Cavity Field Control for High-Intensity Linear Proton Accelerators

Troeng, Olof

2017

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Troeng, O. (2017). *Cavity Field Control for High-Intensity Linear Proton Accelerators*. Department of Automatic Control, Lund Institute of Technology, Lund University.

Total number of authors:
1

Cavity Field Control for High-Intensity Linear Proton Accelerators

Olof Troeng

LUND UNIVERSITY

Department of Automatic Control

Licenciate Thesis
ISRN LUTFD2/TFRT--3273--SE
ISSN 0280–5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by MediaTryck.
Lund 2017

Abstract

The European Spallation Source will, once fully operational in 2025, be the world's brightest neutron source. The neutrons will be generated by bombarding a tungsten target with protons accelerated to 96% the speed of light by electromagnetic fields confined in 155 radio-frequency cavities along the world's most powerful linear accelerator.

This thesis has been motivated by the strict control specifications on the amplitudes and phases of the accelerating electromagnetic fields. By considering the field control problem from an automatic control perspective, the thesis aims at improving the understanding of the problem and to explain important aspects of the control design.

Throughout the thesis it has been helpful to model the cavity and RF system by complex-coefficient single-input single-output systems. The complex-coefficient representation was particularly useful for discussing: (1) the control design for cavities with parasitic resonance modes; (2) the effect of loop phase variations on feedback stability; (3) the directionality of the disturbances and the objective function.

The thesis presents a non-standard parametrization and derivation of the cavity field dynamics that make it easier to relate the physical cavity process to the model, and simplifies the understanding of how the cavity parameters affect the achievable control performance.

The control performance of simple PI(D)-controllers and general linear time-invariant controller was compared using the Youla parametrization and convex optimization; it was found that PI(D)-controllers in many cases achieved performance similar to the more general linear time-invariant controller; this indicates that simple PI(D)-controllers in many cases are an excellent choice for cavity field control.

Lastly, the energy-optimal strategy to build up the electromagnetic cavity fields is derived, allowing the sustainability of the European Spallation Source to be further improved.

Acknowledgments

Oh Great Supervisor Bo Bernhardsson! Thank you for acting as a beacon of light throughout the first half of my PhD studies! With your thorough understanding of both control theory and control applications, your helpfulness in practical matters, and your optimism, you have managed to keep me motivated, and made my studies worthwhile. Your sharp eye for detail and great talent in finding errors has significantly improved the quality of this thesis. I am looking forward to continue the second half of my PhD studies under your supervision. I am also thankful to my co-supervisors Anders J Johansson and Rolf Johansson. Anders has led the LLRF project at Lund University with great enthusiasm, organized visits to other accelerator facilities, been supportive of my work, and carefully reviewed this thesis. Rolf has dealt with contractual matters for my involvement with ESS and carefully reviewed many of my manuscripts.

Before starting my PhD studies I had no prior experience of particle accelerators or LLRF systems, so the helpfulness and openness of the LLRF experts around the world have been invaluable to me in my work on this licentiate thesis. I would like to thank the following LLRF experts that have invited me to their labs, and shared their experience with me: Mark Crofford and his colleagues at SNS, Sven Pfeiffer and his colleagues at DESY, Philippe Baudrengien at CERN, Thomas Schilcher at PSI, and Chang V together with his colleagues at the Institute of Modern Physics, Chinese Academy of Sciences.

I am also grateful to the following LLRF experts: Claudio Rivetta for interesting discussions on complex-coefficient transfer system and for co-authoring an article with me; and Larry Doolittle for sharing his LLRF know-how with me.

Thank you all members of the RF group at ESS for your helpfulness and for many interesting discussions.

I would like to express my gratitude to Mamad Eshraqi for that he has taken the time to explain and discuss beam physics with me.

I would like to thank Fredrik Kristensen, Anders Svensson and Rihua Zeng for designing and implementing firmware and hardware for the ESS LLRF system with great skill, for many enlightening discussions, and for being good travel companions.

I am thankful to Daniel Sjöberg, Anders Karlsson, Gabriele Constanza and Gerhard Kristensson at the Department of Electrical and Information Technology, Lund University, for many insightful discussions on the dynamics of electromagnetic cavity fields.

I am grateful to all my colleagues at the Department of Automatic Control for the excellent working environment and all the fun that we have had together, it is You that have made my time as a PhD student into an awesome experience. Special mention goes to the administrative staff for making things run smoothly, to Leif Anderson for his expert typographical advise on the typesetting of this thesis, and to my office mates Victor Millnert and Martin Karlsson for many good times.

Last, but not least, I would like to thank my friends, my parents, my sister and Sara. You are what really matters to me.

Financial Support

The author is a member of the LCCC Linnaeus Center, supported by the Swedish Research Council, and the ELLIIT Excellence Center, supported by the Swedish Government. The author's attendance at the American Control Conference 2017 was supported by Ericsson's Research Foundation.

Contents

Abbreviations

Abbreviation	Meaning
ADC	Analog to Digital Converter
BCM	Beam Current Monitor
DAC	Digital to Analog Converter
DTL	Drift-Tube Linac
EPICS	Experimental Physics and Industrial Control System
ESS	European Spallation Source
FPGA	Field Programmable Gate Array
IF	Intermediate Frequency
ILC	Iterative Learning Control
IOC	Input Output Controller (EPICS specific term)
IQ	In-phase and Quadrature (i.e., the real and imaginary parts of a complex number)
LLRF	Low-Level Radio Frequency (System)
LO	Local Oscillator
LTI	Linear Time-Invariant
MIMO	Multiple-Input Multiple-Output
MO	Master Oscillator
NIQ	Near/Non IQ
PI(D)	Proportional-Integral-(Derivative) (Controller)
RF	Radio Frequency
RFQ	Radio-Frequency Quadrupole
RMS	Root-Mean Square
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
TITO	Two-Input Two-Output

Remark on nomenclature

In this thesis, control refers to control as in automatic control, and not as in industrial control system. Stability refers to feedback stability, and not that a quantity is kept constant over time. When we say field error or control error we mean what accelerator physicists call jitter. We will use ”reference” for referring to the phase reference provided by the reference oscillator, and set-point for the reference signal to the controller.

1

Introduction

1.1 The European Spallation Source

The European Spallation Source will, once fully operational in 2025, be the world's brightest neutron source. Its high intensity and long pulses have been said to bring a paradigm shift to neutron imaging that will ultimately lead to scientific break-throughs in pharmacology, chemistry, solid-state physics, engineering sciences, material sciences, archeology and other fields [ESSa, 2017].

Background

The history of the European Spallation Source (ESS) goes back to an OECD study from the early 1990s that recommended that megawatt-class spallation sources be built in Europe, Asia and North America. Early designs for the European Spallation Source were presented in 2002 and 2003. In 2009 it was decided that the ESS would be constructed in Lund, Sweden, in a collaboration between 17 European partner countries. Preparatory design efforts culminated in 2012, with the completion of the ESS Technical Design Report [Peggs et al., 2013], and in 2014 the construction began outside of Lund.

ESS is committed to become a sustainable research facility. This amounts to, among other things, that only renewable energy will be used and that waste heat from the cooling system will be recycled to the Lund district-heating network.

Neutron imaging

The ESS can be likened to a large microscope that uses neutrons instead of light. Neutrons interact strongly with hydrogen atoms, but weakly with large atoms, which allows otherwise invisible features in the studied samples, see Figure 1.1. ESS will not only have the highest neutron flux of all neutron sources in the world, but also relatively long neutron pulses (Figure 1.2), which will allow researchers to get an unparalleled insight into the time-evolution of biological process on a molecular scale.

(a) X-rays.   (b) Neutrons.

Figure 1.1 Neutrons interact weakly with large atoms, but strongly with hydrogen, which allows them to reveal features that are invisible to X-rays. Photo credit: Paul Scherrer Institute.

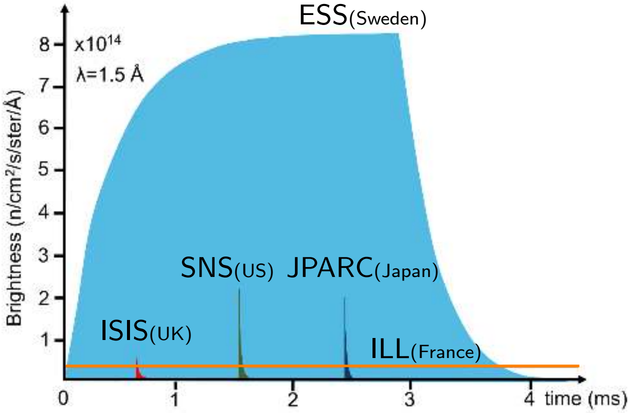


Figure 1.2 The expected brightness of the neutron pulse of ESS, compared to today's leading neutron sources. Source: ESS Activity Report 2015

Overview of ESS

The neutrons at ESS will be obtained through a nuclear reaction called spallation by bombarding a rotating tungsten target with protons, see Figure 1.3. The emitted neutrons are guided to the 16 different state-of-the-art neutron instruments by neutron guides. The protons are accelerated to 96 % the speed of light by the worlds most powerful linear accelerator. The main components of ESS are illustrated in Figure 1.4, for more details see the ESS Technical Design Report [Peggs et al., 2013].

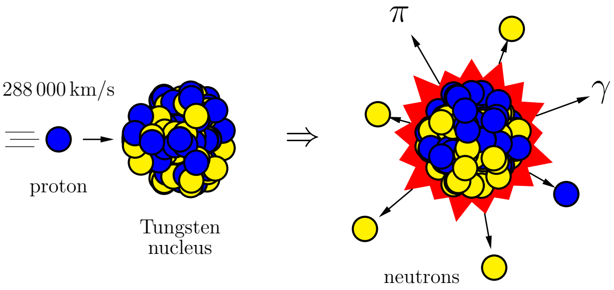


Figure 1.3 The spallation process used to generate free neutrons for the ESS; put simply, high-velocity protons are smashed into tungsten nuclei, which make them emit neutrons; for more details on spallation, see [Russell, 1990].

Particle acceleration

The protons will be accelerated in bunches, with about one billion protons each, by oscillating electromagnetic fields confined in metal structures called RF cavities; in total there will be 155 RF cavities of six different types along the linear accelerator. The acceleration of the bunches and how well they are kept together, depend on the amplitudes of the electromagnetic fields, as well as their phases relative the bunches—if these two quantities are not accurately controlled, the proton bunches get the wrong velocities.

Protons with incorrect velocities are not deflected correctly by the focusing magnets, and will crash into the cavity walls, inducing radioactivation. If the accelerator becomes too radio-activated it needs significant time to cool down, before maintenance can be done, making it impossible to meet the top-level requirement on ESS of 95% availability.

The ESS accelerator will become the most powerful linear accelerator in the world (Figure 1.5), and with its high beam current (62.5 mA), and high duty-factor (4 %), only very small field errors can be tolerated before the radioactivation becomes too severe. The ESS requirements on the allowed field errors are unprecedentedly hard for a high-intensity proton accelerator, and it is the challenges of meeting these requirements that have motivated this thesis.

An analogy: “The equivalent broom model”

An analogous model of the acceleration process, which perhaps is more understandable and familiar to the layman, is illustrated in Figure 1.6. The proton bunches are modeled by dust balls, and the electromagnetic cavity fields are modeled by brooms hinged on low friction bearings. The brooms are free to swing in the direction from the dust source to the dust target,

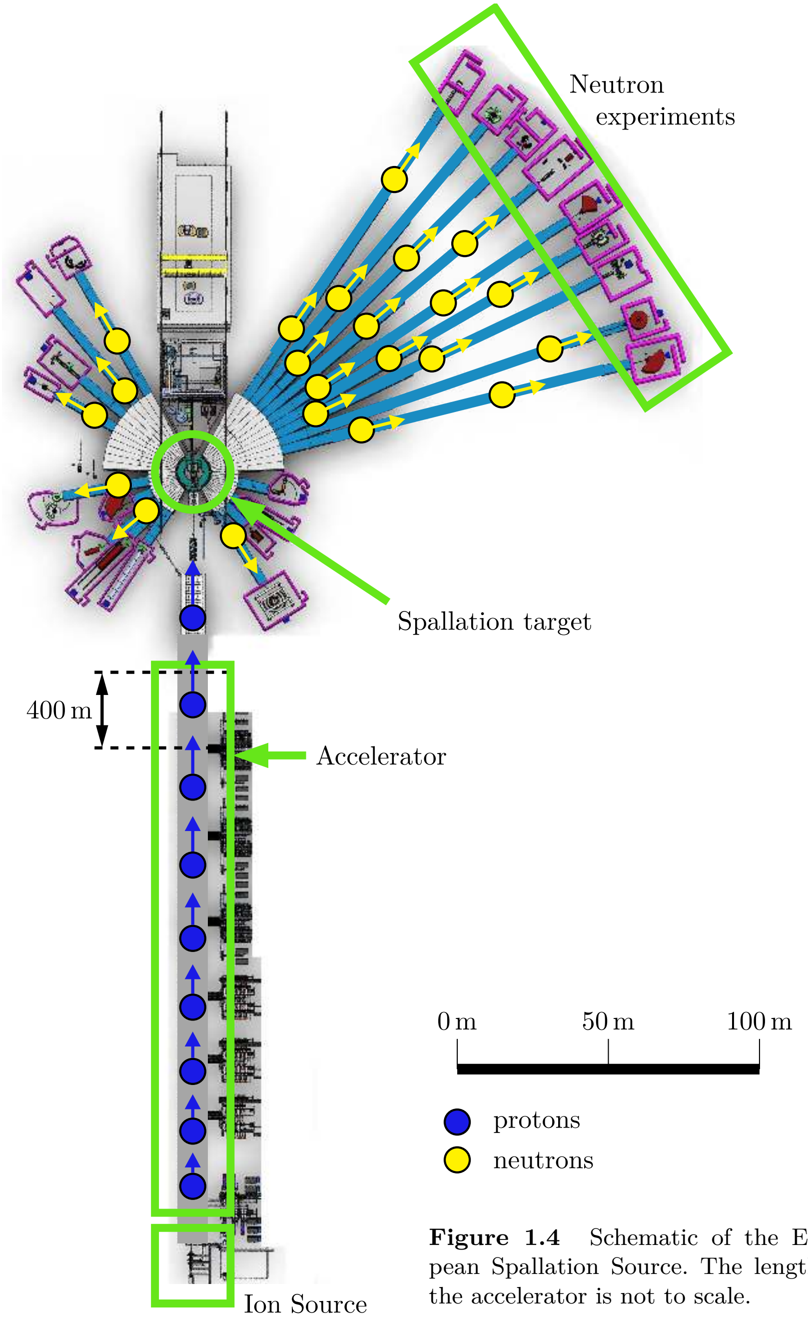


Figure 1.4 Schematic of the European Spallation Source. The length of the accelerator is not to scale.

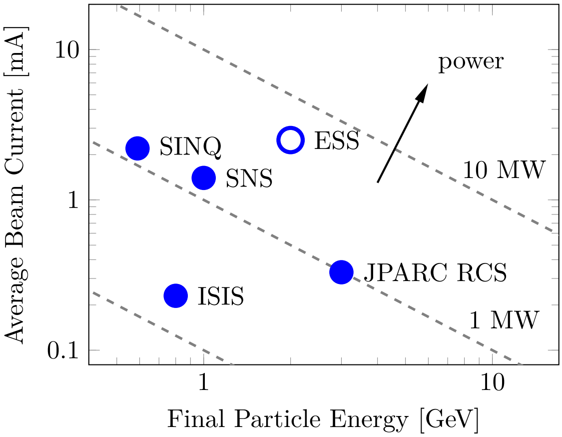


Figure 1.5 How the 5 MW average power of the ESS accelerator compares to the currently most powerful linear accelerators in the world.

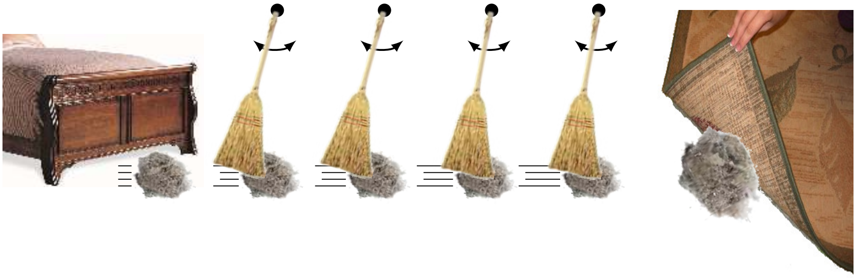


Figure 1.6 Equivalent broom model, the dust is accelerated from the source (under the bed), to the target (under the carpet).

and energy is provided to the brooms by (rather weak) electric motors, which provide torque in phase with the oscillation. The broom energies (potential + kinetic) correspond to the energy stored in the cavity fields.

The movements of the brooms are synchronized to the dust balls, ensuring that they are swept towards the target. The brooms provide an accelerating force on the dust balls, while the dust balls provide a retarding force on the brooms. Due to the large moment of inertia of the brooms, many dust balls are needed to perturb their oscillation.

If the brooms are not properly synchronized to the dust balls, they will spread dust in the room, which results in poor air quality.

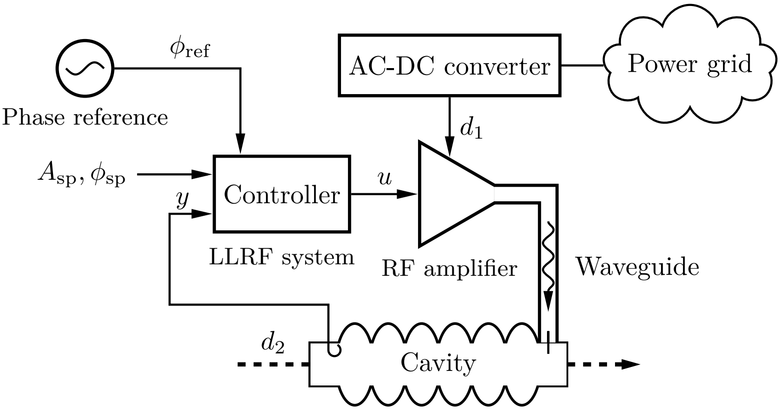


Figure 1.7 Typical field control loop for controlling the amplitude and phase of the electromagnetic field in an accelerator cavity. The amplitude and the phase (relative ϕ_{ref}) of y are controlled to the set-points A_{sp} and ϕ_{sp}. Ripple from the power converter, d_1, and variations in the beam current, d_2, act as disturbances on the system. The controller is implemented in the Low-Level RF (LLRF) system.

The field control problem

The field control loop for a single cavity is shown in Figure 1.7. The control problem is to design a controller so that the amplitude and phase of the cavity field are kept at their desired set-points despite disturbances that act on the system. The controller also needs to be robust to variations in the RF system.

Designing a good controller is one thing, but equally important during the design of an accelerator is to specify (and not overspecify, since this leads to increased costs) the performance of the RF subsystems so that sufficiently small field errors can be achieved.

Particular field control challenges for the ESS accelerator are:

- The large number of field control loops (155), and that there will be six types of cavities, and five types of RF amplifiers, each with their unique characteristics.
- Certain cavity types have parasitic resonance modes that are close in frequency to the accelerating mode (the closest parasitic mode of the medium-β cavities is 700 kHz away from the accelerating mode). These modes need to be carefully considered to avoid instability and robustness problems in the field control loop.
- The beam current is very high for a linear accelerator, which means that disturbances from beam ripple and amplifier ripple will have a

large impact on the cavity fields.

- The DC supply voltage that powers certain RF amplifiers (klystrons), will have significant ripple due to switching in the AC-DC converters. The ripple couples to the field control loop via the RF amplifier, and is expected to contain frequencies around 90 kHz, where the disturbance attenuation of the field control loop is poor. Since the switching of the AC-DC converters will not be locked to the timing system, it will also be hard to cancel this disturbance by feedforward or iterative learning control.

