of 0.5 mm r.m.s. The corresponding orbit attenuation is shown in Figure 6.7. For comparison, the maximum arc corrector strength is about 80 μrad at 7 TeV.

becomes very large for eigenvalues ($\#\lambda_{SVD} > 430$) corresponding to near-singular orbit eigenvectors describing local bumps inside the experimental insertions. This is also visible in the rapidly increasing COD strength shown in Figure 6.8. It is visible that using $\#\lambda_{SVD} \approx 50$ eigenvalues only, the resiudal orbit r.m.s. can be corrected and attenuated by a factor of about ten. The first 50 eigenvalues are global betatron pattern like oscillations. The eigenvalues between $\#\lambda_{SVD} \approx 50$ and $\#\lambda_{SVD} \approx 450$ correspond to larger bumps spreading over a given arc segment or non-experimental long straigt sections (IR3, IR4, IR6 and IR7).

Controllability and Observability

Evidently, the cut parameter has to at least remove eigenvalues equal to zero or those close to the limits of the numeric number representation. Strictly speaking, since all eigenvalues are well above about 0.1, the LHC orbit response matrix and its pseudo-inverse does not contain numerical singularities. Since the orbit is sampled at every quadrupole, the absence of true singularities in the beam response matrix is sufficient to guarantee the controllability of the orbit, as the highest betatron values and thus orbit excursions are always found and constrained at the location of the quadrupoles. Thus in case all BPMs, CODs and eigenvalues are used, the control problem itself does not contain 'unstable poles'.

Sometimes it is favourable to deliberately remove eigenvalues in order to minimise the locality of the orbit correction and to reduce sensitivity to noise and failures of individual BPMs. These removed eigenvalues and orbit eigenvectors cause accumulative orbit distortions that are not corrected by the

feedback system. In the case of the LHC orbit response eigenvalue spectrum (Figure 6.9), one may choose $\lambda_{cut} \approx 80$ to remove those eigenvalues with indices larger than about $\#\lambda_{SVD} = 480$, corresponding to solutions located at the experimental insertions (see Annexe C). The rise time of these special, removed orbit perturbation patterns are usually slower compared to those with larger eigenvalues. In practise, these small drifts are tolerated and corrected using all eigenvalues once given tolerances are exceeded.

Numerical Condition

Almost all orbit correction schemes are implemented digitally and thus have to handle quantised representations of the real beam parameters. An additional source of errors may be induced if the optimal varies from the computed correction due to a reduced number of *bits* in the binary representation.

The *condition number* is a common measure to quantify the susceptibility to numerical errors and is used as an indication of the required dynamic range during computation. The ratio between the largest $\lambda|_{max}$ and smallest eigenvalue $\lambda|_{min}$ can be used as an estimate of the *condition* C_R of the matrix ([153]):

$$C_R \approx \left|\mathbf{R}^{-1}\underline{z} \cdot (\mathbf{R}\underline{\delta})^{-1}\right| \tag{6.27}$$

$$C_R \approx \left|\frac{\lambda_{max}}{\lambda_{min}}\right| \tag{6.28}$$

The eigenvalue spectrum of the decomposition of LHC injection and collision orbit response matrices including both beams is exemplarily shown in Figure 6.9. The eigenvalues cover a range of about six

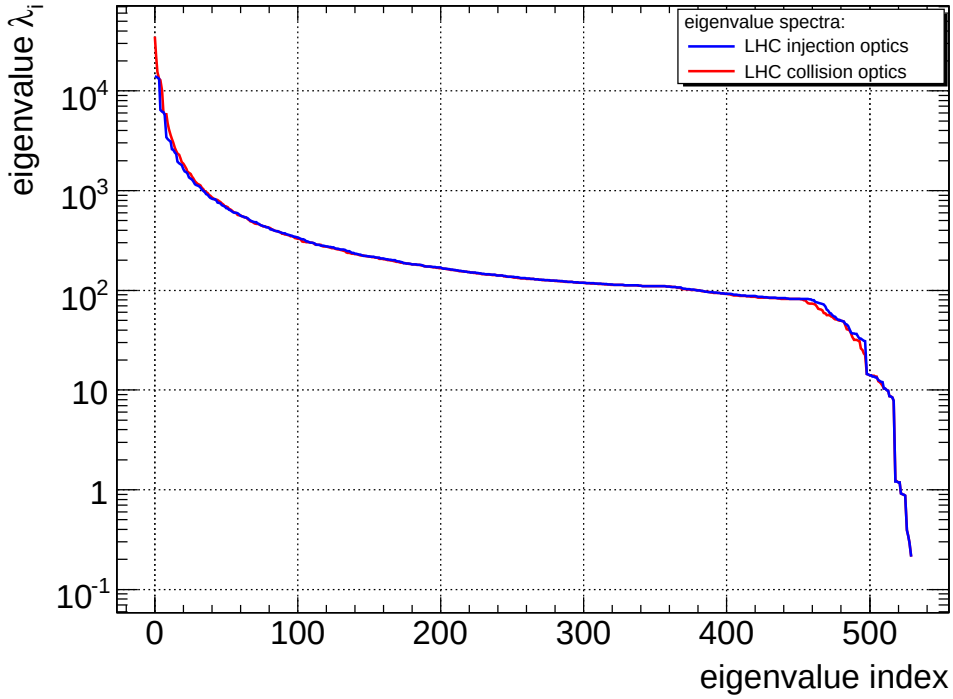


Figure 6.9: Eigenvalue spectrum of the COD-to-BPM response matrix for the LHC injection and collision optics. The matrix takes into account the response for Beam 1 and 2. Near singular eigenvalues are located above about 460 eigenvalues.

orders of magnitude. As shown in [153], the logarithm with base '2' of the condition number

$$n_{bits} \approx \log_2(C_R) \tag{6.29}$$

can be used to estimate the number of digits n_{bits} that are lost during the inversion of the matrix. In case of the SVD of the response matrix that describes the LHC orbit using all available CODs and BPMs in the given plane, roughly 20 significant bits are lost during the decomposition. Thus all of the SVD computations and transformations are performed in double precision (64 bit floating point).

Numerical Complexity

The numerical complexity and the involved computation affects the overall feedback performance through introducing delays. This analysis uses the SVD implementation by 'Golub and Reinsch'. As shown

in [151], the complexity for the Gauss elimination and SVD algorithm is approximately given by:

$$\mathcal{O}_{Gauss} = \frac{1}{2}mn^2 + \frac{1}{6}n^3 \tag{6.30}$$

$$\mathcal{O}_{SVD} = 2mn^2 + 4n^3 \tag{6.31}$$

Assuming square matrices ($m = n$), it is visible that the SVD is about seven times more expensive than the traditional Gauss elimination method. Depending on the computing architecture, the SVD algorithm may become limeted either by memory bandwidth or by the processor's maximum floating point processing capabilities. For example, an Intel Xeon processor with 2.8 GHz CPU and 400 MHz memory clock frequency requires about 15 s for an inversion of a full LHC orbit response matrix ($m \times n = 1056 \times 530$) which is long compared to the targeted feedback sampling frequencies of about 25 to 50 Hz. Once the pseudo-inverse response matrix is computed, the actual correction can be performed by a simple matrix-vector multiplication that requires $n \times m$ multiplications. The optimised computation of the matrix-vector product requires about 12 ms on the above machine specified which makes the SVD correction much faster than the MICADO algorithm.

In case only a small fraction of the most dominant eigenvalues are used for the computation of the steady-state COD deflections, the following alternative description of equation 6.23 can be used

$$\underline{\delta}_{ss} = \sum_{i=0}^{\#\lambda_{SVD}} \frac{a_i}{\lambda_i}\underline{v}_i \qquad \text{with} \qquad a_i = \underline{u}_i^T \cdot \Delta\underline{z} \tag{6.32}$$

with an approximate numerical complexity of:

$$\mathcal{O} = n \cdot m \qquad \longrightarrow \qquad \mathcal{O} = \#\lambda_{SVD} \cdot m \tag{6.33}$$

As visible in equation 6.32, the individual corrector kick strength is approximately proportional to the inverse of the eigenvalue λ_i – since $|\underline{v}_i| = 1$ – stressing the relationship between small eigenvalues and large corrector strengths.

6.1.4 Sensitivity to Machine Imperfections

Besides the convergence and residual orbit, it is important to evaluate the robustness of the orbit correction algorithm in the presence of model uncertainties such as errors in the accelerator lattice (beta-beat), noise of the beam position measurement and failures of individual elements particularly dipole corrector magnets and beam position monitors.

Using a perfect orbit response matrix, it was shown that the pseudo-inverse response matrix yields optimal steady-state corrector deflections for a given number of used eigenvalues. Consequently, any further orbit correction gives the same corrector settings and residual orbit independent of the number of iterations performed as shown below. Assuming the orbit perturbations are constant in time, the residual r_n orbit, given by equation 6.5, can for the first and the $n-th$ consecutive orbit correction steps be written as:

$$\begin{aligned}
r_1 &= |\Delta\underline{z}_0 - \Delta\underline{z}_{corr}(0)| \qquad (6.34)\\
r_1 &= |\Delta\underline{z}_0 - \mathbf{R}\underline{\delta}_0| \\
r_1 &= \left|\Delta\underline{z}_0 - \mathbf{R}\tilde{\mathbf{R}}^{-1}\Delta\underline{z}_0\right| \\
r_1 &= \left|(\mathbf{1} - \mathbf{R}\tilde{\mathbf{R}}^{-1})\Delta\underline{z}_0\right| = |\Delta\underline{z}_1|
\end{aligned}$$

$$r_2 = \left|\Delta\underline{z}_1 - \mathbf{R}\tilde{\mathbf{R}}^{-1}\Delta\underline{z}_1\right| \tag{6.35}$$

$$[\ldots]$$

$$r_n = \left|(\mathbf{1} - \mathbf{R}\tilde{\mathbf{R}}^{-1}))^n \cdot \Delta\underline{z}_0\right| \tag{6.36}$$

with $\Delta\underline{z} = \Delta\underline{z}_0$ the difference of actual to the reference orbit in the '0-th' iteration.

$$r_n = \left|(\mathbf{1} - \mathbf{R}\tilde{\mathbf{R}}^{-1}))^n \cdot \Delta\underline{z}\right| \tag{6.37}$$

$$= \left|(\mathbf{1} - \mathbf{U}\lambda\mathbf{V}^T \cdot \mathbf{V}\tilde{\lambda}\mathbf{U}^T)^n \cdot \Delta\underline{z}\right| \tag{6.38}$$

$$= \left|(\mathbf{1} - \mathbf{U}\lambda\tilde{\lambda}\mathbf{U}^T)^n \cdot \Delta\underline{z}\right| \tag{6.39}$$

In equation 6.38, the orthogonality of the eigenvector matrix $\mathbf{V}$ was used. The product of the eigenvalue and pseudo-inverse eigenvalue matrix is given by

$$\lambda_i \tilde{\lambda}_i^{-1} = \begin{cases} 1 & \text{if} \quad i < \#\lambda_{SVD} \\ 0 & \text{if} \quad i \geq \#\lambda_{SVD} \end{cases} \tag{6.40}$$

which vanishes for the removed singular orbit solution, and is equal to one for the others. Accordingly the orbit eigenvectors corresponding to singular or deselected near-singular solutions are projected onto the null space of the range of the total beam correction transfer matrix $\mathbf{R}\tilde{\mathbf{R}}^T$. Thus consecutive application of the SVD orbit correction leads to the same residual orbit independent of the number of iterations.

A deviation of the real accelerator orbit response matrix, due to optics imperfection or calibration errors from the response matrix used to compute the pseudo-inverse response matrix may cause a decrease in the achievable correction gain. Further, the orthogonality relations of the SVD between the decomposed eigenvalues and eigenvalue matrices are not necessarily preserved between the ideal and perturbed (primed) lattice orbit response:

$$\lambda'_i \neq \lambda_i \tag{6.41}$$

$$\mathbf{V}' \cdot \mathbf{V}^T \neq \mathbf{1} \tag{6.42}$$

$$\mathbf{U}' \cdot \mathbf{U}^T \neq \mathbf{1} \tag{6.43}$$

Accordingly, the actual SVD correction gain depends on the number of orbit correction iteration steps. Provided that the individual correction steps are convergent, the limes of the residual after an infinite number of orbit corrections eventually converge to the same final COD state (settings) as for the case using the ideal response matrix. The orbit residual usually differs though.

The SVD-based inversion algorithm can also be seen as a gradient-based minima search on an n-dimensional hyper-surface, as illustrated in Figure 6.1. If the true orbit response matrix differs from the one used for the correction, the error gradient may point off the true minimum as illustrated in Figure 6.10. The contour lines correspond to constant values of the residual orbit error r. The correction is based on the LHC arc lattice and shows the projection of two selected COD settings. It is visible that

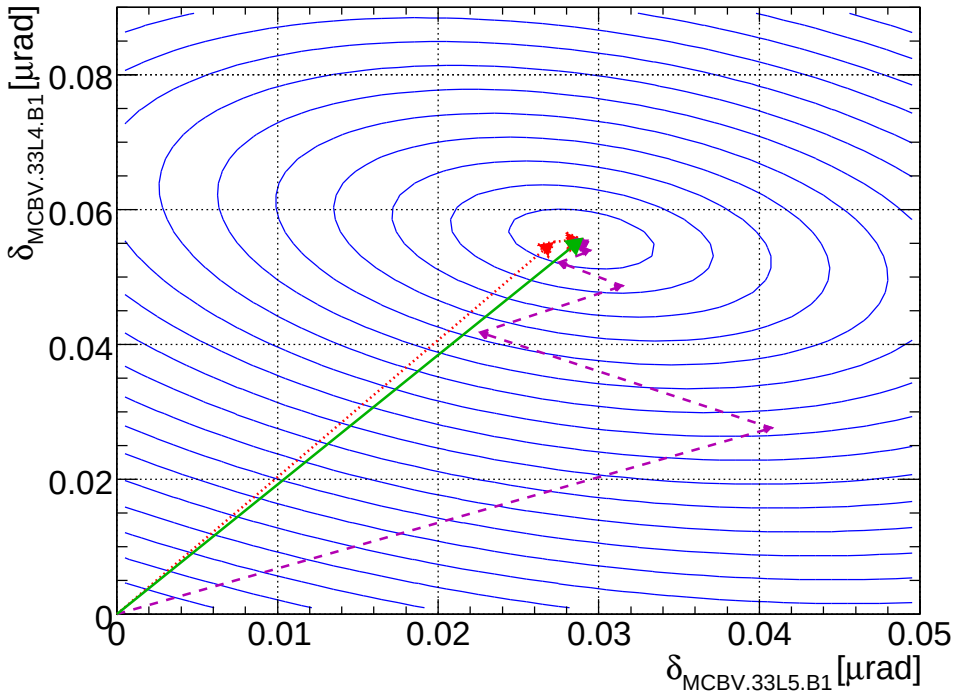


Figure 6.10: SVD error contour: The perfect correction (solid green), the correction with optics error (dotted red) and COD calibration errors (dashed violet) are indicated. The contours represent constant values of the residual orbit error. Independent of the type of error, the corrections converge to the same optimal setting. However, the number of iterations required varies.

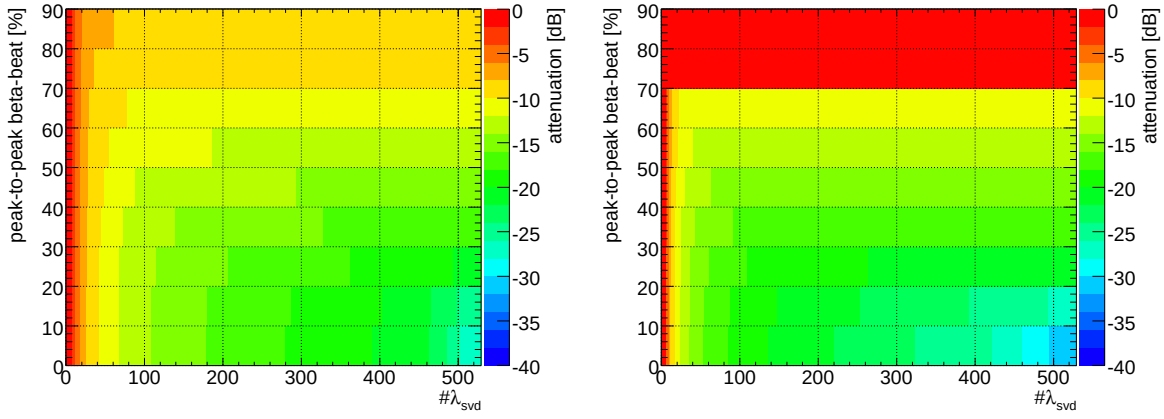
independent of the type of error, the corrections converge to the same steady-state setting. Using the

matched machine response for the correction, the gradient vector points to the global minimum after the first iteration. For an imperfect machine lattice with 20% global peak-to-peak beta-beat, about two iterations are required. In case of a calibration error of two of the used CODs (+20% and -20%, respectively), about 7 SVD iterations are required.

The number of iterations required to reach the same level of convergence estimates the effective reduction of the correction bandwidth. Since this bandwidth reduction is independent of the BPM detection and ramping speed of the CODs, it can be improved by simply increasing feedback sampling frequency.

Sensitivity to Beta-Beating

Figures 6.11 and 6.12 show the disturbance rejection (attenuation) dependence as a function of BPM measurement noise and as a function of beta-beat model errors. The simulated lattice errors include beta-beat of the accelerator lattice and final focus squeeze inside the experimental insertions. The attenuation improves with increasing number of used eigenvalues $\#\lambda_{SVD}$ and decreases with peak-to-peak beta-beating. The orbit correction converges up to a level of 100% peak-to-peak beta-beat for the injection optics. The maximum peak-to-peak beta beating defines a lower limit since MADX, used to compute the lattice optics parameter, was not able to generate or produce converging lattices with more than about 100% beta-beating. For the collision optics, the SVD correction is stable in both cases up to lattices with about 70% peak-to-peak beta-beating, although, with reduced effective bandwidth. The simulations were performed with a seed of 10^5 orbits per eigenvalue and 10% steps of beta-beating. Figure 6.12 shows a



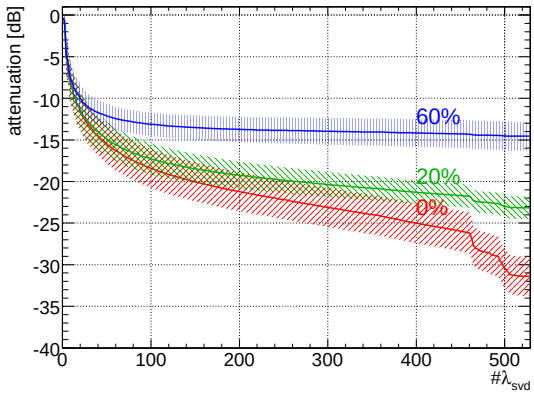
(a) injection optics (b) collision optics

Figure 6.11: Attenuation of the SVD based orbit correction as a function of used eigenvalues $\#\lambda_{SVD}$ and peak-to-peak beta-beat. The attenuation for the injection (left) and collision (right) optics is plotted for the vertical plane. The correction is performed for both beams using all available correctors and monitors.

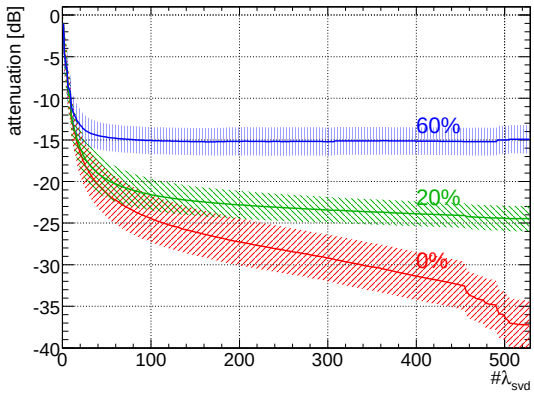
detailed overview of the orbit attenuation evaluated for given regions in the accelerator. The attenuation improves with the number of selected eigenvalues $\#\lambda_{SVD}$. Assuming orbit corrections with large peak-to-peak beta-beat and a large $\#\lambda_{SVD}$ (collision optics), the error due to spurious orbit correction inside the experimental IRs propagates to the whole machine for orbit corrections based on more than about 450 eigenvalues and is visible in the arc and collimation regions. However, this effect is comparatively small and only relevant for peak-to-peak beta-beat errors above 50%. In practise, the correction is still converging in those regions. The steady-state of the attenuation for a given number of eigenvalues can be explained by the fact that highest eigenvalues correspond to orbit eigenvectors that are specific and locally confined to the IRs providing particle collisions, while leaving the other regions unaffected.

Sensitivity to BPM Noise and Failures

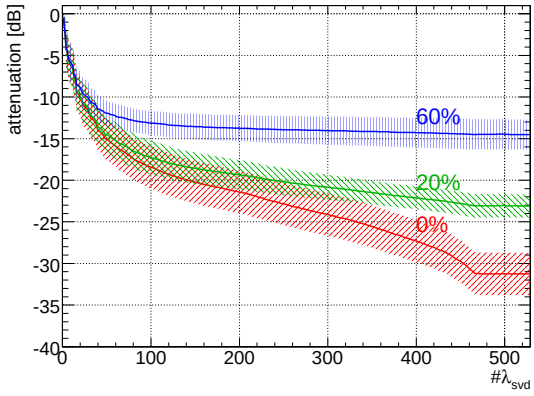
Measurement noise and systematic errors of the BPMs are intrinsic to real accelerator operation and may deteriorate and determine the maximum stability possible. Figure 6.13 shows the effect of the spurious BPM noise propagation onto the orbit and is shown as a function of $\#\lambda_{SVD}$ and peak-to-peak beta-beating error for the injection and collision optics. The sensitivity to noise is fairly independent of the present beta-beat and increases with the number of used eigenvalues. Figure 6.14 shows the BPM noise propagation onto the orbit due to over all BPM equal distributed noise ('white noise'), noise of a single BPM at a location with a large ($\beta \approx 175\,\mathrm{m}$) and small ($\beta \approx 39\,\mathrm{m}$) value of the betatron function. The



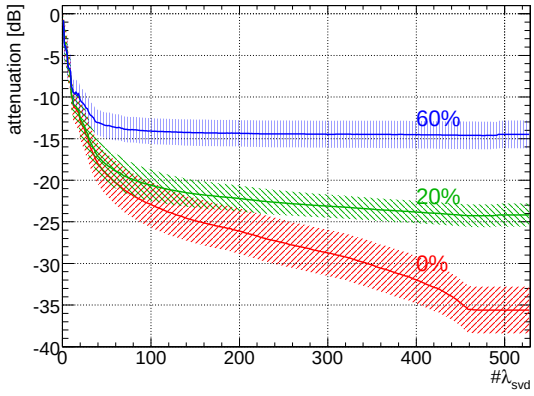
(a) LHC, injection optics



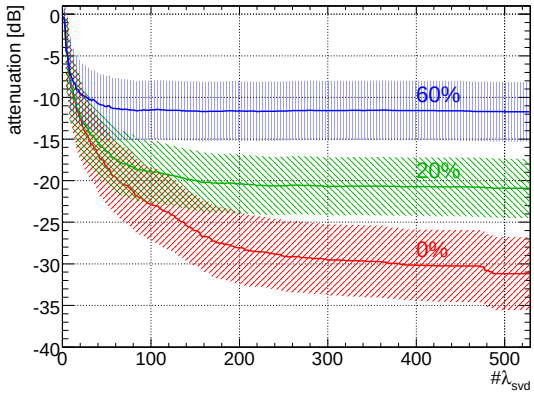
(b) LHC, collision optics



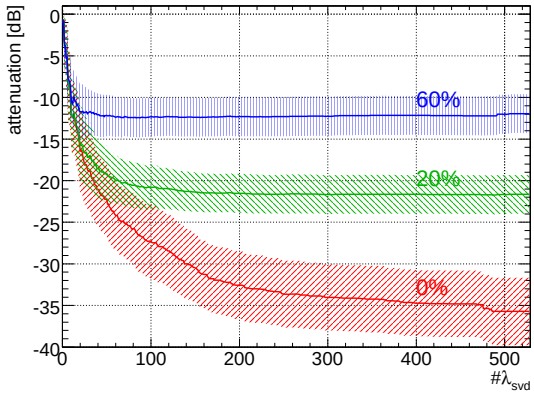
(c) arc, injection optics



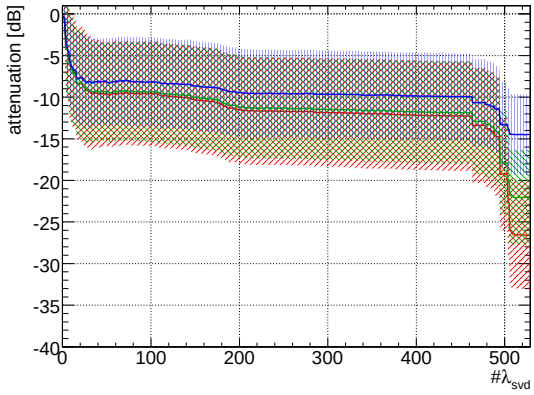
(d) arc, collision optics



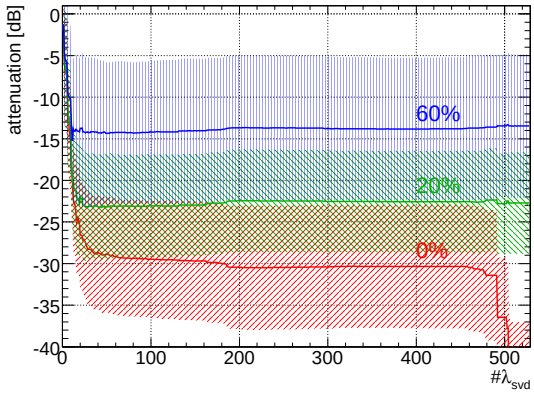
(e) betatron collimation, injection optics



(f) betatron collimation, collision optics

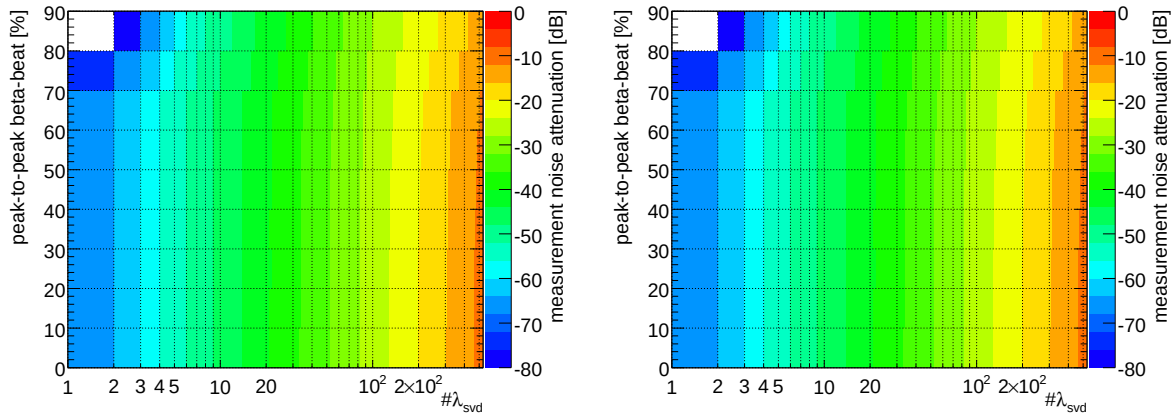


(g) CMS, injection optics



(h) CMS, collision optics

Figure 6.12: Attenuation of SVD-based orbit correction as a function of used eigenvalues $\#\lambda_{SVD}$ for 0, 20 and 60% peak-to-peak beta-beat. The correction is based on all available correctors and monitors.