1.2 Previous work

The proceedings of the CERN Accelerator School give a good overview of accelerator physics, accelerator technology and RF systems; cavity field control is covered by [Baudrenghien, 2000; Schilcher, 2007]. Much work on field control and low-level RF hardware has been presented in the dedicated Low-Level RF Workshop held bi-annually, and also on the more general International Particle Accelerator Conference (IPAC) which is held annually.

Two high-intensity proton linacs, which share many characteristics and challenges with the one at ESS, are the linac of the Spallation Neutron Source (SNS) [UT-Battelle, 2006] and Linac4 at CERN [Arnaudon, 2006]. For discussions of the field control problem at these facilities, see [Ma et al., 2006; Baudrenghien et al., 2014].

The best way to learn specific details of cavity field control is to talk to experts in the field. Second to that, PhD theses from leading accelerator facilities provide detailed investigations of different aspects of cavity field control. It should however be kept in mind that the challenges of the field control problem strongly depend on the considered accelerator type (high-intensity proton linac, electron linac, synchrotron, etc).

Particularly much research has been done on field control for electron linacs for free-electron lasers (FELs), due to the challenging requirements on the field errors [Schilcher, 1998; Brandt, 2007; Hoffmann, 2008; Schmidt, 2010; Pfeiffer, 2014; Doolittle, 2015; Rezaeizadeh, 2016].

We conclude by mentioning some interesting control applications to particle accelerators, which are not related to the work in this thesis, but surely enjoyable for aficionados of feedback theory. For extremely narrowband cavities, where the Lorentz force detuning could be several cavity bandwidths, the self-excited loop introduced in [Delayen, 1978] is the only viable choice. In field control for circular machines, the dominant challenge is the coupled dynamics of the cavity field and the circulating beam, see [Baudrenghien, 2000]. A classic result on feedback stability for circular accelerators is *Robinson's* *stability theorem* [Robinson, 1964]. For applications of extremum-seeking control to a linear particle accelerator, see [Scheinker, 2012].

1.3 Contributions

My aim with this licentiate thesis has been to describe how to analyze and design the field control loops for a high-intensity linear proton accelerator such as the ESS accelerator. I have tried to make the text accessible to accelerator engineers that are not working directly with cavity field control, and to present sufficient details for colleagues from automatic control to understand the topic of cavity field control for linear accelerators. The main contributions of this licentiate thesis are:

- A discussion of how complex-coefficient systems can be used for control analysis, and how this applies to the field control problem.
- A presentation of an alternative, more direct, derivation of the cavity field dynamics.
- An in-depth discussion of the field control problem for high-intensity proton accelerators.
- An investigation showing that for many field control problems, simple PI- and PID-controllers have performance close to the optimal linear controller.
- A demonstration of different control strategies for cavities with parasitic resonance modes, and how a complex-coefficient system representation makes the design more intuitive.
- A discussion of directionality in the control problem, illustrated by "instantaneous" phasor diagrams.
- A derivation of how to minimize the energy needed to build up the fields in RF cavities.

Published material on which this licentiate thesis is based is listed below. Most of its content has not been published elsewhere, but some are loosely based on technical reports produced by O. Troeng and B. Bernhardsson for the critical design review of the ESS LLRF system.

The formalism with complex-coefficient transfer functions, which is the foundation for much of the work in this thesis, was introduced in

Troeng, O., B. Bernhardsson, and C. Rivetta (2017). "Complex-coefficient systems in control". In: *Proceedings of the 2017 American Control Conference.* (Seattle, WA, May 24–26, 2017).

This publication grew out of the realization of B. Bernhardsson, that the cavity field control problem can be analyzed using complex-coefficient transfer function. C. Rivetta independently had similar ideas. The paper was written by O. Troeng who collected the theory and applications of complex-coefficient transfer functions, and realized that also Cartesian feedback linearization could beneficially be analyzed by complex coefficient transfer functions. B. Bernhardsson also contributed many valuable suggestions and comments. The publication can be found in Appendix E.

The content of Chapter 9 on directionality in the field control problem was first presented (not subject to peer-review) in

Troeng, O., B. Bernhardsson, A. J. Johansson, and R. Johansson (2015). *Cavity field control for the European Spallation Source*. Low-Level Radio Frequency Workshop 2015, Shanghai, Nov. 3–6, 2015.

The ideas are due to O. Troeng. Anders J Johansson and Bo Bernhardsson contributed valuable comments on the presentation.

The content of Chapter 10 on how to minimize the energy required to build up the electromagnetic fields in an RF cavity is based on the following conference contribution

Troeng, O. and B. Bernhardsson (2017). “Optimal excitation of radio-frequency cavity”. In: *Proceedings of the 20th World Congress of the International Federation of Automatic Control*. (Toulouse, France, July 9–14, 2017), Accepted.

The main ideas are due to O. Troeng, who also wrote the manuscript. B. Bernhardsson contributed the self-contained optimality proof and many valuable suggestions that significantly improved the paper.

1.4 Overview of the thesis

Chapter 2, Basics of Automatic Control, provides an overview of basic concepts from automatic control that will be used throughout the thesis. The chapter also introduces complex-coefficient system, which will be used extensively throughout the thesis, and discusses their differences to real-coefficient systems.

Chapter 3, The ESS Accelerator, gives an overview of the ESS accelerator and its RF system, and considers the field control problem from some different perspectives.

Chapter 4, Subsystem Models for the ESS Accelerator, discusses dynamic modeling of RF subsystems that are relevant for the field control performance.

Chapter 5, Modeling of Accelerator Cavities, presents a derivation of the baseband dynamics of an accelerator cavity. Parasitic modes are also accounted for in the model, and a normalization suitable for control design is introduced.

Chapter 6, The Field Control Loop, introduce a model for the dynamics of the field control loop based on the results from the two preceding chapters. Desirable behavior of the field control loop is discussed, and a mathematical formulation of the control problem is presented based on that discussion. The chapter also classifies the disturbances that affect the field control loop, and considers how loop phase variations impact feedback stability.

Chapter 7, Limits of Field Control Performance, investigates the optimal field control performance that can be achieved by a general, linear time-invariant controller, and how that performance compares to the performance of PI(D)-controllers.

Chapter 8, Field Control for Parasitic Cavity Modes, presents different strategies to deal with parasitic cavity modes in the field control design.

Chapter 9, Directionality in the Control Problem, considers how disturbances affect the cavity field in particular directions, and how also the objective function has a directional dependence.

Chapter 10, Energy-Optimal Cavity Filling, presents the energy-optimal strategy to build up the electromagnetic cavity fields at the beginning of each RF pulse.

Chapter 11, Conclusions and Future Work

2

Basics of Automatic Control

In the first three sections of this chapter we give recaps/short introductions to basic concepts in automatic control that will be used extensively throughout this licentiate thesis: transfer functions, feedback, Nyquist's stability criterion, and norms of signals and systems. For a more thorough introduction to these topics, see the Chapters 1–4 of [Åström and Hägglund, 2006].

In Section 2.4, we give an introduction to using linear systems with *complex-coefficients* for control design; the complex-coefficient representation will allow us to simplify the analysis and design of the field control loop.

2.1 Transfer functions

The dynamics of linear, time-invariant (LTI) systems can be represented by linear differential equations in the time domain, or by transfer functions in the frequency domain. The two representations are related via the Laplace transform. In this thesis we will mostly use the frequency domain representation since it gives more insight, and is well-suited for analyzing the field control problem.

An important property of (stable) linear time-invariant systems is that for a sinusoidal input signal, then also the output signal becomes sinusoidal after transients. The amplitudes and phases of the two sinusoids are related by the transfer function of the system; see Figure 2.1.

Example

Assume that the linear, time-invariant, first-order system $\mathcal{S}$ has the following relationship between its input signal $u(t)$ and output signal $y(t)$,

$$y'(t) + 10y(t) = 10u(t), \tag{2.1}$$

then by taking the one-sided Laplace transform (for zero initial conditions) we obtain

$$sY(s) + 10Y(s) = 10U(s), \tag{2.2}$$

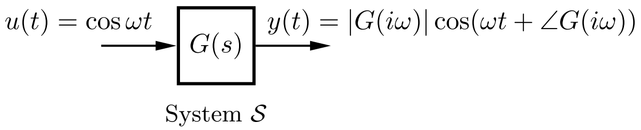


Figure 2.1 If $\mathcal{S}$ is a linear time-invariant system, then its transfer function $G(s)$ relates the Laplace transforms of the input and output signals via $Y(s) = G(s)U(s)$.

where $U(s) = \mathcal{L}u$ and $Y(s) = \mathcal{L}y$. By introducing the transfer function from u to y,

$$G(s) = \frac{10}{s+10}$$

we can re-write (2.2) as

$$Y(s) = G(s)U(s).$$

Transfer function of time delay

A system given by a time delay of length L,

$$y(t) = u(t - L),$$

has the transfer function

$$G(s) = \mathrm{e}^{-sL}.$$

Bode diagrams

Transfer functions are conveniently visualized with so-called Bode diagrams, see Figure 2.2a for the Bode diagram of the transfer function of the system

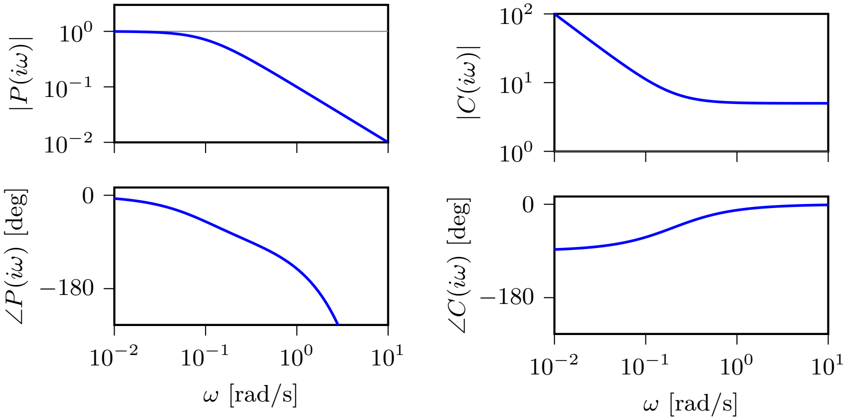
$$P(s) = \frac{10}{s+10}\mathrm{e}^{-s}, \tag{2.3}$$

which corresponds to the system (2.1) in series with a time delay of 1 s.

It is seen from the magnitude curve in Figure 2.2a that low-frequency variations ($\omega < 0.1$ rad/s) in the input signal give rise to output variations of the same amplitude, but that variations of higher frequencies are attenuated.

2.2 Feedback

Given a *stable* system, we could in principle make its output y equal to a constant set-point y_{sp} by simply choosing the input signal u appropriately, in the case of the system (2.3) by choosing $u = y_{\mathrm{sp}}$. There are several problems with this approach: (1) it requires perfect knowledge of the process dynamics



(a) Process $P(s) = \mathrm{e}^{-s}/(10s+1)$.

(b) Controller $C(s) = 5 + 1/s$.

Figure 2.2 Bode diagram for the process $P(s)$ in (2.3) and the PI-controller (2.4) with $K = 5$ and $T_i = 5$. The upper diagram shows the amplitude of the transfer function $P(i\omega)$ as a function of frequency, and the lower diagram shows the phase of the transfer function.

$P(s)$, (2) if there are disturbances acting on the process, see Figure 2.3a, there will be errors on y, and (3) if the system dynamics are slow it will take a long time for y to converge to y_{sp}. All these problems can be mitigated by introducing a feedback controller $C(s)$ that adjusts the input-signal u based on the control error $e = y_{\text{sp}} - y$; see Figure 2.3b.

To illustrate the advantages of feedback, we consider a so called proportional-integral (PI) controller, for which the control signal is given by

$$u(t) = K\left(e(t) + \frac{1}{T_i}\int_{-\infty}^{t} e(t)\,dt\right).$$

The first term is proportional to the control error and the second term is proportional to the integral of previous control errors.

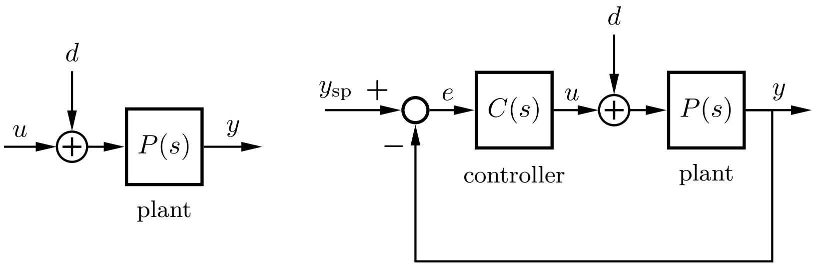
In the Laplace-domain the PI-controller takes the form

$$C(s) = K\left(1 + \frac{1}{sT_i}\right) \tag{2.4}$$

see Figure 2.2b for the Bode diagram of the controller (2.4) with $K = 5$ and $T_i = 5$.

Sensitivity function

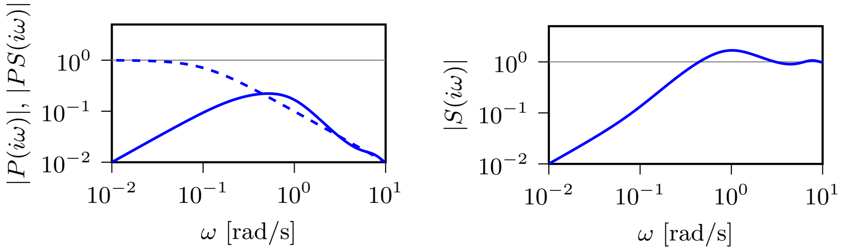
Without feedback (Figure 2.3a) the transfer function from disturbances d to control errors e is given by $P(s)$; by introducing a feedback controller $C(s)$



(a) No feedback.

(b) With feedback.

Figure 2.3 Illustration of control without feedback (open-loop control), and with feedback (closed-loop control).



(a) Transfer function from disturbances d to the measured signal y; *dashed line:* P (without feedback), *solid line:* PS (with feedback).

(b) Sensitivity function $S = 1/(1 + PC)$.

Figure 2.4 Bode magnitude diagrams for two closed-loop transfer functions when the process P and controller C are given by the transfer functions in Figure 2.2; the phases of the transfer functions are not shown since it is rarely of interest for closed-loop transfer functions.

(Figure 2.3b) the transfer function from disturbances to the control errors becomes instead $P/(1+PC)$; Figure 2.4a shows the improved attenuation of low-frequency disturbances that is provided by the feedback controller.

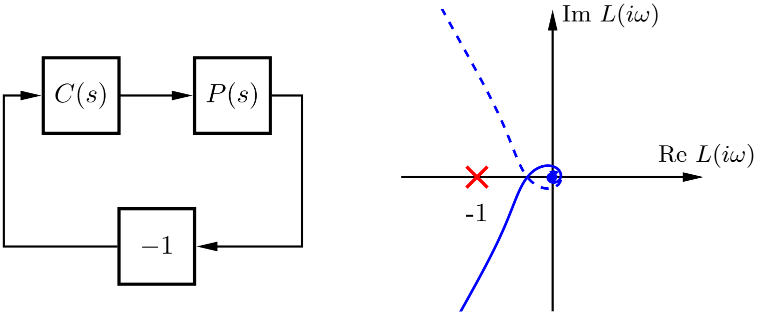
The difference in disturbance attenuation from introducing feedback is quantified by the so-called sensitivity function

$$S = \frac{1}{1+PC}, \tag{2.5}$$

see Figure 2.4b.

Stability

Stability is a fundamental concern for systems involving feedback; in simple terms, instability results if the controller compensates control errors too ag-



(a) Considered feedback interconnection.

(b) The Nyquist curve of the open-loop system $L(s) = C(s)P(s)$.

Figure 2.5 Illustration of the Nyquist criterion for the for the process P and controller C in Figure 2.2. From (b) it is seen that the Nyquist curve is to the right of the point -1, thus, according to the Nyquist criterion, the closed-loop system in (a) is stable.

gressively. In practice, instability leads to that the system starts to oscillate uncontrollably or gets destroyed. For the applications in this thesis, closed-loop stability is conveniently verified by the Nyquist stability criterion.

THE NYQUIST STABILITY CRITERION (SIMPLIFIED VERSION)
Let $L(s) = C(s)P(s)$ be the open-loop transfer function and assume that $L(s)$ is strictly stable (i.e., has no poles in the right half-plane). Then the closed-loop system is stable if and only if the Nyquist curve $L(i\omega), \omega \in (-\infty, \infty)$ does not encircle and stays to the right of the critical point -1. ☐

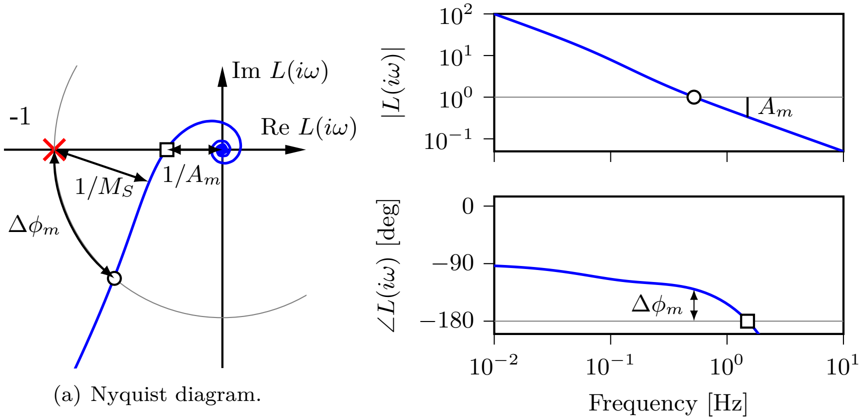
See Figure 2.5 for an illustration of the Nyquist stability criterion.

Robustness

A feedback system is stable when its Nyquist curve is to the right of the point -1; not surprisingly, the shortest distance between the Nyquist curve and the point -1 is a good indication of how far the closed-loop system is from instability, i.e., how robust the system is to modeling errors and process variations. Since the distance in the complex-plane between the Nyquist curve and the point -1 is given by $|L(i\omega) - (-1)| = 1/\left|S(i\omega)\right|$, we realize that the maximum value of the sensitivity function (2.5),

$$M_S := \sup_{\omega} |S(i\omega)| \tag{2.6}$$

is good measure for the robustness of a feedback interconnection; the smaller the value of M_S, the more robust is the feedback interconnection.



(a) Nyquist diagram.

(b) Bode diagram.

Figure 2.6 Illustration of different robustness measures in the Nyquist diagram and the open-loop Bode diagram.

Amplitude and phase margins Two traditional robustness measures are the *amplitude margin* A_m: how much the gain of the open-loop transfer function can be increased before instability occurs, and the *phase margin* $\Delta\phi_m$: how much the phase of the open-loop transfer function can be reduced before instability occurs; see Figure 2.6.

Even if the amplitude and phase margins are large, the closed-loop system could have poor robustness [Åstrom and Murray, 2010, Figure 9.11]. On the other hand, a small value of M_S always guarantees reasonable amplitude and phase margins [Skogestad and Postlethwaite, 2007, (2.39)],

$$
\begin{aligned}
A_m &\geq \frac{M_S}{M_S - 1} \\
\Delta\phi_m &\geq 2 \arcsin \frac{1}{2M_S}.
\end{aligned}
$$

For example, $M_S = 1.6$ guarantees, according to the inequalities above, an amplitude margin *of at least* 2.6, and a phase margin of *at least* 36°. Thus, a robustness specification in terms of M_S is more general than amplitude and phase margins, and typically it is also more convenient to work with.

Fundamental limitations

Before continuing, we recall that there are *fundamental* limitations on the achievable control performance. If both the process and the controller are linear and time invariant, then one such limitation is given by Bode's integral formula (assuming a stable open-loop system with sufficient roll-off),

$$\int_{-\infty}^{\infty} \log |S(i\omega)| \, d\omega = 0, \tag{2.7}$$

where S is the sensitivity function (2.5).

Bode's integral formula can be considered as a conservation law of sensitivity—if the feedback pushes down the sensitivity below 1 at some frequencies, it will unavoidably pop up above 1 at other frequencies; fittingly, this is known as the waterbed effect[1]. The waterbed effect is seen in Figure 2.4a, where the control system reduces disturbances at low frequencies, but amplifies them around 1 rad/s.

2.3 Useful norms

Norms quantify the size of mathematical objects. In what follows it will be convenient to use the 2-norm of signals to quantify their energy/rms-value and the H_∞-norm of transfer functions to quantify robustness of feedback interconnections. Below we provide the definitions of these norms for the single-input single-output case; for the multi-input multi-output case, as well as for more details, see [Skogestad and Postlethwaite, 2007, 2.8.1, 4.10, A.5].

2-norm of signal

The 2-norm of a scalar signal y with finite support $[0, T]$, and with one-sided Laplace transform Y is defined as

$$||y||_2 = ||Y||_2 := \sqrt{\int_0^T |y(t)|^2 \, dt} \tag{2.8a}$$

$$= \sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} |Y(i\omega)|^2 \, d\omega.} \tag{2.8b}$$

From (2.8a) we see that the rms-value of y equals $||y||_2 / \sqrt{T}$.

If y is a stochastic signal from a stationary process with spectrum Y, the norm (2.8b) instead corresponds to the *expected* rms value of the signal as $T \to \infty$.

The 2-norm will allow a compact representation of control performance—for example, if y is the amplitude error of the cavity field, then $||Y||_2$ is the rms value of y, which is what needs to be kept within specifications.

---

[1] In the context of field control, the waterbed effect has been previously mentioned in [Schilcher, 2007]. See [Stein, 2003, figs 3, 4] for a nice illustration of the waterbed effect.

H_∞-norm of a system

The H_∞-norm (H-infinity norm) of a linear, time-invariant, single-input single-output system $G(s)$ is given by

$$||G||_\infty = \begin{cases} \sup_\omega |G(i\omega)| & \text{if } G \text{ is stable} \\ \infty & \text{if } G \text{ is unstable} \end{cases} . \tag{2.9}$$

Using the H_∞-norm we can now re-define the robustness measure (2.6) as

$$M_S := ||S||_\infty ,$$

after which a small value of M_S implies that the system is both stable and robust.

2.4 Complex-coefficient systems

In this section we go through some results on complex-coefficient systems that will be used throughout this thesis; for more details see Appendix E.

Complex-coefficient systems from baseband transformations

For a so-called passband system $G_{\mathrm{PB}}(s)$ that has narrow support around some frequency ω_c (Figure 2.7a), the analysis is often facilitated by considering the system in the baseband, i.e., by shifting the system's spectrum by $-i\omega_c$. The transfer function for the equivalent baseband system, $G(s) = G_{\mathrm{PB}}(s+i\omega_c)$, then typically has complex coefficients, see Figure 2.7b.

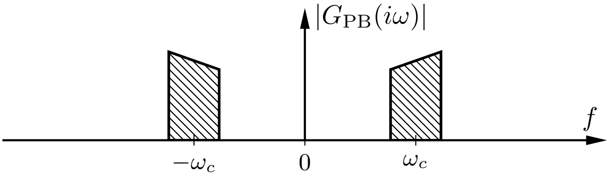
The input signal $u_{\mathrm{PB}}(t)$ and output signal $y_{\mathrm{PB}}(t)$ of $G_{\mathrm{PB}}(s)$, have in cases of practical interest, a narrowband spectrum around ω_c. They can thus be considered as sinusoidal signals with (angular) frequency ω_c, whose amplitudes and phases are slowly modulated. The corresponding (complex-valued) signals in the baseband are then given by $u(t) = A_u(t)\mathrm{e}^{i\phi_u(t)}$ and $y(t) = A_y(t)\mathrm{e}^{i\phi_y(t)}$, where A_u (A_y) is the slowly varying amplitude, and ϕ_u (ϕ_y) is the slowly varying phase of u_{PB} (y_{PB}).

The main points

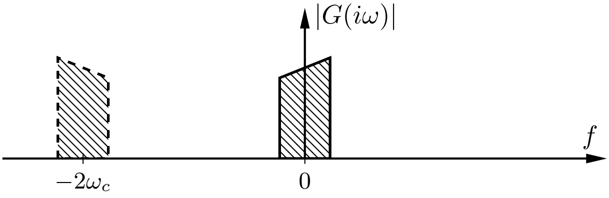
The frequency response of complex-coefficient systems $G(s)$ is not conjugate symmetric[2]with respect to 0, which implies that the frequency response needs to be considered at both positive and negative frequencies in analyses. Given that this is done, that *Hermitian* transposition is used instead of transposition, and that double-side integration is used in the frequency domain, then

---

[2] Conjugate symmetry means that $G(i\omega) = \overline{G(-i\omega)}$; this is a property that holds for systems with real coefficients, but typically not for systems with complex coefficients.



(a) Original passband system.



(b) Baseband transformed system, the passband at $-2\omega_c$ is neglected.

Figure 2.7 Illustration of baseband transformation, $s \mapsto s - i\omega$, of the system $G_{\mathrm{PB}}(s)$, resulting in a baseband system $G(s) = G_{\mathrm{PB}}(s + i\omega_c)$.

most formulae and results from standard control theory apply. In particular, Nyquist's theorem still holds, which allows intuitive frequency-domain design methods to be used.

Care is required when working with complex-coefficient systems in Matlab since the current toolbox versions do not handle them correctly.

Comparison to two-input two-output representation

The transfer function $G(s)$ of a complex-coefficient, single-input single-output (SISO) system can be written uniquely as

$$G(s) = G_{\mathrm{Re}}(s) + iG_{\mathrm{Im}}(s), \tag{2.10}$$

where $G_{\mathrm{Re}}(s)$ and $G_{\mathrm{Im}}(s)$ have real coefficients[3]. If the complex input signal is written

$$u(t) = u_{\mathrm{Re}}(t) + iu_{\mathrm{Im}}(t), \tag{2.11}$$

then $G(s)$ can be represented as a real-coefficient, two-input two-output (TITO) system

$$\mathbf{G}(s) = \begin{bmatrix} G_{\mathrm{Re}}(s) & -G_{\mathrm{Im}}(s) \\ G_{\mathrm{Im}}(s) & G_{\mathrm{Re}}(s) \end{bmatrix} \tag{2.12}$$

---

[3] If the impulse response of $G(s)$ is separated in its real and imaginary parts, $g(t) = g_{\mathrm{Re}}(t) + ig_{\mathrm{Im}}(t)$, then $G_{\mathrm{Re}}(s)$ is the Laplace transform of $g_{\mathrm{Re}}(t)$, and $G_{\mathrm{Im}}(s)$ is the Laplace transform of $g_{\mathrm{Im}}(t)$. Alternatively, we from (E.8) that

$$G_{\mathrm{Re}}(s) = \frac{G(s) + G^*(\bar{s})}{2}, \qquad G_{\mathrm{Im}}(s) = \frac{G(s) - G^*(\bar{s})}{2i}.$$

acting on real, vector-valued signals

$$\begin{bmatrix} u_{\mathrm{Re}}(t) \\ u_{\mathrm{Im}}(t) \end{bmatrix}. \tag{2.13}$$

The real-coefficient representation (2.12) is in many respects similar to the complex-coefficient representations (2.10), however the complex-coefficient representation has some advantages:

- The insight and simplicity of working with SISO systems is maintained. Especially stability and robustness analysis in the frequency domain becomes more intuitive since it is not necessary to resort to the general (MIMO) Nyquist criterion.
- Less computations are required. Consider as an example, multiplying the frequency response of two TITO systems $\mathbf{G_1}(i\omega)$ and $\mathbf{G_2}(i\omega)$ of the form (2.12). If the structure of the matrix is not exploited, eight multiplications and four additions of complex numbers would be required. If structure is exploited, the work is reduced to four multiplications and two additions. The complex-coefficient SISO representation requires only 2 complex multiplications (also the frequency response at $-i\omega$ needs to be considered).
- There is no need to explicitly impose the rotationally invariant structure of (2.12) when doing systems identification and control design, since that structure is implicit in the complex-coefficient representation.

Norms for complex vs real representation The norms introduced in Section 2.3 do not depend on whether the complex representation (2.10–2.11), or the real representation (2.12–2.13) is used, i.e.,

$$||u||_2 = \left|\left| \begin{bmatrix} u_{\mathrm{Re}} \\ u_{\mathrm{Im}} \end{bmatrix} \right|\right|_2 \qquad \text{and} \qquad ||G||_\infty = ||\mathbf{G}||_\infty,$$

for details, see Appendix E.

Example: Baseband transformation of a time delay

It is illustrative to consider the baseband equivalent of a time delay. Let

$$G_{\mathrm{PB}}(s) = \mathrm{e}^{-sL}, \tag{2.14}$$

and then the equivalent baseband system is given by

$$G(s) = G_{\mathrm{PB}}(s + i\omega_c) = \mathrm{e}^{-i\omega_c L}\mathrm{e}^{-sL}.$$

We see that the baseband model contains the same time delay as (2.14), but that there is also an additional complex factor $\mathrm{e}^{-i\omega_c L}$. Typically $\omega_c L$ is large

(2000–4000 for the field control loops at ESS), so even small changes of the time delay L give significant phase changes in $G(s)$, which has a large impact on the control performance, as discussed in Section 6.5; for example, with $f_c = \omega_c/(2\pi) = 704\,\text{MHz}$, then a delay variation $\Delta L = 100\,\text{ps}$ corresponds to a phase change of 25°.

3

The ESS Accelerator

In this chapter we give an overview of the ESS accelerator and its RF system. We then discuss the field control requirements and give some comments from an automatic control perspective.

3.1 Overview of the ESS accelerator

An overview of the accelerator is shown in Figure 3.1.

Free protons (i.e., hydrogen ions) are generated in the ion source by ionizing hydrogen gas with microwave radiation. The protons drift through the low-energy beam transport (LEBT, not shown in the figure), to the radio-frequency quadrupole (RFQ). The RFQ is a special type of RF cavity whose electromagnetic field both accelerates the beam and forms it into bunches; the bunching is what allows the beam to be accelerated by the oscillating electromagnetic fields in the downstream cavities.

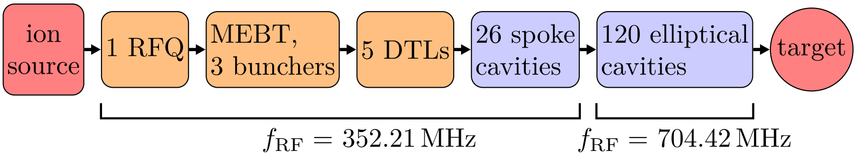


Figure 3.1 Block diagram of the linear proton accelerator at ESS. Depending on the proton velocity, different accelerator cavities are used. The first 9 cavities (orange color; 1 radio-frequency quadrupole, 3 buncher cavities, and 5 drift-tube linacs) are made of copper, and are operated at room temperature, the succeeding 146 cavities (blue color) are made of niobium and kept at 2 K, which makes them superconducting; this reduces the resistive losses and allows stronger electric fields. There are two types of elliptical cavities, medium-β and high-β, that are optimized for different proton velocities.

After the RFQ, the three buncher cavities in the medium-energy beam transport (MEBT) keep the beam bunched, before five drift tube linacs (DTLs) accelerate the bunches from from 9 % to 41 % of the speed of light.

These first 9 cavities (RFQ, bunchers, DTLs) are made of copper, and are operated at room temperature. All the succeeding cavities are made of niobium and are cooled to 2 K. This makes them superconducting, which gives very low electric losses and allows strong electric fields. Three different superconducting cavities are used depending on the velocity of the proton bunches: spoke cavities, medium-β elliptical cavities, and high-β elliptical cavities. The RF frequency of the RFQ, bunchers, DTLs and spoke cavities is 352 MHz and for the elliptical cavities it is doubled to 704 MHz[1]. Images of three different types of cavities are shown in Figure 3.2.

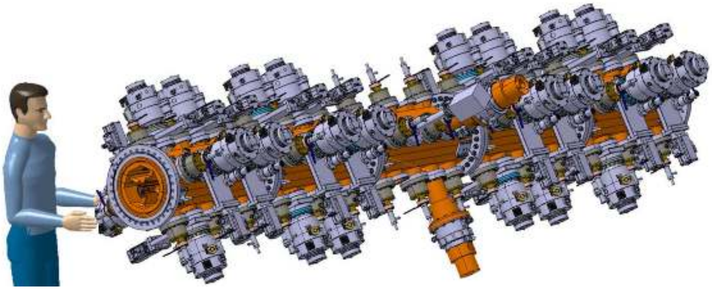
Pulse structure The proton beam of the ESS accelerator will be pulsed at 14 Hz, with 2.86 ms long beam pulses, corresponding to a 4 % duty cycle. The beam current during the pulses will be 62.5 mA, which is quite high for a linear accelerator. Each pulse contains about one million bunches, and every bunch consists of approximately one billion protons. The RF fields need to be built up prior to the beam pulse, so the RF amplifiers and their power supplies are turned on some 300 µs before the arrival of the proton bunches (Figure 3.3).

The motivation for the pulsed operation is to allow higher peak intensities and to give a time separation of the neutron energies that arrive at the experiment stations. The specific pulse rate of 14 Hz is related to the time it takes for the neutrons to travel from the target, through the neutron guides, to the instrument stations.

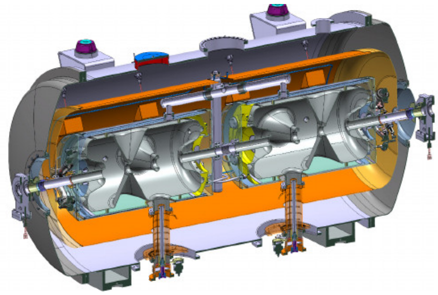
More details on the ESS accelerator The accelerator cavities and the RF system are the focus of this licentiate thesis, but to give a sense of the complexity of the ESS accelerator, we mention some of its other subsystems: beam instrumentation to observe properties of the proton beam, a vacuum system to give the protons a free path to the target, a cryo system to keep the superconducting cavities cold, magnets to steer the beam, water cooling to remove heat, choppers to shape the beam, protection systems to avoid that things get destroyed, and an industrial control system which allows all systems to be monitored and controlled from the control room. The reader who wants to learn interesting things about these systems is encouraged to read Chapter 4 of the ESS Technical Design Report [Peggs et al., 2013]. A good introduction to linear accelerators is [Wangler, 2008].

---

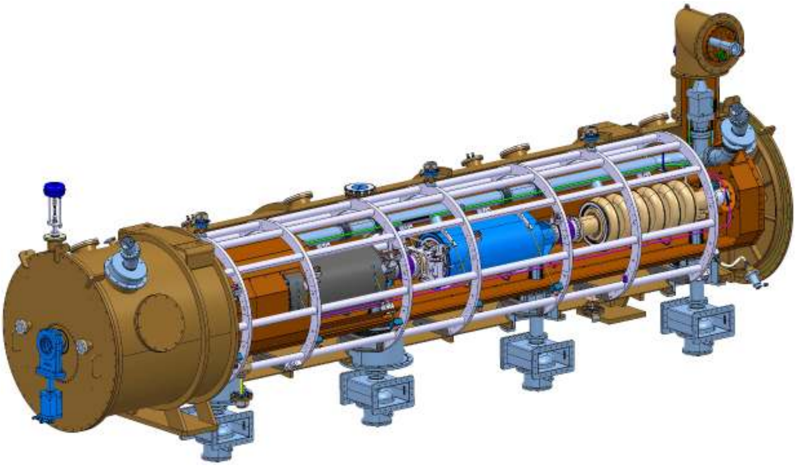
[1] 352.21 MHz and 704.42 MHz to be exact.



(a) Radio-frequency quadrupole with vacuum pumps and tuners [Alberi and Lacroix, 2015].



(b) Cryomodule with two spoke cavities [Bousson et al., 2014].



(c) Cryomodule with four elliptical medium-β cavities [ESSb, 2017].

Figure 3.2 Three of the six RF cavity types that will be used for the ESS accelerator. The spoke cavities and the elliptical cavities are kept in cryomodules to allow them to be cooled to 2 K, which makes them superconducting. Image credit: ESS.

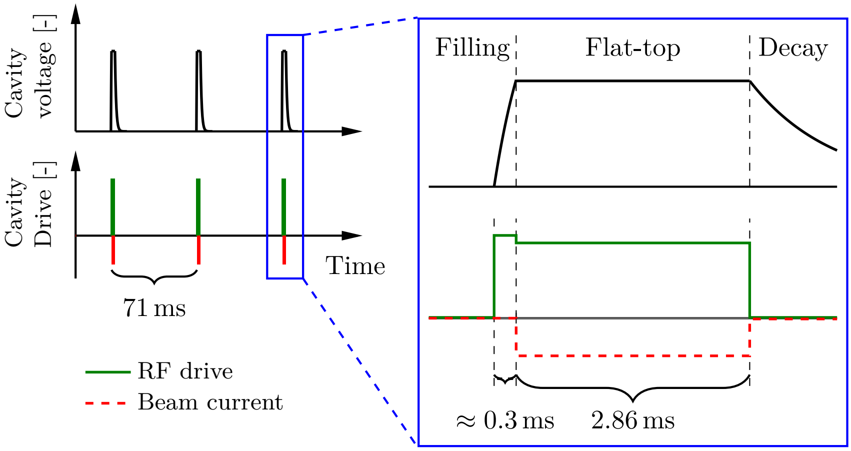


Figure 3.3 Timing of the 14 Hz pulses of the ESS. The RF amplifiers are turned on some 300 µs before the beam, since the RF amplifiers need to build up the cavity fields to their nominal level before the beam can be accelerated; this part of the pulse is called *filling*. With all cavity fields at their nominal levels, the beam is turned on; while energy is provided by the cavity fields to the proton beam, the RF amplifiers provide power to the cavity fields to maintain them at their nominal levels; this part of the pulse is called *flat-top*. Finally, the RF amplifiers are turned off together with the beam, and the cavity fields decay by resistive losses, and via the waveguide, to the RF load. The terminology for the different parts of a pulse is a de-facto standard, see for example [Brandt, 2007].

3.2 The RF system of the ESS accelerator

The energy in the electromagnetic cavity fields that accelerate the protons is provided by the RF system. Since the radiation from beam losses in the RF cavities is high enough to destroy ordinary electronics, the accelerating cavities will be located in a tunnel several meters below ground, while the RF amplifiers and RF electronics are kept in a ground level gallery above the tunnel; see Figure 3.4 for a high-level schematic of the RF system.