(a) injection optics (b) collision optics

Figure 6.13: Orbit correction sensitivity to white noise as a function of $\#\lambda_{SVD}$ and peak-to-peak beta-beat. The vertical orbit attenuation for the injection (left) and collision (right) optics is plotted.

corrections were performed with a perfect machine optic. Not surprisingly, the noise sensitivity increases

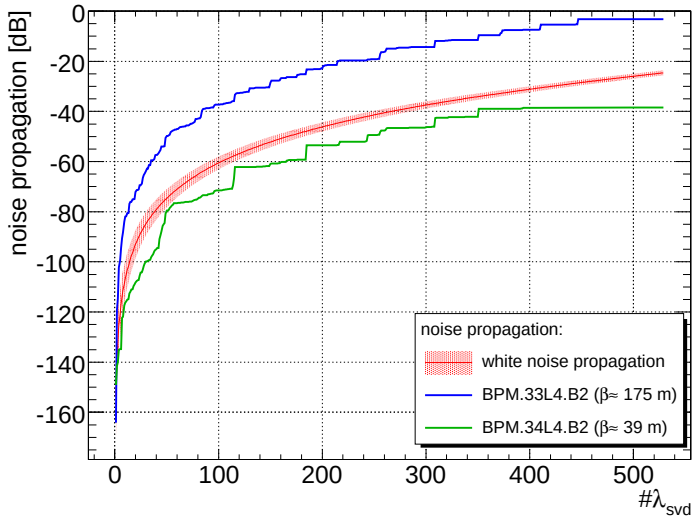


Figure 6.14: BPM noise propagation onto the orbit as a function of $\#\lambda_{SVD}$. The more eigenvalues are used for the correction, the higher the sensitivity to measurement noise.

with the number of used eigenvalues, explained by the localised perturbation correction required large $\#\lambda_{SVD}$ and increases with the value of the betatron function at the location of the BPM. The robustness of the correction is given less by the type of orbit correction than the longitudinal sampling of the BPMs. Due to the approximate $\pi/4$ BPM sampling in the LHC, naturally occurring betatron oscillation or small bumps are sampled by at least three BPMs, which reduces the effect of a single spurious BPM reading.

The previous example assumes that the BPM errors are randomly distributed from one sample to the next, which is common for noise. Another type of failure is systematic errors due to BPM misalignments, calibration errors and electronic systematics is discussed in Section 5.1 and which are approximately constant for a given time interval. Figure 6.15 shows the spacial distribution of the residual normalised steady-state orbit as a result of spurious orbit correction of an arc BPM (large $\beta(s)$) with fixed position reading. The 'ridge' is at the location of the BPM giving spurious position readings. The spurious orbit perturbation correction becomes more local with increasing number of used eigenvalues. The worst-case BPM error propagation, ameliorated by the high longitudinal oversampling, is about 80% with respect to the initial spurious BPM position reading.

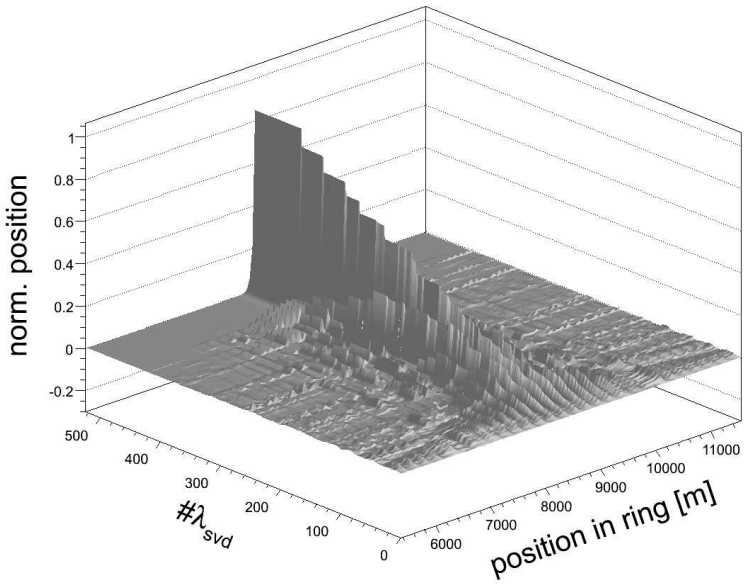


Figure 6.15: Steady-state orbit due to the correction of a single spurious BPM offset. The reproduction of the spurious orbit improves and becomes more localy with increasing $\#\lambda_{SVD}$. The BPM error is 'smeared' out throughout its neighbours if a low number of eigenvalues is used and becomes more localised for larger number $\#\lambda_{SVD}$ used.

Sensitivity to Final Focus Squeeze

A special case of the beta-beat perturbations is the deliberate change of optics inside the experimental insertions in order to produce the final focusing required for minimising the beam sizes at the interaction point. In this special case, the quadrupole strength is systematically changed over a period of about 20 minutes. Small beam and quadrupole misalignments are unavoidable and lead to additional dipole kicks that can be written according to equation 2.66 as:

$$\delta_{xi} = \Delta kl_{mag} \cdot x_s$$

According to alignment targets specified in [122], the quadrupole misalignments are aimed to be less than 0.2 mm r.m.s. globally and less than about 0.1 mm r.m.s. when measured over a range of 10 neighbouring accelerator cells. In order to estimate the maximum drift rates, a worst case magnet misalignment of 0.5 mm r.m.s. is assumed globally. Since the orbit perturbation effects are fairly linear, the results below can be easily re-normalised to the targeted alignments given in [122].

The maximum expected r.m.s. orbit drifts are expected to be about 30 mm. It is clear that such large beam position excursions cannot be tolerated during normal machine operation and thus the residual orbit transient is a combined effect of underlying orbit perturbation and orbit correction. A systematic analysis has been performed to further assess the orbit transient under more realistic conditions. The simulations are based on a seed of 10^5 different quadrupole misalignments that have been normalised to 0.5 mm r.m.s. BPMs are assumed to be perfectly aligned to the quadruple centre and only Beam 1 was used for correction. The resulting orbit perturbation has been corrected before and after the squeeze with varying number of eigenvalues.

Figure 6.16 summarises the expected transients of the difference orbit between corrected orbit before and after the squeeze in IR5 (CMS) as observed in the specified LHC regions. The r.m.s. transient without orbit correction is about 30 mm and improves with the number of used eigenvalues. The transient significantly improves as soon as orbit eigenvectors that are specific for IR5 are used. Further, the median of the transient BPM distribution converges faster with respect to the r.m.s. and maximum residual orbit, which illustrates the locality of the perturbation source and its correction.

Figure 6.17(a) shows a detailed overview of the expected maximum orbit transient seen in the beta-cleaning collimation region during each individual squeeze step as a function of $\#\lambda_{SVD}$. The largest transients are expected during the β^* change from 17 to 9 m and from 1.5 to 1.1 m. The maximum transient per step ranges between 450 μm for $\#\lambda_{SVD} = 300$ and 300 μm for $\#\lambda_{SVD} = 400$ as shown in Figure 6.16. The collimation system requires an approximate orbit stability of a few ten micrometers as discussed in Section 4.5.1. In order to meet these tight orbit stability requirements in the presence of a worst case quadrupole misalignment of 0.5 mm r.m.s., it is visible that nearly 425 eigenvalues are required

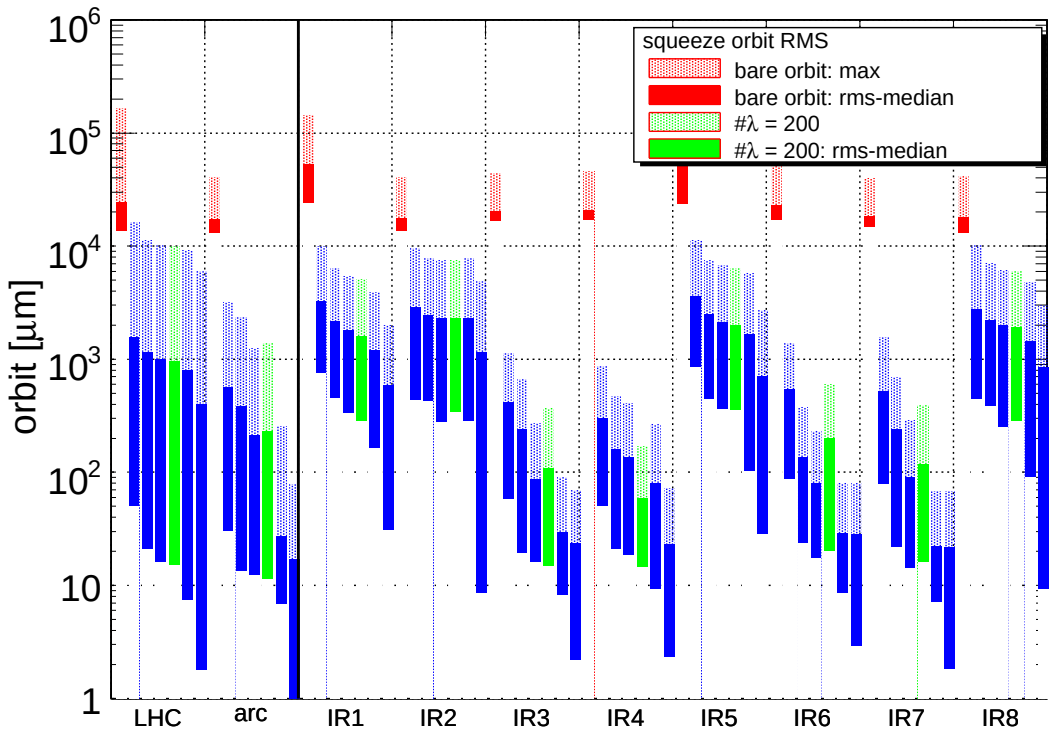
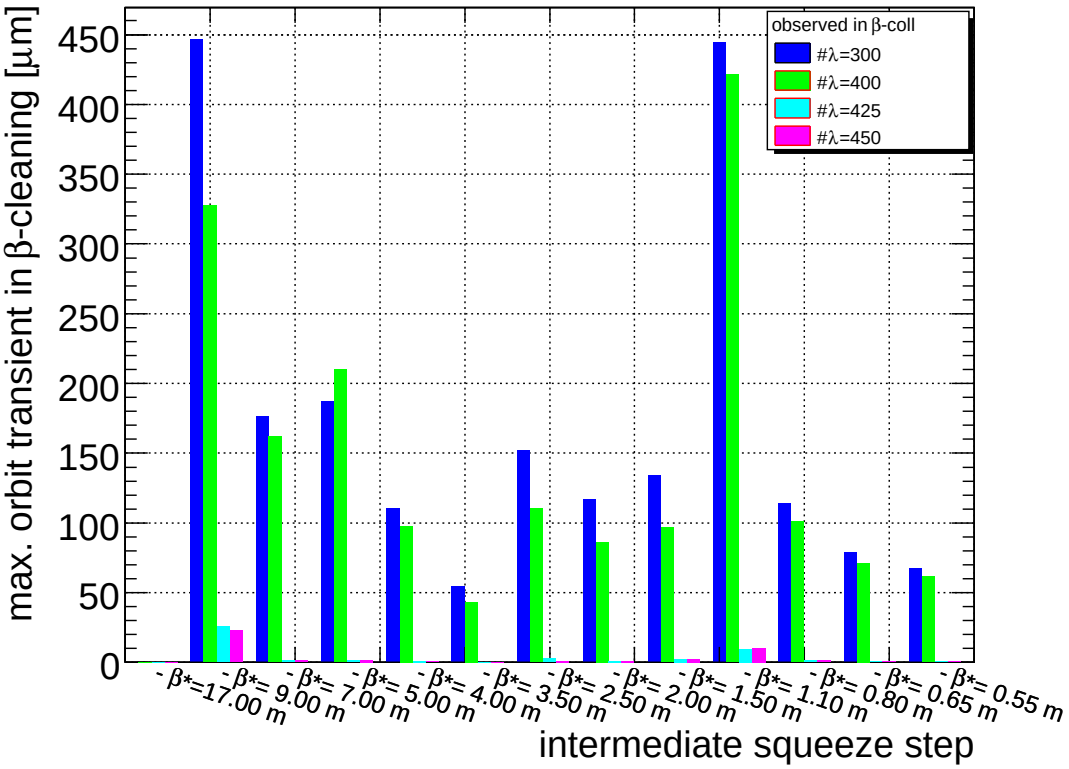


Figure 6.16: Residual maximum orbit transient for different regions in the LHC due to squeeze from $\beta^* = 19\,\mathrm{m}$ to $\beta^* = 0.55\,\mathrm{m}$ in IP5: Each bar shows the 'median' as the lower bound of the filled bar, the 'r.m.s.' as the upper bound of the filled bar and the 'maximum' transient as the upper bound of the hatched surface of the difference orbit. The orbit difference is evaluated between the corrected orbit after and the corrected orbit before the squeeze. The correction is performed using SVD and $\#\lambda_{SVD}$ eigenvalues. The plotted bars per region correspond to $\#\lambda_{SVD} = 0$, 50, 100, 150, 200, 250 and 270 eigenvalues. The correction is performed for both beams using all available correctors and monitors. The convergence for an increasing number of used eigenvalues is visible in each region. The difference between the individual IPs is due to the difference lattice layouts and crossing schemes inside the long straigt sections.

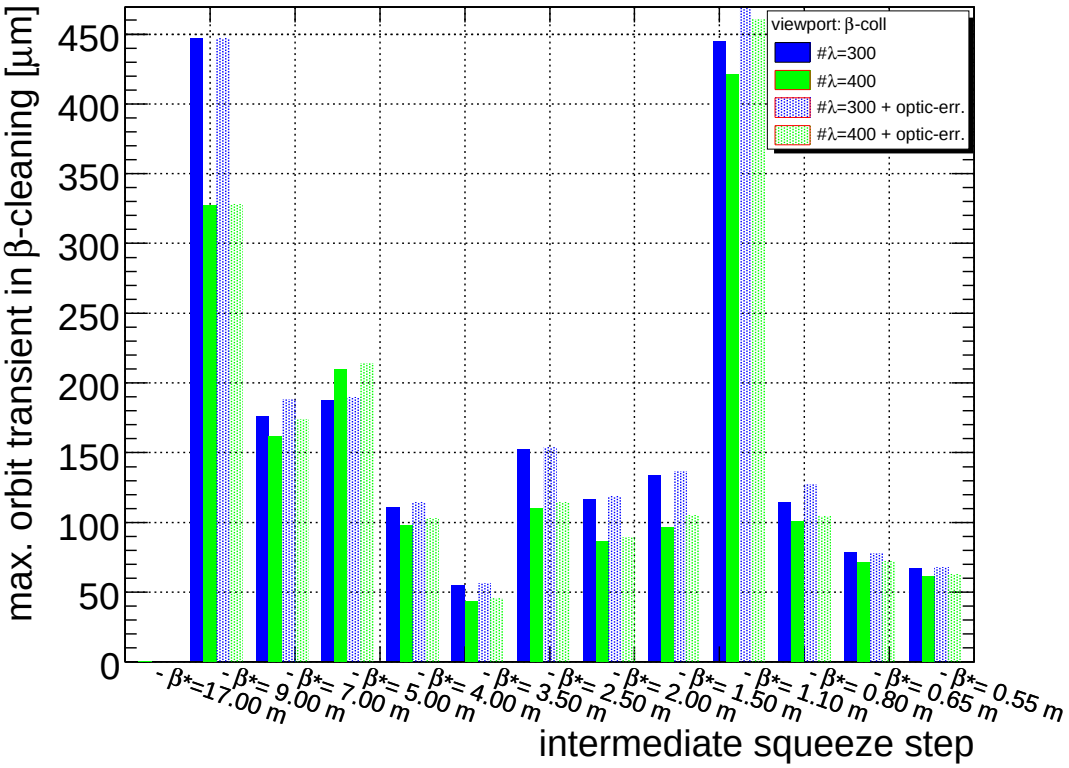
to keep the maximum observed transient inside the betatron cleaning insertion due to the squeeze in IP5 below the target. It is important to note: The assumed quadrupole misalignment can be considered as a worst case assumption compared to the targeted 0.1-0.2 mm r.m.s. described in [122].

Using a more optimistic assumption, the presented result can be scaled down by a factor of about 2.5 for a more optimistic alignment scenario. These drifts are mainly important for the first squeeze; further squeeze attempts may benefit from the knowledge of the first orbit correction and incorporate these into feed-forward tables in order to minimise the need of an orbit feedback. However, based on the expected fill-to-fill variation discussed in Section 4.4.2, the remaining random orbit perturbation component during the squeeze that has to be compensated by the orbit feedback is at least between about $300-600\,\mu$mr.m.s. distributed over about 20 minutes. Similar to LEP [72], the alignment and drift of the quadrupole girdes as a significant impact on the closed orbit inside the LHC.

In order to further assess the sensitivity on optics imperfections the correction was performed either using the initial optics valid before the squeeze till the squeeze step is completed and vice versa. Figure 6.17(b) shows the sensitivity on the optics mismatch for orbit corrections using 300 and 400 eigenvalues. It is visible that the transient is only slightly impacted by the choice of optics during the transition but favours the use of the initial optics.



(a) Maximum orbit transient during squeeze



(b) Maximum orbit transient during squeeze and optics mismatch

Figure 6.17: Residual maximum orbit transient observed in betatron cleaning insertions as a function of squeeze steps in IP5 (CMS) and the number of eigenvalues used for the orbit correction. See text for details.

Sensitivity to Single Corrector Failures

Based on early estimates of the mean-time-between-failure, one out of 1060 CODs are expected to fail about every 105 hours during LHC operation with beam. This is equivalent to a power converter failure during operation of every 10th fill. As discussed in Section 5.2.3, the expected average COD quadrupole misalignment compensation corresponds to an average deflection of 8 μrad and an orbit perturbation of 0.9 mm in the arc at a location of $\beta = 170$ m. If uncompensated, this would easily break the tolerances of the collimation system on orbit stability by one order of magnitude.

However, in case of a failing MCBH(V) magnet, the deteriorating effect on the global orbit can be compensated and kept locally using available CODs in the vicinity of the failing COD. The minimum number of correctors required is two, in order to have enough degree of freedom to match amplitude and phase of the perturbation created by the decaying dipole field as illustrated in Figure 6.18. It is visible that the perturbation of the single failing COD is perfectly matched at large distances and only a small transient is left at the location of the failing COD. For robustness it is favourable to distribute the required compensation pattern over more than two correctors in order to minimise the individual maximum required strength and increase the possible maximum ramping speed. For long term compensation, the failing COD has to be deselected and the pseudo-inverse response matrix to be recalculated. This can, however, only happen on a slower time scale than the time scale of a COD power converter failure. Thus for a faster intermediate correction, the pre-computed compensation pattern can be applied in place of the failing COD and while disabling the BPMs in the vicinity to minimise potential cross-talk between this feed-forward and the nominal feedback loop.

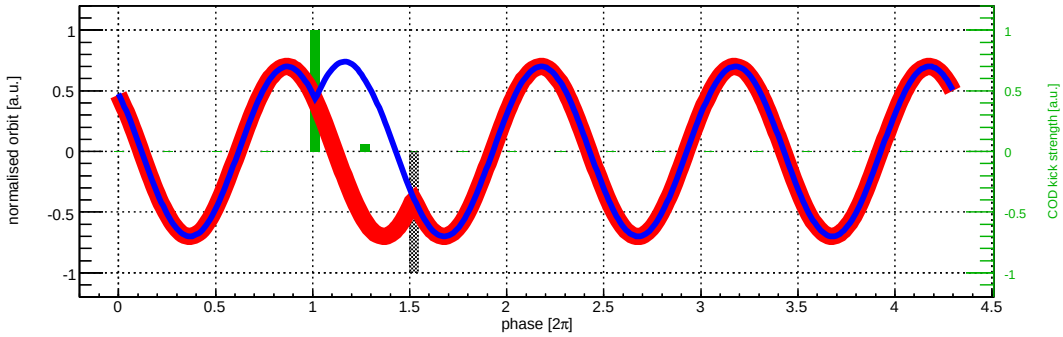


Figure 6.18: Schematic COD replacement scheme. The betatron response (thick red line) of a single failing COD (black bar) can be approximated (blue line) by two (green bars) or more neighbouring CODs. The residual transient is confined to the location of the failing COD.

Figure 6.19 shows the orbit attenuation of a single COD failure as a function of used eigenvalues that the failing COD is unknown and under the assumption that the COD is known and compensated for. It is visible that the attenuation improves for both cases for a low number ($\#\lambda_{SVD} < 20$) of used eigenvalues. However, in case the failing COD is unknown, the attenuation quickly deteriorates as soon as more than 20 eigenvalues are used in contrast to the scheme, assuming knowledge of the failure. In order to be most effective and to minimise the transient it is important that the delay between the power converter failures until the start of the compensation is short. A delay of 1 second corresponds to a residual ripple of about 3% whereas a delay of 0.1 second would minimise the ripple to about 10^{-3}. In order to guarantee a short notification delay, it is foreseen that the power converter gateway will send the status of the COD periodically, through an active real-time link at the same rate as the BPM data.

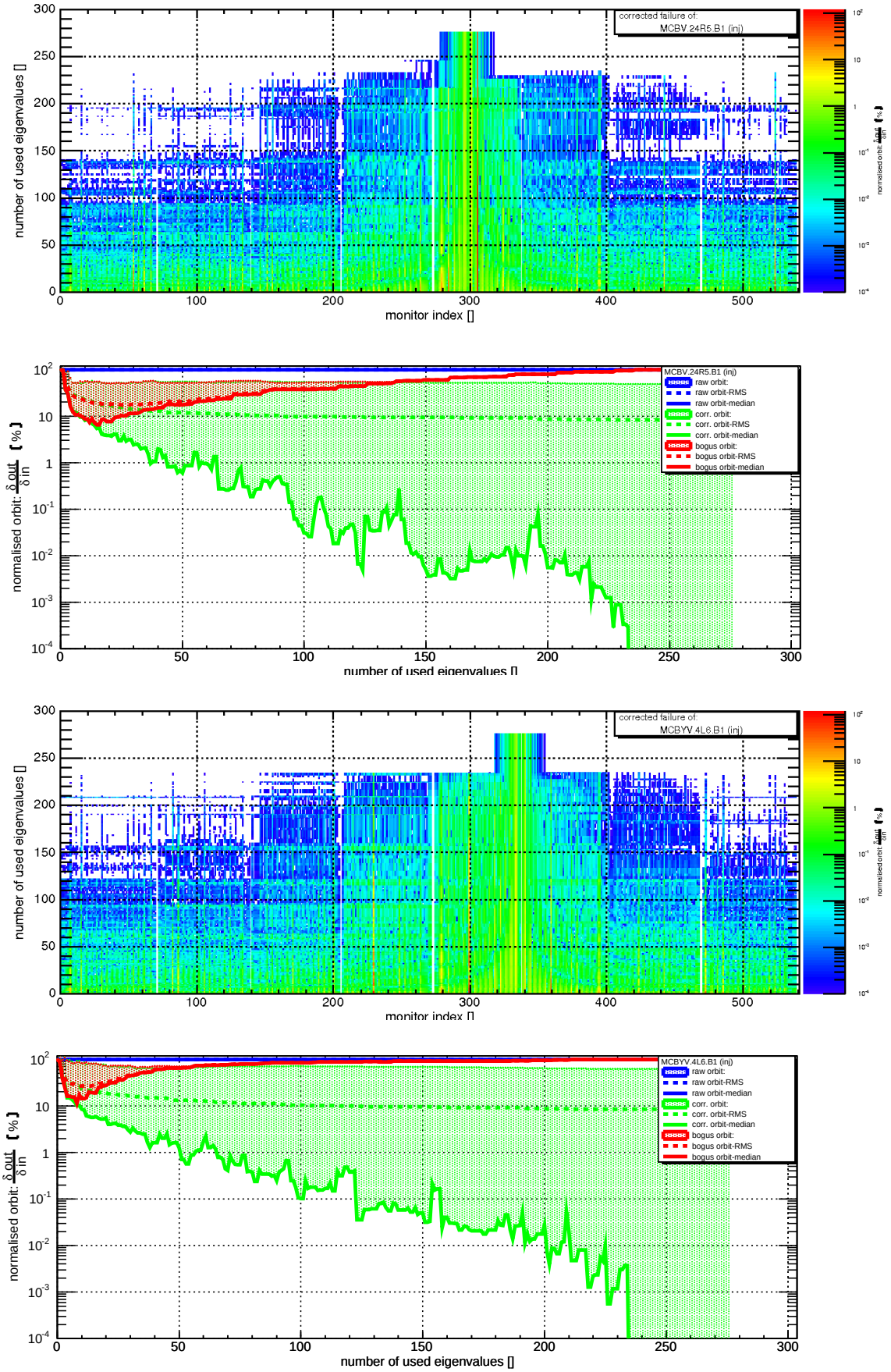


Figure 6.19: Attenuation of an orbit perturbation caused by the failure of MCBV.24R5.B1 (top two) and MCBYV.4L6.B1 (bottom two) during injection. The orbit attenuation with (red) and without (green) are indicated. Orbit correction was performed on only one beam. The region of the transient becomes more localised the more eigenvalues are used for the compensation. Orbit correction was performed on only one beam.

6.1.5 Local Refinement of the Global Orbit Correction

It has been shown in previous sections that the trade-off between maximum possible attenuation of orbit perturbations and robustness against BPM noise and element failures is intrinsic to orbit correction. In case of the SVD-based algorithm, for example, this trade-off is steered through the choices of the number $\#\lambda_{SVD}$ of eigenvalues used for the correction. Traditionally, the local and global orbit constraints were fulfilled using two or more independent loops with either only a few locally selected or all available BPMs and CODs. The local loops are usually operated at a much higher frequency than the global feedbacks in order to minimise the cross-talk. As constraints of global and local constraints became more similar, the distinction of separate local and global feedbacks becomes less evident and the local constraints need to be incorporated into a single fast global orbit feedback loop. The following section shows possibilities of how to enforce stricter local constraints on the orbit using a global type feedback in regions where required.

Local SVD

The SVD type of correction can be easily adjusted to steer orbit only locally while leaving the rest of the machine unaffected, demonstrated and described for example in [154].