The RF amplifiers generate electromagnetic waves, which build up and maintain the electromagnetic fields in the cavities. The most suitable type of RF amplifier for a given application depends on the required output power, the duty cycle, and the RF frequency. A high-intensity proton accelerator, such as the one of ESS, has several different cavity types, which calls for different RF amplifier technologies. At ESS it is planned to use four different amplifier types: klystrons, solid-state amplifiers, tetrodes and inductive output tubes (IOTs), see Table 3.1. For an introduction to RF amplifiers, see

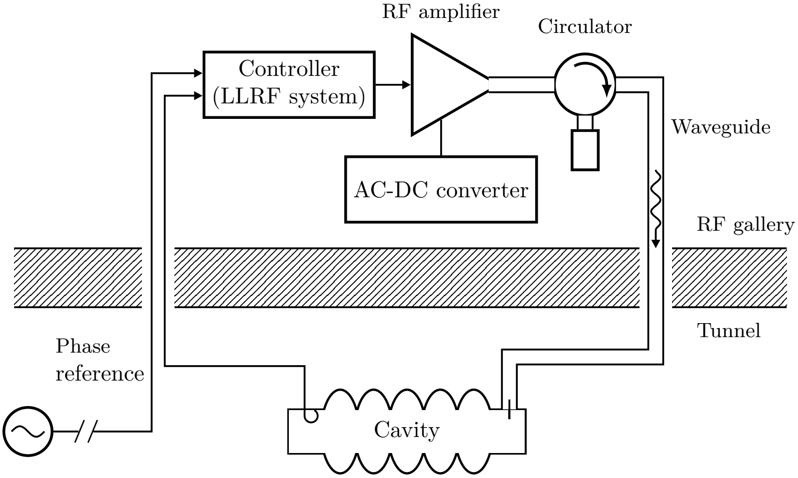


Figure 3.4 Schematic of a typical RF system.

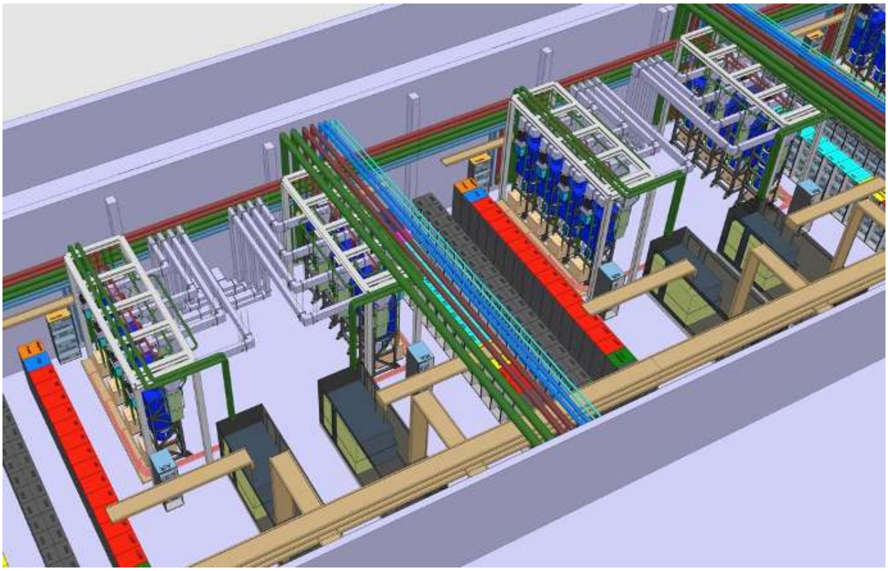


Figure 3.5 Two RF cells for the medium-β section of the ESS accelerator. Each cell drives eight medium-β cavities, and consists of eight klystrons, two AC-DC converters (so-called modulators), and more than twenty racks with electronics for control and monitoring. Image credit: ESS.

Table 3.1 The RF amplifiers that will be used for the ESS accelerator. It will be the first accelerator to use multi-beam inductive output tubes.

Cavity Type (No. Units)	$\boldsymbol{f_{\mathrm{RF}}}$ [MHz]	**Output Power** [kW]	**Amplifier Type**
RFQ (1)	352.21	3000	Klystron*
Buncher (3)	–”–	30	Solid-State Amplifier
DTL (5)	–”–	3000	Klystron*
Spoke (26)	–”–	2×200	Tetrode
Medium-β (36)	704.42	1500	Klystron*
High-β (84)	–”–	1200	Inductive Output Tube†

*The klystrons will have 200 W–300 W preamplifiers.

†The IOTs will have of 15 kW–20 kW preamplifiers.

[Carter, 2010].

All RF amplifiers for the ESS accelerator need a DC power supply, so AC-DC converters[2] are needed to convert the AC grid voltage. The energy required during the pulses needs to be drawn continuously from the regional power grid, and stored in capacitors banks in the AC-DC converters, since a pulsed draw of about[3] 200 MW, would cause flicker on the grid.

The energy from the RF amplifiers is transferred to the cavities in rectangular waveguides, which in appearance are similar to ventilation ducts. As we will see in Chapter 5, the waveguides will also carry waves traveling backwards from the cavities. These backward waves can destroy the RF amplifiers, and to avoid this, special RF devices, called circulators, re-direct the backward waves to RF loads, where they are absorbed. The waveguides, circulators and loads are the main components of the *RF distribution system.*

As we mentioned in the introduction, the phases of the electromagnetic fields in the RF cavities relative the phase of the proton bunches need to be accurately controlled. This is done by the *low-level RF system* which adjusts the drive signals to the high-power amplifiers, based on measurements of the cavity fields; additional measured signals of, e.g., the amplifier output, or the beam current, could also be used by the controller to improve the control performance.

Since sufficiently fast and accurate measurements of the bunch phase are not feasible, the phases of the cavity fields are controlled with respect to a very stable reference oscillator, called the master oscillator (MO). This essentially corresponds to control with respect to the bunch phase, since the

---

[2] Which are called modulators if the RF amplifier it drives is a klystron or an inductive output tube.

[3] The power provided to the beam equals 125 MW, but there are losses in the conversion of the electrical grid power.

bunch phase is determined by the phase of the RFQ.

The signal from the master oscillator is distributed along the accelerator tunnel in a coaxial line, which is temperature stabilized to avoid phase drifts from thermal expansion/contraction. The cables for the cavity field signal and the phase-reference signal, from the tunnel to the RF gallery, will not be temperature stabilized, but will be kept close together so that phase drifts cancel out. The master oscillator and the temperature-stabilized phase reference line constitute the *phase reference system*.

To avoid damage to the RF amplifiers and other systems there will be a *local protection system*, that shuts down the RF station if, for example, overheating of RF amplifiers or vacuum leaks in the klystrons are detected.

3.3 Field control requirements

Protons that deviate too much from their nominal velocity will be incorrectly deflected by the transversal focusing magnets, which might cause them to crash into the cavity walls and induce radioactivity.

According to the top-level technical specification for ESS, the availability of the facility shall be better than 95 %. To meet this requirement, “hands on maintenance” of the accelerator, in case something breaks, must be possible within a reasonable amount of time—for this, the accelerator must not be too radioactive. For proton accelerators it has been concluded that beam losses of 1 W/m is “a reasonable limit for hands-on maintenance” [Mokhov and Chou, 1999].

Beam Physics

Starting from the beam loss requirement of 1 W/m, a highly specialized cadre of physicists, known as beam physicists, designed the ESS accelerator layout and derived requirements on mechanical tolerances, alignment errors, magnet field errors and cavity field errors. The requirement of 1 W/m translates to that only 1–2 out of 100 000 particles can be lost, which was challenging to achieve in the design; in fact, the ESS accelerator is *beam-loss limited*, i.e., the requirements on many subsystems have been driven by the need for low beam losses [Levinsen et al., 2016].

The design of beam-loss-limited accelerators is a formidable task since so-called halo formation around the central core of the bunches is a major cause of beam loss. To accurately account for halo formation requires that $10^5 - 10^6$ macro particles (each representing some 1000 protons) are simulated. The simulations are numerically demanding due to the large number of particles and the nonlinear Coulomb forces between them.

To arrive at tolerance requirements for the ESS accelerator, the beam losses for different error distributions were studied; hundreds or thousands of simulations had to be performed for each set of error distributions before conclusions could be drawn from statistical analyses of the simulation results. Each simulation took about one hour on a high-performance computer, so the complete error studies took many months. The primary computer code used for the beam-physics simulations of the ESS accelerator was Tracewin [Uriot and Pichoff, 2015].

After many thorough error studies, the beam physics group at ESS set the cavity field error requirements to 0.2 % and 0.2° rms for warm cavities and 0.1 % and 0.1° rms for cold cavities [Eshraqi and Levinsen, 2016; Levinsen et al., 2016]. The different error levels are motivated by that it is easier to obtain good field stability for superconducting cavities, which we will see in Chapter 7.

3.4 Beam-loss objective function

The proton bunches pass the entire accelerator in less than 4 µs, and after the second DTL tank they travel 1 m in less than 5 ns; the cavity field errors vary much more slowly, so the expected[4] beam losses per time unit is a static function of the field errors, which we introduce as

$$\Psi_{BL}(\Delta A_1, \Delta\phi_1, \Delta A_2, \Delta\phi_2, \ldots) := \begin{array}{l}\text{expected instantaneous}\\ \text{beam loss}\end{array}. \tag{3.1}$$

The objective function that should be minimized is thus

$$J_{BL} := \begin{array}{l}\text{expected average}\\ \text{beam loss}\end{array} = \mathbb{E} \int_0^{t_f} \Psi_{BL}(\Delta A_1, \Delta\phi_1, \Delta A_2, \Delta\phi_2, \ldots)\, dt, \tag{3.2}$$

where the field errors are considered stochastic, and the expectation is taken with respect to their distribution.

The beam physicists try to make J_{BL} small by a suitable design of Ψ_{BL} (i.e., the accelerator lattice); cavity field control, which is discussed in this thesis, is about how to make J_{BL} small by making the variances of ΔA_k and $\Delta\phi_k$ small.

Discussion of field error correlation

When evaluating (3.2) it should be remembered that the field errors $\Delta A_1, \Delta\phi_1, \Delta A_2, \Delta\phi_2, \ldots$ are correlated; beam-current variations give rise to errors correlated between all cavities, and similarly, there will be correlations between the disturbances from amplifiers supplied by the same AC-DC converter.

---

[4] The initial particle distribution can be considered to be stochastic.

In Chapter 9 we will see that there will be significant correlation between the errors ΔA_k and $\Delta\phi_k$ for a fixed k, and discuss consequences of this. The full understanding of the effects due to correlation between the field errors requires more studies.

3.5 A perspective from process control

Taking a step back from the field control problem, we note that the European Spallation Source has many similarities with the plants considered in process control. The ESS will with tens of thousands of measured signal and controlled output, all which will be interfaced to an industrial control system.

There will be many thousands control loops, from very fast ones, such as those for cavity field control, which will have bandwidths around 100 kHz, to rather slow ones, such as those for the temperature control of the phase reference line, which will have a bandwidth of around 0.001 Hz. There will also be even slower feedback in terms of set-point calibration of the amplitudes and phases of the electric fields, where the calibration procedures take several hours and are performed on the time scale of months monthly.

From a control perspective, the ESS is one enormous, multi-input multi-output, nonlinear system, with a complicated, nonlinear objective function. Just as in process control it is not feasible to implement a central controller with thousands of measured signals and control signals, but rather the process is designed so that the control to a large degree can be decentralized, and so that the majority of the loops can be closed by relatively simple and robust controllers that only use local information.

The cavity field controllers used for linear accelerators are typically completely decentralized, but the benefits of using centralized field control for the electron accelerators of free-electron lasers have been studied in [Pfeiffer, 2014; Rezaeizadeh, 2016]. To quantify the benefits of a centralized control for a high-intensity, linear proton accelerators, such as the one of ESS, it is necessary to have a good knowledge of the function Ψ_{BL} in (3.1); however, it is likely that the benefits are outweighed by increased complexity of the controller and its implementation.

3.6 Comparison to accelerators for free-electron lasers

To give some perspective on the field-control challenges for a high-intensity proton accelerator such as the ESS accelerator, we make some comparisons to the field control problem for the electron accelerators used for free-electron lasers (FELs); this is particularly relevant since there is relatively much literature on field control for this type of accelerators.

For free-electron lasers, the emittance of the electron bunches (which determines the brilliance of the laser pulses), and the arrival time of the bunches at the undulator (which determine the timing of the laser pulse), are the primary design objectives [Pfeiffer, 2014; Rezaeizadeh, 2016].

Beam losses are typically not an issue for electron linacs due to four reasons: (1) the *average* beam current in FELs is much lower than in a high-intensity proton accelerator; (2) the Coulomb repulsion between the particles is less significant since they are moving at the speed of light (a relativistic effect); (3) there is less transversal defocusing due to the electric field[5]; (4) the type of radioactivation induced by electrons is less severe, so significantly higher beam losses can be tolerated.

To achieve the low emittance required for high-brilliance laser pulses, very high demands have to be put on the field stability. For the European XFEL [Altarelli, 2007] and LCLS-II [Doolittle, 2015] the requirements are 0.01 % and 0.01° rms, a factor 10–20 tighter than for ESS. However, the average beam current and the corresponding beam loading are significantly lower for FEL accelerators than for high-intensity proton accelerators, which implies that the disturbances they experience are much less severe.

Since the allowed field errors for FEL accelerators are of the same magnitude as the noise from the detection electronics, a great amount of care needs to be taken in the electronics design. Also the direct effect of measurement noise on the field errors should be carefully accounted for in the control design. Relative to the field error specifications for the ESS accelerator, this direct effect of measurement noise is negligible, although the impact of noise on the control signal activity still needs to be considered.

Another aspect is that many different cavity types and RF amplifier technologies are used for high-intensity proton accelerators which makes the control problem more heterogeneous; FEL accelerators on the other hand mostly use identical elliptical cavities (and a few special cavity types, the RF gun and 3rd harmonic cavities).

Then there are challenges that are unique to specific FEL accelerators: at the European XFEL each RF amplifier drives, not one, but *thirty-two* cavities [Brandt, 2007; Hoffmann, 2008; Schmidt, 2010; Pfeiffer, 2014], at the SwissFEL the pulses are too short for intra-pulse feedback, and the normal conducting accelerating structures require accurate temperature control [Rezaeizadeh, 2016], and at LCLS II the cavities have bandwidths of around

---

[5] From Maxwell's equations we have that the electric field $\boldsymbol{\mathcal{E}}$ in the cavity satisfies

$$0 = \nabla \cdot \boldsymbol{\mathcal{E}} = \frac{\partial \boldsymbol{\mathcal{E}}}{\partial x} + \frac{\partial \boldsymbol{\mathcal{E}}}{\partial y} + \frac{\partial \boldsymbol{\mathcal{E}}}{\partial z}.$$

To achieve longitudinal focusing it is necessary that the particle bunches experience a negative gradient in the z-direction, $\partial \boldsymbol{\mathcal{E}} / \partial z < 0$, which implies that at least one of the last two terms need to be positive, which corresponds to transversal defocusing.

10 Hz, which calls for field control by self-excited loops, and makes control of the resonance frequencies of the cavities a major concern [Doolittle, 2015].

4

Subsystem Models for the ESS Accelerator

In this chapter we discuss the subsystems of the ESS accelerator whose dynamics and characteristics are relevant for the field control performance.

4.1 Ion source

The ESS will have a pulsed Microwave Discharge Ion Source [Neri et al., 2014; Celona et al., 2016] that generates free protons (H^+) by ionizing hydrogen gas (H_2) to a plasma. From the plasma, which is confined by magnetic fields, a proton current of 62.5 mA is extracted by a 75 keV potential.

Plasma instabilities lead to variations of the extracted proton current, which subsequently act as load disturbances on the cavity field; it should be remembered that the impact of these disturbances on the field control performance is strongly dependent on the spectral properties of the beam-current variations.

The beam-current variations are arguably the most important characteristic of the ion source from a *field-control perspective*, but for overall accelerator performance it is typically more important with a low beam emittance.

4.2 RF distribution system

Waveguides

The propagation velocity of an electromagnetic wave in a waveguide depends on the frequency of the wave and the waveguide dimensions [Pozar]; the propagation velocity is to a small extent also affected by the temperature, the humidity and the pressure of the air in the waveguide. The waveguides at ESS will have dimensions of 58 cm × 29 cm (29 cm × 15 cm) for distribution at 352 MHz (704 MHz), and in both cases the propagation velocity is 68 %

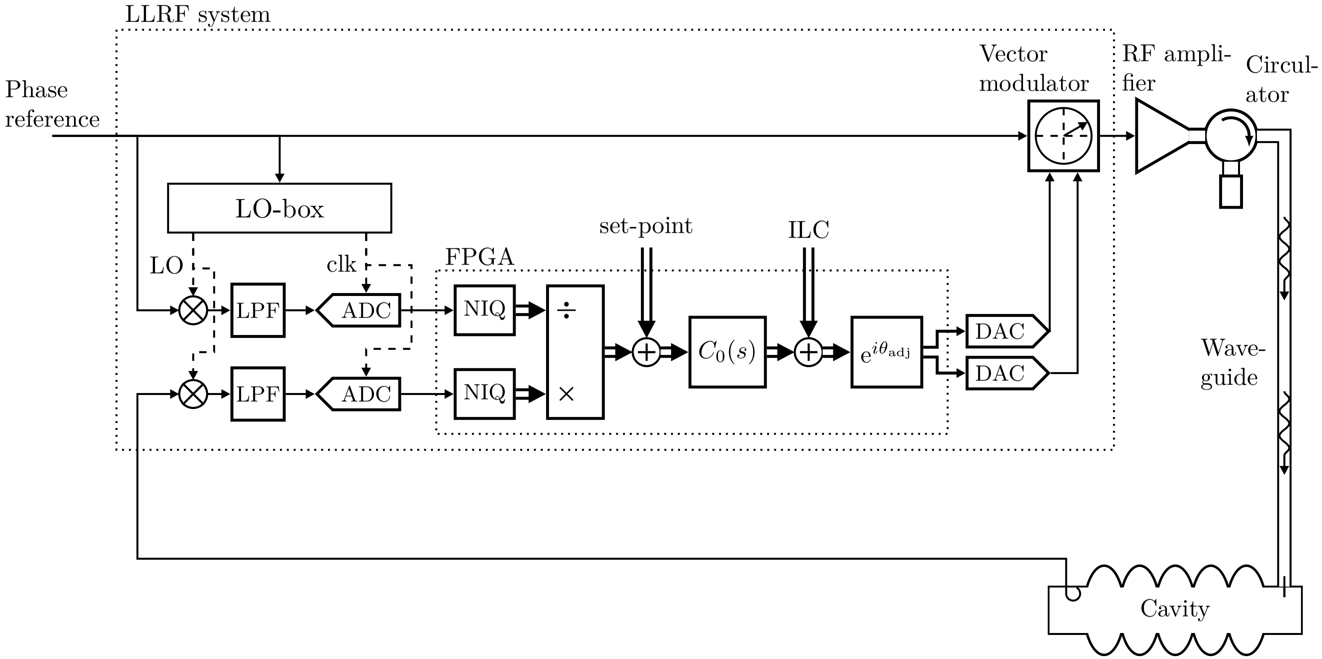


Figure 4.1 Overview of the field-control loop which illustrates: components for the complex-envelope detection (LO, mixers, LPF, NIQ; see Appendix A), the controller C_0 (Chapter 6), the phase adjustment factor $e^{i\theta_{adj}}$ (Section 6.5), the RF amplifier (Section 4.3), and the cavity (Chapter 5); double lines (⇒) indicate complex signals (implemented by two real signal in the FPGA).

the speed of light. The waveguides from the RF amplifiers to the cavities will be up to 40 m, which corresponds to a propagation time of 200 ns.

Circulators

We briefly discussed in Section 3.2 how the circulators are crucial for redirecting the backward waves from the cavities to RF loads. From a cavity field control perspective however, they are wideband enough to have negligible impact on the field control loop.

4.3 High-power RF amplifiers

There are no ideal amplifiers, and the imperfections that are most relevant for field control performance are:

- Limited bandwidth
- Nonlinearity from that the gain and phase shift of the amplifier depend on the input amplitude
- Gain and phase dependence on DC supply voltage (klystrons)
- Noise and spurious components (solid-state amplifiers)

Limited bandwidth

We will model the limited bandwidth of the RF amplifiers by a first-order low-pass filter as in [Schilcher, 1998],

$$P_{\mathrm{amp}}(s) = \frac{\omega_{\mathrm{amp}}}{s + \omega_{\mathrm{amp}}}, \tag{4.1}$$

where ω_{amp} is the 3dB bandwidth in radians per second.