For this type of correction, only the CODs and BPMs in the region of interest are selected and the pseudo inverse calculated with usually the full number of non-singular eigenvalues. The effect on the rest of the machine can be estimated if the position and deflection at the border of the corrected region is matched to zero as illustrated in Figure 6.20 by using half of a four-corrector bump scheme described in Annexe B.1. The residual position and angle at the border of the region to be corrected are simply

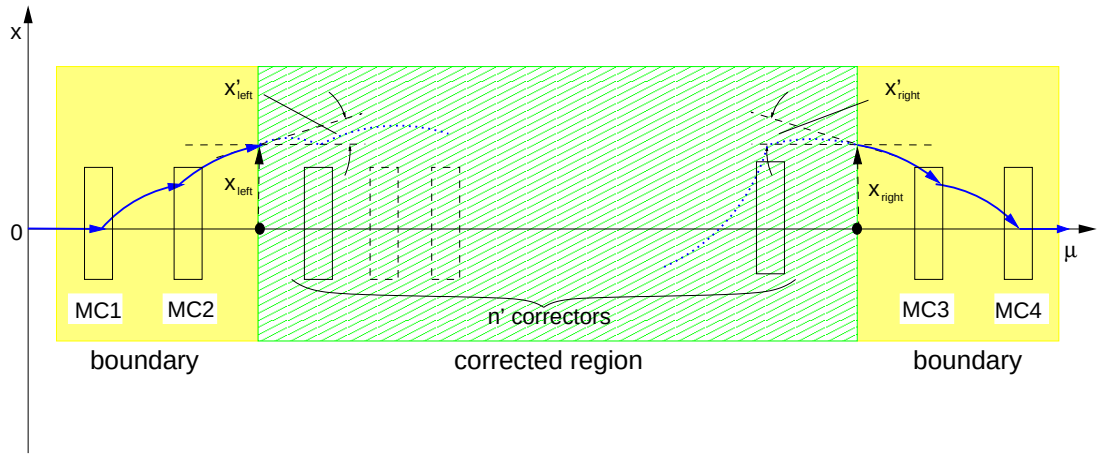


Figure 6.20: Local correction scheme: The boundary condition of the corrected region is matched using two correctors on each side. These correctors ensure that the position and angle at the boundary are matched to zero to avoid a propagation of the local solution to the global orbit.

given by the superimposition of the closed orbit and angle response given in equations 2.71 and 2.72. Since SVD correction and the boundary closure are linear operations, both operations can be cast into the same local SVD orbit correction matrix $\tilde{\mathbf{R}}_L^{-1}$ with closed boundaries. Figure 6.21 shows the difference orbit between a randomly perturbed orbit and an orbit that has been locally corrected with the above described method. It is visible that the correction is confined to the selected region. In principle, both types of corrections, global and local, can be achieved in the same machine. Traditionally, the loops were decoupled in time, where the local loop is running at a much higher speed than the interleaved global loop. For orbit corrections where the global and local corrections are performed at similar rates, this scheme is less favourable due to the inherently possible residual transients between global and local correction loops switching since the global loop does not necessarily obey the constraints of the local loops.

A more robust scheme can be achieved if both corrections are merged to a hybrid correction with coarse but robust global correction features and less robust but better orbit attenuation in regions where required. Since both global and local SVD-based orbit corrections are purely linear operations that are expressed by a dense global orbit correction matrix $\tilde{\mathbf{R}}_G^{-1}$ and spare local orbit correction matrix $\tilde{\mathbf{R}}_L^{-1}$, the superimposed result can be cast into a combined correction matrix $\tilde{\mathbf{R}}_C^{-1}$ that can be written as

$$\tilde{\mathbf{R}}_C^{-1} = \left[\tilde{\mathbf{R}}_G^{-1} + \tilde{\mathbf{R}}_L^{-1} \cdot \left(\mathbf{1} - \mathbf{R}\tilde{\mathbf{R}}_G^{-1}\right)\right] \tag{6.44}$$

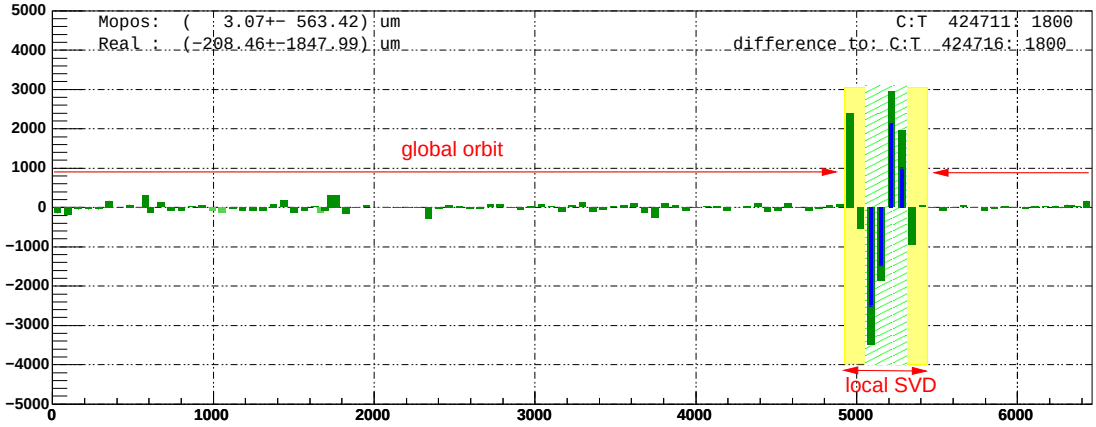


Figure 6.21: Example of a local SVD correction in the SPS: The difference orbit between the closed orbit before and after correction is shown in units in micrometres versus the position inside the SPS in units of meters. The regular SPS BPMs measuring the global orbit (green) and four LHC prototype BPMs (blue) located in BA5 are shown. The locality of the correction is visible. Further, the difference of the orbit change measured by the two different BPM systems was found to be due to a systematic 20% scaling error of the LHC prototype versus SPS BPM system [154].

with $\mathbf{1}$ the identity matrix and $\mathbf{R}$ the global orbit response matrix. The local SVD orbit correction matrix $\tilde{\mathbf{R}}_L^{-1}$ is assumed to be padded with zeros in order to have the same dimension as the global orbit correction matrix. There are following two main disadvantages of this scheme:

- The possible large orbit and angle excursions in the closure boundary regions that cannot be minimised without deteriorating the local correction.

- The additional sensitivity to beta-beating since this scheme requires a forecast of the orbit response matrix that may differ from the real beam response. The perturbation of the machine and predicted orbit response adds up linearly and effectively doubles the beta-beating seen by the controller in the region where localised orbit control is applied.

Assuming that the dynamic beta-beating is controlled better than 20% (see [48] for details), the latter might be accepted as part of the general stability versus robustness trade-off.

Local Correction through BPM Weighting

Another possibility to improve the local orbit control is by weighting the SVD entries of given BPMs in the regions of interest.

$$\tilde{\mathbf{R}}\omega\omega^{-1} := \mathbf{R}_W\omega^{-1} \tag{6.45}$$

with diagonal weighting matrix $\omega = diag(\omega_1, ..., \omega_m)$ and its inverse $\omega^{-1} = diag(1/\omega_1, ..., 1/\omega_m)$. The SVD inverse is then performed on $\tilde{\mathbf{R}}$ and the resulting orbit correction matrix multiplied with ω in order to ensure the proper calibration of the BPM reading. This weighting process effectively produces specific eigenvectors with larger eigenvalues for the targeted local regions. Test with the LHC lattice and correction of both beams showed that the BPM weighting between $\omega_i = 10$ and 100 seems to be suitable to improve the local orbit attenuation inside the cleaning insertion.

This method intrinsically widens the spread of eigenvalues which leads to an increased condition number and thus increased number of lost bits, as expressed by equation 6.29. For weighting coefficients ω_i much larger than 1000, the numerical stability was compromised by the weighting and caused failures of the singular-value-decomposition in the given floating point representation (*double*).

The disadvantage of this method is the uncertainty of the total number of eigenvalues that have to be selected in order to include localised solutions that have been emphasised by the weighting. In practise, the number of eigenvalues to be selected is uncertain, depends strongly on the weighting coefficients, and nearly always requires manual intervention: the larger the weighting factors, the easier the selection of the threshold $\#\lambda_{SVD}$, but also the correction becomes more prone to numerical errors and BPM noise.

Selective SVD

A more advanced scheme can be achieved if the eigenvalues are not selected simply according to their size or location in the eigenvalue spectra (which may vary depending on the BPM, COD and lattice configuration) but selected according to orbit eigenvectors that have significant entries in those regions that have to be specially stabilised.

The following steps should be applied for the proposed method:

- Selection of the first $\#\lambda_{SVD} = 200$ to 300 eigenvalues in order to provide a coarse but robust global orbit correction.
- Of the remaining eigenvalues select those that
  1. are non-singular ($\lambda_i > 20$) and
  2. have one or more orbit eigenvector entries $\underline{u}_i$ (see equation 6.16) that are larger than $\mu_{local} >$ 1..2% in the region with special stability requirements.

The latter utilises the ortho-normality of the orbit eigenvectors. Thus the cut-parameter μ_{local} can be kept constant under different COD, BPM and lattice configurations. As visible in equation 6.32, this method selects those eigenvalues that yield a non-vanishing

$$a_i = \underline{u}_i^T \cdot \Delta\underline{z} \tag{6.46}$$

contribution for regions where both the measured difference orbit $\Delta\underline{z}$ and orbit eigenvector $\underline{u}_i$ are non-vanishing. The local correction can be further improved if the BPMs are weighted, favouring the creation of localised orbit eigenvector solutions in the regions of interest.

As summarised in table 4.9, the magnitude of the different LHC orbit requirements are similar. Most of the requirements are distributed around the machine and a distinction between local and global requirements is less obvious. Future LHC operation may show whether a distinction between local and global orbit correction is necessary or will possibly become obsolete in favour of a simpler single orbit correction that stabilises the beam globally to the best required level.

6.1.6 Energy and Dispersion Orbit Correction

In the previous sections, it was assumed that orbit perturbations are driven by dipoles only. As shown and described in sections 2.4.2, 4.2 and 4.4.1, solar and lunar tides as well as the systematic decay and snap-back of the main dipoles b_1 field contribution may change the machine momentum, which contributes through the dispersion $D(s)$ at a given location to an additional orbit offset

$$\Delta x_{mom} = D(s) \cdot \frac{\Delta p}{p} \tag{6.47}$$

that is dependent on the systematic momentum shift $\frac{\Delta p}{p}$. Solar and lunar tides (see 4.2 for details) may cause relative peak-to-peak momentum variation of about $1.3 \cdot 10^{-4}$ corresponding to an orbit up to about $250\,\mu$m at a location with $D(s) = 2\,$m. Though these perturbations are very slow with a timeconstant in the order of about 6 hours for the tides and 1000 seconds for the b_1 decay, these momentum drifts need to be compensated in order to match the SPS-LHC injection momentum and to fulfil the orbit stability while colliding beams at 7 TeV.

During injection, it is required that the momenta given by the SPS and LHC are matched in order to optimise the injection and to minimise the longitudinal emittance blow-up and particle losses. Theoretically, limiting to only energy effects, one could use the RF frequency of either the SPS or LHC to match the energies. However, in order to provide a constant phase synchronisation between bunch and RF bucket, it is essential that the frequencies of the SPS and LHC RF systems are constant and at a fixed ratio with respect to each other during injection.

Maintaining a constant RF-frequency during injection, and using equations 2.51 and 2.46, leaves the following possibilities to match the SPS and LHC beam momenta: Adjustment of the beam momentum before

1. SPS extraction using the SPS main dipole magnets
2. LHC injection using the horizontal LHC dipole corrector magnets
3. LHC injection using the LHC main dipole magnets
4. SPS extraction using the horizontal SPS dipole corrector magnets

It is usually preferred to compensate perturbations locally or close to where they occur. Thus, initially, adjusting the LHC main dipole reference field seemed to be favourable. However, even small powering changes may trigger a snap-back of the decayed fields and also further perturb powering history making the fill-to-fill reproducibility more complicated to predict. Thus, during injection, it is favourable to use the horizontal corrector circuits in the dispersive regions to change the momentum. As stated in equation 6.4, the momentum change through the orbit corrector dipoles alone can be expressed as:

$$\frac{\Delta p}{p} \approx \left(\frac{1}{\gamma} - \frac{1}{\alpha_p}\right) \sum_{i=0}^{n} \frac{D_z(s_i)}{C} \delta_i \tag{6.48}$$

with δ_i the deflection and $D_z(s_i)$ the dispersion function at the i-th corrector, and C the average machine circumference. For reference, the momentum change due to a relative main dipole field change $\frac{\Delta B}{B}$ is given by:

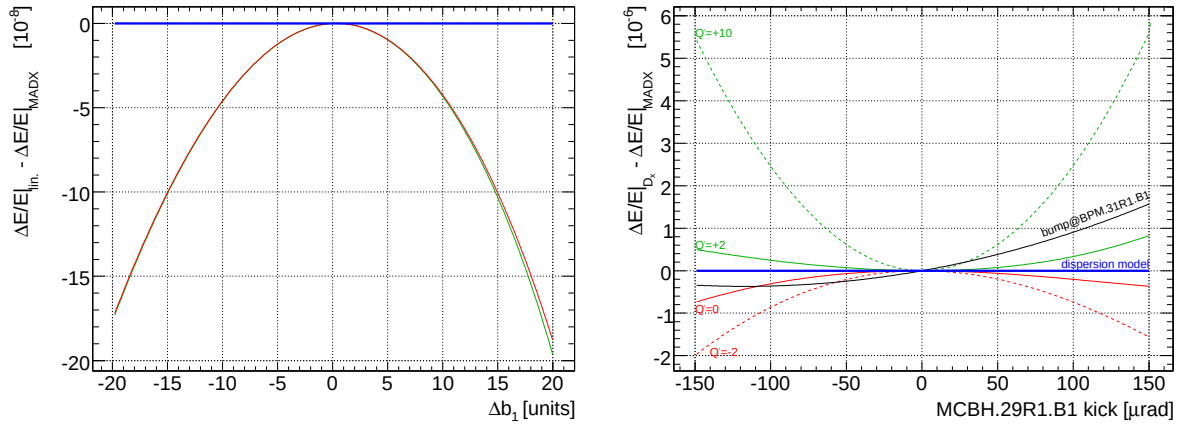
$$\frac{\Delta p}{p} \approx \frac{\Delta B}{B} \tag{6.49}$$

The total integral field around the ring given by corrector dipoles that is much smaller than the main dipole contribution and the fact that orbit correctors, compared to the main dipoles, are only sparsely distributed over the machine explains the dipole correctors' slightly different momentum perturbation response. As a second order effect, the usage of orbit correctors for momentum correction also causes a small residual orbit perturbation. This perturbation can be compensated either in advance by superimposing the momentum and correction of the residual orbit perturbation. Another compensation is done intrinsically by the automated orbit feedback that compensates for the dipolar but not for momentum driven orbit perturbation.

Figure 6.22 shows the error due to the linearisation of the perturbation effect of main dipole and orbit corrector on the beam momentum in comparison to a full machine model (MADX) that takes the non-linearities and feed-down effects due to imperfect orbit into account. As part of the simulation, the chromaticity has been varied between -2 and +10 units and the perturbed orbit corrected to about 100 mm r.m.s. The relative error on the beam energy between the linear and non-linear models is within about $2 \cdot 10^7$ for main dipole field imperfection of up to 20 units[3]. The variations on the momentum are

[3]Reminder: One 'unit' corresponds to a relative main dipole field change of 10^{-4}.

symmetric and fairly insensitive to the chromaticity. In case of the orbit correctors, the error between linear and non-linear models reach up to $6 \cdot 10^{-5}$ for maximum deflections of up to 150 μrad, which are about twice as big as what a MCB type magnet can produce at 7 TeV.



(a) main bend (b) orbit corrector

Figure 6.22: Relative energy error between the full (MADX) and linear energy perturbation model for the main dipole (left) and orbit correction (right) magnets. The errors are plotted for operational scenarios with Q' between -2 and +10 units of chromaticity, as well as closed orbit bump.

While the absolute measurement of the beam momentum is more complicated (see [120, 155] for details), the relative beam momentum change around a given (but unknown) reference momentum can be measured quickly with sufficient accuracy. There are two possibilities: The SPS to LHC injection momentum mismatch can be measured through the injection oscillations. Figure 6.23 shows the Landau damping attenuated injection oscillation. The time constants are typically about a few hundred turns. It is

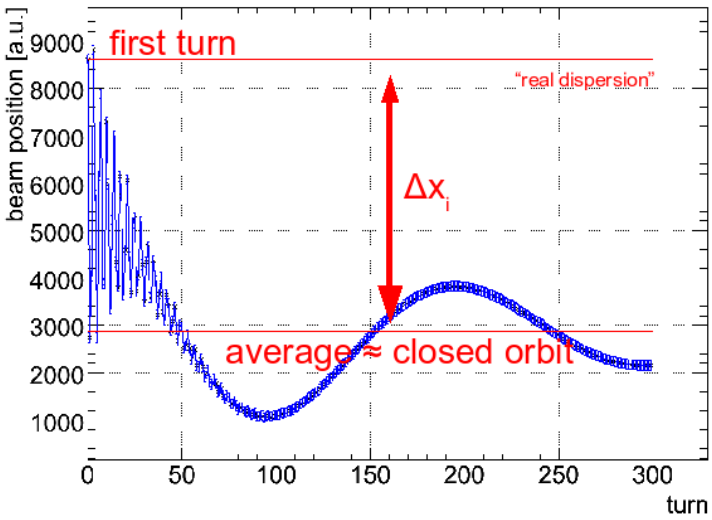


Figure 6.23: Schematic injection oscillation due to energy mismatch. The momentum mismatch $\Delta p/p$ is proportional to the difference Δx between first turn amplitude and closed orbit after the energy oscillation has been attenuated.

important to note that the individual average particle amplitude is unaffected by this damping but merely the visible coherent oscillation is attenuated due to the de-phasing of the individual particle oscillations. The momentum mismatch $\Delta p/p$ is proportional to the difference Δx between first turn amplitude and closed orbit after the energy oscillation has been attenuated. An other possibility to measure a change of momentum is through monitoring the difference of the actual to a given reference orbit.

As given by equation 2.73, the effect of dipole and momentum perturbation superimpose to an effective closed orbit. While the dipole perturbation create orbits that in first order oscillate sinusoidal around the

reference trajectory, the dispersion orbit is systematically shifted either outwards or inwards to the centre of the machine, which can be used to estimate the momentum shift. This analysis tests the following two estimates:

1. Classically the relative momentum shift is estimated by

$$\left.\frac{\Delta p}{p}\right|_{estimate} = \frac{\sum iD_i \cdot \Delta x_i}{\sum iD_i^2} \tag{6.50}$$

   with D_i the dispersion and Δx_i the orbit difference reading between either the injection trajectory or the reference and the actual measured closed orbit. Assuming a constant dispersion, it is visible that the momentum estimate reduces to

$$\frac{\Delta p}{p} \approx \frac{1}{D} \cdot \Delta z$$

   which is about the direct inverse of equation 6.47. The dispersion based weighting improves the signal to noise ratio as it gives preference to BPMs with dispersions where also the measurable effect is largest.

2. The second possibility is to treat the dispersion detection as an additional pseudo corrector-like entry in the orbit response matrix with the response equal to the dispersion function at each BPM and to invert the matrix as done for the SVD-based orbit correction described in the previous sections.

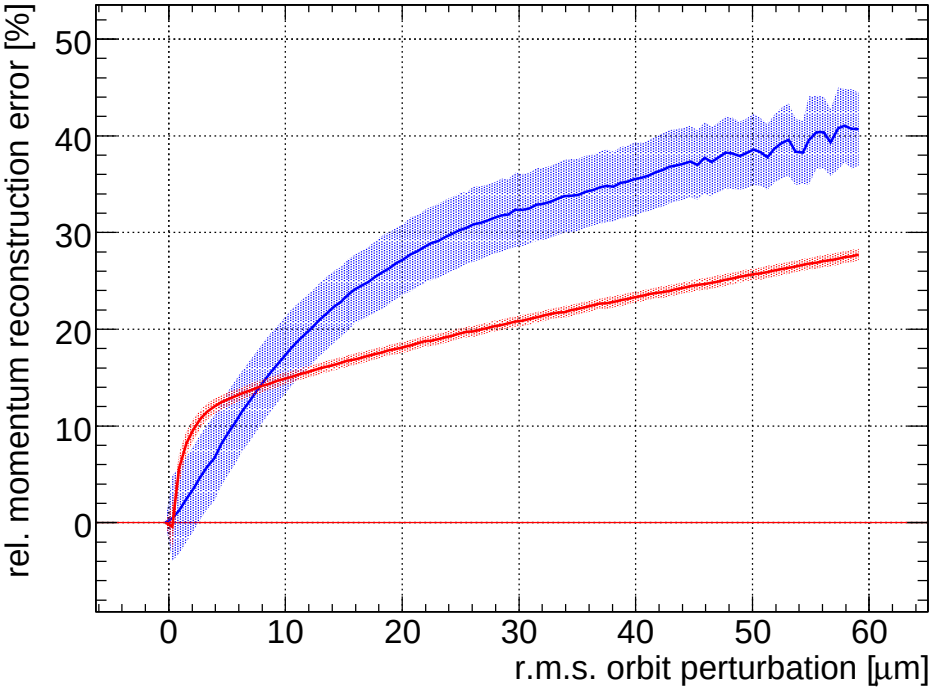
In both cases, the reference dispersion function can easily be measured using equation 6.47 and a momentum modulation through a known RF frequency shift. Measuring the difference orbit before and after the frequency change and dividing by the known momentum change gives a dispersion estimate that is essentially limited by the stability of the orbit and BPM system during the measurement.

The estimated momentum either of the classic or the SVD-based method can then be used to subtract the predicted dispersion orbit prior to orbit correction. This improves the orbit correction's robustness, and most importantly, minimises the possible cross-talk between the orbit and momentum compensation required to measure the chromaticity.

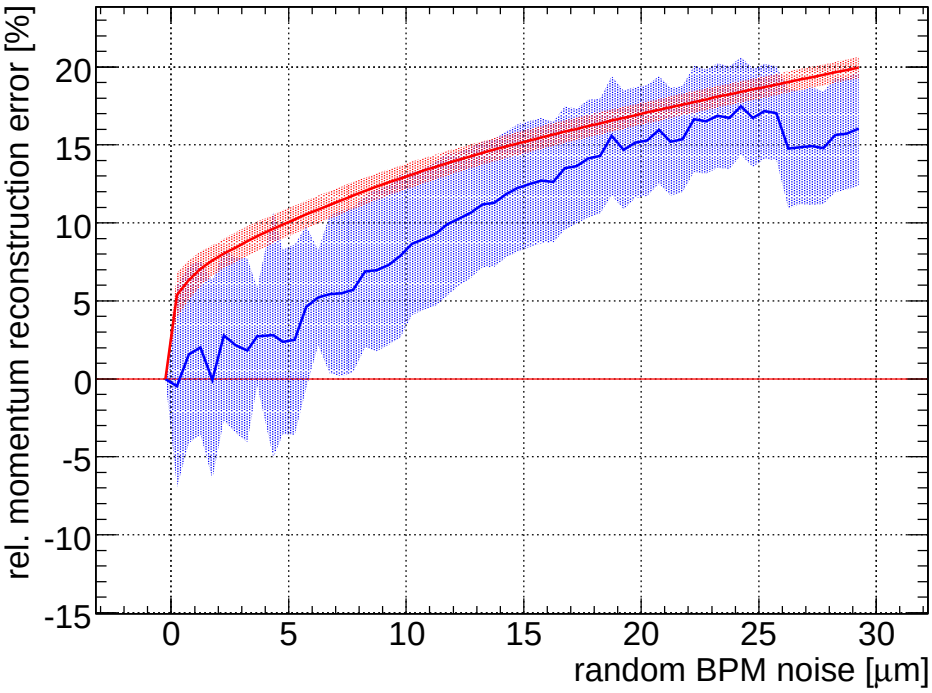
In real-life, the measured dispersion orbit is, however, a superimposition of the momentum related effect, dipolar orbit disturbances and BPM noise. A simulation has been prepared to assess the sensitivity of the two momentum estimators. As summarised in Table 4.9, the required momentum stability is better than 10^{-4}. In order to have a sufficient margin for detection and possible feedback, the simulation assumed a ten-time smaller fixed minimum momentum perturbation of 10^{-5}. If required in this linear approximation, the results can be rescaled to an arbitrary higher relative momentum offset. Figure 6.24(a) and 6.24(b) summarise the reconstruction error as a function of superimposed dipolar orbit perturbation and random BPM r.m.s. noise level. The reconstruction error $\epsilon_{recon.}$ is defined by the estimated $\left.\frac{\Delta p}{p}\right|_{est}$ and reference relative momentum shift $\left.\frac{\Delta p}{p}\right|_{ref}$ to:

$$\epsilon_{recon.} = \frac{\left.\frac{\Delta p}{p}\right|_{est} - \left.\frac{\Delta p}{p}\right|_{ref}}{\left.\frac{\Delta p}{p}\right|_{ref}} \tag{6.51}$$

The reconstruction error of both estimates increases with rising noise level and reaches a level of less than 40% corresponding to an relative momentum error of about $4 * 10^{-6}$ for r.m.s. orbit or BPM noise level of $60\,\mu$m. The SVD estimate has a generally higher r.m.s. spread and seems to be slightly more robust against BPM noise. In case of dipolar orbit perturbations, the SVD based estimate has a slightly smaller systematic reconstruction error for small orbit perturbation (less than about $7\,\mu$m) and a larger systematic reconstruction error for larger orbit perturbations than the classic method. For large orbit perturbations, the classic momentum estimate shows a better reconstruction with a significantly smaller measurement distribution. The systematic offset can be slightly improved in case of BPM noise either through averaging or through shifting the measured estimate based on the estimated BPM noise from continuous time data. The choice for the classic or SVD based method depends on the assumption of the varying orbit r.m.s.: While setting up the machine (large betatron orbit perturbations) the classic method is more robust and precise. If one continuously correct the random orbit perturbations to better than $7\,\mu$m one may choose to use the SVD based method instead.



(a) orbit perturbation



(b) BPM noise

Figure 6.24: Relative energy reconstruction error of a set momentum shift of $\Delta p/p \approx 10^{-5}$ as a function of superimposed r.m.s. closed orbit perturbations (top) and random BPM noise (bottom). The systematic error of the classic reconstruction (red) and the SVD-based method (blue) and corresponding 1σ error surfaces are indicated.

Since the momentum reconstruction errors at the chosen working point ($\frac{\Delta p}{p} = 10^{-5}$) are at least one order of magnitude larger than the known model linearisation error, the choice of using the linearised model seems to be acceptable.

6.2 Time Domain

After orbit and momentum correction described in the previous section, the n coupled differential equations describing the full orbit dynamic (equations 6.4 and 6.4) are first decoupled into n independent one-dimensional differential equations describing the dynamic behaviour of the individual corrector circuit.

The following sections describe the optimisation of the transition in the time domain from the initial corrector to the desired final-state corrector setting while keeping constraints such as minimising the time required for the transition, minimal or no overshoot, robustness with respect to measurement noise and model uncertainties.

For controlling the linear part of the circuit response, the LHC orbit feedback controller is based on a discrete PID implementation. Saturation and current rate-limiter of the power converter as well as computational and transmission delays inside the feedback system can introduce significant non-linearities that may break the decoupling of the n corrector circuits. Their optimal compensation is discussed in Section 6.2.2. The feedback performance of the complete orbit feedback loop is shown in Section 6.3.

6.2.1 Discrete Controller Representation

The PID controller[4] is commonly used to steer linear second order processes and has been extensively studied [56,57]. Nowadays, feedback controllers usually implement discrete algorithms that evaluates and holds the result of the last correction for the duration of a sampling period T_s (ZOH, Zero-Order-Holder). Based on an idea of Euler, the first derivative of a continuous function can on the time scale T_s be approximated by the following difference equation as:

$$\frac{\partial f(t)}{\partial t} \approx \frac{f(k) - f(k-1)}{T_s} \tag{6.52}$$

The continuous PID controller given in equation 3.45 can thus be rewritten to the following discrete form

$$\begin{aligned} u(k) &= u(k-1) \\ &\quad + \left(K_p + K_i T_s + K_d \frac{1}{T_s}\right) e(k) - \left(K_p + 2K_d \frac{1}{T_s}\right) e(k-1) + \left(K_d \frac{1}{T_s}\right) e(k-2) \\ &= u(k-1) + K_p \cdot [e(k) - e(k-1)] + K_i T_s \cdot e(k) + K_d \frac{1}{T_s} \cdot [e(k) - 2e(k-1) + e(k-2)] \end{aligned} \tag{6.53}$$

with T_s the sampling period, K_p the proportional, K_i the integral and K_d the derivative gain. Here, the control signal $u(k)$ at the sample time k depends on the previous control signal $u(k-1)$ and on the change of the feedback error signal $e(k)$. Thus, this form is also called the *velocity* PID algorithm in contrast to the *positional* PID algorithm form

$$u(k) = u(0) + K_p \cdot e(k) + K_i \frac{1}{T_s} \cdot \sum_{i=0}^{k} e(i) + K_d T_s \cdot [e(k) - e(k-1)] \tag{6.54}$$

that depends on the initial controller state $u(0)$ and all error signals since the start of the controller. Both forms are mathematically equivalent but have implications on the implemenation. The velocity form is usually preferred in terms of easier feedback commission, since it does not require knowledge of the initial state of the correctors when the feedback is started and thus needs to keep track of only the last three samples.

As discussed in Section 3.3, high frequency measurment noise may deteriorate the robustness of the loop. Thus it is favourable and common to modify the derivative controller to include a first order low-pass filter, given in equation 3.44, that low-pass filters the error signal prior to evaluating the derivative term. Accordingly, equation 6.53 can be rewritten and the measured error signal $e(k)$ in the derivative term is replaced with the low-pass filtered error signal $\tilde{e}(k)$:

$$u(k) = u(k-1) + K_p \cdot [e(k) - e(k-1)] + K_i T_s \cdot e(k) + K_d \frac{1}{T_s} \cdot [\tilde{e}(k) - 2\tilde{e}(k-1) + \tilde{e}(k-2)] \tag{6.55}$$

The filtered $\tilde{e}(k)$ is given by the first order filter with T_s the sampling period, τ_D the filter constant and the filter bandwidth $\frac{1}{\tau_D}$, respectively:

$$\tilde{e}(k) = \frac{\tau_D}{T_s + \tau_D} \tilde{e}(k-1) + \frac{T_s}{T_s + \tau_D} e(k) \tag{6.56}$$

[4] PID: Proportional Integral Derivative

It is visible that for $\tau_D = 0$ the filtering vanishes and the derivative term will accordingly act directly on the measured signal.

6.2.2 Optimal Non-Linear Control

Summarising the results discussed in Section 5.2, the open-loop response for an individual corrector circuit can be split into three terms

$$G(s) = G_0(s) \cdot e^{-\lambda s} \cdot G_{NL}(s) \tag{6.57}$$

with $G_0(s)$ describing the linear response, λ the constant open loop lag and $G_{NL}(s)$ the non-linear response due to the power converter's saturation and current rate-limiter. For a given excitation amplitude which does not saturate or exceed the current rate limit of the power converter, the open loop response can be described by a linear first order response

$$G_0 = \frac{K_0}{\tau s + 1} \tag{6.58}$$

with τ the circuits time constant which is a function of the excitation amplitude as shown in Figure 5.19. The expected pertubation scale during collision is about $15\,\mu$m corresponding to a current oscillation amplitude of about 0.1 A and an open-loop time constant τ of about $\tilde{\tau} \approx 1\,$s (1 Hz bandwidth). The factor K_0 is a constant for normalising the open loop response for low frequencies to one or, in other terms, to treat conversions between different physical units. For further discussion of the time domain controller it assumed that the calculations are performed in terms of currents in units of 'amperes' and thus K_0 equals one.

Ignoring the non-linearities (delays, saturation and rate-limiter) in the power converter's the open loop response given in equation 6.57, the optimal controller design is identical to the first-order process example discussed in Section 3.4.1. Requiring that the closed loop has a similar response to that of a first order low-pass filter, it is shown using YOULA's affine parameterisation that the optimal controller has the form of a PI controller (equation 3.66

$$D(s) = \frac{\tau \cdot s + 1}{K_0 \alpha s} = \frac{\tau}{K_0 \alpha} + \frac{1}{K_0 \alpha s}$$

with proportional K_p and integral gain K_i given by (equation 3.67):

$$K_p = \frac{\tau}{K_0 \alpha} \qquad \text{and} \qquad K_i = \frac{1}{K_0 \alpha}$$

Here, the parameter $\frac{1}{\alpha}$ corresponds to the effective closed loop bandwidth. The corresponding closed-loop frequency response is shown in Figure 3.7.

As discussed and summarised in Section 4.6, the majority of orbit and energy perturbations are expected to have relative slow drift rates during most machine operation phases while the maximum drift rates of up to $25\,\mu$/s are only reached at the start of the ramp and during the β^* squeeze. Due to the affine parameterisation $\frac{1}{\alpha}$, it is possible to improve the closed loop performance and its robustness by dynamically adapting $\frac{1}{\alpha}$ to the operational machine mode and requirements. For example, during injection or physics $\frac{1}{\alpha}$ could be adjusted to a low effective closed loop bandwidth in favour of being less sensitive to measurement noise and increased at the start of the ramp, during the squeeze or whenever a fast feedback response is required.

Further, a significant reduction of the closed loop bandwidth during collision is also essential to minimise the cross-talk between the orbit and a possible luminosity feedback, since the orbit feedback will otherwise steer down to the noise of the BPMs surrounding the IPs that can be as large as $5\,\mu$m and, if not minimised, cause an imperfect beam-beam overlap of about $0.3\,\sigma^*$ at the interaction point (σ^* being the r.m.s. beam size at the IP).

The maximum effective closed loop bandwidth is apriori a free parameter and limited only by non-linearities that constrain the maximum available excitation signal. As shown in Section 5.2, the use of, for example, the current rate-limiter introduces an additionally non-linear phase contribution once the desired excitation amplitude and frequency product exceeds (equation 5.59)

$$\Delta x \cdot f \overset{!}{\lesssim} \begin{cases} 202\ \mu m/s & \text{at } p = 450\,\text{TeV} \\ 13\ \mu m/s & \text{at } p = 7\,\text{TeV} \end{cases}$$

The following sections discuss the loop deteriorating effects and their compensation of non-linearities introduced by pure computational (transport) delays, saturation and current rate-limiter.

Computational and Transport Delays

Pure time delays can deteriorate the closed-loop response if the inverse delay is in the order or larger than the closed loop bandwidth causing transient oscillations that may drive the loop to be unstable due to the reduced phase margin.

The actual per-sample SVD-based orbit correction can be cast into a simple matrix-vector multiplication. The dimension of the LHC orbit correction matrix alone is 530×1056 per plane. Due to the $\mathcal{O}_{SVD}(n^2)$ complexity, the critical time consuming parts of the controller algorithms are essentially memory bandwidth limited. The total orbit correction and computation of new corrector strengths and further proposed compensation schemes require about $4 \cdot 10^6$ double floating point operations that can on an up-to-date computer[5] and for speed optimised algorithms[6] be processed within less than 20 ms.

The additional delays intrinsic to the beam position acquisition system, COD power converter, and technical network are summarised in Table 6.1. The expected total open-loop delay is in the order of 65-75 ms assuming a nominal global feedback sampling frequency of 25 Hz or ultimate sampling frequency of 50 Hz. The total open-loop delay becomes dominated by the controller's sampling delays for feedback sampling frequencies less than 10 Hz.

Beam Position Monitor System:	
acquisition (255 samples @f_{rev} 11 kHz)	~ 12 ms
Processing and Sending of the Data	~ 5 ms
Technical Network	< 1 ms
Feedback Controller:	
Network Inbound (1GBit/s)	~ 2 ms
Data Processing (matrix multiplication...)	< 20 ms
Resampling to 50 resp. 25 Hz	$10 - 20$ ms
Network Outbound (1GBit/s)	~ 2 ms
Technical Network	< 1 ms
COD power converter:	
Network Inbound (1GBit/s)	~ 2 ms
WorldFIP sampling clock and RST algorithm (50 Hz)	~ 10 ms
Total:	$\approx 65 - 75$ ms

Table 6.1: Summary of expected computational and transport delays. According to equation 3.74, the effective sampling delays is half of the sampling period.

Some of these delays are concurrent and thus the open-loop is usually longer than the feedback sampling delay. Still, it has to be ensured that each involved device is capable of processing its data within the targeted feedback delay. In order to guarantee a reproducible deterministic response, it is necessary that the open-loop delay is constant within limits. In computer science, this is referred to as 'real-time' requirement [7].

The analysis and feedback test were performed on a commodity system using standard personal computers running a real-time enhanced Linux operating system [142, 143]. Though providing worst case latencies in the order of few ten microseconds, small delay uncertainties due to the operating system's drivers and hardware involved are inevitable. In order to ensure a predefined response and loop performance, it is important that the absolute upper bound of the open-loop delay is known. In the LHC, the precise fixed open-loop delay and synchronisation between the subsystems will be ensured by the slow timing system that guarantees a LHC wide synchronisation better than 90 ps. For details on the LHC timing system, see [162, 163]. Using the timing system it is possible to trigger and synchronise the individual processing steps and define a "delay budget" in addition to the expected open-loop delay, which can be used to relax the hard timing constraints on the individual involved sub-systems. As a working assumption, the total deterministic open-loop delay is assumed to be 100 ms.

[5] HP Proliant DL 585 with 8.8 GFlops/s peak performance, 11.2 GB/s memory bandwidth and HP Proliant DL 380, with four AMD Opteron (875) processors. Courtesy to E. Genuardi and Hewlett-Packard.

[6] ATLAS, Automatically Tuned Linear Algebra Software [156–159]

[7] *A process is referred to as a 'real-time process' if the process acts to a certain stimuli within some small fixed upper limit of response time τ_m.*

The term 'real-time' does not imply that the process has to run very fast or with a high repetition rate but that the process in any case must not last longer than the specified upper limit τ_m after it is invoked[8].

A more formal mathematical approach of this definition can be found in [136]. More information on the topic of real-time aspects of operating systems can be found in [161].

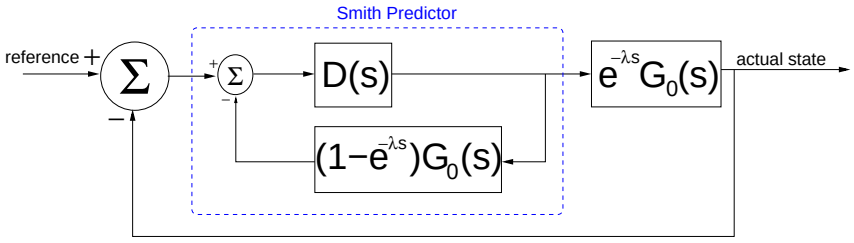


Figure 6.25: Smith Predictor scheme in its traditional form, first introduced in [53]. The internal feedback path introduces an inhibitor signal that compensates for integral wind-up due to a pure time delay λ only. Using this scheme, the controller $D(s)$ design depends only on the un-delayed process $G_0(s)$.

Smith Predictor

The closed loop robustness can in the presence of large open-loop delays be increased by reducing the effective bandwidth ($\frac{1}{\alpha}$) which increases the available phase margin at the closed loop bandwidth frequency. A more elegant approach with better performance can be achieved using the linear controller that has been optimised for the perfect un-delayed plant and including an inhibitor signal prior to the controller as first proposed and discussed by Smith [53]. The principle is based on the idea that an additional signal prior to the controller reduces the error signal seen by the controller until the previous corrections have been applied. Figure 6.25 shows the modified control loop including the so-called *Smith-Predictor*. The Smith-Predictor is given in its traditional form [53]. Smith found his predictor through a semi-analytic-empirical approach. Using modern control theory techniques, the controller can be also derived using Youla's affine parameterisation. The delayed plant process model can be written as

$$G(s) = G_0(s) \cdot e^{-\lambda s} \tag{6.59}$$

with λ being the pure transport lag. Due to the causality, the delay term cannot be inverted, thus using the same first-order roll-off function $F_Q(s)$ as in equation 3.65 yields the following Q parameterisation:

$$Q(s) = F_Q(s) \cdot G_0^{-1}(s) \tag{6.60}$$

Inserting equation 6.60 into equation 3.46 and using equation 6.59 yields the following controller parameterisation

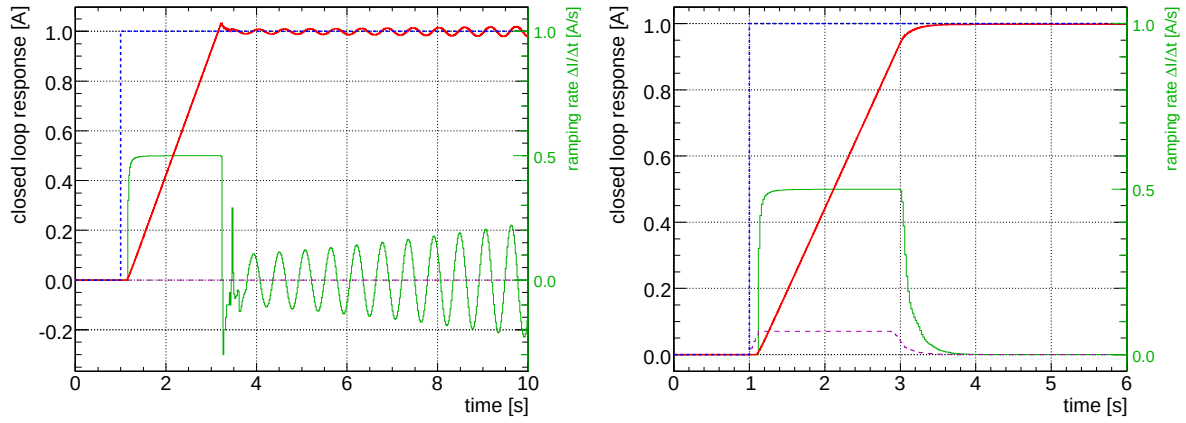
$$D(s) = \frac{D_0(s)}{1 + D_0(s)(1 - e^{-\lambda s})G_0(s)} \tag{6.61}$$

with $D_0(s)$ the linear controller for the un-delayed plant. Equation 6.61 has the same controller structure found by Smith and depicted in Figure 6.25. Inserting equation 6.60 into the closed loop transfer function

$$T_0(s) = Q(s)G_0(s) \cdot e^{-\lambda} \tag{6.62}$$

shows that the total closed loop transfer function is given by the closed loop response and controller design $Q(s)$ and $D(s)$ respectively, which is based on the un-delayed process $G_0(s)$ but shifted by a time constant λ. Figure 6.26 shows the comparison of step responses with and without Smith Predictor compensation. The growing oscillatory transient is visible. The second compensated response also includes the anti-windup scheme discussed in the next section. The simulation of the orbit corrector response and RST controller algorithm ([130]) has been performed with a step size of 1 ms. The actual orbit feedback control loop assumes a step size of 20 and 40 ms. Details on the choice of feedback sampling frequency are found in further sections.

It is important to note: Sometimes the Smith-Predictor is erroneously branded to be sensitive to modelling errors of the process $G(s)$. However, as visible in equations 6.62 and 3.57, the stability and robustness of the controller is given by the choice and robustness of $Q(s)$, whose design is based on the plant response without the delay. The strength of the Smith predictors is that it makes the choice of $Q(s)$ and $D(s)$, respectively, independent of the time delay. Further information can be found in [52]. Figure 6.27 shows the deteriorating effect of mismatches or inaccuracies in the open loop delays. The loop is more stable for an underestimation of the delay compensation which is in agreement with the causality constraint and available phase margins. Still, the results with large mismatched Smith predictors show more stable responses than without delay compensation. In case of jitter, the loop shows stable responses for delay errors of up to at least $\pm 100\,\%$.



(a) no Smith predictor

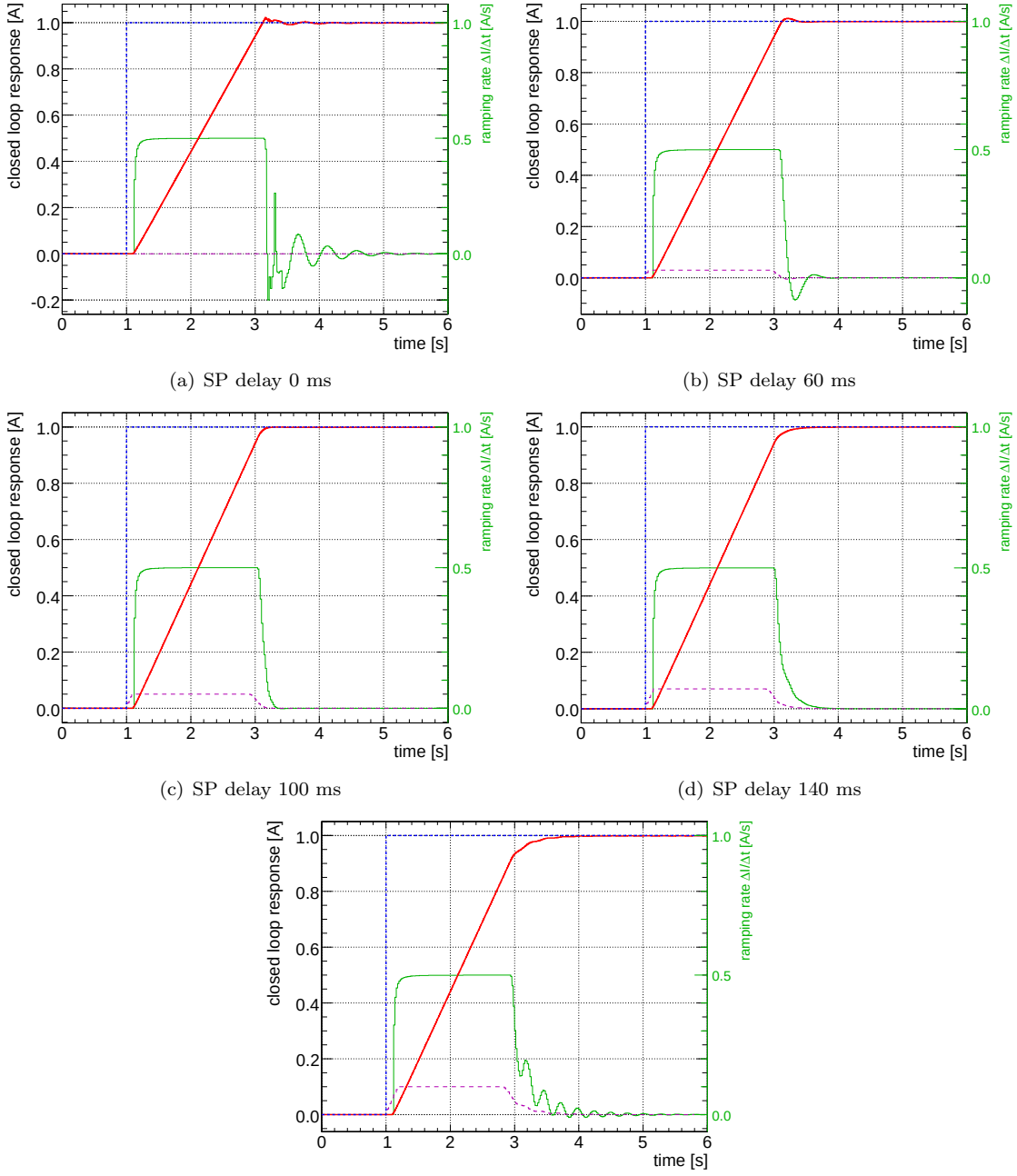
(b) with Smith predictor

Figure 6.26: Closed loop feedback response to a reference step function without (left) and with Smith predictor (right). The reference function (blue), the current response (red), current change rate (green) and differential controller signal (violet) are plotted.

In case of the orbit feedback controller, both – the PID controller and Smith Predictor – are implemented as a sampled system. It is clear that the Smith Predictor can only generate inhibit signals for delays that are larger than the sampling frequency. For the digital application, the actual continuous delay λ can be split into two parts: a multiple NT_s of the sampling period and a remaining rest delay λ_s

$$e^{-\lambda s} = e^{-NT_s s} \cdot e^{-\lambda_f s} = e^{-NT_s s} \Delta G_\lambda(s) \tag{6.63}$$

that has to be treated as a multiplicative process model error $\Delta G_\lambda(s)$. The actual controller design has to take this model error into account and round off the bandwidth, using for example the parameter $\frac{1}{\alpha}$, before the model error becomes significant.



(a) SP delay 0 ms

(b) SP delay 60 ms

(c) SP delay 100 ms

(d) SP delay 140 ms

(e) SP delay 200 ms

Figure 6.27: Effect of the the Smith predictor under various compensated delay mismatches. The true open-loop delay is 100 ms.

Rate Limiter and Anti-Windup Loop

In order to improve and achieve faster response times, it is possible to momentarily overdrive and supply higher excitation signals (voltages) to the corrector circuits than required for the steady-state. However, the maximum available excitation signal is usually limited by saturation and current rate-limits described in Section 5.2.1. In case the integral signal is too large and increases faster than the circuit is able to follow, the closed loop response shows the so-called *windup* phenomena illustrated in Figure 6.28. In

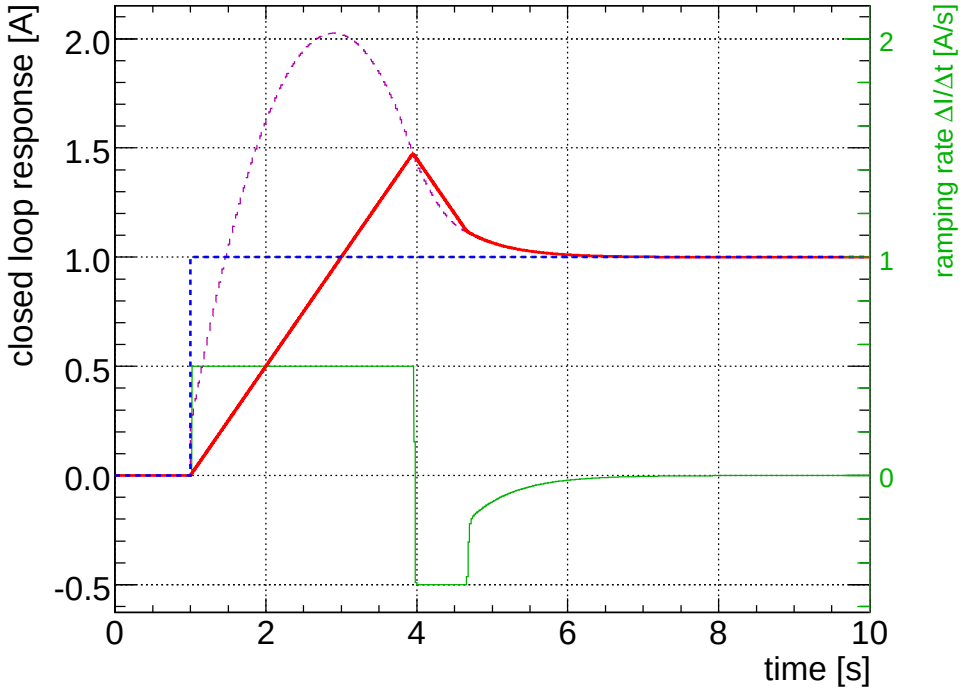


Figure 6.28: Integral windup due to COD rate limiter. The transient overshoot of the requested (dotted purple) and actual current (solid red) is visible.

first order, the PI controller is optimised for a maximum current excitation amplitude of 0.1 A and can thus achieve bandwidths of about 1 Hz. If the requested current is 100 times larger, the open-loop time constant is significantly increased and the PI controller cannot reproduce the same loop bandwidth as for the smaller current working point. While the actual current is lagging behind due to the rate limiter, the requested current builds up in order to optimise the rise time, which works for linear-only systems. Until the actual current reaches its steady state value, the integral controller signal buildup is already so large causing the actual current to perform a transient overshoot.

A first order approach would be to minimise the integral control to a level that respects the rate limit for a given operational range and thus avoids the windup. However, using this scheme, the closed loop bandwidth may become arbitrarily small and the closed loop ineffective, possibly useless. A more robust approach can be achieved through the deployment of an *anti-windup* scheme based on a similar principle as the Smith predictor. An additional internal controller feedback loop simulates the effect of the rate limiter and adds an inhibitor signal prior to the actual controller. In case the current rate limit is exceeded, it prevents further increase of the integral contribution of the PI controller.

The windup of a single circuit is especially an issue in a MIMO system since a windup of one circuit can cause coupling within the system and weak overall instability. Thus, in case a single circuit exceeds the rate limit, the other correctors' strength changes need to be rescaled by the same factor as the single circuit in order to provide a coherent solution. An example for the windup compensated response is shown in Figure 6.26.

6.3 LHC Orbit Feedback Controller Performance

The compensation for delays and non-linearities can in principle be performed in the same predictor path. It was decided to keep the anti-windup and Smith Predictor schemes separate, in order to be able to perform independent studies on both of them. Figure 6.29 shows the combined control scheme of the SVD-based orbit correction, the linear PI controller, Smith Predictor and anti-windup scheme discussed in the previous sections.

The basic closed loop structure containing the process $G(s) \cdot e^{-\lambda s}$, monitor response $M(s)$ and controller $D_{PID}(s)$ is visible. It has to be noted that arrows illustrating the signal flow are actually multi-dimensional. The block labled 'SVD' decouples its m inputs to n COD current outputs that are forwarded to n independent $D_{PID}(s)$ controllers with each an individual Smith predictor path. The rate-limter is evaluated for each invidual circuit. In case a single corrector saturates or exceeds its current rate limit, D_{norm} normalises the requested current strengths of all correctors down, in order to respect the speed of the slowest corrector and to preserve the decoupling of the n loops. The term $F_0(s)$ compares the predicted of the power converter with its actual state and adds a feed-forward prediction, discussed in Section 6.1.4, to the controller output in order to mitigate the effect of failing or unresponsive power converters.