RF amplifiers typically have bandwidths that are greater than 1 MHz, about ten times greater than the cross-over frequency of a typical field control loop for ESS; thus, the amplifier dynamics has a small impact on the field control loop, but it is prudent to account for it in field control analysis since it has a slight negative impact on performance.

An alternative way to model the RF amplifier dynamics is by a pure time delay equal to the group delay of the amplifier.

The truth probably lies somewhere in between; at least in the case of klystrons, where a more detailed model accounts for that it contains several high-bandwidth cavities, each with baseband dynamics of the form $1/(s\tau_i + 1)$, coupled via an electron beam; such systems can be accurately approximated by a first-order system in series with a time delay [Skogestad, 2003; Åström and Hägglund, 2006, (2.23)].

Nonlinearity

For reasons of sustainability and economy, the amplifiers at ESS will be operated close to saturation, with roughly 25 % power overhead, corresponding to 12 % overhead in amplitude, so saturation and nonlinear affects *will* have an impact. The harmonics due to the nonlinearity, which is a major concern for telecommunications due to adjacent channel interference, is of little concern for field control, due to the small bandwidth of the cavity. The nonlinearities could *in principle* be inverted[1] by the LLRF system[Omet, 2014], mitigated by an inner loop around the klystron [Baudrenghien et al., 2014], or simply ignored. We will not consider this issue further in this thesis.

Gain and phase dependence on DC supply voltage (klystrons)

The gain and phase shift of klystrons depend on both the DC supply voltage via so-called *pushing*[2]. Variations ΔV_c in the DC (cathode) voltage V_c that powers a klystron lead to gain variations [Gilmour, 2011]

$$\frac{\Delta G_{\mathrm{kly}}}{G_{\mathrm{kly}}} = \frac{5}{4}\frac{\Delta V_c}{V_c}.$$

and phase variations

$$\Delta\phi_{\mathrm{kly}} = \xi\frac{\Delta V_c}{V_c},$$

where $\Delta\phi_{\mathrm{kly}}$ is measured in radians, and ξ is the so-called *(phase) pushing factor*. For the klystrons that will be used at ESS, $\xi \approx$ 10–15°/% $\approx$ 17–25 rad, depending on the type and the operating point[3]. How the pushing factor ξ depends on the physical parameters of the klystron is discussed in [Hara et al., 1998].

If the ideal klystron output is given by $\mathbf{F_g}$, and the deviation from the nominal output due to cathode voltage variations is denoted by $\mathbf{F_g^d}$, we have

$$\begin{aligned}\mathbf{F_g} + \mathbf{F_g^d} &= \left(1 + \frac{\Delta G}{G}\right) \mathrm{e}^{i\xi\Delta\phi_{\mathrm{kly}}}\mathbf{F_g}\\ &= \left(1 + \frac{5}{4}\frac{\Delta V_c}{V_c}\right)\exp\left(i\xi\frac{\Delta V_c}{V_c}\right)\mathbf{F_g}. \qquad (4.2)\end{aligned}$$

Assuming that ΔV_c is small compared to V_c and that $\mathbf{F_g}$ is close to the nominal value $\mathbf{F_g^0}$, then we can approximate $\mathbf{F_g^d}$ by linearizing (4.2),

$$\mathbf{F_g^d} \approx \left(\frac{5}{4} + i\xi\right)\mathbf{F_g^0}\cdot\frac{\Delta V_c}{V_c}. \qquad (4.3)$$

---

[1] In telecommunications this is known as pre-distortion.

[2] Gain variations and phase shifts from a varying load impedance, so-called *pulling*, is negligible since the amplifier always sees a matched load because of the circulator.

[3] Personal communication with Morten Jensen, ESS.

Since $\xi \approx 20$, we see that the amplitude variations are small compared to the phase variations. See Figure 4.2 for an illustration of the klystron model.

It should be kept mind that the impact of variations ΔV_c (typically from switching in the AC-DC converter, a.k.a. modulator), on cavity field errors, is determined by the spectral properties of the variations.

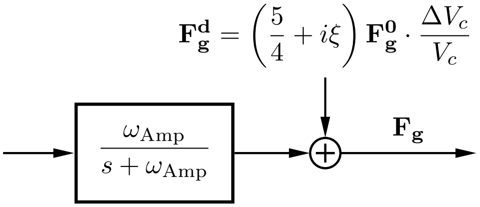


Figure 4.2 Klystron model which accounts for a limited bandwidth and the impact of cathode voltage variations.

Noise and spurious components (solid-state amplifiers)

The spectrum of the disturbance $\mathbf{F_g^d}$ added by solid-state amplifiers contains both *flicker* noise, with a $1/f$-dependence in the baseband, spurious peaks originating from combinations of switching and nonlinear phenomena, and to a minor extent also white, thermal noise. To evaluate the impact of these disturbances requires a spectral model of them. There are also slower effects such as droop and phase drifts.

4.4 Low-level RF system

The LLRF system for the ESS accelerator will be based on components that are compliant with the MicroTCA.4 standard [MTCA.4 2017], which was developed as a subspecification of the MicroTCA standard [Jamieson, 2006], to allow a flexible carrier-grade platform for big science projects.

A schematic of the LLRF system for ESS is shown in Figure 4.4, and typical components that have been considered are listed in Table 4.1 and shown in Figure 4.3. For an introduction to the RF components used in the LLRF system, such as vector modulators, down converters, etc, see [A. Gallo, 2010].

Figure 4.3 MicroTCA compliant for a typical LLRF system. *Left*: Digitizer SIS8300-KU from Struck. *Middle*: Signal conditioning board DWC8VM1 from Struck. *Right*: 12-slot MicroTCA4 crate from ELMA. Photo Credit: Struck Innovative Systeme GmbH, Elma Electronic Inc.

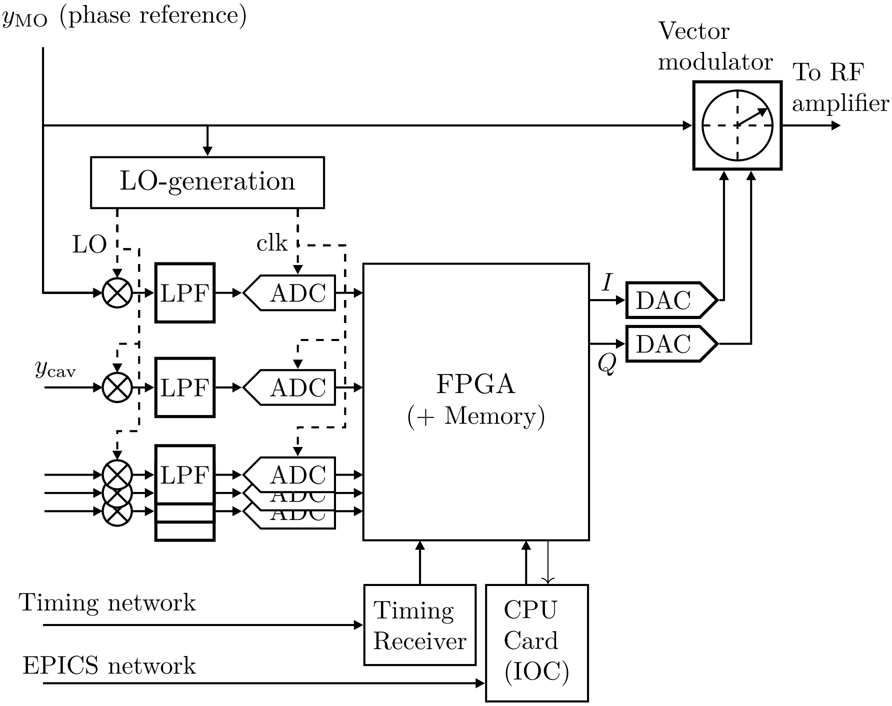


Figure 4.4 Overview of the LLRF system for ESS. Interfaces to the interlock system and cavity tuning system are not shown.

Table 4.1 Planned components for the ESS LLRF system.

Component	**Model/Comments**
12-slot chassis	TBD
MicroTCA Carrier Hub	NAT-MCH-PHYS
Power Supply Unit	Wiener MTCA.4 1000W PS
Digitizer (AMC*)	Struck SIS8300-KU: Kintex Ultrascale FPGA 10 ADCs, 16-bit, 125 MS/s 2 high speed DACs 2 GByte DDR4 Memory
Downconversion & vector modulator (RTM†)	Struck DWC8VM1(-LF): 8 down-conversion channels 1 up-conversion channel (vector modulator)
Timing receiver (AMC*)	Module from Micro Research Finland
CPU (AMC*)	TBD
LO-generation 352.21 MHz (stand-alone unit)	In-kind contribution form ESS Bilbo
LO-generation 704.42 MHz (RTM†)	In-kind contribution from the Polish Electronic Group‡

[*] Advanced Mezzanine Card, inserted from the front of the chassis.

[†] Rear Transition Module, inserted from the back of the chassis.

[‡] National Centre for Nuclear Research, TU Warsaw, and TU Łódź.

5

Modeling of Accelerator Cavities

5.1 Introduction

For analysis and control design for the field control loop it is essential to have a baseband model of the electromagnetic cavity fields. In existing particle accelerator literature, such models are derived by considering an equivalent electric circuit [Schilcher, 1998; Tückmantel, 2011][1].

While the equivalent circuit approach may seem natural to the electrical engineer, it arguably obfuscates the connection between the derived model and the actual cavity field dynamics. In this chapter we will use an alternative approach from the optoelectronics literature [Haus, 1983; Suh et al., 2004], that allows the cavity dynamics to be derived directly from Maxwell's equations using fundamental laws of physics. In his book, Haus has the following to say about using equivalent circuits for modeling cavity dynamics:

> .../Aside from the fact that such a procedure would be more cumbersome, the derivation would be model dependent and thus would lack generality/...

The only difference between the cavity model that we derive in this chapter, and those that result from consideration of an equivalent circuit model, is that the model parameters are chosen somewhat differently; in particular, the mode amplitudes are quantified by the square root of the mode energies, rather than by the effective accelerating voltage of the modes. Arguably, the parameterization proposed in this chapter simplifies the understanding of how the cavity parameters affect the achievable field control performance.

---

[1] Often the emphasis is on steady-state relations [Padamsee et al., 2008; Wangler, 2008], which is not enough to enable control design. Section 5.7 of [Wangler, 2008] contains an interesting approach to gain intuition for the dynamics of a waveguide-coupled cavity. See also the classic references [Montgomery et al., 1948; Slater, 1950; Pedersen, 1975].

Outline

In Section 5.2 we derive the dynamics of a waveguide-coupled cavity using an approach inspired by [Haus, 1983]. For an accelerator cavity it is necessary to include the effects of beam-interaction, so in Section 5.3 we introduce a suitable parameter to describe this interaction. After that, in Section 5.4, we put together the differential equation for a single cavity mode and transform the dynamics to the baseband.

From the baseband model of the accelerating cavity mode we find the optimal coupling and detuning of the cavity in Section 5.5, as well as investigate the relative size of the terms that affect the cavity field; both these aspects are then illustrated by a somewhat novel type of phasor diagram in Section 5.6; then in Section 5.7 we normalize the cavity equation to a form suitable for control design.

In Section 5.8 we consider multiple cavity modes (the accelerating mode plus parasitic modes), and illustrate the system dynamics with a block digram.

Finally in Section 5.9 we compare the modeling approach of this chapter to the equivalent circuit approach.

5.2 Dynamics of waveguide-coupled cavity

In this section we present a derivation for the dynamics of a waveguide-coupled cavity that is based on [Haus, 1983, Section 7.2]. Attempts have been made to make the presentation self-contained and aligned with the rest of this thesis. The reader is advised to seek out the original reference for more details.

Outline of the derivation

We start by assuming that the cavity is closed and lossless, which allows us to describe the electric cavity field as a superposition of orthogonal eigenmodes: $\boldsymbol{\mathcal{E}}(\boldsymbol{r},t) = \sum e_n(t)\boldsymbol{E}_n(\boldsymbol{r})$. We then introduce cavity-waveguide coupling, assuming that the coupling is sufficiently weak to allow the coupling between the modes $\boldsymbol{E}_n$ to be neglected. Starting from that the electric field in the waveguide can be considered as a superposition of a forward wave and a backward wave, we argue, as in [Haus, 1983], that due to linearity of Maxwell's equations the cavity modes couple linearly to the backward and forward waves. By using the time reversibility of Maxwell's equations and conservation of energy, we obtain a relation between the constant for how the forward wave couples to the cavity field and the constant for how the cavity field couples to the backward wave. Then in Section 5.4, resistive losses are introduced, again under the assumption that they are sufficiently small to allow mode coupling to be neglected; in connection with this step also the beam-interaction is included.

Mode expansion of the cavity field

From Maxwell's equations it follows that the electric field $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}(\boldsymbol{r}, t)$ inside a vacuum-filled cavity satisfies

$$\nabla^2 \boldsymbol{\mathcal{E}} - \epsilon_0 \mu_0 \frac{\partial^2}{\partial t^2} \boldsymbol{\mathcal{E}} = -\mu_0 \frac{d}{dt} \boldsymbol{J}. \tag{5.1}$$

If the cavity has perfectly electrically conducting walls, the electric field can be expanded as a sum of orthogonal eigenmodes

$$\boldsymbol{\mathcal{E}}(\boldsymbol{r}, t) = \sum_{n=a,1,2,\ldots}^{\infty} e_n(t) \boldsymbol{E}_n(\boldsymbol{r}) \tag{5.2}$$

where each $\boldsymbol{E}_n$ satisfies the eigenrelation

$$\nabla^2 \boldsymbol{E}_n = -k_n^2 \boldsymbol{E}_n$$

as well as the boundary conditions. The symbol a in the sum in (5.2) corresponds to the mode used for particle acceleration, while 1, 2, . . . correspond to parasitic modes with increasing frequency offset from the accelerating mode. The eigenfunctions are assumed to be normalized so that

$$\frac{\epsilon_0}{2} \iiint_V |\boldsymbol{E}_n|^2 \, dV = 1. \tag{5.3}$$

This normalization makes the oscillation amplitude of $|e_n(t)|^2$ equal the mode energy, which we will denote by U_n. From inserting the eigenmode expansion (5.2) into (5.1) it follows that

$$\frac{d^2 e_n(t)}{dt^2} + \omega_n^2 e_n(t) = -\frac{1}{2} \frac{\partial}{\partial t} \iiint_V \boldsymbol{J}(\boldsymbol{r}, t) \cdot \boldsymbol{E}_n(\boldsymbol{r}) \, dV, \tag{5.4}$$

where $\omega_n = c k_n$ is the resonance frequency of mode n.

For increased clarity, we will up until Section 5.8 only consider the accelerating mode, so we will drop the mode indexes except when there is a risk of ambiguity, e.g., for $\boldsymbol{E}_a$ and ω_a. The discussion in the next two sections applies to any cavity mode, while sections 5.5 to 5.7 are only relevant for the accelerating mode.

Complex-valued representation of mode dynamics

We recognize the differential equation (5.4) for the mode dynamics as an undamped second-order system; and as we will see, it will be driven by a signal of frequency $\omega_{\mathrm{RF}} \approx \omega_a$ (with slowly varying amplitude and phase). Thus also the mode amplitude will have a similar sinusoidal dependence, and we can write $e(t) = A(t)\sin(\omega_{\mathrm{RF}} t + \phi(t))$, where $A(t)$ and $\phi(t)$ vary slowly; for this reason it will be convenient to use a complex-valued representation of the differential equation (5.4). To this end, we let $\xi_1 = e_a$, $\xi_2 = \dot{e}_a/\omega_a$, and $\Gamma_a(t) = \frac{1}{\omega_a}\frac{\partial}{\partial t}\iiint_V \boldsymbol{J}(\boldsymbol{r},t)\cdot\boldsymbol{E}_a(\boldsymbol{r})\,dV$, after which (5.4) can be written as two coupled, first-order differential equations

$$\dot{\xi}_2 = -\omega_a\xi_1 - \frac{1}{2}\Gamma_a(t) \tag{5.5a}$$

$$\dot{\xi}_1 = \omega_a\xi_2. \tag{5.5b}$$

By introducing the complex variable

$$\mathbf{a} = e_a - i\frac{1}{\omega_a}\dot{e}_a = \xi_1 - i\xi_2, \qquad \left[\sqrt{\mathrm{J}}\right] \tag{5.6}$$

where $\left[\ \cdot\ \right]$ indicates the units of the equation, we can write (5.5) as

$$\frac{d}{dt}\mathbf{a} = i\omega_a\mathbf{a} + \frac{i}{2}\Gamma_a(t), \tag{5.7}$$

where the real part of the equation corresponds to (5.5b) and the imaginary part to (5.5a). The beam-interaction term $\Gamma(t)$ on the right hand side is evaluated in Appendix D.1, but up until Section 5.4 we will assume it to be zero.

Electric fields in the waveguide

To transport energy between the RF amplifier and the cavity, they are both coupled to a waveguide. For most accelerator applications, the waveguide dimensions are chosen so that there is only one mode whose cut-off frequency is below the RF frequency, and hence there is only one *propagating mode* that is able to transport energy from the RF amplifier to the cavity.

From Maxwell's equations it follows that the amplitude of the electromagnetic field in the waveguide along the propagation direction (z-direction), satisfies the electromagnetic wave equation (equation 5.1 with $\boldsymbol{J} \equiv 0$, and assuming that the relative permittivity $\epsilon_r = 1$)

$$\nabla^2\boldsymbol{\mathcal{E}} - \epsilon_0\mu_0\frac{\partial^2}{\partial t^2}\boldsymbol{\mathcal{E}} = 0. \tag{5.8}$$

From (5.8) it can be shown that the electric field in a waveguide with uniform cross-section is a superposition of one *forward* wave traveling from the RF amplifier to the cavity, and one *backward* wave traveling away from the cavity. Typically a so-called circulator is inserted after the RF amplifier, which lets

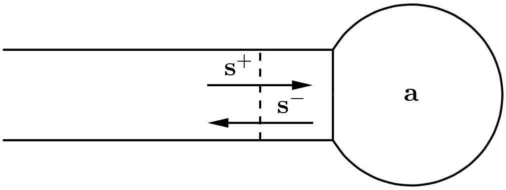


Figure 5.1 Illustration of a cavity coupled to a waveguide.

the forward wave pass through unaffected, but re-directs the backward wave to an RF load where it is absorbed. This is done to protect the RF amplifier from the backward wave. See Figure 5.1 for an illustration.

We describe the amplitude and phase of the forward and backward waves at suitably chosen reference planes in the waveguide, close to the cavity, by the complex signals $\mathbf{s}^+(t)$ and $\mathbf{s}^-(t)$. We assume that the $\mathbf{s}^+$, and $\mathbf{s}^-$ are defined, in units of $\sqrt{\mathrm{J/s}}$, so that

$$\begin{aligned}\left|\mathbf{s}^+\right|^2 &= \text{power carried by the forward wave,}\\ \left|\mathbf{s}^-\right|^2 &= \text{power carried by the backward wave.}\end{aligned}$$

Throughout this section we assume that the waveguide is lossless in order to derive a relationship between the constants for the cavity-waveguide coupling. After that we could introduce waveguide losses if we would like to, but from a field control perspective they would just correspond to a constant factor which would disappear after normalization.

Waveguide coupling

Due to the waveguide coupling, the equation (5.4) for the lossless cavity needs to be modified to take into account:

1. The energy of the cavity field will decay through the coupling mechanism, and travel away from the cavity in the backward wave.
2. A (non-zero) forward wave in the waveguide will excite the cavity field.