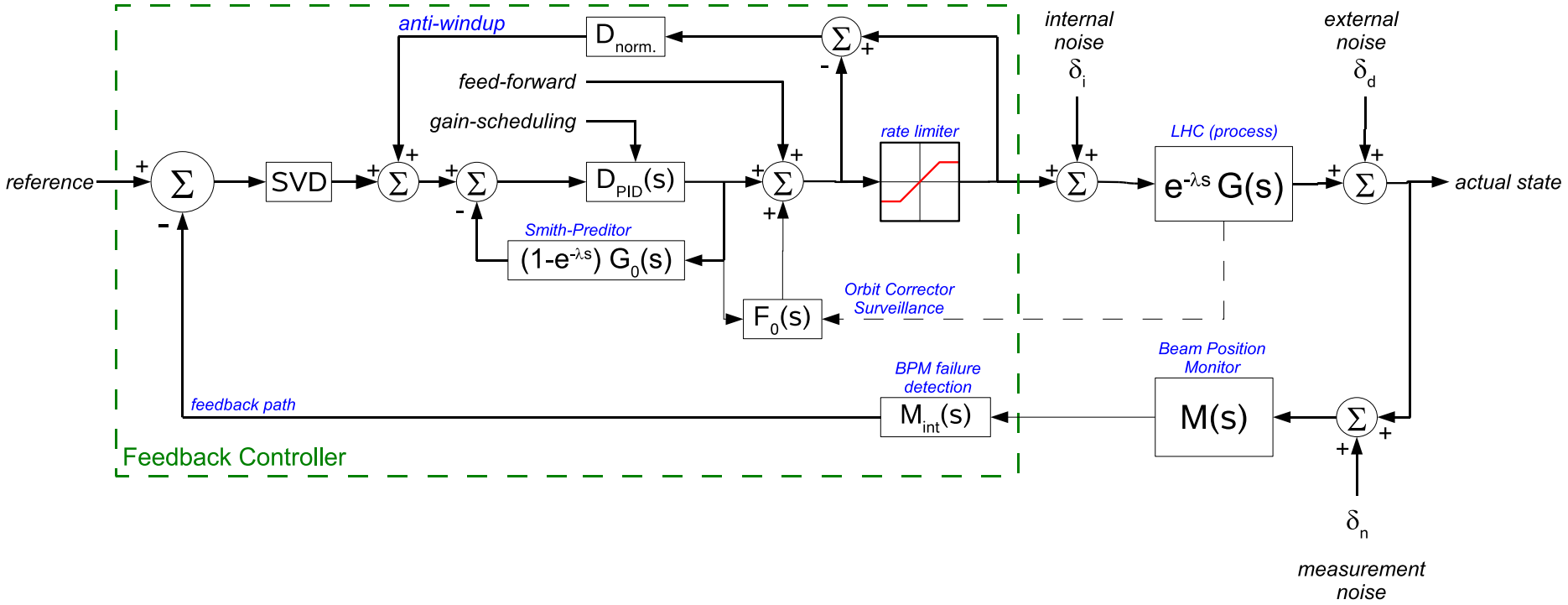


Figure 6.29: Full LHC orbit feedback block diagram. $G(s)$ abreviated $G(s) = G_0(s) \cdot G_{NL}(s)$ for better readablitiy. The function of the different branches are explained Sections 6.2.2 and 6.2.2.

6.3.1 Orbit Feedback Test-System

In order to test the feedback structure presented in Figure 6.29, its performance under long-term continuous operation, it is important to test the controller under real load and in realistic conditions prior to LHC completion.

A testbed complementary to the feedback controller was developed, which simulates a realistic open loop orbit response in order to simulate effects due to delays, BPM and magnet responses and to test the OFC with a realistic load. To simplify debugging and commissioning of the feedback with the real LHC, the testbed has essentially the same control interfaces as the real machine to make it transparent for the controller. The testbed operates in two distinguished modes:

- Real-time with a reduced closed loop simulation about eight times faster than the feedback sampling frequency.
- Sampled mode with discrete 1 ms time steps that is used to simulate computational complex effects such as eddy currents, power converter specifics and inter-sample responses. In particular, though the feedback sampling was limited to 50 Hz, perturbation sources with frequencies up to about 200 Hz were simulated in order to assess the effect of aliasing onto the feedback response.

Figure 6.30 shows the schematic block diagram of the test system. Based on equations described in Chap-

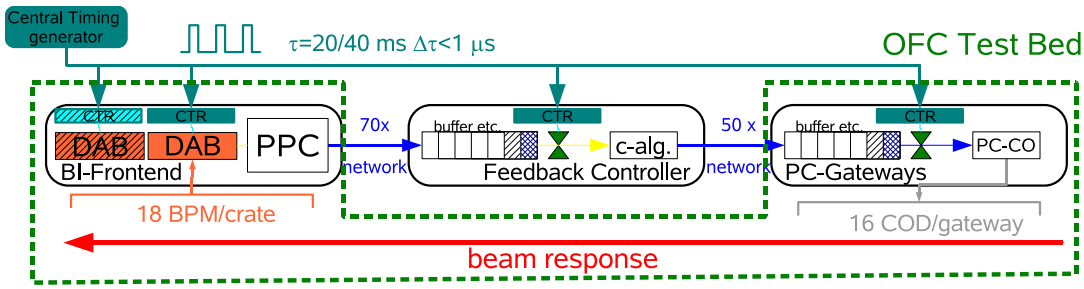


Figure 6.30: Orbit Feedback Controller (OFC) testbed architecture. The system mimics the COD, beam response and BPM behaviour while implementing the same control interfaces that will be used in the LHC.

ters 2 and 3, the simulation includes effects such as dipole deflections of orbit corrector and main dipole magnets, dipolar quadrupole feed-downs, eddy currents, field decays, noise and systematic measurement errors as well as element failures. Since the system implements the same data exchange mechanisms as for the LHC system, it is possible to assess controller performance and stragegies. The system can operate in a manual clocked and real-time mode in which the loop response is simulated in real time.

It is foreseen to use the testbed beyond the feedback development and tests for verification and quality assurance of new control schemes and software revisions prior to their application to the real LHC machine.

6.3.2 Effect of Sampling Frequency

In most digital approximations, analogue systems are sampled at least 10 times higher than the analogue bandwidth. Since the deteriorating effect due to beam response matrix uncertainties can be mitigated through a higher sampling frequency (see section 6.1.4 for details), it is usually favourable to sample much higher than a factor of 10. In case of synchrotron light sources, typical sampling to bandwidth ratios are about 40 or more [77, 164, 165].

Figure 6.31 shows the effect of different choices of sampling frequencies onto the closed loop response. The nominal sensitivity and disturbance rejection are plotted. The designed closed loop bandwidth of 0.8 Hz ($\alpha = 0.2$) is essentially independent of the sampling frequency. However, the higher the sampling frequency the better the digital representation of the analogue first-order system and the smaller the transient overshoot of the loop response for frequencies in the range of about 1 to 10 Hz.

The initial LHC orbit feedback requirements suggested a reference sampling frequency of 10 Hz [118]. Tests at the SPS have shown that this frequency can a priori be increased significantly. At 7 TeV/c the CODs are able to produce sinusoidal orbit perturbations with amplitudes of 13 μm at 1 Hz, which is well above the noise level of $\approx 5\,\mu$m of the BPMBs. To take advantage of the power converter's performance, it may be interesting to increase the sampling frequency for the following reasons:

1. A higher sampling frequency improves the performance due to a smaller sampling delay and better digital approximation of the analogue process.

2. The loss of one data sample becomes less critical.

3. For operation, it is desirable to have a certain safety margin in order to cope with plant inaccuracies due to lattice imperfections discussed in Section 6.1.4.

The available infrastructure is capable of supporting sampling frequncies of 25 Hz and has been tested up to 100 Hz. The maximum sensible frequency is however limited by the COD control system to be less than 50 Hz. The loop responses without model uncertainties are essentially the same for sampling frequencies above 25 Hz. Thus, further figures have been generated for 25 Hz sampling.

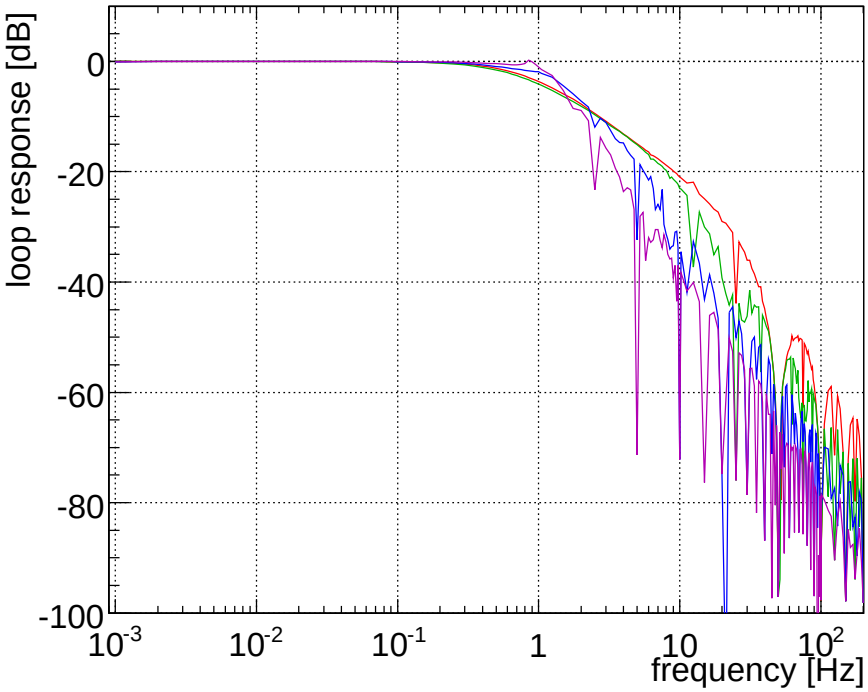
6.3.3 Feedback Performance Summary

This section summarises the total orbit feedback performance which has been obtained using the testbed simulation. Most results are given with respect to operation at beam momentum of 7 GeV. The bandwidths or maximum possible excitation amplitude, respectively, for operation during injection (450 GeV) scale with energy by a factor of $\gamma_{coll}/\gamma_{inj} \approx 15.6$. The following figures illustrate the complete feedback performance by showing the closed loop response (nominal sensitivity), which is the response to a change of the reference and the closed loop sensitivity which is the response to an external perturbation input. Figure 6.32 shows the nominal feedback response and disturbance rejection for a closed loop running close to the open loop bandwidth limit of 1 Hz for an oscillation amplitude of 13 μm.

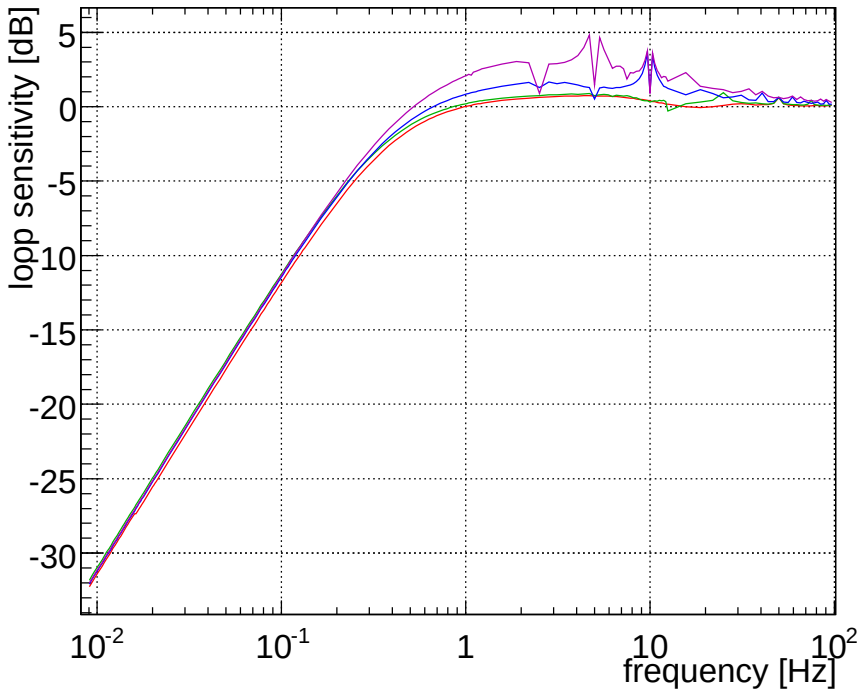
Figure 6.33 shows the effect the controller parameterisation parameter α on the closed loop sensitivity. The perturbation of 0.1 A was used corresponding to an open loop bandwidth of about 1 Hz. It is visible that the closed loop bandwidth increases linearly with decreasing value of α. For a closed loop bandwidth of about 0.8 Hz ($\alpha = 0.2$) the transient overshoot reaches up to 26% (2 dB) for perturbation frequencies of about 4 Hz. The overshoot is due to the the remaining sub-sample delay that cannot be fully compensated by the Smith predictor path as described in Section 6.2.2.

Figure 6.34 shows the measurement noise suppression for various noise levels. The minimum noise reduction is about a factor of two for a signal-to-noise ratio of 20 dB.

Figure 6.35 shows the step response and effect of the feedback loop containing the Smith predictor for a small (1.3 μm) and large step (7 mm). It is visible that the Smith predictor response produces qualitatively the same result without transient overshoot, independent of the amplitude of the excitation current. It is visible that the loop respects the intrinsic limits control sensitivity by limiting the maximum current change to less than 0.5 A/s. Figure 6.36 shows the closed loop step response for an $\pm$50% open loop COD calibration error. Figure 6.37 shows exemplary step responses with different amplitudes.

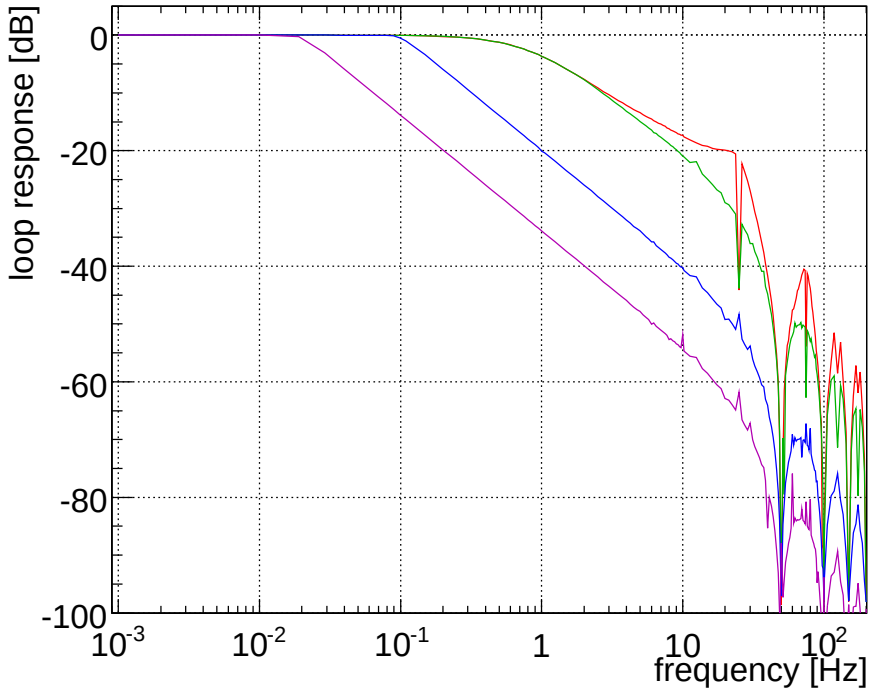


(a) loop response

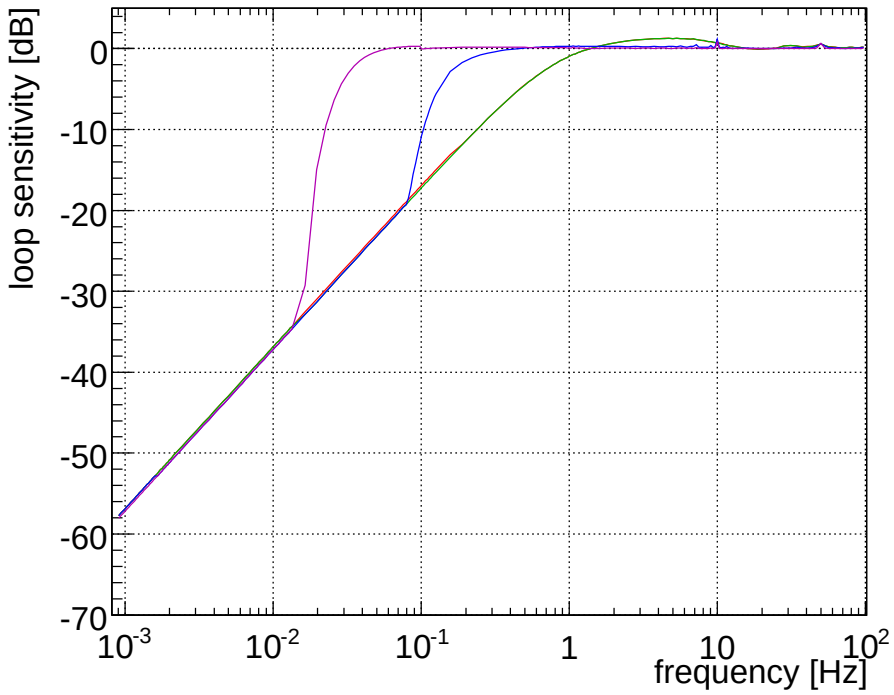


(b) loop sensitivity

Figure 6.31: Nominal loop response (top) and sensitivity (bottom) as a function of sampling frequency (red: 50 Hz; green: 25 Hz; blue: 10 Hz; purple: 5 Hz).



(a) nominal feedback transfer function T_0



(b) nominal feedback disturbance rejection S_{d0}

Figure 6.32: Bode plot of the simulated transfer function $|T_0(i2\pi f)|$ and disturbance rejection $|S_{ds}(i2\pi f)|$ of the nominal orbit feedback loop. The individual curves correspond to requested reference, disturbance signals with amplitudes of 0.01, 0.1, 1.0 and 5 A, respectively. It is visible that the closed loop bandwidth decreases for increasing currents. For comparison: a MCBH/V current of 0.1 A corresponds to an orbit oscillation amplitude of about 13 μm at 7 TeV.

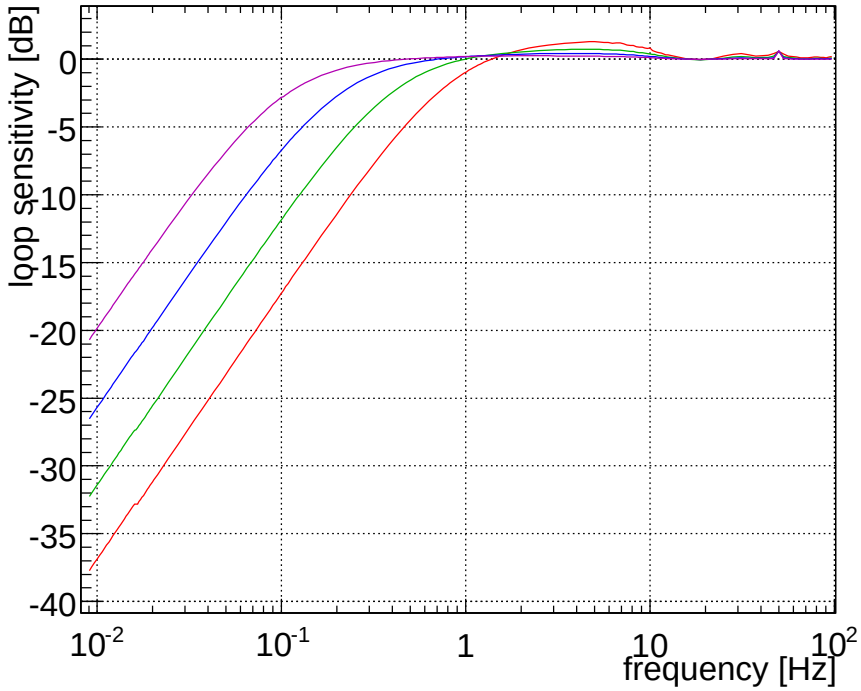
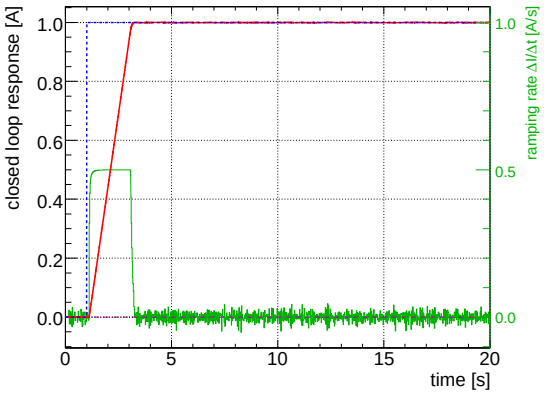
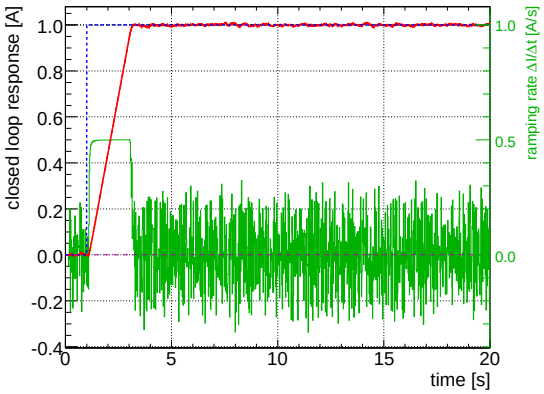


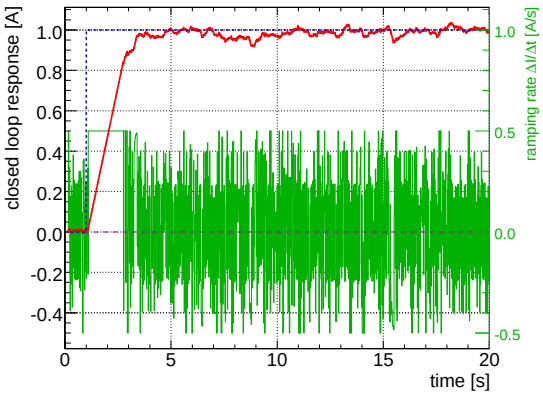
Figure 6.33: Effect of closed loop bandwidth parameterisation. The sensitivity functions for $\alpha = 1.6$, 0.8, 0.4 and 0.2 are shown (from left to right).



(a) 0.001A noise

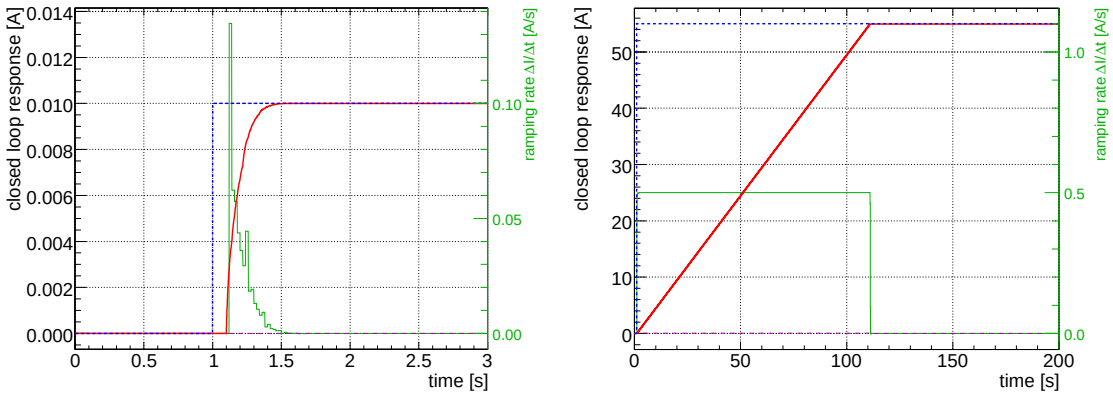


(b) 0.01A noise



(c) 0.1A noise

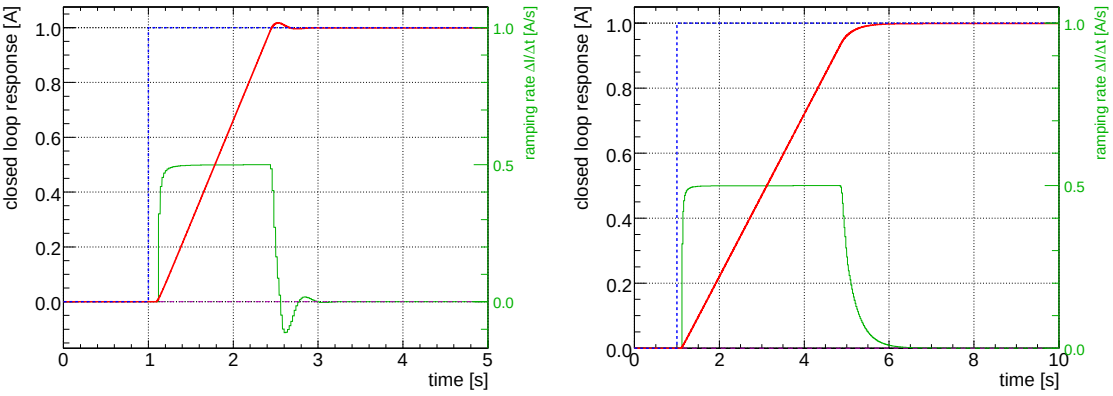
Figure 6.34: Measurement noise rejection. The equivalent BPM noise levels are $\pm 13\,\mu$ (0.1 A), $\pm 1.3\,\mu$ (0.01 A) and $\pm 0.13\,\mu$ (0.001 A). The minimum relative noise reduction is about a factor of two for a signal-to-noise ratio of 20 dB.



(a) 0.01 A step response

(b) 55.0 A step response

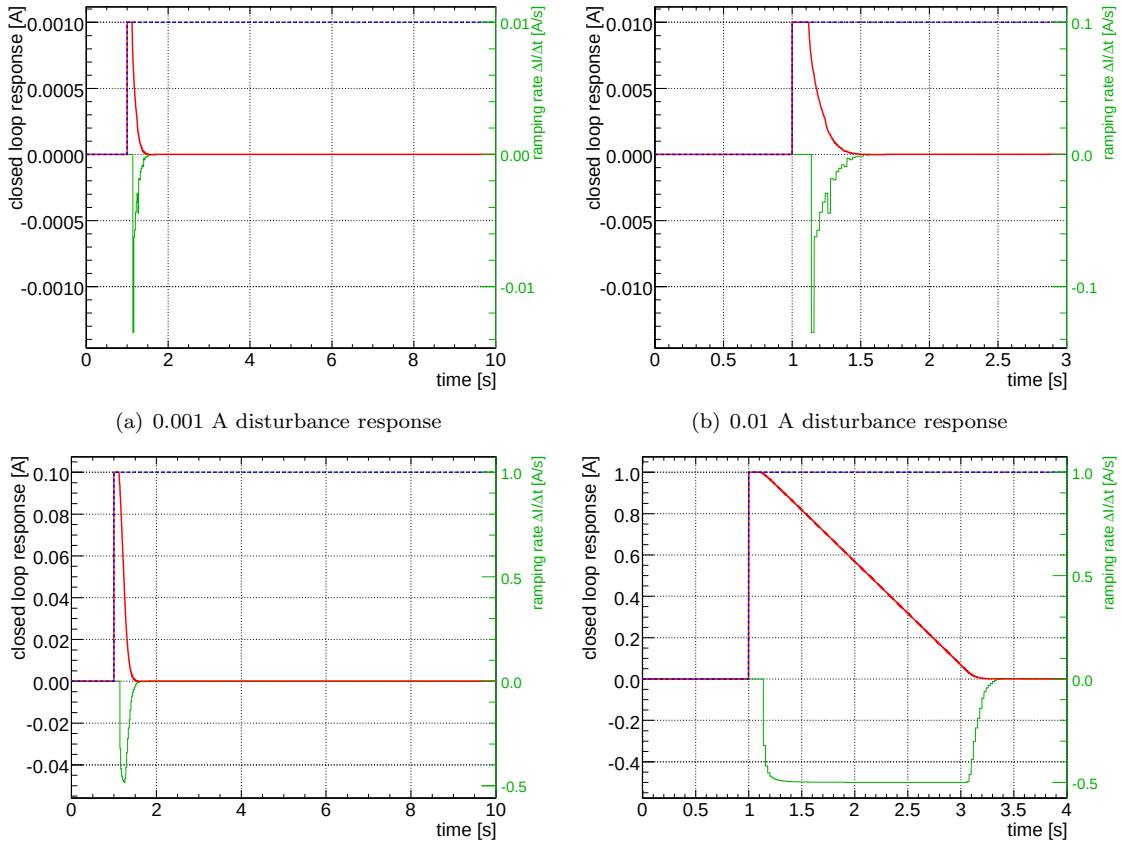
Figure 6.35: Effect of the Smith predictor on a small and large reference current change.



(a) SP delay 0 ms, +50% COD error

(b) SP delay 0 ms, -50% COD error

Figure 6.36: Effect of the Smith predictor including a 50% negative (left) and 50% positive systematic COD calibration error.



(a) 0.001 A disturbance response

(b) 0.01 A disturbance response

(c) 0.1 A disturbance response

(d) 1.0 A disturbance response

Figure 6.37: Disturbance rejection of the simulated nominal orbit feedback loop to a step response. The responses to external step disturbances with 0.001, 0.01, 0.1 and 1 A amplitude are plotted. For comparison: an MCBH/V current of 0.1 A corresponds to an orbit oscillation amplitude of about 13 μm at 7 TeV.

6.4 Outlook

From the design point of view, the proposed feedback scheme is optimal with respect to the assumed BPM and COD responses. However, this scheme is fundamentally based on the working assumption that decouples orbit control into a spatial and temporal component in order to facilitate adjustments in case of element failures and simply commissioning procedures. For the time being, as an implication of this trade-off, the fastest local orbit correction is determined by the amplitude of the largest global orbit perturbation. A more elaborate but also more complex design would focus on a direct global solution of equations 6.3 and 6.4 with the draw-back of significantly reduced flexibility in case of element failure and required loop adjustment as well as limited flexibility of incorporating classic but well-established orbit correction techniques.

The LHC is expected to be commissioned at the end of 2007. In contrast to schemes based on feed-forward only that require a thorough understanding of the lattice parameters, feedbacks (where applicable) can operate already with a coarse knowledge of the accelerator lattice and in the presence of large errors through using continuous beam measurement as direct observable to minimise the parameter residual error. However, it is clear that the feedback performance is limited by the quality of involved monitors and corrector circuits which might be only partially or with limited range of function available. It was recognised [166, 167], that for LHC operation, the merit of feedbacks on beam parameters is not limited to nominal operation but also beneficial when used during an early commissioning stage, when the knowledge on magnetic fields, external perturbations and other effects is usually poor. It is hoped, that the feedbacks proposed and discussed in this thesis will be helpful in these efforts.

Appendix A

Acronyms and Abbreviations

A.1 Acronyms

ADC	Analogue to Digital Converter
B1	LHC beam rotating clockwise
B2	LHC beam rotating counter-clockwise
BPM	Beam Position Monitor
BLM	Beam Loss Monitor
BW	Bandwith
CERN	originally: *Conseil Europeen pour la Recherche Nucleaire* now: European Organisation for Nuclear Research
COD	Closed Orbit Dipole Corrector Magnet
CL	Closed Loop
DAB	Digital Acquisition Board
DAC	Digital to Analogue Converter
D1-D2	Dipole pair joining the two beams before and separating them after the IP
D3-D4	Dipole pair increasing and decreasing the beam separation in IR3, IR4 and IR7
FB	Feedback
FF	Feed-Forward
IP	Interaction Point
IR	Interaction Region
LEP	Large Electron Positron Collider at CERN
LHC	Large Hadron Collider at CERN
LSS	Long Straight Section
MB	Main dipole magnet (bending magnet)
MCB	Cold arc orbit corrector dipole magnet
MCBX	Cold orbit corrector dipole magnet found in the final-focus triplets
MKE	Beam Extraction Kicker Magnet
MKI	Beam Injection Kicker Magnet
MQ	(Lattice) Quadrupole Magnet
MQT	Trim Quadrupole Magnet
MQS	Skew Quadrupole Magnet
MS	(Lattice) Sextupole Magnet
MSD	(Dump) Extraction Septum Magnet
MSI	Injection Septum Magnet
PC	Power Converter
RHIC	Relativistic Heavy Ion Collider at Brookhaven Laboratories
RF	Radio Frequency (cavity)
SSS	Short Straight Section
TAN	First absorber ouside the IP that intercepts particles escaping along the beam axis
TCDQ	Extraction protection absorber in IR6
TCP	Primary collimator of the LHC Cleaning System
TCS	Secondary collimator of the LHC Cleaning System
TDI	Injection protection absorber in IR2 and IR8
WBTN	Wide-Band-Time-Normaliser

A.2 List of Symbols

α_c	momentum compaction factor	
$\alpha(s)$	$\alpha(s) := -\frac{1}{2}\beta'(s)$	
$\beta(s)$	betatron function	[m]
β	relativitic beta function $\beta = \frac{v}{c}$	
β^*	betatron function at the interaction point	[m]
$\gamma(s)$	$\gamma(s) := \frac{1+\alpha^2(s)}{\beta(s)}$	
γ	relativistic gamma function $\gamma = \frac{1}{\sqrt{1-\beta^2}}$	
$\kappa(f)$	lattice amplification factor, see equation 4.9	
$\Delta\mu(s)$	betatron phase advance	[rad]
ϵ	beam emittance	$[\mu mrad]$
ϵ_n	normalised beam emittance $\epsilon_n = \frac{\epsilon}{\gamma}$	$[\mu mrad]$
$\sigma(s)$	r.m.s. beam size $\sigma = \sqrt{\frac{\epsilon\beta(s)}{\gamma}}$	$[\mu m]$

Appendix B

Linear Beam Dynamics

B.1 Three and Four Corrector Bump

In contrast to global orbit correction, it is possible to perturb the orbit only locally, using an appropriate linear combination of dipole corrector responses, given in equation 2.71. The corresponding localised orbit change is commonly referred to as a *bump* and is often used to change the position at a single device of interest while leaving the rest of the orbit in the machine unaffected.

The simplest 'bump' can be constructed by superimposing the orbit perturbation of two dipole kicks that are spaced by a phase advance of π. Depending on the number of correctors, it is possible to not only adjust the position Δx but also the deflection $\Delta x'$ as well as meet other constraints of the beam at a given reference point p between the selected kick positions.

The two-corrector bump is often impractical due to the required phase advance of π and the reduced degree of freedom of having the peak of the bump always at the centre between the correctors. As one of the most frequently used closed orbit steering operations, the *three-corrector-bump* can steer the beam position x_p using only three orbit correctors with a possible arbitrary phase advances in between as illustrated in Figure B.1. Using the following boundary conditions for a perturbation having an amplitude

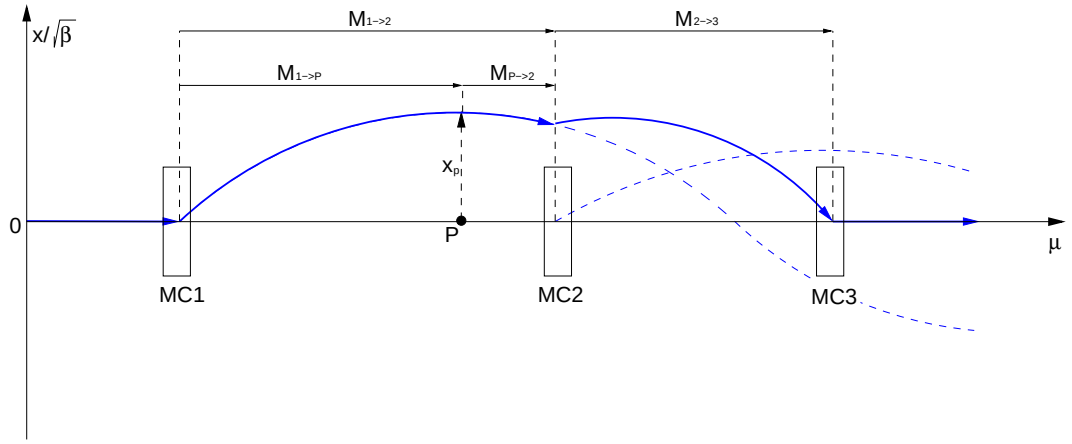


Figure B.1: Three corrector bump

Δx at the location p in between the selected correctors and

$$\underline{x}_1 = \begin{pmatrix} 0 \\ \delta_1 \end{pmatrix} \qquad \underline{x}_p = \begin{pmatrix} \Delta x \\ x'_p \end{pmatrix} \qquad \underline{x}_2 = \begin{pmatrix} x_2 \\ x'_2 + \delta_2 \end{pmatrix} \qquad \underline{x}_3 = \begin{pmatrix} 0 \\ \delta_3 \end{pmatrix} \tag{B.1}$$

$$\mathbf{M}_{13}\underline{x}_1 + \mathbf{M}_{23}\underline{x}_2 + \underline{x}_3 = 0 \tag{B.2}$$

that essentially states that the position vanishes and the beam angle is given by the corrector deflection at the first and last dipole. Using the above boundary conditions and the beam transfer matrix, one gets the following equations:

$$\mathbf{M}_{13}\underline{x}_1 = \underline{x}_3 \qquad \mathbf{M}_{12}\underline{x}_1 = \underline{x}_2 \qquad \mathbf{M}_{1p}\underline{x}_1 = \underline{x}_p \qquad \mathbf{M}_{p2}\underline{x}_p = \underline{x}_2 \tag{B.3}$$

A straight-forward solution of the above equations yields the following expressions for δ_1, δ_2, and δ_3.

$$\delta_1 = +\frac{1}{\sqrt{\beta_1\beta_p}\sin(\mu_{1p})}\cdot\Delta x \tag{B.4}$$

$$\delta_2 = -\sqrt{\frac{\beta_1}{\beta_2}}\frac{\sin(\mu_{13})}{\sin(\mu_{23})}\cdot\delta_1 \tag{B.5}$$

$$\delta_3 = +\sqrt{\frac{\beta_1}{\beta_3}}\frac{\sin(\mu_{12})}{\sin(\mu_{23})}\cdot\delta_2 \tag{B.6}$$

A similar procedure can be applied for a *four-corrector bump* which, compared to three-corrector bump, can also match the angle of the beam $\Delta x'$ at the desired location of the bump, as illustrated in Figure B.2. Here the desired location of the bump is located between the second and third dipole corrector.

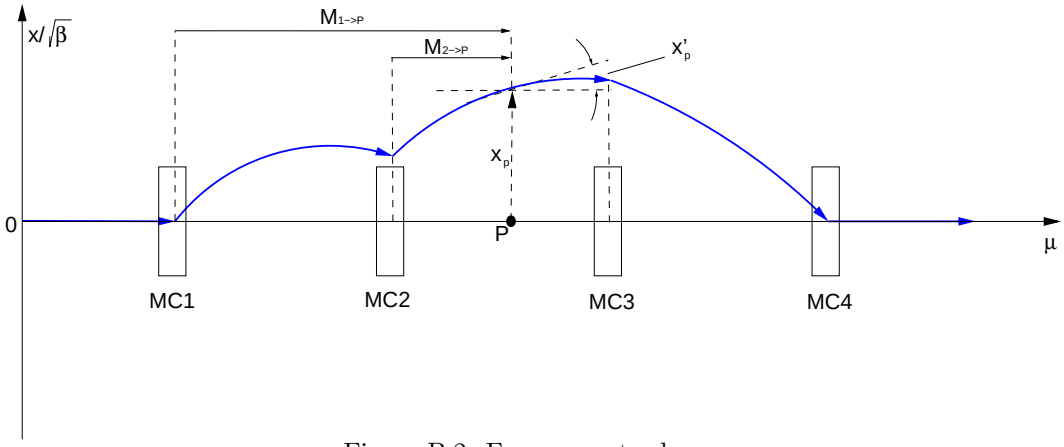


Figure B.2: Four corrector bump.

From Figure B.2, one can deduce the following boundary conditions for the first two correctors:

$$\underline{x}_p = \begin{pmatrix} x_p \\ x'_p \end{pmatrix} = \mathbf{M}_{1p}\underline{x}_1 + \mathbf{M}_{2p}\underline{x}_2 \tag{B.7}$$

$$\underline{x}_1 = \begin{pmatrix} 0 \\ \delta_1 \end{pmatrix} \quad \text{and} \quad \underline{x}_2 = \begin{pmatrix} 0 \\ \delta_2 \end{pmatrix} \tag{B.8}$$

Due to the symmetry of the problem, δ_1 and δ_2 corresponds to δ_4 and δ_3 of the correctors at the right of the reference point with:

$$\alpha \longrightarrow -\alpha \tag{B.9}$$

$$x'_p \longrightarrow -x'_p \tag{B.10}$$

The four dipole deflections required to specify the beam position Δx and beam angle $\Delta x'$ at the reference point p are given by:

$$\delta_1 = +\frac{1}{\sqrt{\beta_1\beta_p}}\frac{[\cos(\mu_{2p}) - \alpha_p\sin(\mu_{2p})]}{\sin(\mu_{12})}\cdot\Delta x - \frac{\beta_p}{\beta_1}\frac{\sin(\mu_{2p})}{\sin(\mu_{12})}\cdot\Delta x' \tag{B.11}$$

$$\delta_2 = -\frac{1}{\sqrt{\beta_2\beta_p}}\frac{[\cos(\mu_{1p}) - \alpha_p\sin(\mu_{1p})]}{\sin(\mu_{12})}\cdot\Delta x + \frac{\beta_p}{\beta_2}\frac{\sin(\mu_{1p})}{\sin(\mu_{12})}\cdot\Delta x' \tag{B.12}$$

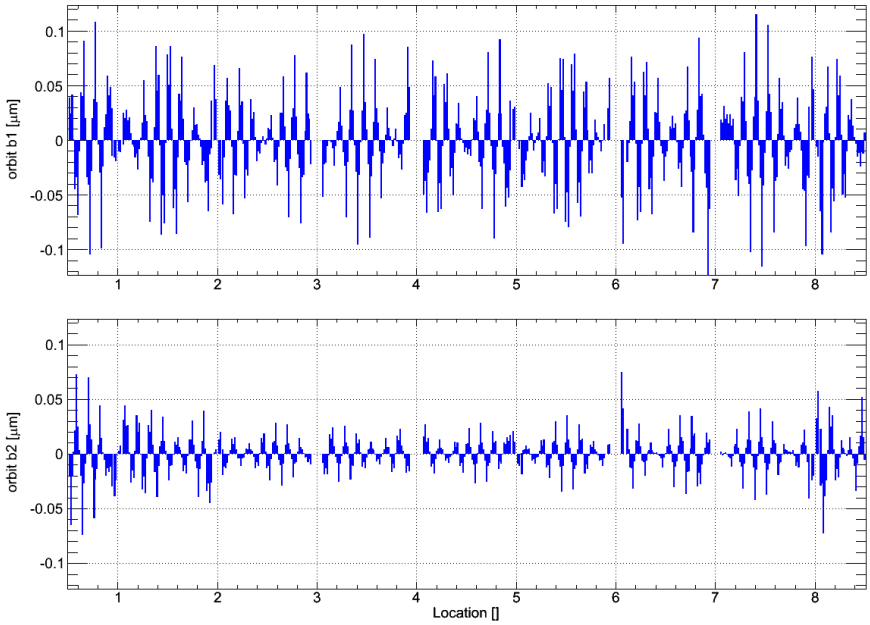
$$\delta_3 = -\frac{1}{\sqrt{\beta_3\beta_p}}\frac{[\cos(\mu_{p4}) - \alpha_p\sin(\mu_{p4})]}{\sin(\mu_{34})}\cdot\Delta x - \frac{\beta_p}{\beta_3}\frac{\sin(\mu_{4p})}{\sin(\mu_{34})}\cdot\Delta x' \tag{B.13}$$

$$\delta_4 = +\frac{1}{\sqrt{\beta_4\beta_p}}\frac{[\cos(\mu_{p3}) - \alpha_p\sin(\mu_{p3})]}{\sin(\mu_{34})}\cdot\Delta x + \frac{\beta_p}{\beta_4}\frac{\sin(\mu_{3p})}{\sin(\mu_{34})}\cdot\Delta x' \tag{B.14}$$

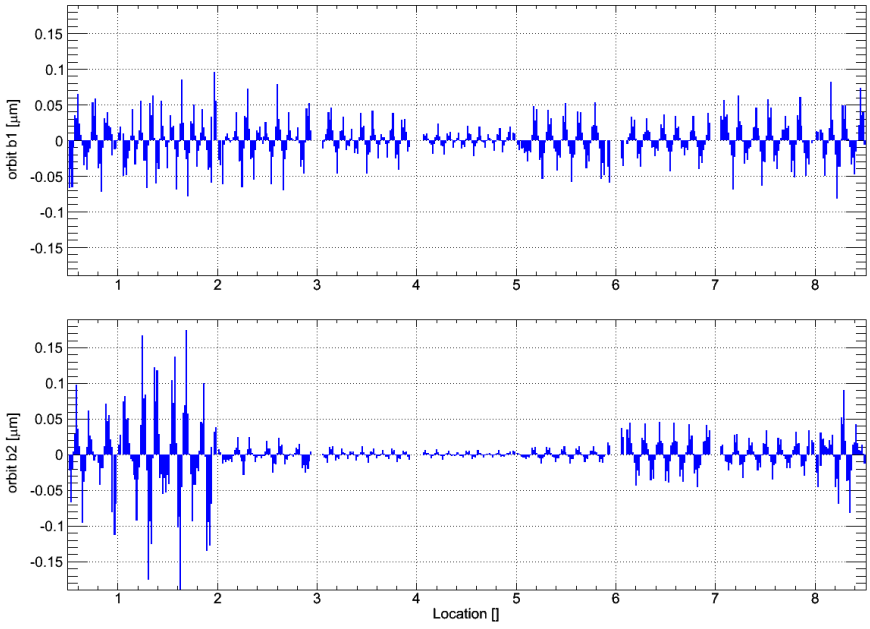
The left and right half of the four-corrector bump can be used to close an arbitrary local correction in order to leave the global orbit outside the selected region unaffected.

Appendix C

LHC SVD Orbit Eigenvectors

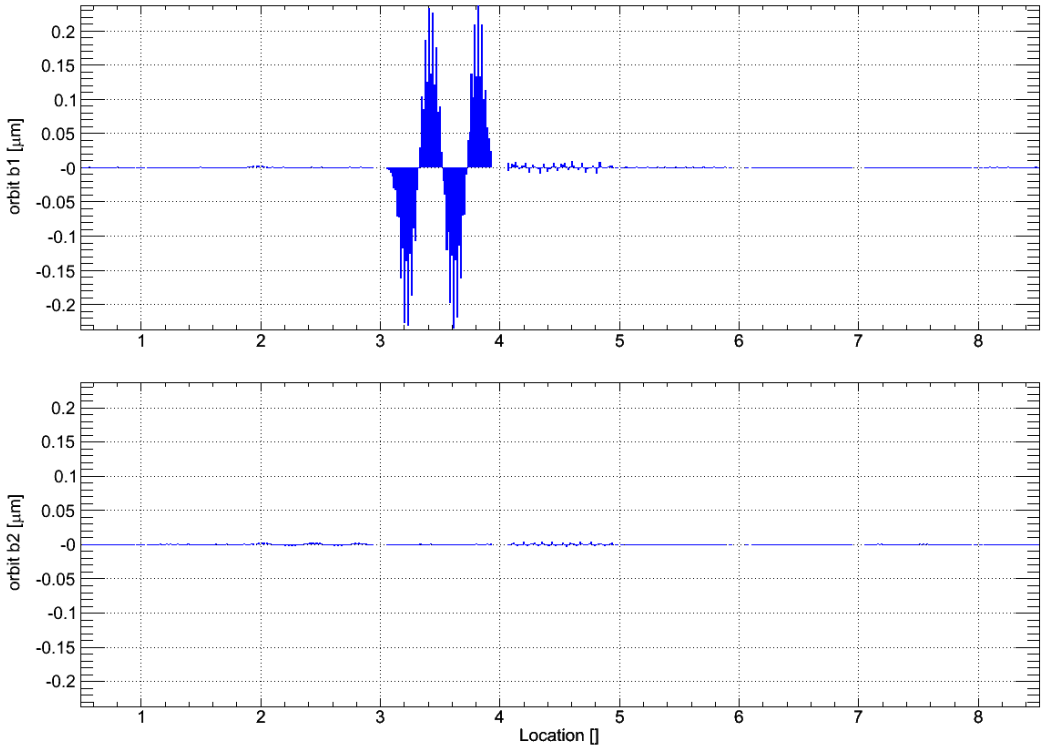


(a) $\#\lambda_{50} = 668.9$

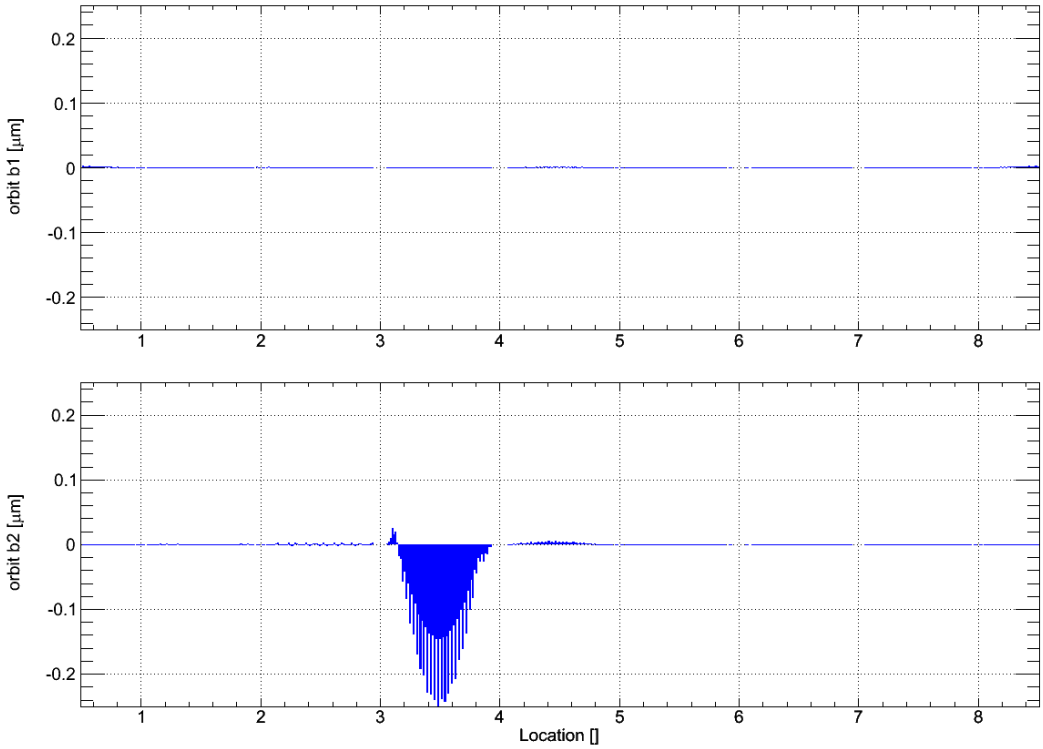


(b) $\#\lambda_{100} = 339.3$

Figure C.1: LHC orbit eigenvector for various eigenvalues and collision optics. See section 6.1.3 for details.

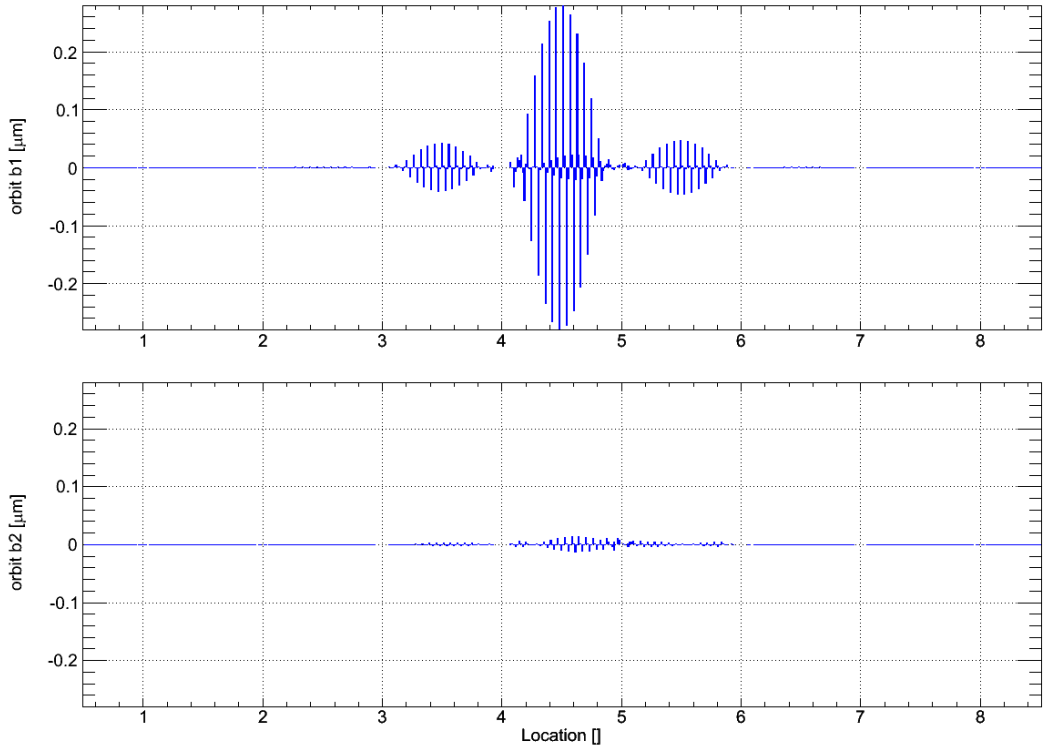


(a) $\#\lambda_{291} = 121.3$

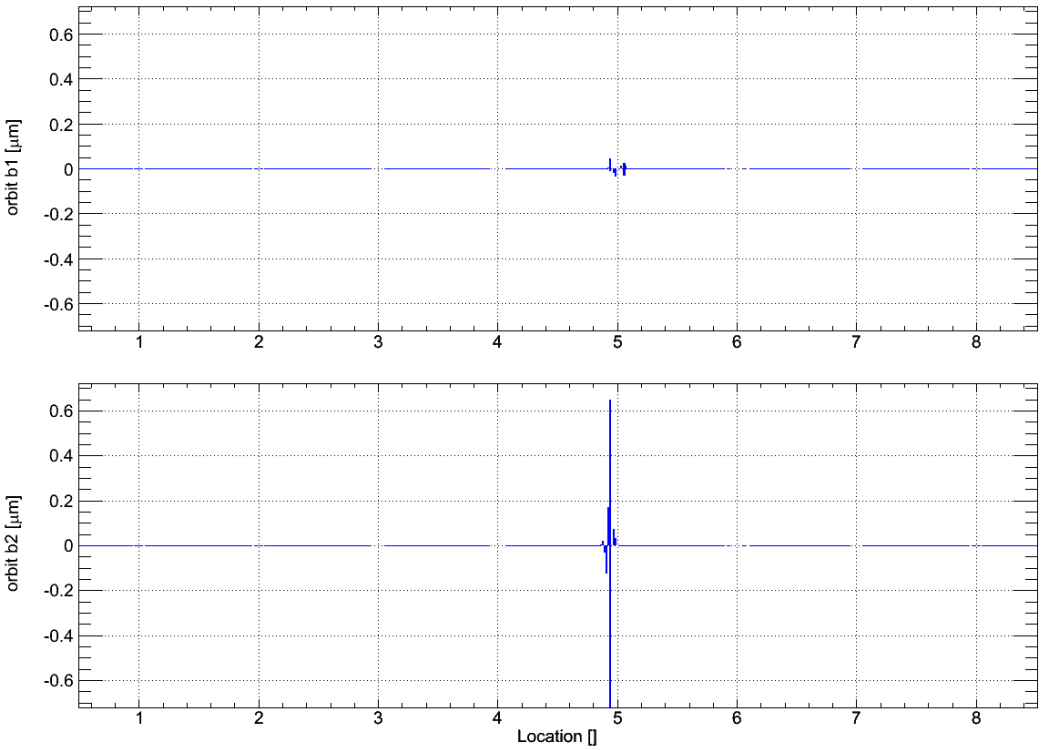


(b) $\#\lambda_{350} = 110.0$

Figure C.2: LHC orbit eigenvector for various eigenvalues and collision optics. See section 6.1.3 for details.