From Maxwell's equations it follows that both these effects are linear, and since the coupling could be both electric and/or magnetic, both the decay and the excitation from the forward wave could directly affect both $\dot{e}$ and $\ddot{e}$.

By denoting the rate at which the mode energy decays via the waveguide coupling mechanism by γ_{ext}, and by quantifying the coupling of the forward wave $\mathbf{s}^+$ to the cavity field by κ, which in principle could be complex, we get

$$\frac{d}{dt}\mathbf{a} = i\omega_a\mathbf{a} - \gamma_{\text{ext}}\mathbf{a} + \kappa\mathbf{s}^+. \qquad \left[\frac{\sqrt{\mathrm{J}}}{\mathrm{s}}\right] \tag{5.9}$$

Time-reversibility of Maxwell's equations

Time-reversibility means that if $\boldsymbol{\mathcal{E}}(\boldsymbol{r},t)$ is a solution to the waveguide-coupled cavity system, and hence satisfies the electromagnetic wave equation (5.8) both in the cavity and the waveguide, then so does the time-reversed solution $\widetilde{\boldsymbol{\mathcal{E}}}(\boldsymbol{r},t) = \boldsymbol{\mathcal{E}}(\boldsymbol{r},-t)$. This is immediate from (5.8).

We denote the complex amplitude of the forward wave of the time-reversed solution by $\tilde{\mathbf{s}}^+$, and it is clear that $|\tilde{\mathbf{s}}^+(-t)| = |\mathbf{s}^-(t)|$.

For the eigenmode expansion of the time-reversed solution inside the cavity we have that time-varying coefficients in (5.2) are given by $\tilde{e}_n(t) = e_n(-t)$. By introducing the time-reversed, complex-valued mode amplitude

$$\tilde{\mathbf{a}} = \tilde{e}_a - i\frac{1}{\omega_a}\dot{\tilde{e}}_a$$

analogously to (5.6), it follows that it also satisfies the equation (5.9),

$$\frac{d}{dt}\tilde{\mathbf{a}} = i\omega_a\tilde{\mathbf{a}} - \gamma_{\text{ext}}\tilde{\mathbf{a}} + \kappa\tilde{\mathbf{s}}^+,$$

since $\tilde{\mathbf{a}}$, just like $\mathbf{a}$, corresponds to a valid solution to Maxwell's equations inside the cavity.

Relation between κ and γ_{ext}

It seems reasonable that the decay rate of the cavity field via the coupling mechanism (quantified by γ_{ext}) and how easily the cavity is excited by the forward wave (quantified by κ), are related; we now, based on the same ideas as in [Haus, 1983], derive such a relationship from fundamental laws of physics.

Consider the time evolution of $\mathbf{a}$ for $t \geq 0$, from the initial state $\mathbf{a}(0) = 1$, when there is no forward wave exciting the cavity, i.e, $\mathbf{s}^+(t) = 0$, $t \geq 0$. We denote this solution $\mathbf{a_1}(t)$, and thus have

$$\begin{aligned} &\frac{d}{dt}\mathbf{a_1} = i\omega_a\mathbf{a_1} - \gamma_{\text{ext}}\mathbf{a_1}, \quad t \geq 0 \\ &\mathbf{a_1}(0) = 1. \end{aligned} \tag{5.10}$$

Since there is no forward wave exciting the cavity, the mode energy U will decay through the coupling mechanism at rate

$$\frac{dU}{dt} = \frac{d}{dt}\left|\mathbf{a_1}\right|^2 = \mathbf{a_1}^*\frac{d\mathbf{a_1}}{dt} + \frac{d\mathbf{a_1}^*}{dt}\mathbf{a_1} = -2\gamma_{\text{ext}}\left|\mathbf{a_1}\right|^2 \geq 0.$$

Conservation of energy gives that the energy that escapes the cavity is carried away by the backward wave, i.e.,

$$\left|\mathbf{s_1^-}\right|^2 = 2\gamma_{\text{ext}}\left|\mathbf{a_1}\right|^2, \qquad t \geq 0. \tag{5.11}$$

Now, consider the time-reversed solution $\tilde{\mathbf{a}}_{\mathbf{1}}$ of $\mathbf{a}_{\mathbf{1}}$ which is defined for $t < 0$, and satisfies $|\tilde{\mathbf{a}}_{\mathbf{1}}(-t)| = |\mathbf{a}_{\mathbf{1}}(t)|$. As discussed in the previous section, it satisfies

$$\frac{d}{dt}\tilde{\mathbf{a}}_{\mathbf{1}} = i\omega_a\tilde{\mathbf{a}}_{\mathbf{1}} - \gamma_{\text{ext}}\tilde{\mathbf{a}}_{\mathbf{1}} + \kappa\tilde{\mathbf{s}}_{\mathbf{1}}^{+}, \qquad t \leq 0. \tag{5.12}$$

From (5.10) we see that $\mathbf{a}_{\mathbf{1}}(t) = \exp(i\omega_a t)\exp(-\gamma_{\text{ext}}t)$, and since the backward wave depends linearly on the mode amplitude, we also have that $\mathbf{s}_{\mathbf{1}}^{-} \propto \exp(i\omega_a t)\exp(-\gamma_{\text{ext}}t)$. From time reversal it then follows that we have

$$\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(t) = \tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0)\exp(i\omega_a t)\exp(\gamma_{\text{ext}}t), \qquad t \leq 0. \tag{5.13}$$

By substituting (5.13) into (5.12), we can compute $\tilde{\mathbf{a}}_{\mathbf{1}}(t)$ as a function of $\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(t)$,

$$\tilde{\mathbf{a}}_{\mathbf{1}}(t) = \int_{-\infty}^{t} \kappa \mathrm{e}^{i\omega_a(t-\tau)-\gamma_{\text{ext}}(t-\tau)}\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0)\mathrm{e}^{i\omega_a\tau+\gamma_{\text{ext}}\tau}\,d\tau, \quad t \leq 0.$$

For $t = 0$, we get

$$\tilde{\mathbf{a}}_{\mathbf{1}}(0) = \kappa\int_{-\infty}^{0} \mathrm{e}^{2\gamma_{\text{ext}}\tau}\,d\tau\ \tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0) = \kappa\frac{\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0)}{2\gamma_{\text{ext}}},$$

from which it follows that

$$|\tilde{\mathbf{a}}_{\mathbf{1}}(0)| = \frac{|\kappa|}{2\gamma_{\text{ext}}}\left|\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0)\right|. \tag{5.14}$$

We also have that

$$\left|\tilde{\mathbf{s}}_{\mathbf{1}}^{+}(0)\right|^2 = \left|\mathbf{s}_{\mathbf{1}}^{-}(0)\right|^2 = 2\gamma_{\text{ext}}\left|\mathbf{a}_{\mathbf{1}}(0)\right|^2 = 2\gamma_{\text{ext}}\left|\tilde{\mathbf{a}}_{\mathbf{1}}(0)\right|^2, \tag{5.15}$$

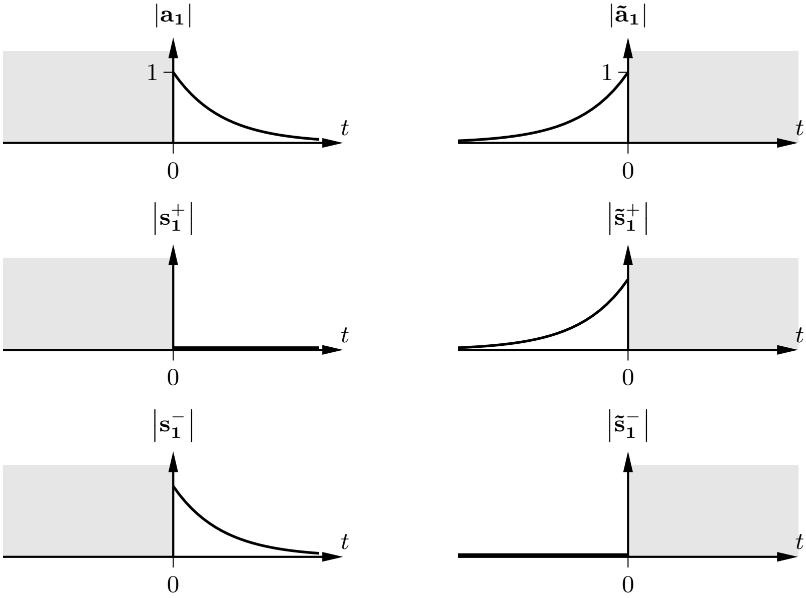
where the first and last equalities follow from equality of the original and the time-reversed solution, and the middlemost equality follows from (5.11). Comparing (5.14) and (5.15) we see that

$$|\kappa| = \sqrt{2\gamma_{\text{ext}}},$$

and by choosing the reference plane, relative to which $\mathbf{s}^{+}$ is defined, we can assume that κ is positive and real, so that (5.9) becomes

$$\frac{d}{dt}\mathbf{a} = i\omega_a\mathbf{a} - \gamma_{\text{ext}}\mathbf{a} + \sqrt{2\gamma_{\text{ext}}}\mathbf{s}^{+}. \tag{5.16}$$

The magnitudes of the original and the time-reversed cavity field amplitudes and the forward/backward waves are illustrated in Figure 5.2.



$$\frac{d}{dt}\mathbf{a_1} = i\omega_a\mathbf{a_1} - \gamma_{\mathrm{ext}}\mathbf{a_1} + \kappa\mathbf{s_1^+} \qquad \frac{d}{dt}\tilde{\mathbf{a}}_\mathbf{1} = i\omega_a\tilde{\mathbf{a}}_\mathbf{1} - \gamma_{\mathrm{ext}}\tilde{\mathbf{a}}_\mathbf{1} + \kappa\tilde{\mathbf{s}}_\mathbf{1}^+$$

Figure 5.2 Left: Time evolution of the cavity equation corresponding to (5.10). Right: The corresponding time-reversed solution.

Backward wave

From the linearity of Maxwell's equations we know that the backward wave depends linearly on the forward wave and the mode amplitude:

$$\mathbf{s}^- = C_s\mathbf{s}^+ + C_a\mathbf{a},$$

where C_s and C_a are complex constants. We already know from (5.11) that $|C_a| = \sqrt{2\gamma_{\mathrm{ext}}}$, and by choosing the reference plane that $\mathbf{s}^-$ is defined with respect to, we can make C_a real and positive, giving that $C_a = \sqrt{2\gamma_{\mathrm{ext}}}$.

Next we derive an expression for C_s. Conservation of energy gives that

$$\left|\mathbf{s}^+\right|^2 - \left|\mathbf{s}^-\right|^2 = \frac{d}{dt}\left|\mathbf{a}\right|^2, \tag{5.17}$$

and from (5.16) it follows that

$$\frac{d}{dt}\left|\mathbf{a}\right|^2 = -2\gamma_{\mathrm{ext}}\left|\mathbf{a}\right|^2 + \sqrt{2\gamma_{\mathrm{ext}}}\Big(\mathbf{a}^*\mathbf{s}^+ + \mathbf{a}(\mathbf{s}^+)^*\Big). \tag{5.18}$$

Adding (5.17) and (5.18) gives

$$\left|\mathbf{s}^+\right|^2 - \left|\mathbf{s}^-\right|^2 = -2\gamma_{\text{ext}}\left|\mathbf{a}\right|^2 + \sqrt{2\gamma_{\text{ext}}}\Big(\mathbf{a}^*\mathbf{s}^+ + \mathbf{a}(\mathbf{s}^+)^*\Big), \tag{5.19}$$

and by making the substitution $\mathbf{a} = (\mathbf{s}^- - C_s\mathbf{s}^+)/\sqrt{2\gamma_{\text{ext}}}$ in (5.19) we obtain

$$\begin{aligned}\left|\mathbf{s}^+\right|^2 - \left|\mathbf{s}^-\right|^2 = &-\Big(\left|\mathbf{s}^-\right|^2 - C_s(\mathbf{s}^-)^*\mathbf{s}^+ - C_s^*\mathbf{s}^-(\mathbf{s}^+)^* + \left|C_s\right|^2\left|\mathbf{s}^+\right|^2\Big)\\ &+\Big((\mathbf{s}^-)^*\mathbf{s}^+ - C_s^*\left|\mathbf{s}^+\right|^2 + \mathbf{s}^-(\mathbf{s}^+)^* - C_s\left|\mathbf{s}^+\right|^2\Big),\end{aligned}$$

which can be re-written as

$$\left|(1 + C_s)\mathbf{s}^+ - \mathbf{s}^-\right|^2 = \left|\mathbf{s}^-\right|^2.$$

For this equality to hold for all $\mathbf{s}^+$ and $\mathbf{s}^-$, we must have that

$$C_s = -1.$$

Thus the backward wave is given, as in [Haus, 1983, (7.36)] by

$$\mathbf{s}^- = -\mathbf{s}^+ + \sqrt{2\gamma_{\text{ext}}}\mathbf{a}.$$

5.3 Parameters for cavity field-beam interaction

In this section we introduce two parameters which are important for quantifying the cavity-beam interaction: the effective accelerating voltage, and the ratio of the effective accelerating voltage to the square root of electromagnetic energy stored in the cavity mode[2].

Effective accelerating voltage

Denote the electric field of the *accelerating* cavity mode (we only consider one mode so that we can drop the mode indexes), in the direction of the accelerated particles (the z-direction), by $\mathcal{E}_z(z,t)$. Then the energy transferred to a particle with charge q as it crosses the cavity along a trajectory $z(t)$ is given by

$$\Delta W = q\int_0^L \mathcal{E}_z(z, t(z))\, dz, \tag{5.20}$$

Under the assumption that the mode energy U_a is constant as the particle crosses the cavity, we have from the normalization (5.3) that

$$\mathcal{E}_z(z,t) = e_a(t)\sqrt{U_a}E_z(z) = \cos(\omega_{\text{RF}}t)\sqrt{U_n}E_z(z), \tag{5.21}$$

---

[2] We refrain from introducing the transit-time factor and the axial RF voltage since they are not needed for this presentation.

where $E_z(z)$ is the z-component of $\boldsymbol{E}_a(\boldsymbol{r})$ along the axis of the accelerated particles, and where the phase of $e_a(t)$ was taken to equal 0, since we will assume that the accelerating mode will nominally have zero phase.

We will also assume that the accelerated particles travel through the cavity with constant velocity v, in which case we can write

$$z = vt - \frac{v}{\omega_{\mathrm{RF}}}\phi, \tag{5.22}$$

where ϕ is a phase offset between the particle bunches and the electric field. Strictly speaking, it is phase of the electric field which is chosen based on the phase of the particle bunches, but we will consider it to be the other way around, since this will simplify the exposition; the two viewpoints are equivalent.

Under the assumption that the mode energy and particle velocity are constant, we can use (5.21) and (5.22) to re-write (5.20), as

$$\Delta W = q\int_0^L \cos(\omega_{\mathrm{RF}}\frac{z}{v} + \phi)\sqrt{U_a}E_z(z)\,dz. \tag{5.23}$$

As shown in [Wangler, 2008, Exercise 2.12], there exists a value $\phi = \phi_0$ for which the energy transferred to the accelerated particle is maximized. Using the value ϕ_0 we define the effective accelerating voltage

$$V_{\mathrm{eff}} := \int_0^L \cos(\omega_{\mathrm{RF}}\frac{z}{v} + \phi_0)\sqrt{U_a}E_z(z)\,dz. \tag{5.24}$$

Thus, the most energy that the cavity can transfer to a particle with charge q is given by qV_{eff}.

For linear proton accelerators it is desired that the particle bunches arrive with a phase offset ϕ_b before what corresponds to maximal acceleration. This is to achieve longitudinal focusing that keeps the bunches together. With a phase offset ϕ_b from ϕ_0 it can be shown [Wangler, 2008] that the energy transferred to the bunches is given by $\Delta W = qV_{\mathrm{eff}}\cos\phi_b$.

Relation between effective voltage and stored energy

The effective accelerating voltage (5.24) depends on the mode amplitude, so it is convenient to introduce

$$\alpha := \frac{V_{\mathrm{eff}}}{\sqrt{U_a}} = \int_0^L \cos(\omega_{\mathrm{RF}}\frac{z}{v} + \phi_0)E_z(z)\,dz, \qquad \left[\frac{\mathrm{V}}{\sqrt{\mathrm{J}}}\right] \tag{5.25}$$

which only depends on the geometry of the cavity.

A more common parameter for relating the effective accelerating voltage and the energy stored in a cavity mode is [Wangler, 2008]

$$\frac{r}{Q} := \frac{V_{\mathrm{eff}}^2}{\omega_{\mathrm{RF}}U_n},$$

We see that α and r/Q are related via $\alpha^2 = \omega_{\mathrm{RF}}(r/Q)$.

5.4 Baseband dynamics of the accelerating cavity mode

Starting from (5.16), and including the term $\Gamma(t)$ from equation (5.7), and adding a term $-\gamma_0\mathbf{a}$ for resistive losses in the cavity walls, gives

$$\frac{d\mathbf{a}}{dt} = i\omega_a\mathbf{a} - (\gamma_0 + \gamma_{\mathrm{ext}})\mathbf{a} + \sqrt{2\gamma_{\mathrm{ext}}}\mathbf{s}^+ + \frac{i}{2}\Gamma_a(t). \tag{5.26}$$

In Appendix D.1 it is shown that

$$\Gamma_a(t) \approx -i\alpha\mathbf{I_b}(t)\mathrm{e}^{i\omega_{\mathrm{RF}}t} \tag{5.27}$$

where α, defined in (5.25), quantify the coupling between the cavity field and the beam, and the definition of $\mathbf{I_b}$ in (D.5) is given by

$$\mathbf{I_b}(t) = I_{\mathrm{DC}}(t)\mathrm{e}^{i(\pi-\phi_b(t))}.$$

By letting

$$\mathbf{a}(t) = \mathbf{A}(t)\mathrm{e}^{i\omega_{\mathrm{RF}}t} \qquad \left[\sqrt{\mathrm{J}}\right] \tag{5.28a}$$

$$\mathbf{s}^+(t) = \mathbf{F_g}(t)\mathrm{e}^{i\omega_{\mathrm{RF}}t} \qquad \left[\sqrt{\mathrm{J}}/\mathrm{s}\right] \tag{5.28b}$$

and substituting these expressions and (5.27) into (5.26), and dividing by $\mathrm{e}^{i\omega_{\mathrm{RF}}t}$, we obtain

$$i\omega_{\mathrm{RF}}\mathbf{A} + \frac{d\mathbf{A}}{dt} = i\omega_a\mathbf{A} - (\gamma_0 + \gamma_{\mathrm{ext}})\mathbf{A} + \sqrt{2\gamma_{\mathrm{ext}}}\mathbf{F_g} + \frac{\alpha}{2}\mathbf{I_b}. \tag{5.29}$$

What we have done is essentially a baseband transformation in the time domain. By defining

$$\Delta\omega := \omega_a - \omega_{\mathrm{RF}},$$

we can rewrite (5.29) as,

$$\frac{d\mathbf{A}}{dt} = i\Delta\omega\mathbf{A} - (\gamma_0 + \gamma_{\mathrm{ext}})\mathbf{A} + \sqrt{2\gamma_{\mathrm{ext}}}\mathbf{F_g} + \frac{\alpha}{2}\mathbf{I_b}. \qquad .$$

By introducing the combined decay rate of the cavity field from both waveguide coupling and resistive losses in the cavity walls,

$$\gamma = \gamma_0 + \gamma_{\mathrm{ext}}$$

we finally get

$$\boxed{\frac{d\mathbf{A}}{dt} = (-\gamma + i\Delta\omega)\mathbf{A} + \sqrt{2\gamma_{\mathrm{ext}}}\mathbf{F_g} + \frac{\alpha}{2}\mathbf{I_b}.} \qquad \left[\frac{\sqrt{\mathrm{J}}}{\mathrm{s}}\right] \tag{5.30}$$

When the cavity dynamics are derived from an equivalent circuit model, the field is usually represented by a phasor for the effective voltage. We arrive at such an equation by multiplying with α (recall that $\mathbf{V} = \alpha\mathbf{A}$),

$$\frac{d\mathbf{V}}{dt} = (-\gamma + i\Delta\omega)\mathbf{V} + \alpha\sqrt{2\gamma_{\mathrm{ext}}}\mathbf{F_g} + \frac{\alpha^2}{2}\mathbf{I_b}(t). \quad \left[\frac{\mathrm{V}}{\mathrm{s}}\right] \tag{5.31}$$

Note that (5.30) and (5.31) only differ by a constant factor, so after normalization there will not be any difference

Comments on the choice of parameters

The parameters and variables in the resulting differential equation (5.30) are somewhat different from when the derivation is based on an equivalent circuit approach; below we provide some justification for the choices made, and also discuss them further in Section 5.9.