(a) $\#\lambda_{449} = 81.6$



(b) $\#\lambda_{521} = 1.2$

Figure C.3: LHC orbit eigenvector for various eigenvalues and collision optics. See section 6.1.3 for details.

Appendix D

Empirical PID Tuning Methods

Ziegler and Nichols [56], and later Cohen and Coon [57], studied different types of processes and developed approximative tuning rules for PID gains based on the closed and open loop step response. The presented methods should be considered as a fall-back option in case the process model is not available with sufficient precision. A more rigorous optimisation should rely on model based approaches, such as the affine parameterisation presented in section 3.4.

The first, Ziegler-Nichols oscillation-based method uses the closed loop response to a step and a proportional control only. The proportional gain K_p is increased and the step repeated till the process becomes instable and starts to oscillate. The critical proportional gain K_c and oscillation period P_c are noted at this point. As visible in the frequency response in Figure 3.5, this method essentially decreases the gain margin to zero till a phase lag of π is reached and the process is under verge of becoming instable. The optimal stable P, PI and PID gain parameter that Ziegler and Nichols found are summarised in table D.1.

	Ziegler-Nichols Oscillation Method		
	K_p	K_i	K_d
P	$0.50 \cdot K_c$		
PI	$0.45 \cdot K_c$	$0.54\frac{K_c}{P_c}$	
PID	$0.60 \cdot K_c$	$\frac{1.2K_c}{P_c}$	$\frac{3K_cP_c}{40}$

Table D.1: Optimal PID gains using the Ziegler-Nichols oscillation method. K_c is the critical proportional gain at which the closed loop plant starts to oscillate and P_c its period [56].

The second type of tuning, performed by Ziegler-Nichols and Cohen-Coon, is based on the reaction curve of the open loop system to an arbitrary reference step change from u_0 to u_{ss}. An exemplary reaction curve of a second order process is shown in figure D.1.

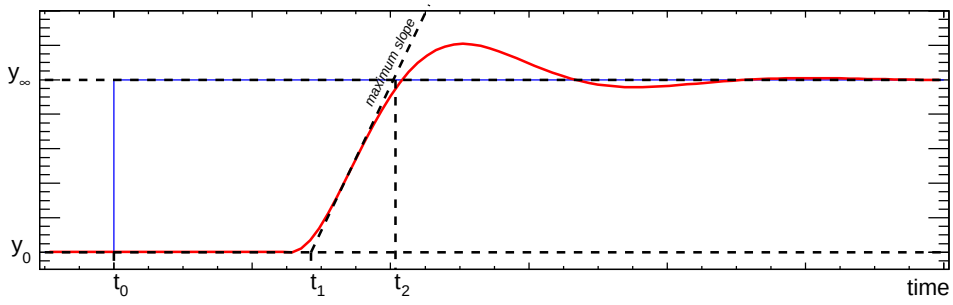


Figure D.1: Reactive curve of a second order process. The step response starts at t_0. The initial and final state are given by y_0 and y_∞, respectively. The start t_1 and end t_2 of the ramp are determined by the maximum slope tangent crossing with y_0 and y_∞.

One defines the following parameter

$$K_0 := \frac{y_{ss} - y_0}{u_{ss} - u_0}, \qquad \tau_0 = t_1 - t_0, \qquad \text{and} \qquad \nu_0 = t_2 - t_1 \tag{D.1}$$

with y_0 the initial and y_∞ the final state and t_1, t_2, t_3 as denoted in figure D.1. The gains according to Ziegler-Nicols and Cohen-Coon optimal P, PI and PID are summarised in table D.2 and D.3 Both

	Ziegler-Nichols		
	K_p	K_i	K_d
P	$\frac{\nu_0}{K_0\tau_0}$		
PI	$\frac{0.9\nu_0}{K_0\tau_0}$	$\frac{1}{3\tau_0}\cdot K_p$	
PID	$\frac{1.2\nu_0}{K_0\tau_0}$	$\frac{1}{2\tau_0}\cdot K_p$	$\frac{\tau_0}{2}\cdot K_p$

Table D.2: Optimal PID gains using the reaction curve and Ziegler-Nichols [56]. The parameter K_0, τ_0 and ν_0 are defined in equation D.1 and illustrated in figure D.1.

	Cohen-Coon		
	K_p	K_i	K_d
P	$\frac{\nu_0}{K_0\tau_0}\left[1+\frac{\tau_0}{3\nu_0}\right]$		
PI	$\frac{\nu_0}{K_0\tau_0}\left[0.9+\frac{\tau_0}{12\nu_0}\right]$	$\frac{9\nu_0+20\tau_0}{\tau_0\left[30\nu_0+3\tau_0\right]}\cdot K_p$	
PID	$\frac{\nu_0}{K_0\tau_0}\left[\frac{4}{3}+\frac{\tau_0}{4\nu_0}\right]$	$\frac{13\nu_0+8\tau_0}{\tau_0\left[32\nu_0+6\tau_0\right]}\cdot K_p$	$\frac{11\nu_0+2\tau_0}{4\tau_0\nu_0}\cdot K_p$

Table D.3: Optimal PID gains using the reaction curve and Cohen-Coon optimisation method [57]. The parameter K_0, τ_0 and ν_0 are defined in equation D.1 and illustrated in figure D.1.

optimisation methods require that the open loop response is stable with a proportional control only. The oscillation method is relatively fast and easy to perform and is essentially limited to systems where the driven instability does not have a undesired impact to the machine. If instability of the process is unwanted, the reaction curve is the preferred choice. The Cohen-Coon method seems generally to produce solutions with less overshoot compared to the Ziegler-Nichols reactive curve method.

Appendix E

LHC Technical Network Optical Fibre Lengths

The optical fibre lengths of the Technical Network are given in table E.1

Tunnel									
	SR1	SR2	SR3	SR4	SR5	SR6	SR7	SR8	CCC
SR1		4.2	9.2	14.2	17.2	13	8.6	4.2	3.8
SR2	4.2		5	10	14.2	17.2	12.8	12.8	8
SR3	9.2	5		5	9.2	13.4	17.8	13.4	13
SR4	14.2	10	5		4.2	8.4	12.8	17.2	18
SR5	17.2	14.2	9.2	4.2		4.2	8.6	13	21
SR6	13	18.4	13.4	8.4	4.2		4.4	8.8	16.8
SR7	8.6	13.2	17.8	12.8	8.6	4.4		4.4	12.4
SR8	4.2	8.6	13.4	17.2	13	8.8	4.4		8
CCC	3.8	8	13	18	21	16.8	12.4	8	
Surface									
	SR1	SR2	SR3	SR4	SR5	SR6	SR7	SR8	CCC
SR1		9	8	13.1	13.8	11.3	10.8	3.8	3.8
SR2	9		3.1	14.5	15.2	12.7	12.2	12.2	5.2
SR3	8	3.1		13.5	14.2	11.7	11.2	11.2	4.2
SR4	13.1	14.5	13.5		19.3	16.3	16.3	16.3	9.3
SR5	13.8	15.2	14.2	19.3		4.7	7.3	7.3	10
SR6	11.3	12.7	11.7	16.8	4.7		4.8	12.5	7.5
SR7	10.8	12.2	11.2	16.3	7.3	4.8		12	7
SR8	3.8	10.2	9.2	14.3	15	12.5	12		5
CCC	3.8	5.2	4.2	9.3	10	7.5	7	5	

Table E.1: LHC technical network optical fibre length in km. Courtesy D. Francart

Appendix F

Amplification and Statistics of Un-correlated Ground Motion

F.1 Analytic Optical Amplification Approximation

Let ϵ be the r.m.s. quadrupole displacement error. The orbit position shift Δx_i due to a displacement ϵ of the j-th quadrupole with strength k_j and length l_j at the i-th observation point is described by:

$$\Delta x_i = \frac{\sqrt{\beta_i \beta_j} k_j l_j}{2\sin(\pi Q)} \cdot \cos(\Delta\mu_{ij} - \pi Q) \cdot \epsilon \tag{F.1}$$

Here $\Delta\mu_{ij}$ is the positive phase advance between the quadrupole and the point of observation. β_j and β_i are the values of the betatron function at the location of the quadrupole and the point of observation. Q is the machine tune. The effects of N_q quadrupoles at monitor i add up linearly. If the misalignments are uncorrelated, the variance of the shift is given by,

$$(\Delta x_i)^2 = \sum_j^{N_q} \frac{\beta_i \beta_j (k_j l_j)^2}{(2\sin(\pi Q))^2} \cos^2(|\Delta\mu_{ij}| - \pi Q) \cdot \epsilon^2 \tag{F.2}$$

The orbit r.m.s. $(\Delta x)^2$ over N_m monitors (observation points):

$$(\Delta x)^2 = \frac{1}{N_m} \sum_{ij} \frac{\beta_i \beta_j (k_j l_j)^2}{(2\sin(\pi Q))^2} \cos^2(|\Delta\mu_{ij}| - \pi Q) \cdot \epsilon^2 \tag{F.3}$$

and considering the frequency dependence $S(f) = \frac{d\epsilon^2}{df}$:

$$\frac{d\,(\Delta x)^2}{df} = \underbrace{\frac{1}{N_m} \sum_{ij} \frac{\beta_i \beta_j (k_j l_j)^2}{(2\sin(\pi Q))^2} \cos^2(\Delta\mu_{ij} - \pi Q)}_{\kappa^2(f)} \cdot \frac{d\epsilon^2}{df} \tag{F.4}$$

$$S_{orbit} = \kappa^2(f) \cdot S_{gm}(f); \tag{F.5}$$

it is possible to perform a rough estimate of κ using the following approximations:

- All quadrupoles have about the same integrated field strengths: $k_j l_j \approx kl = \text{const.}$
- The machine is composed of a regular FODO lattice: focusing quadrupoles have $\beta_{QF} = \text{const.}$ and defocusing quadrupoles have $\beta_{QD} = \text{const.}$ and $\beta_{eff} = \sqrt{\beta_{QF}^2 + \beta_{QD}^2}$.
- Monitors are close to quadrupoles: $\beta_i \approx \beta_{eff}$

leading to the following simplified equation:

$$\kappa^2 = \frac{\beta_{eff}^2 (kl)^2}{(2\sin(\pi Q))^2} \cdot \frac{1}{N_m} \sum_{ij} \cos^2(\Delta\mu_{ij} - \pi Q) \tag{F.6}$$

For large machines, the phase advance between monitors and quadrupoles modulo 2π is evenly distributed so that the sum of the $\cos^2$-term over i is approximately $\frac{1}{2}$. Summing over the number N of FODO cells and taking the square root yields:

$$\boxed{\kappa = \frac{kl\beta_{eff}}{4\sin(\pi Q)} \cdot \sqrt{N}} \tag{F.7}$$

F.2 Statistical Properties of a Gaussian Process

The movement of a particle with mass m due to a randomly varying force $F(t)$ in a viscous (α being the damping coefficient) medium can be described through STOKE's equation. $F(t)$ shall be symmetric around '0' and is, for instance, GAUSSIAN distributed. For simplicity, without constraining the universality, we assume that the movement is uni-dimensional along the x axis. The differential equation describing the Gaussian process can be written as:

$$m\ddot{x} + \alpha\dot{x} = F(t) \tag{F.8}$$

Re-arranging Equation F.8 and multiplying with x yields,

$$m\ddot{x} = -\alpha\dot{x} + F(t) \tag{F.9}$$

$$mx\ddot{x} = -\alpha x\dot{x} + xF(t) \tag{F.10}$$

$$\texttt{with} \quad x\ddot{x} = \frac{d}{dt}(x\dot{x}) - \dot{x}^2 \tag{F.11}$$

$$m \cdot \left[\frac{d}{dt}(x\dot{x}) - \dot{x}^2\right] = -\alpha x\dot{x} + xF(t) \tag{F.12}$$

$F(t)$ being a non-regular randomly varying force, it is clear that there cannot be an analytical solution for equation F.12. However, it is still possible to derive some statistical properties, such as the mean and root-mean-squared, of the process. Taking the time average $\langle\rangle_t$ leads to:

$$\left\langle m \cdot \left[\frac{d}{dt}(x\dot{x}) - \dot{x}^2\right]\right\rangle_t = \langle -\alpha x\dot{x} + xF(t)\rangle_t \tag{F.13}$$

$$m\frac{d}{dt}\langle x\dot{x}\rangle_t - m\langle\dot{x}^2\rangle_t = -\alpha\langle x\dot{x}\rangle_t + \langle xF(t)\rangle_t \tag{F.14}$$

The force being randomly distributed, which is equivalent to X and $F(t)$ being fully uncorrelated, the following terms vanish.

$$\langle xF(t)\rangle_t = \langle x\rangle_t \langle F(t)\rangle_t = 0 \tag{F.15}$$

One can identify the term describing the average kinetic energy that in kinetic gas-theory is related to the temperature T of the particle through the BOLTZMANN's constant k

$$\frac{1}{2}m\langle\dot{x}^2\rangle_t = \frac{1}{2}kT \tag{F.16}$$

resulting in the following equations

$$m\frac{d}{dt}\langle x\dot{x}\rangle_t = kT - \alpha\langle x\dot{x}\rangle_t \tag{F.17}$$

that can be rearranged to:

$$\left(\frac{d}{dt} + \frac{\alpha}{m}\right)\langle x\dot{x}\rangle_t = \frac{kT}{\alpha} \tag{F.18}$$

The following ansatz can be used to solve the differential equation F.18:

$$\langle x\dot{x}\rangle_t = C \cdot e^{-\frac{\alpha}{m}t} + \frac{kT}{\alpha} \tag{F.19}$$

with C being a constant that has to be adjusted in order to satisfy the initial conditions

$$\langle x\dot{x}\rangle_{t=0} := 0 \leftrightarrow 0 = C + \frac{kT}{\alpha} \tag{F.20}$$

using the relation

$$\langle x\dot{x}\rangle_t = \frac{1}{2}\frac{d}{dt}\left\langle x^2\right\rangle_t \tag{F.21}$$

that leads to:

$$\frac{1}{2}\frac{d}{dt}\left\langle x^2\right\rangle_t = \frac{kT}{\alpha}\left(1 - e^{-\frac{\alpha}{m}t}\right) \tag{F.22}$$

After integration, one gets the following important relation for $\left\langle x^2\right\rangle_t$:

$$\left\langle x^2\right\rangle_t = \frac{2kT}{\alpha}\left[t - \frac{m}{\alpha}\left(1 - e^{-\frac{\alpha}{m}t}\right)\right] \tag{F.23}$$

One can easily see that after applying the time average on Equation F.8 directly, the mean $\mu = \langle x\rangle_t$ exponentially vanishes for timescales:

$$\boxed{\lim_{t\to\infty}\mu \sim \lim_{t\to\infty} e^{-\frac{\alpha}{m}t} = 0} \tag{F.24}$$

The root-mean-squared of the process is defined as:

$$\sigma^2_{rms} = \left\langle x^2 - \mu^2\right\rangle_t = \left\langle x^2\right\rangle_t - \langle x\rangle_t^2 \tag{F.25}$$

since the mean μ vanishes for Gaussian processes:

$$\sigma_{rms} = \sqrt{\langle x^2\rangle_t} \tag{F.26}$$

Expanding the exponential dependence of Equation F.23, one can derive the following two solutions for the r.m.s. depending on the parameter $\frac{m}{\alpha}$:

1. for $t \gg \frac{m}{\alpha}$ (small timescales)

$$\sigma_{rms} = \sqrt{\frac{kT}{m}}\cdot t \tag{F.27}$$

2. for $t \ll \frac{m}{\alpha}$ (large timescales)

$$\boxed{\sigma_{rms} = \sqrt{\tfrac{2kT}{\alpha}}\cdot\sqrt{t}} \tag{F.28}$$

F.3 Correlation Length in the LEP Tunnel as a Function of Frequency - REPRINT

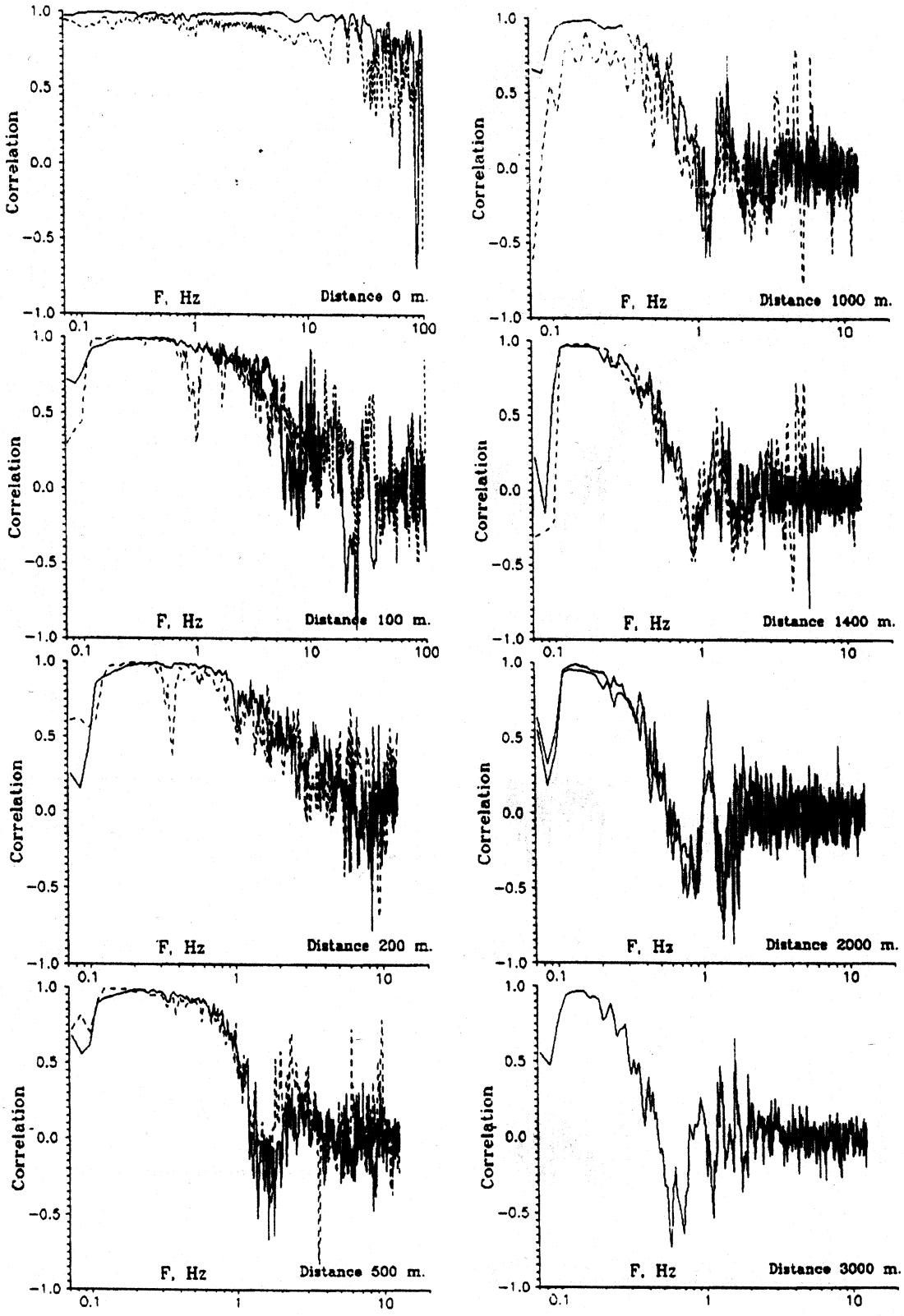


Figure F.1: Correlation between two horizontal probes oriented transverse to the LEP tunnel. Probes are 0 m, 100 m, 200 m, 500 m, 1000 m, 1400 m, 2000 m and 3000 m apart. Reprint of reference [65]. It is visible that the correlation decreases with the probe distance.

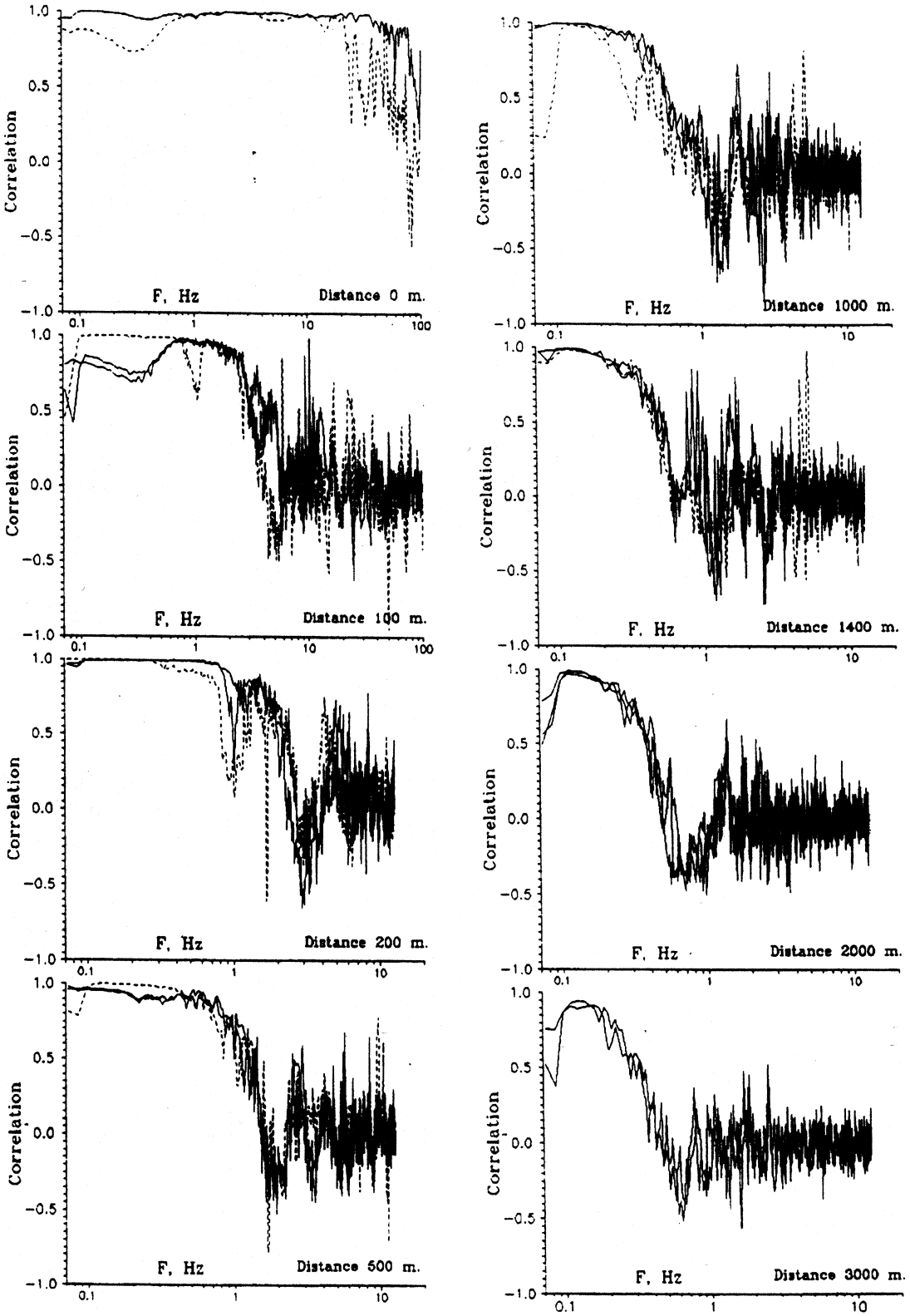


Figure F.2: Correlation between two vertical probes oriented transverse to the LEP tunnel. Probes are 0 m, 100 m, 200 m, 500 m, 1000 m, 1400 m, 2000 m and 3000 m apart. Reprint of [65]. It is visible that the correlation decreases with the probe distance.

Bibliography

[1] Olivier Sim Brüning et al., editor. *LHC Design Report*, volume 1. CERN - European Organization for Nuclear Research, 2004.

[2] J.R. Cudell et al. Benchmarks for the Forward Observables at RHIC, the Tevatron-Run II, and the LHC. *Phys. Rev. Lett.*, 89(20), 2002. 201801-1.

[3] CERN. LHC - The Large Hadron Collider. Technical report, CERN - European Organization for Nuclear Research, 2006. *http://lhc.web.cern.ch/lhc* and references therein.

[4] S. Eidelman et al. Review of Particle Physics. *Physics Letters B*, 592:1+, 2004.

[5] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13(16):508, 509, October 1964.

[6] F. Englert and R. Brout. Broken Symmetry and the Mass of Vector Mesons. *Phys. Rev. Lett.*, 13(9):321 – 323, August 1964.

[7] M. Pieri. Searches for Higgs Bosons at LHC. In *HCP 2005*, 2005. CMS CR 2005/31.

[8] LEP Higgs Working Group. Search for the Standard Model Higgs Boson at LEP. *Phys. Lett. B*, 565:61–75, 2003. CERN-EP/2003-011.

[9] S. Abdullin et al. Summary of the CMS Potential for the Higgs Boson Discovery. Technical report, CERN - European Organization for Nuclear Research, 2003. CMS-Note-2003-033.

[10] ATLAS Collaboration. ATLAS Technical Proposal. Technical report, CERN - European Organization for Nuclear Research, 1994. CERN/LHCC 94-43.

[11] CMS Collaboration. CMS Technical Proposal. Technical report, CERN - European Organization for Nuclear Research, 1994. CERN/LHCC 94-38.

[12] TOTEM Collaboration. Total cross section, elastic scattering and diffraction dissociation at the lhc - technical proposal. Technical report, CERN - European Organization for Nuclear Research, 1999. CERN/LHCC 99-7.

[13] ALICE Collaboration. ALICE Technical Proposal. Technical report, CERN - European Organization for Nuclear Research, 1995. CERN/LHCC 95-71.

[14] LHCb Collaboration. LHCb Technical Proposal. Technical report, CERN - European Organization for Nuclear Research, 1998. CERN/LHCC 98-4.

[15] Roberto Saban. Quality Assurance Definition - Equipment Naming Conventions. Technical report, CERN - European Organization for Nuclear Research, 2005. LHC-PM-QA-204 rev 3.1 and EDMS number 103369.

[16] V. Kain et al. The Expected Performance of the LHC Injection Protection System. In *Proceedings of EPAC'04*. CERN - European Organization for Nuclear Research, 2004. LHC Project Report 746.

[17] Klaus Wille. *The Physics of Particle Accelerators - an introduction*. Oxford, 2000.

[18] Florian Sonnemann. *Resistive Transition and Protection of LHC Superconducting Cables and Magnets*. PhD thesis, RWTH-Aachen University, 2001. CERN-THESIS-2001-004.

[19] P. Estop et al. Magnetization and critical current density of ultra-fine multifilamentary superconducting wires. *IEEE Transactions on Magnetics*, 27(2):2170–2173, 1991.

[20] Markus Haverkamp. *Decay and Snapback in Superconducting Accelerator Magnets*. PhD thesis, Twente University, 2003. CERN-THESIS-2003-030.