- The reason for denoting the bandwidth/decay rate of the cavity by γ, rather than the more common notation $\omega_{1/2}$, which evokes thoughts of a bandwidth, is that the discussion in Section 5.9 of how cavity parameter variations impact the field control performance becomes more intuitive when γ is thought of as a decay rate. We will however also freely call γ a bandwidth, when considering the accelerating mode as a first-order system.
- In (5.30) the generator drive is modeled by a power flow (actually $\sqrt{\text{power}}$), so it would not make sense to represent it with the standard symbol for the general drive, $\mathbf{I_g}$, which would be assumed to have units of Ampere. Instead the letter **F** as in forward wave, field and forcing term, was chosen, while keeping the subscript **g** for generator.
- It should also be noted that the magnitude of $\mathbf{I_b}$ is the DC-current; there is no need to introduce an "equivalent RF current".

5.5 Optimal coupling and detuning

In this section we compute the optimal detuning and coupling for minimization of the generator power at steady state. How to do this is well known [Schilcher, 1998; Wangler, 2008], but we go through the computations to show what they look like with the parametrization of (5.30), and because we need the results in the next section.

From the cavity dynamics (5.30) of the accelerating mode, in steady state, at the nominal operating point $(\mathbf{A}, \mathbf{I_b}, \mathbf{F_g}) = (\mathbf{A}_0, \mathbf{I_b^0}, \mathbf{F_g^0})$, it follows that

$$0 = (-\gamma + i\Delta\omega)\mathbf{A}_0 + \sqrt{2\gamma_{\mathrm{ext}}}\mathbf{F_g^0} + \frac{\alpha}{2}\mathbf{I_b^0}. \tag{5.32}$$

Assume that the nominal phasor for the accelerator voltage is real and positive, and given by αa_0, so that $\mathbf{A}_0 = a_0$, and also assume that the nominal beam phasor is given by

$$\mathbf{I_b^0} = I_b \mathrm{e}^{i(\pi-\phi_b)} = -I_b\cos(\phi_b) + iI_b\sin(\phi_b),$$

where[3] $-\pi/2 \le \phi_b \le 0$. The generator power at steady-state then equals

$$\begin{aligned}P_g^0 &= \left|\mathbf{F_g^0}\right|^2 \\ &= \frac{1}{2\gamma_{\text{ext}}}\left|(-\gamma + i\Delta\omega)a_0 + \frac{\alpha}{2}\mathbf{I_b^0}\right|^2 \\ &= \frac{1}{2\gamma_{\text{ext}}}\left[\left((\gamma_0+\gamma_{\text{ext}})a_0 + \frac{\alpha}{2}I_b\cos\phi_b\right)^2 + \left(\Delta\omega a_0 + \frac{\alpha}{2}I_b\sin\phi_b\right)^2\right],\end{aligned}$$

and we see that the detuning that minimizes the generator power is given by

$$\Delta\omega^{(\text{opt})} = -\frac{\frac{\alpha}{2}I_b\sin\phi_b}{a_0}. \tag{5.33}$$

With this optimal detuning, the generator power becomes,

$$P_g\Big|_{\Delta\omega=\Delta\omega^{(\text{opt})}} = \frac{1}{2\gamma_{\text{ext}}}\left((\gamma_0+\gamma_{\text{ext}})a_0 + \frac{\alpha}{2}I_b\cos\phi_b\right)^2,$$

and it is clear that power-optimal coupling corresponds to[4]

$$\gamma_{\text{ext}}^{(\text{opt})} = \gamma_0 + \frac{\frac{\alpha}{2}I_b\cos\phi_b}{a_0}. \tag{5.34}$$

Thus, for an optimally tuned and coupled cavity the required generator power equals

$$P_g = 2\gamma_0 a_0^2 + \alpha I_b\cos(\phi_b)\cdot a_0,$$

that is, all energy in the forward wave is either dissipated in the cavity walls or transferred to the particle beam, and no power is wasted in the backward wave.

Note that the second quantity in (5.34),

$$\frac{\frac{\alpha}{2}I_b\cos\phi_b}{a_0} =: \text{“}\gamma_{\text{beam}}(a_0, \mathbf{I_b})\text{”}$$

---

[3] This condition is necessary in order for the beam to be both accelerated and longitudinally focused [Wangler, 2008, p. 179].

[4] We wish to minimize $2f(x) = (ux+v)^2/x = u^2x+2uv+v^2/x$ wrt $x > 0$. Differentiation gives $2f'(x) = u^2 - v^2/x^2$, and we find the stationary (and optimal) point $x^* = v/u$, at which $f(x^*) = 2vu$

corresponds to the decay rate of the cavity field due to beam loading, at the nominal operating point and with optimal detuning. Thus, for optimal external coupling, the external decay rate at the nominal operating point equals the sum of the decay rate from resistive losses and the decay rate from beam loading.

For the combined decay rate, which also accounts for resistive losses in the cavity walls, we have

$$\gamma = 2\gamma_0 + \frac{\frac{\alpha}{2} I_b \cos\phi_b}{a_0}. \tag{5.35}$$

Comparison of the different terms

Consider the cavity equation at steady state (5.32), for an optimally tuned and optimally coupled cavity at the nominal operating point. To get an idea of the relative size of the terms, we compare them relative to the term for resistive and external decay $\gamma \mathbf{A}_0 = \gamma a_0$.

First, for the term that depends on the mode amplitude we get

$$\frac{(-\gamma + i\Delta\omega)\mathbf{A}_0}{\gamma a_0} = -1 + i\frac{\Delta\omega}{\gamma}. \tag{5.36a}$$

For the beam loading term, we have

$$\frac{\frac{\alpha}{2}\mathbf{I_b^0}}{\gamma a_0} = \frac{-\frac{\alpha}{2} I_b \cos\phi_b + i\frac{\alpha}{2} I_b \sin\phi_b}{\gamma a_0}, \tag{5.36b}$$

and using (5.33) we see that the imaginary part is the negative of that in (5.36a), since the detuning was chosen to be optimal (5.33). We also see that the real part of (5.36b) is between -1 and 0 (using (5.35)). For a superconducting cavity ($\gamma_0 \approx 0$ and $\gamma \approx (\alpha/2) I_b \cos\phi_b / a_0$), we have that

$$\frac{\frac{\alpha}{2}\mathbf{I_b^0}}{\gamma a_0} \approx -1 + i\frac{\sin\phi_b}{\cos\phi_b}.$$

By solving for the generator term in (5.32) and using (5.36a) and (5.36b), we obtain

$$\frac{\sqrt{2\gamma_{\text{ext}}}\mathbf{F_g^0}}{\gamma a_0} = 1 + \frac{\alpha I_b \cos\phi_b}{2\gamma_0 + \alpha I_b \cos\phi_b}, \tag{5.36c}$$

which we see is between 1 and 2, and approximately equal to 2 for a superconducting cavity.

5.6 Phasor diagrams

In the previous section we considered the relative size of the different terms in equation (5.30), which we restate for convenience,

$$\frac{d\mathbf{A}}{dt} = (-\gamma + i\Delta\omega)\mathbf{A} + \sqrt{2\gamma_{\text{ext}}}\mathbf{F}_{\mathbf{g}} + \frac{1}{2}\alpha\mathbf{I}_{\mathbf{b}}.$$

To gain more insight into the dynamics of the accelerating mode, which will be helpful for the control design, we will illustrate the different terms of (5.30) in phasor diagrams such as in Figure 5.3, where the coloring coincides with that in the equation above. To avoid clutter, the phasor for the cavity field is drawn separately from the phasors that affect its derivative (field decay, generator drive and beam loading).

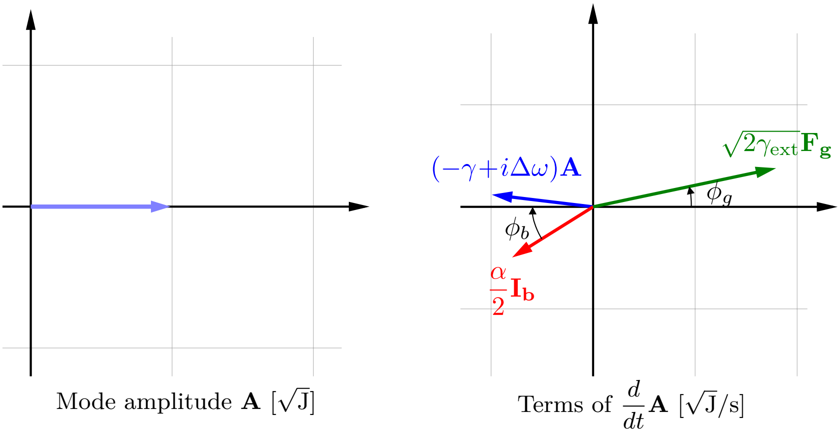


Figure 5.3 Phasor diagrams for visualizing the terms in the differential equation for the accelerating cavity mode (5.30). *Left:* Phasor for the mode amplitude. *Right:* Terms that affect the time derivative of the mode amplitude.

If the cavity is optimally tuned and optimally coupled, then the nominal generator phasor (green) will lie on the positive real axis, and the imaginary part of nominal decay phasor (blue) and the imaginary part of the beam loading phasor (red) will have equal magnitude put opposite signs. We realize that the phasor diagram in Figure 5.3 does not correspond to an optimally tuned or coupled cavity, but that the one in Figure 5.4 does.

Remark on phasor diagrams

Phasor diagrams are common in literature on cavity field control and accelerator physics, but to my knowledge they have only been used for illustrating

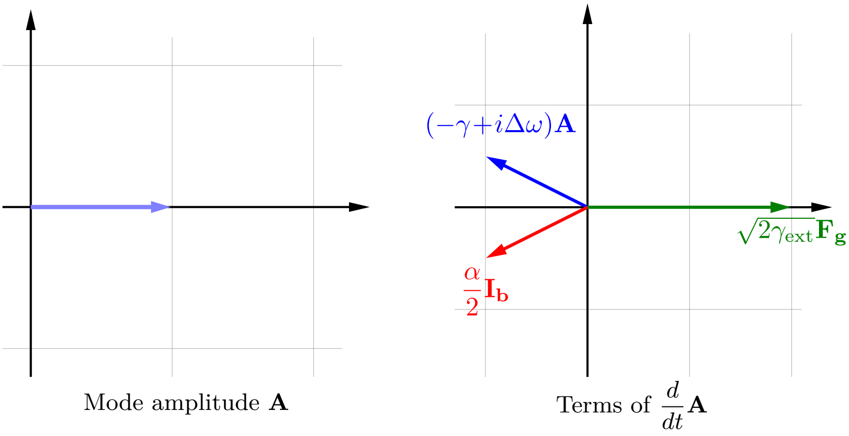


Figure 5.4 Phasors diagram for a superconducting cavity that is both optimally tuned and optimally coupled.

the nominal phasors at steady state. Typically, the phasors for the generator drive and the beam loading, together with the voltages that they induce, i.e., their steady-state effects on the cavity field

$$\begin{aligned}\mathbf{V_g} &= \frac{\alpha}{\gamma - i\Delta\omega}\sqrt{2\gamma_{\text{ext}}}\mathbf{F_g}\\ \mathbf{V_b} &= \frac{\alpha}{\gamma - i\Delta\omega}\frac{\alpha}{2}\mathbf{I_b},\end{aligned}$$

are shown. In this thesis we will use phasor diagrams to understand how *variations* in the driving terms $\mathbf{F_g}$ and $\mathbf{I_b}$ affect the cavity field, and we will not consider induced voltages.

In previous literature, the beam loading phasor $\mathbf{I_b}$ is typically oriented along the negative real axis. This is not unreasonable, since after all, the beam phase is the reference relative to which the cavity field should be controlled. However, at least from a field control perspective, where the goal is to keep the cavity field close to a set-point, and beam variations act as disturbances, it seems natural to orient the phasor for the cavity field along the positive real axis. With this convention, we get a nice symmetry in the phasor diagrams for optimally tuned cavities (see Figure 5.4), with the nominal phasors for the cavity field and generator drive lying along the real axis. Thus, for these two phasors we have that amplitude variations correspond to variations of the real part, and small phase variations correspond to variations of their imaginary parts.

5.7 Relative impact of disturbances

Requirements on field control performance, and specifications on the amplitude of amplifier ripple and beam ripple are typically given in relative terms, i.e., on the form $x\%$ rms and x° rms. By normalizing the dynamics from disturbances to field errors (of the accelerating mode), it will be easy to compute the relative field errors from a relative disturbance specifications.

To simplify the control design we would also like to normalize the static gain from the control signal to the measured signal (i.e., the mode amplitude) to one.

For these reasons we introduce the relative variations of the accelerating mode as[5] $\boldsymbol{z}$, the relative variations of the generator drive *due to disturbances* as $\boldsymbol{d_g}$, the relative beam variations as $\boldsymbol{d_b}$, and the normalized control signal $\boldsymbol{u}$, via

$$\mathbf{A} = \mathbf{A}_0(1+\boldsymbol{z}) \tag{5.37a}$$

$$\mathbf{F_g} = \mathbf{F_g^0}(1+\boldsymbol{d_g}) + \frac{\gamma}{\sqrt{2\gamma_{\text{ext}}}}\mathbf{A}_0\boldsymbol{u} \tag{5.37b}$$

$$\mathbf{I_b} = \mathbf{I_b^0}(1+\boldsymbol{d_b}) \tag{5.37c}$$

As an example, $\boldsymbol{z} = 0.01 + 0.02i$ corresponds to that the accelerating mode has an amplitude error of $1\,\%$ and a phase error of $0.02\,\text{rad} \approx 1.1^\circ$.

By inserting (5.37) into the cavity equation (5.30), and using that the nominal phasors sum up to zero in stationarity (5.32) we get

$$\dot{\boldsymbol{z}}\mathbf{A}_0 = (-\gamma + i\Delta\omega)\mathbf{A}_0\boldsymbol{z} + \gamma\mathbf{A}_0\boldsymbol{u} + \sqrt{2\gamma_{\text{ext}}}\mathbf{F_g^0}\boldsymbol{d_g} + \frac{\alpha}{2}\mathbf{I_b^0}\boldsymbol{d_b}, \tag{5.38}$$

and dividing by $\mathbf{A}_0$,

$$\dot{\boldsymbol{z}} = (-\gamma + i\Delta\omega)\boldsymbol{z} + \gamma\boldsymbol{u} + \frac{\sqrt{2\gamma_{\text{ext}}}\mathbf{F_g^0}}{\mathbf{A}_0}\boldsymbol{d_g} + \frac{\frac{\alpha}{2}\mathbf{I_b^0}}{\mathbf{A}_0}\boldsymbol{d_b}. \quad \left[\frac{1}{\text{s}}\right] \tag{5.39}$$

By introducing the dimension-less coefficients

$$\boldsymbol{K_g} = \frac{1}{\gamma}\frac{\sqrt{2\gamma_{\text{ext}}}\mathbf{F_g^0}}{\mathbf{A}_0}$$

$$\boldsymbol{K_b} = \frac{1}{\gamma}\frac{\frac{\alpha}{2}\mathbf{I_b^0}}{\mathbf{A}_0}$$

---

[5] The notation z for the relative field errors of the accelerating mode was selected since in the automatic-control literature z is commonly used for denoting *controlled variables*, i.e., the variables whose variation is of interest, as opposed to *measured variables* that are denoted by y.

we can re-write (5.39) as

$$\dot{\boldsymbol{z}} = (-\gamma + i\Delta\omega)\boldsymbol{z} + \gamma\boldsymbol{u} + \gamma\boldsymbol{K_g}\boldsymbol{d_g} + \gamma\boldsymbol{K_b}\boldsymbol{d_b}. \tag{5.40}$$

From the discussion in the second half of Section 5.5, we know that typically

$$\begin{aligned} 1 \le \operatorname{Re}\boldsymbol{K_g} \le 2 \\ 0 \le \operatorname{Re}\boldsymbol{K_b} \le 1. \end{aligned}$$

By Laplace transforming and re-arranging (5.40), we obtain

$$\begin{aligned} Z(s) &= \frac{\gamma}{s+\gamma-i\Delta\omega}\Big[U(s) + \boldsymbol{K_g}D_g(s) + \boldsymbol{K_b}D_b(s)\Big] \\ &= P_a(s)\Big[U(s) + \boldsymbol{K_g}D_g(s) + \boldsymbol{K_b}D_b(s)\Big], \end{aligned} \tag{5.41}$$

where

$$P_a(s) = \frac{\gamma}{s+\gamma-i\Delta\omega}.$$

We will use equation (5.41) in Section 6.1 when we formulate a normalized model for the RF system.

5.8 Parasitic modes

When additional (i.e., parasitic) modes, need to be considered in addition to the accelerating mode, the parameters for the accelerating mode will be labeled with a, those for the parasitic mode closest in frequency to the accelerating mode by 1, those for the second closest parasitic mode by 2, and so on. Subscripts will be used when possible, and superscripts when necessary.

It follows directly from Section 5.4 how the beam and RF generator output affect the mode amplitudes, which we illustrate in Figure 5.5; the figure also shows how the cavity modes affect the beam, and that the voltage induced in the cavity pickup antenna, $\mathbf{V}_{\mathrm{pu}}$, is a linear combination of the mode amplitudes.

In general the parameters $\gamma_{\mathrm{ext}}^{k}$, γ_k, and the pickup gains $\mathbf{c}_k$ for the different modes are not related. However for elliptical cavities, which are commonly used and where parasitic modes often need to be considered, such relations exist for the modes closest to the accelerating mode, i.e., in the so-called fundamental passband.