[21] R. Wolf. The Decay of the Field Integral in Superconducting Accelerator Magnets wound with Rutherford Cables. Technical report, CERN - European Organization for Nuclear Research, 1998. LHC Project Report 175.

[22] N. Sammut L. Bottura, T. Pieloni. Scaling Laws for the Field Quality at Injection in the LHC Dipoles. *IEEE Trans. Appl. Supercond.*, 2(2):1217–20, 2005. LHC Project Note 361.

[23] L. Bottura. The Use of Magnetic Measurements for LHC Operation. In *Proceedings of Chamonix XIV*, pages 235–243. CERN - European Organization for Nuclear Research, 2005.

[24] J. Wenninger. What systems request a beam dump? In *Proceedings of Chamonix XV*, pages 85–87. CERN - European Organization for Nuclear Research, 2006.

[25] V. Kain. Damage Levels: Comparison of Experiments and Simulation. Technical report, CERN - European Organization for Nuclear Research, 2005. CERN-AB-2005-014.

[26] M. Zerlauth et al. Detecting Failures in Electrical Circuits leading to very fast Beam Losses in the LHC. In *Proceedings of EPAC'04*. CERN - European Organization for Nuclear Research, 2004. LHC Project Report 749.

[27] F. Bodry et al. Machine Protection for the LHC: Architecture of the BEAM and Powering Interlock Systems. Technical report, CERN - European Organization for Nuclear Research, 2001. LHC Project Report 521.

[28] R. Schmidt et al. Machine Protection Systems(s) - Overview. In *Proceedings of Chamonix XIV*, page 255 ff. CERN - European Organization for Nuclear Research, 2005.

[29] R. Assmann. Status of the LHC Collimation Project. In *LHC Performance Workshop - Chamonix XIII*. CERN - European Organization for Nuclear Research, 2004.

[30] M. Tigner A.W. Chao, editor. *Handbook of Accelerator Physics and Engineering*. World Scientific, 2002.

[31] John David Jackson. *Classic Electrodynamics - second edition*. John Wiley and Sons, 1975.

[32] M.G. Minty and F. Zimmermann. *Measurement and Control of Charged Particle Beams*. Springer, 2003.

[33] G. Guignard. The General Theory of all Sum and Difference Resonances in a Three-Dimensional Magnetic Field in a Synchrotron. Technical report, CERN - European Organization for Nuclear Research, 1976. CERN 76-06 ISR.

[34] Donald W. Kerst. The acceleration of electrons by magnetic induction. *Phys. Rev.*, 60:47–53, 1941.

[35] K. Takayama et al. First Observation of the Acceleration of a single bunch by using the induction device in the KEK Proton Synchrotron. KEK - High Energy Accelerator Research Organization, Japan, 2004. arXiv:physics/0412006 v1.

[36] K. Torikai et al. Induction Acceleration System for a Proton Bunch Acceleration in the KEK Proton Synchrotron. In *Proceedings of PAC05, Knoxville/TN,USA*, 2005.

[37] O. Brüning et al., editor. *LHC Design Report*, volume 1, chapter RF Systems & Beam Feedback, page 548 ff. CERN - European Organization for Nuclear Research, 2004.

[38] George William Hill. On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon. *Acta Mathematica*, 8:1–36, 1886.

[39] W. Magnus and S. Winkler. *Hill's Equation*. Dover Pubns, 1979.

[40] Donald W. Kerst. Electronic orbits in the induction accelerator. *Phys. Rev.*, 60:53–58, 1941.

[41] H. Grote and F. Iselin. The MAD (Methodical Accelerator Design) Program. Technical report, CERN - European Organization for Nuclear Research, 1996. CERN/SL/90-13 Rev. 5.

[42] W. Herr and F. Schmidt. *A MADX Primer*. CERN - European Organization for Nuclear Research, 2004. CERN-AB-2004-027-ABP.

[43] Ch. Iselin F. Rothacker Karl L. Brown, D.C. Carey. TRANSPORT - A Computer Program for Designing Charged Particle Beam Transport Systems. Technical report, CERN - European Organization for Nuclear Research, 1980. CERN 73-16.

[44] Richard Q. Twiss and N. H. Frank. Orbital stability in a proton synchrotron. *Rev. Sci. Instr.*, 20(1):1–17, January 1949.

[45] CERN. Design Performance. In *The Large Hadron Collider Home Page*. CERN - European Organization for Nuclear Research, 2003. *http://cern.ch/lhc-new-homepage/*.

[46] J. Wenninger. Orbit corrector magnets and beam energy. Technical report, CERN - European Organization for Nuclear Research, 1997. SL-Note 97-06 OP.

[47] A. Hoffmann. Transporting the energy calibration across optics. In *Proceedings of Chamonix III*, pages 303–306. CERN - European Organization for Nuclear Research, 1993.

[48] O. Brüning S. Fartoukh. Field Quality Specification for the LHC Main Dipole Magnets. Technical report, CERN - European Organization for Nuclear Research, 2001. LHC Project Report 501.

[49] J.P. Koutchouk S. Fartoukh. On the Measurement of the Tunes, [..] in LHC. Technical report, CERN - European Organization for Nuclear Research, 2001. LHC-B-ES-0009 et EDMS # 463763.

[50] Marek Gasior and Rhodri Jones. The principle and first results of betatron tune measurement by direct diode detection. Technical report, CERN - European Organization for Nuclear Research, 2005. LHC Project Report 853.

[51] Y. Luo R. Jones, P. Cameron. Torwards a robust phase locked loop tune feedback system. Technical report, Brookhaven National Laboratory, 2005. C-A/AP/#204.

[52] M. Salgado G. Goodwin, S. Graebe. *Control System Design*. Prentice Hall, 2000.

[53] Otto J.M. Smith. *Feedback Control Systems*. McGraw-Hill, 1958.

[54] A. Emami-Naeimi G. Franklin, J. Powell. *Feedback Control of Dynamic Systems*. Adison Wesley, 1995.

[55] M. Workman G. Franklin, J. Powell. *Digital Control of Dynamic System*. Adison Wesley, 1997.

[56] G. Ziegler and N. B. Nichols. Optimum settings for automatic controllers. *Trans. A.S.M.E.*, 64:759–765, 1942.

[57] Coon Cohen. Theoretical consideration of retarded control. *Trans. A.S.M.E.*, 75:827–834, 1953.

[58] D. C. Youla et al. Modern Wiener-Hopf Design of Optimal Controllers. *IEEE Transactions on Automatic Control*, 21(1):3–13 & 319–338, 1976.

[59] C. Deoer et al. Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis. *IEEE Transactions on Automatic Control*, 25(3):399–412, 1980.

[60] R. J. Steinhagen. Analysis of Ground Motion at SPS and LEP, Implications for the LHC. Technical report, CERN - European Organization for Nuclear Research, 2005. CERN-AB-2005-087.

[61] R. Pitthan. LEP Vertical Tunnel Movements - Lessons for Future Colliders. Technical report, CERN - European Organization for Nuclear Research, 1999. CERN CLIC-Note 422, LCC-0028 and SLAC-PUB-8286.

[62] M Güler et al. An appearance experiment to search for $\nu_\mu \leftrightarrow \nu_\tau$ oscillations in the CNGS beam : experimental proposal. Technical report, CERN - European Organization for Nuclear Research, 2000. CERN-SPSC-2000-028; LNGS-2000-25; SPSC-P-318.

[63] A. Sommaruga F. Marillier, U. Eichenberger. Seismic Synthesis of the Swiss Molasse Basin. Technical report, Institute de Géophysique - UNI Lausanne, Switzerland, 2004.

[64] G. Amberger. *La Molasse du Bassin Genevois in "Cahiers de la Faculté des Science"*. UNIGE - Geneva, Switzerland, 1987.

[65] V.M. Juravlev et al. Investigations of Power and Spatial Correlation Characteristics of Seismic Vibrations in the CERN LEP Tunnel for Linear Collider Studies. Technical report, CERN - European Organization for Nuclear Research, 1993. CERN-SL/93-53 and CLIC-Note 217.

[66] Eberhard Keil. Effect of Plane Ground Waves on the Closed Orbit in Circular Machines. Technical report, CERN - European Organization for Nuclear Research, 1997. CERN SL/97-61 (AP).

[67] Matteo Frigo and Steven G. Johnson. FFTW - The Fastest Fourier Transform in the West. Massachusetts Institute of Technology, 1999. *http://www.fftw.org/*.

[68] Daniel Cocq. The Wide Band Normaliser - A New Circuit to Measure Transverse Bunch Position in Accelerators and Colliders. *NIMA 416, Elsevier*, 1998.

[69] Masakazu Yashioka-Yasunori Takeuchi Shigeru Takeda, Hiroshi Matsumoto. Incoherent Ground Motion. In *Proceedings of EPAC 2000*, page 2394 ff., Vienna, Austria, 2000.

[70] V. Shiltsev. Space-Time Ground Diffusion: The ATL Law for Accelerators. Technical report, DESY - Deutsches Elektronenen Synchrotron Hamburg, 1995. DESY-MEA-95-02.

[71] A. Wolski and N.J. Walker. A Model of ATL Ground Motion For Storage Rings. In *Proceedings of the PAC Conference, Portland,USA*. DESY and LBNL (Berkely), 2003.

[72] F. Tecker. Closed Orbit Feedback from Low-Beta Quadrupole Movements at LEP. Technical report, CERN - European Organization for Nuclear Research, 1997. CERN SL-97-026.

[73] J. Wenninger et al. Beam Stability of the LHC Beam Transfer Line TI 8. In *Proceedings of PAC05, Knoxville/TN,USA*, 2005. LHC Project Report 821.

[74] L. Arnaudon et al. Effects of Terrestrial Tides on the LEP Beam Energy. Technical report, CERN - European Organization for Nuclear Research, 1994. CERN SL/94-07 (BI).

[75] J. Wenninger. Observations of Radial Ring Deformations using Closed Orbits at LEP. In *Proceedings of PAC'99*, New York City, NY, 1999. CERN - European Organization for Nuclear Research. CERN-SL-99-025-OP.

[76] Frank D. Stacey. *Physics of the Earth*. Wiley, New York, 1977.

[77] M. Böge. Achieving Sub-Micron Stability in Light Sources. In *Proceedings of EPAC 2004*, Lucerne, Switzerland, 2004.

[78] O. Brüning et al., editor. *LHC Design Report*, volume 1, chapter Cooling and Ventilation, page 548. CERN - European Organization for Nuclear Research, 2004.

[79] J. Dwivedi et al. The Alignment Jacks of the LHC Cryomagnets. Technical report, CERN - European Organization for Nuclear Research, 2004. LHC Project Report 725.

[80] G. Decker. Orbit Stabilization at the Advanced Photon Source. DOE-BES Review, Argonne National Laboratory, May 2005.

[81] J.C. Denard et al. Overview of the Diagnostics Systems of SOLEIL and Diamond. In *Proceedings of DIPAC 2002*, Mainz, Germany, 2002.

[82] S. Zelenika et al. The SLS Storage Ring Support and Alignment Systems. *NIM A 467-467, Elsevier*, 2001.

[83] J. Inigo-Golfin. Expected Air Temperatures in the LHC Tunnel. private communications, May 2006.

[84] T. Schilcher et al. Orbit Control at the SLS Storage Ring. In *Proceedings of the 26th Advanced ICFA Beam Dynamics Workshop*, 2002.

[85] Masaru Takao and Taijai Shimada. Long-Term Variation of the Circumference of the Spring-8 Storage Ring. In *Proceedings of EPAC'00*, 2000.

[86] R. Pitthan. Why Does the LEP Tunnel Move the Way it Moves. In *Proceedings of the 22nd Advanced ICFA Beam Dynamics Workshop*, 2002.

[87] L. Bottura. Cold Test Results: Field Aspects. In John Poole, editor, *Proceedings of Chamonix XII*. CERN - European Organization for Nuclear Research, 2003.

[88] L. Bottura. Superconducting Magnets on Day I. In John Poole, editor, *Proceedings of Chamonix XI*. CERN - European Organization for Nuclear Research, 2002.

[89] E. Tedesco et al. An Estimate of Multipolar Errors in the LHC Dipoles. Technical report, CERN - European Organization for Nuclear Research, 2001. LHC Project Report 625.

[90] R. Wolf L. Bottura, P. Burla. LHC Main Dipoles Proposed Baseline Current Ramping. Technical report, CERN - European Organization for Nuclear Research, 1998. CERN-LHC-Project-Report-172.

[91] J. Micallef N. Sammut, L. Bottura. A Mathematical Formulation to Predict the Harmonics of the Superconducting LHC Magnets. Technical report, CERN - European Organization for Nuclear Research, 2005. LHC Project Report 854.

[92] T. Pieloni et al. Field decay and snapback measurements using a fast hall probe sensor. Technical report, CERN - European Organization for Nuclear Research, 2004. LHC Project Report 716.

[93] C. Montag. Observation of Mechanical Triplet Vibrations in RHIC. In *Proceedings of the 22nd Advanced ICFA Beam Dynamics Workshop*, 2002.

[94] W. Venturini. Magnetic Behaviour of LHC correctors: Issues for Machine Operation. In *Proceedings of Chamonix XIV*. CERN - European Organization for Nuclear Research, 2005.

[95] S. Sanfilippo J. P. Koutchouk. Magnetic Issues affecting Beam Commissioning, Session Summary. In *Proceedings of Chamonix XIV*. CERN - European Organization for Nuclear Research, 2005.

[96] F. Bodry. LHC Power Converters - Performance Requirements. In John Poole, editor, *Proceedings of Chamonix XI*. CERN - European Organization for Nuclear Research, 2001.

[97] W. Venturini et al. Hysteresis measurement of a twin aperture MCB orbit corrector. AT-MTM internal Memo, October 2005.

[98] V. Montabonnet A. Cantone. 60A Power Converter Testing in SM18. private communications, 2005.

[99] S. Redaelli. LHC aperture and commissioning of the Collimation System. In *Proceedings of Chamonix XIV*. CERN - European Organization for Nuclear Research, 2005.

[100] R. Assann et al. Collimators and Cleaning, Could this Limit the LHC Performance ? In *Proceedings of Chamonix XII*, 2003.

[101] R. Assmann. Collimation Status Report. Technical report, CERN - European Organization for Nuclear Research, December 2004. LHC Machine Advisory Committee.

[102] J. Wenninger. Orbit feedback for collimation. Technical report, CERN - European Organization for Nuclear Research, July 2005. LHC Machine Advisory Committee.

[103] R. Assmann. Orbit Stability during Stage I LHC Operation. private communication.

[104] R. Assmann. 450GeV commissioning: Increasing the beam intensity - Collimation. Technical report, CERN - European Organization for Nuclear Research, May 2006. LHC-CWG#5.

[105] W. Höfle. Towards a Transverse Feedback System and Damper for the SPS in the LHC Era. Technical report, CERN - European Organization for Nuclear Research. CERN-SL-97-017-RF.

[106] D. Boussard et al. The LHC Transverse Damper (ADT) Performance Specification. Technical report, CERN - European Organization for Nuclear Research. SL-Note-99-055-HRF.

[107] R. Assmann et al. On the Measurement of the Relative Luminosity at the LHC. Technical report, CERN - European Organization for Nuclear Research. EDMS: LHC-B-ES-0007 rev 1.1.

[108] C. Bovet et al. LHC BPM design. Technical report, CERN - European Organization for Nuclear Research, 1997. CERN-SL-97-69 (BI).

[109] O. Brüning et al. Electron Cloud and Beam Scrubbing in the LHC. In *Proceedings of PAC'99*, New York, NY, 1999.

[110] P. Cameron Y. Yuo. Possible phase loop for the global betatron decoupling. Technical report, Brookhaven National Laboratory, 2004. CAD/AP/174.

[111] ATLAS Collaboration. ATLAS Forward Detectors for Luminosity Measurement and Monitoring. Technical report, CERN - European Organization for Nuclear Research, 2004. CERN/LHCC/2005-010.

[112] E. Chapochnikova et al. RF Requirements and constraints for Stage I commissioning. private communication, 2005.

[113] G. Decker. Beam Stability in Synchrotron Light Sources. Technical report, Argonne National Laboratory, Lyon, France, 2005.

[114] R. Littauer. Beam instrumentation. In *Physics of High Energy Particle Accelerator, SLAC Summer School*. Stanford Linear Accelerator Center, 1982. p.902.

[115] G. Vismara. Signal Processing for Beam Position Monitors. Technical report, CERN - European Organization for Nuclear Research, 2000. CERN-SL-200-056 (BI).

[116] D. P. McGinnis. The Design of Beam Pickup and Kickers. In *Proceedings of the 6th Beam Instrumentation Workshop*, Vancouver, Canada, 1994.

[117] R. T. Avery et al. Nonintercepting Emittance Monitor. *IEEE Trans. Nucl. Sci.*, NS(18):920, 1971.

[118] J.-P. Koutchouk. Measurement of the Beam Position in the LHC Main Rings. Technical report, CERN - European Organization for Nuclear Research, 2002. LHC-BPM-ES-0004 rev 2.0, edms #327557.

[119] E. Calvo et al. The LHC Orbit and Trajectory System. In *Proceedings of DIPAC'03*, Mainz, Germany. CERN - European Organization for Nuclear Research.

[120] R. Assmann et al. Non-Linearities in the Response of Beam Position Monitors. In *AIP Conference Proceedings 546*, pages 267–274, Boston, MA, 2000. CERN-SL-2000-055-BI.

[121] R. J. Steinhagen et al. LHC Orbit Stabilisation Tests at the SPS. In *Proceedings of PAC'05*, Knoxville/TN, 2005. CERN-AB-2005-052.

[122] Report on the mini-workshop on the LHC alignment. Technical report, CERN - European Organization for Nuclear Research, 2001. LHC Project Report 625.

[123] J. Matheson et al. Beam Based Alignment at LEP. Technical report, CERN - European Organization for Nuclear Research, 2003. CERN-AB-2003-087 BDI.

[124] I. Barnett et al. Dynamic Beam-based Alignment. Technical report, CERN - European Organization for Nuclear Research, 2003. CERN SL/94-84 (BI).

[125] I. Reichel. Beam position measurement by modulation of quadrupole strengths, 1995. Pitha Nr.94/54, also CERN SL/95-50.

[126] J. Wenninger. Orbit Response Measurements at the SPS . Technical report, CERN - European Organization for Nuclear Research, 2004. CERN-AB-2005-009.

[127] G. Decker. Detection of Spiking BPM readings in the Advanced Photon Surce. private communication during IWBS'04, 2004.

[128] Hewlett-Packard. *http://www.hp.com/* and references therein.

[129] Wiener Plein & Baus GmbH, Burscheid, Germany. *Users Manual: Series 6000 VME, VME64x, VME64xP VXI*, 2006.

[130] F. Bordry and H. Thiesen. RST Digital Algorithm for Controlling the LHC Magnet Current. Technical report, CERN - European Organization for Nuclear Research, 1998. LHC Project Report 625.

[131] D. Nisbet V. Montabonnet. Reliability and Failures of the 60A COD Power Converter. private communications, 2005.

[132] S. Pittet L. Ceccone. Augumentation du Temps de Decharge des Convertisseurs LHC [60A,8V]. Modification du Bloc Dechargeur. LHC Engineering Change Request, EDMS # 458013, 2005.

[133] R. Gupta. Introduction to Lattice QCD. Technical report, Los Alamos National Laboratory, 1998. arXiv:hep-lat/9807028.

[134] T. DeGrand. Lattice QCD at the end of 2003. Technical report, University of Colorado, 2003. arXiv:hep-ph/0312241.

[135] M. Koma. Optimization of Lattice QCD codes for the AMD Opteron processor. Technical report, DESY - Deutsches Elektronenen Synchrotron Hamburg, 2005. arXiv:hep-lat/0510029.

[136] Stefan D. Bruda and Selim G. Akl. Real-Time Computation: A Formal Definition and its Applications. *International Journal of Computers and Applications*, 25(2), 2003.

[137] Enterasys Technology. Specification of the enterasys x-pedition 8600 router. *http://www.enterasys.com/* and references therein.

[138] O. Brüning et al., editor. *LHC Design Report*, volume 2, chapter Communication Networks, page 103 ff. CERN - European Organization for Nuclear Research, 2004.

[139] J. Postel. User Datagram Protocol. Technical report, 1980. also RFC768 standard.

[140] V. Cerfand R. Kahn. A Protocol for Packet Network Intercommunication. *IEEE Transactions on Communications*, COM-22(5):pp 637–648, May 1974. also RFC793 standard.

[141] J. Postel. Internet Control Message Protocol. Technical report, 1981. also RFC792 standard and updated in RFC950.

[142] L. Torvalds et al. The Linux Kernel. *http://www.kernel.org* and references therein.

[143] I. Molnar et al. Real-Time Preemption for the Linux Kernel. *http://people.redhat.com/mingo/realtime-preempt/*. *http://rt.wiki.kernel.org* and references therein.

[144] A. Bland O. Van Der Vossen. Mean-Time between Failure of the LHC Network Infrastructure. private communication, May 2006.

[145] T. Himel et al. Use of Digital Control Theory State Space Formalism for Feedback at SLC. Technical report, Stanford Linear Accelerator Center, San Francisco, CA, 1991. SLAC-PUB-5470.

[146] P. Castro. Betatron function measurements at LEP using the BOM 1000 turns facility. Technical report, CERN - European Organization for Nuclear Research, 1992. SL/Note 92-63-BI.

[147] B. Autin and Y. Marti. Closed Orbit Correction of A.G. Machines using a small Number of Magnets. Technical report, CERN - European Organization for Nuclear Research, 1973.

[148] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, NJ, 1963.

[149] C. Bocchetta. Review of Orbit Control. In *Proceedings of EPAC'98*, Stockholm, Sweden, 1998.

[150] M. Böge et al. Fast Closed Orbit Control in the SLS Storage Ring. In *Proceedings of the 1999 Particle Accelerator Conference, New York,1999*, page 1129 p.p. Paul Scherrer Institute, Switzerland, 1999.

[151] G. Golub and C. Reinsch. *Handbook for Automatic Computation II, Linear Algebra.* Springer, NY, 1971.

[152] K.G.J. Jacobi. Über ein leichtes Verfahren, die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. *Crelle's Journal*, 30 p. 51-94:97–144, 1846.

[153] Jonathan Goodman. *Principles of Scientific Computing*. Courant Institute of Mathematical Sciences, NY University, New York, 2006.

[154] R.J. Steinhagen. Orbit Feedback Control for the LHC - Prototyping at the SPS. Technical report, 2004. CERN-AB-Note-2004-052.

[155] J. Wenninger. Orbit Steering and Central Frequency for LEP2 Energy Calibration. Technical report, CERN - European Organization for Nuclear Research, 2000. SL-Note-2000-009 OP.

[156] A. Petitet R. C. Whaley. Minimizing development and maintenance costs in supporting persistently optimized BLAS. *Software: Practice and Experience*, 35(2):101–121, February 2005.

[157] J. J. Dongarra R. C. Whaley, A. Petitet. Automated empirical optimization of software and the ATLAS project. *Parallel Computing*, 27(1–2):3–35, 2001.

[158] ATLAS homepage. *http://math-atlas.sourceforge.net/*.

[159] R. C. WHALEY. *Automated Empirical Optimization of High Performance Floating Point Kernels*. PhD thesis, Florida State University, FL, 2004.

[160] Technical Committee on Operating Systems and Operational Environments. *Information Technology-Portable Operating System Interface (Api)*. Number ISBN 1-55937-061-0. Inst of Elect and Electronic, October 1990. The POSIX 1003.1 standard.

[161] Manuel Zaera Sanz. *Evaluation of Real-Time Operating Systems using dynamic Scheduling and Management of shared Resources applied to the Monitoring of Industrial Manufacturing Processes*. PhD thesis, Polytechnical University of Valencia, 2003.

[162] D. Dominguez et al. An FPGA based multiprocessing CPU for beam synchronous timing in CERN's SPS and LHC. Technical report, CERN - European Organization for Nuclear Research, 2003. CERN-AB-2003-112 (CO).

[163] P. Alvarez et al. PLL Usage in the General Machine Timing System for the LHC. In *Proceedings of ICALEPS2003*, Gyeonju, Korea, 2003.

[164] M. Böge, editor. *Summary of the 3rd International Workshop on Beam Orbit Stabilization - IWBS2004*. PSI - Swiss Light Source, 2004. SLS-TME-TA-2005-0257.

[165] G. Decker. Beam Stability in Synchrotron Light Sources. In *Proceedings of DIPAC'05*, Lyon, France, 2005.

[166] J. Wenninger. Orbit control for machine operation and protection. Technical report, CERN - European Organization for Nuclear Research, 2003.

[167] J. Poole, editor. *3rd LHC Project Workshop - Chamonix XV*. CERN - European Organization for Nuclear Research, 2006. CERN-AB-2006-014.

Acknowledgements

I thank most sincerly Prof. Dr. Albrecht Böhm for his support and guidance during my doctoral thesis. His famous effectuations – not limited to high energy physics – were most insightful, and influenced not only this work but also my character, for which I am thankful to him.

I extend my gratitude to Prof. Dr. Thomas Hebbeker for his willingness to be the second referee of this thesis and for his kind support on behalf of the 3^{rd} Institute of Physics during my stay at CERN.

My special thanks to Dr. Jörg Wenninger who introduced me to the subject and guided me throughout my studies at CERN. I enjoyed working with him and believe I had the best support a doctoral student can have.

I thank Prof. Dr. Federico Galvez-Durand Besnard for proof-reading this thesis, in particular the many insightful discussions on control theory and implementation details of digital signal processors.

I thank Susan, my beloved wife, for proof-reading this thesis, her patience, interest and willingness to listen and share – in spite of her non-physics "United Nations world" – unusual discussions revolving around galactic rotation curves, quantum-chromo-dynamics, the LHC and many other topics.

While the perturbation analysis and feedback design is based on this work, the 'real-world' implementation required help, relied on preparatory work and insightful comments of the many people in the AB Department who contributed and made this project possible. I specially thank Gianluigi Arduini, Ralph Aßmann, Alastair Bland, Andrea Cantone, Bernd Dehning, Enzo Genuardi, Marek Gasior, Massimo Giovanozzi, Lars Jensen, Rhodri Jones, Verena Kain, Mike Lamont, Quentin King, Valerie Montabonnet, Stefano Redaelli, Rüdiger Schmidt, Walter Venturini Delsolaro and Frank Zimmermann.

I appreciate and thank my section, its leader Dr. Karel Cornelis, the team of SPS operators and many other colleagues for the welcoming environment and helpful discussions that made my stay, time and work at CERN real fun.

Lebenslauf

1980-03-09:	geboren in Bonn als Sohn der Eheleute OTL a.D. Georg Steinhagen und Barbara Steinhagen, geb. Koch
1990-1999:	Thomas-Morus-Gymnasium Daun
1999-2004:	Studium der Physik an der RWTH-Aachen
Feb. 2003:	*Technical Student* am CERN – Europäisches Kernforschungszentrum, Genf, Schweiz
Feb. 2003:	Diplomarbeit am III. Physikalischen Institut A der RWTH Aachen bei Prof. Dr. A. Böhm
2004:	Diplom in Physik
2004-2006:	*Doctoral Student* am CERN – Europäisches Kernforschungszentrum, Genf, Schweiz und III. Physikalischen Institut A der RWTH Aachen bei Prof. Dr. A. Böhm
seit 2006:	*Fellow* am CERN – Europäisches Kernforschungszentrum, Genf, Schweiz