Normalized cavity model with parasitic modes

In the following chapters when we consider the general mode structure in Figure 5.5, we would also like to use the normalization of Section 5.8. Unfortunately the beam loading $\mathbf{I_b}$ and the generator drive $\mathbf{F_g}$ couple differently

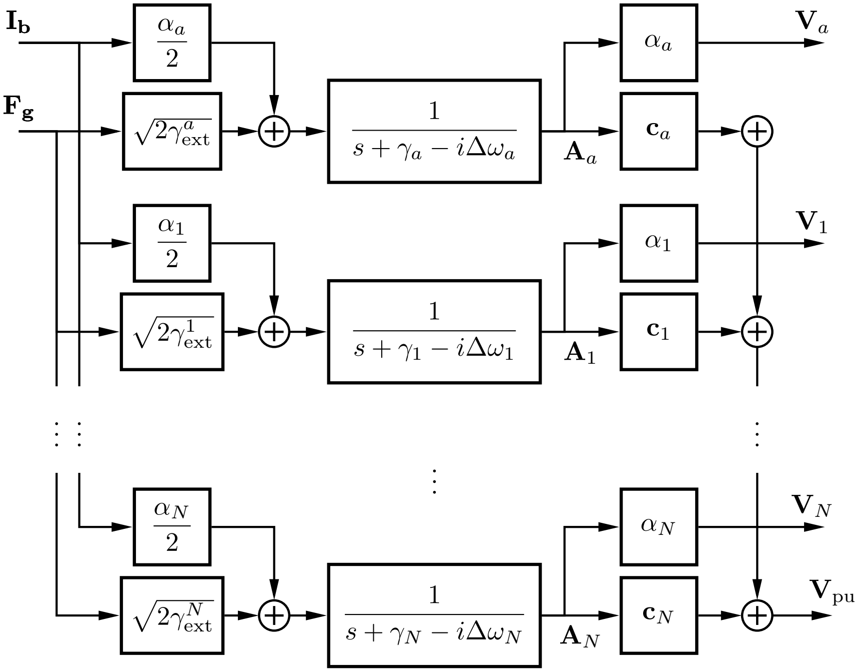


Figure 5.5 Block diagram for how cavity modes couple to the RF system and the beam. $\mathbf{V}_{\text{pu}}$ is the voltage detected by the pickup.

to the modes, so things would become slightly messy if we want to do this in an exact manner. To this end, we introduce the notation

$$P_k(s) = \frac{\sqrt{2\gamma_{\text{ext}}^k}}{\sqrt{2\gamma_{\text{ext}}^a}} \frac{\mathbf{c}_k}{\mathbf{c}_a} \frac{\gamma_a}{s+\gamma_k-i\Delta\omega_k}, \tag{5.42}$$

and consider the block diagram of Figure 5.6.

We see that it *almost* corresponds to Figure 5.5 with the modes normalized around their nominal operating points: the coupling of the modes to generator drive and to the pickup antenna is correct, but the coupling to the beam is only correct for the accelerating mode. However it is the output of the accelerating mode that we really care about, so for our intended use, it will be a good approximation, in particular since the spectral content of the beam is centered around the RF frequency.

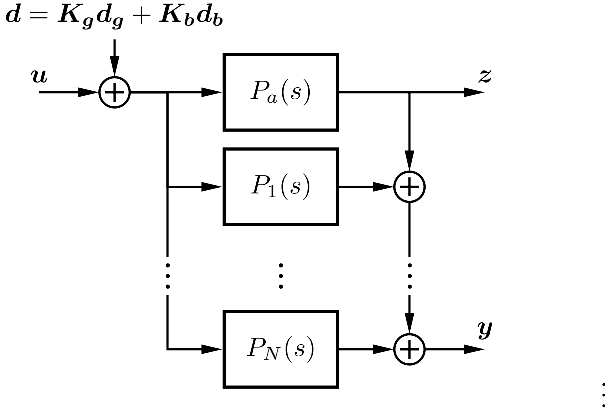


Figure 5.6 Normalized cavity dynamics including parasitic modes, with $P_k(s)$ given by (5.42). Note that the impact of d_b is only approximately correct.

5.9 Comparison to the equivalent circuit model

The presented approach from [Haus, 1983] for deriving the cavity field dynamics is somewhat different from the equivalent circuit approach, which is standard in the accelerator literature. In this section, we discuss advantages of Haus' approach, *for getting insight into the field control problem.* For electrical engineers who are accustomed to electrical circuits, these advantages may not outweigh the familiarity of thinking in terms of a circuit model.

It should be emphasized that the only differences are in the *derivation* (Haus' derivation follows directly from Maxwell's equations), and in the *parametrization* of the differential equation. In particular after normalization the resulting differential equations are *identical.*

Insight and simplicity

The equivalent circuit approach where the cavity dynamics are mapped to an analogous equivalent circuit from which a model is then derived, obfuscates the connection between the cavity and the derived model. For example, it is hard to visualize how particle bunches interact with the voltage over a resistor. The approach taken here, where everything follows from Maxwell's equations, is arguably more convincing and easier to follow.

Choice of parameters

With the equivalent circuit approach considered in [Schilcher, 1998], the complex-coefficient differential equation for the accelerating mode takes the

Table 5.1 Comparison of the parameters and the variables used in this thesis compared to those used in [Schilcher, 1998]. A cross × indicates that the given quantity does not occur in the considered model, and an asterisk * indicates that the parameter occurs in the model (5.43).

Quantity	**[Schilcher, 1998]**	**This thesis**
*Accelerating voltage (phasor)	$\mathbf{V}$	$\mathbf{V}_a = \alpha\mathbf{A}$
*Generator current (phasor)	$\mathbf{I_g}$	$2\sqrt{2\gamma_{\text{ext}}}/\alpha\mathbf{F_g}$
*Beam current (phasor)	$\mathbf{I_{b(RF)}}$	$2\mathbf{I_b}$
Forward wave (phasor)	×, $\sqrt{R/(2\beta)}\mathbf{I_g}$	$\mathbf{F_g}$
*Cavity Bandwidth	$\omega_{1/2}$	γ
*Detuning of acc. mode	$\Delta\omega$	$\Delta\omega$
Resonance frequency of accelerating mode	ω_0	ω_a
Unloaded Q	Q_0	×, $= \omega_a/(2\gamma_0)$
Coupling Factor	β	×, $= \gamma_{\text{ext}}/\gamma_0$
External Q	$Q_{\text{ext}} = Q_0/\beta$	×, $= \omega_a/(2\gamma_{\text{ext}})$
Loaded Q	$Q_L = Q_0/(\beta+1)$	×, $= \omega_a/(2\gamma)$
Normalized Shunt Impedance	r/Q	×, α^2/ω_a
Resistor R	$R = (r/Q)\cdot Q_0/2$	×, $= \alpha^2/(4\gamma_0)$
*Loaded shunt impedance	$R_L = R/(1+\beta)$	×, $= \alpha^2/(4\gamma)$
Detuning Angle	$\tan\psi$	$-\Delta\omega/\gamma$
Forward Power	$P_g = R/(2\beta)\,\lvert\mathbf{I_g}\rvert^2$	$\lvert\mathbf{F_g}\rvert^2$
Beam current (DC)	I_{b0}	I_b

form (the original work considered an equivalent real-valued two-input two-output representation),

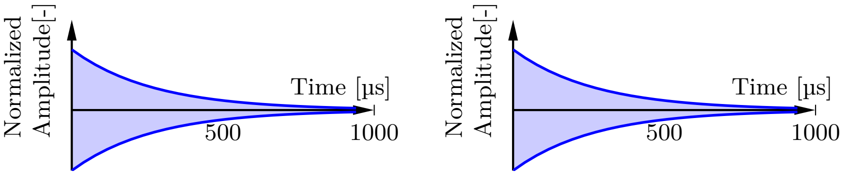
$$\frac{d\mathbf{V}}{dt} = (-\omega_{1/2} + i\Delta\omega)\mathbf{V} + R_L\omega_{1/2}\left(2\mathbf{I_g} + \mathbf{I_{b(RF)}}\right), \tag{5.43}$$

where the parameters and variables of the equations are listed in Table 5.1.

We now compare how the parameters in this differential equations, relate to those in (5.30), which we restate for convenience,

$$\frac{d\mathbf{A}}{dt} = (-\gamma + i\Delta\omega)\mathbf{A} + \sqrt{2\gamma_{\text{ext}}}\mathbf{F_g} + \frac{\alpha}{2}\mathbf{I_b}.$$

Definition of parameters in terms of Q_0 Quantities such as the coupling factor $\beta = Q_0/Q_L - 1$, and the resistor $R = (r/Q)/2 \cdot Q_0$, are defined in terms of the unloaded quality factor, Q_0. For a superconducting cavity Q_0 is very large and may vary significantly due to anomalous losses and field emission, it seems strange that important parameters such as R and β are affected by this variation.



(a) $Q_L = 300000$. (b) $Q_L = 600000$.

Figure 5.7 The impulse responses of two hypothetical cavities with resonance frequencies of 352 MHz (a) and 704 MHz (b). The Q-values for the two cavities differ by a factor of two, but their baseband dynamics are *identical*, which is seen from the figures. The Q-value for (b) is typical for the ESS medium-β cavities, and the Q-value of (a) is *somewhat similar* to the ESS spoke cavities. This examples illustrates that Q-values on their own say very little about the cavity dynamics; they must always be considered together with the resonance frequency of the cavity, and thus, one might just as well use a single parameter, such as the decay rate γ, to quantify the cavity dynamics.

For a perfectly conducting cavity, β and R becomes infinite, and the equations that they occur in become ill-defined; while a lossless RF cavity is not physically possible, it high-lights the problem.

Quantification of generator drive The drive from the RF amplifier is modeled as a current $\mathbf{F_g}$ in the equivalent circuit model. To obtain the single most important property of the generator drive, its power, it is necessary to use the rather cumbersome expression $P_g = R_L/(8\beta)\,|\mathbf{I_g}|^2$, where R_L depends on β, r/Q and Q_0. In Haus' approach, the generator drive is quantified by the amplitude of the electric field in the waveguide (in units of $\sqrt{\mathrm{J/s}}$). This provides a considerable simplification.

Interpretation of bandwidth From (5.43) we see that the transfer function from amplifier ripple and beam ripple to field errors is given by

$$\frac{\omega_{1/2}}{s+\omega_{1/2}} R_L.$$

Thus one is led to believe that by reducing the bandwidth $\omega_{1/2}$, the effect of generator ripple and beam ripple on the cavity field is reduced. This is not correct. The reason is that R_L depends inversely on $\omega_{1/2}$. From (5.30) one sees that the effect of beam ripple is given by the transfer function

$$\frac{1}{s+\gamma}\frac{\alpha}{2},$$

from which we see that the effect of beam ripple is actually increased (at low frequencies) by reducing $\omega_{1/2} = \gamma$. However now one would be led to believe that the effect of generator ripple also is increased when γ is reduced, however this is not quite the case.

Different definitions of r/Q There are two common ways to define the shunt resistance in the equivalent circuit mode, the “circuit definition” and the “linac definition”, that differ by a factor two; care is required to avoid confusion and mistakes [Tückmantel, 2011], [Schilcher, 1998, Eq 3.2]. In this chapter we have instead used the parameter α, which is the ratio of two clearly defined cavity parameters, the effective accelerating voltage and the stored energy; thus, there is is little risk of confusion.

A comment on Q-values Designers of accelerator cavities typically quantify the cavity losses in terms of Q-values, which are generally defined as

$$Q = 2\pi \frac{\text{stored energy}}{\text{energy dissipated per oscillation period}}.$$

Put differently, the Q-value quantify how many *oscillation periods*, measured in radians, it takes for an impulse response to the cavity to decay to $\mathrm{e}^{-1} = 37\,\%$ of the initial amplitude. Depending of the type of accelerator cavity, it could take from some several thousand to several hundred thousand oscillation periods. From a field control perspective, where the interest is in the complex envelope of the cavity field, the number of oscillations within the envelope is of no interest.

Also, from just a Q-value, one can say *nothing* about the cavity dynamics, *unless* the corresponding resonance frequency is known, consider for example the situation shown in Figure 5.7. What actually matters is the corresponding time constant for the decay. Even more insightful is to consider the decay rate, i.e., the inverse of the time constant. Decay rates, e.g., from resistive losses or from decay via the coupling mechanism, also simply add up linearly, $\gamma = \gamma_0 + \gamma_{\text{ext}}$ which is very convenient. Decay rates also go nicely together with the type of phasor diagrams shown in 5.3; in particular, if the phasor diagram for the time-derivate terms is normalized with respect to the magnitude of the cavity field, it is possible to compare the relative impact of disturbances between different cavities.

6

The Field Control Loop

First, let us motivate the need for a field control loop. Without feedback, the effects of beam ripple and amplifier ripple on the accelerating cavity mode have a frequency dependence given by the transfer function $P_a(s)$ in (5.41), which is shown in Figure 6.1. It is seen that the cavity acts as a low-pass filter—high-frequency disturbances are attenuated, but there is no attenuation of slow, low-frequency disturbances.

With a feedback controller, it is possible to attenuate the effect of the low-frequency disturbances. To ensure stability and sufficient control performance, it is important to include dynamics and time delays in the closed-loop analysis. That is why we in the two previous chapters considered the dynamics of different systems in the field control loop. In this chapter we put those pieces together and provide a general discussion of the field control loop.

In all the remaining chapter, except Chapter 10, we will only consider

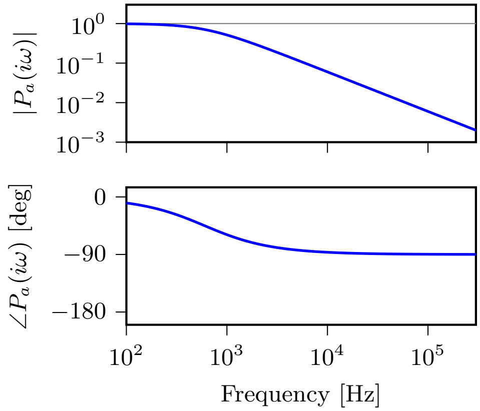


Figure 6.1 Normalized transfer function for the accelerating mode of a superconducting cavity with 600 Hz bandwidth.

feedback, and consider operation around a nominal operating point; the control performance during the filling transient, and for repetitive disturbances, can typically be improved by using iterative learning control, see [Norrlöf, 2000; Rogers et al., 2010].

6.1 The model

By collecting the results from the two previous chapters we get the baseband model of the field control loop in Figure 6.2, where L_{reg} is the delay from signal processing in the LLRF system, L_1 is the propagation time of the electric waves from the RF amplifier to the cavity, and L_2 is the propagation time from the pickup probe in the cavity to the LLRF system. The factors $\mathrm{e}^{-i\theta_1}$ and $\mathrm{e}^{-i\theta_2}$ are phase shifts from considering the propagation delays in the baseband (see Section 2.4), we also consider phase-shifts from the klystron and other components to be lumped into these two factors; the factor $\mathrm{e}^{i\theta_{\mathrm{adj}}}$ is for loop phase adjustment, and can in principle be considered to be a part of the controller (see Section 6.5). For our analysis we assume the controller $C_0(s)$ to be linear in its action on complex signals, i.e., of the form (2.12).

The set-point $\mathbf{V}_{\mathrm{pu}}^0$ for the cavity voltage measured by the cavity pickup probe, is provided by the beam physicists. For $\mathbf{V}_{\mathrm{pu}}$ to be controlled to $\mathbf{V}_{\mathrm{pu}}^0$ it is crucial that $\widehat{\theta}_2$ is calibrated to equal θ_2. This calibration is also handled by the beam physics group. We will not consider this further, and thus assume that $\theta_2 = \widehat{\theta}_2 = 0$.

By using the normalized cavity model from Figure 5.6, introducing the cavity transfer function $P_{\mathrm{cav}}(s) := P_a(s) + P_1(s) + \cdots + P_N(s)$, and then lumping the process dynamics in Figure 6.2 into one transfer function

$$P(s) := P_{\mathrm{cav}}(s)P_{\mathrm{amp}}(s)\mathrm{e}^{-sL}\mathrm{e}^{-i\theta}, \tag{6.1}$$

where $L := L_1 + L_2 + L_{\mathrm{reg}}$, $\theta := \theta_1 + \theta_2$, and lumping the loop-phase adjustment and the nominal controller into a new controller,

$$C(s) := C_0(s)\mathrm{e}^{i\theta_{\mathrm{adj}}},$$

we obtain the simplified and normalized block diagram in Figure 6.3, where the normalized disturbances from the generator and the beam loading term have been added into one signal

$$\boldsymbol{d} = \mathrm{e}^{-sL_1}\boldsymbol{K_g}\boldsymbol{d_g} + \boldsymbol{K_b}\boldsymbol{d_b}. \tag{6.2}$$

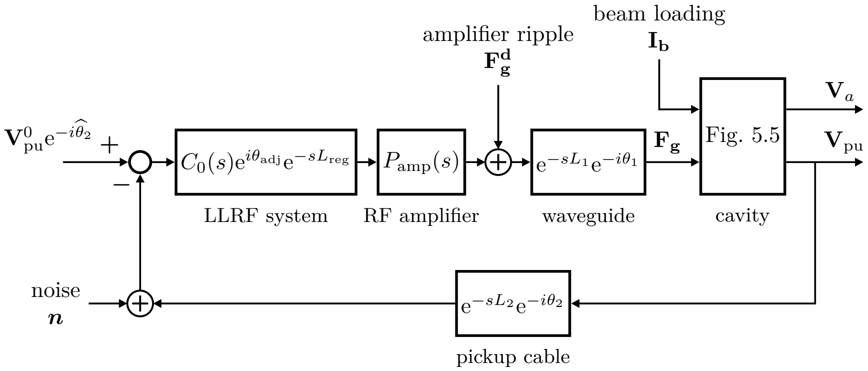


Figure 6.2 Linearized, somewhat simplified, baseband model of the RF system, which will be used for analyzing the field control loop. For improving transient performance and dealing with repetitive disturbances, one could add a feedforward signal to the output of the controller; but since only flat-top operation will be considered in this thesis, that feedforward signal is not shown in the figure.

6.2 Three types of disturbances

We find it useful to think about three classes of disturbances affecting the cavity field stability: ***random disturbances***, that are not feasible to predict, such as

- Beam current fluctuations $\mathbf{\tilde{I}_b}$ due to plasma instabilities in the ion source
- Disturbances $\mathbf{F_g^d}$ on the RF amplifier output due to power supply switching and various other noise sources[1].

These disturbances will inevitably give rise to cavity field errors, and the only way to reduce their effects is by feedback.

Then there are ***repetitive disturbances***, which are more or less identical from RF pulse to RF pulse, for example

- The beam pulse $\mathbf{I_b}$.
- Resonance frequency variations, i.e. variations of $\Delta\omega_a$, from mechanical deformation of the accelerator cavity caused by the Lorentz force[2].

---

[1] If the modulator switching is deterministic and synchronized to the timing system, it *could* effectively be canceled by iterative learning control. Unfortunately there is not plan to do this for the planned modulator at the European Spallation Source, however it has been done at other facilities [Personal correspondence with Thomas Schilcher, Paul Scherrer Institute].

[2] The electromagnetic field acting on the cavity walls.

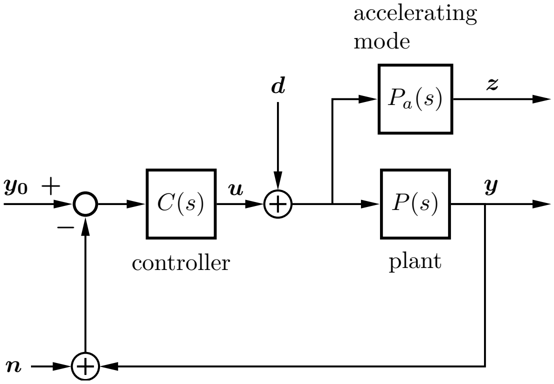


Figure 6.3 Normalized version of the block diagram in Figure 6.2, where the plant dynamics $P(s)$ are given by (6.1), and the normalized disturbances $\boldsymbol{d}$ are given by (6.2). We will for the most part assume that θ_{adj} is calibrated so the $\theta_{\text{adj}} = \theta$, in order to simplify the exposition.

Since these disturbances can be anticipated, it is in principle possible to cancel them completely by adding a suitable signal $u_{\text{ILC}}(t)$, which is synchronized to the start of the beam, to the control signal. The technique of updating the compensation signal $u_{\text{ILC}}(t)$ based on the errors from the previous iteration (i.e. RF pulse), is called iterative learning control (ILC)[3]. We will not consider this topic in this thesis, and refer to [Norrlöf, 2000; Bristow et al., 2006; Rogers et al., 2010].

Finally we find it useful to think about ***model changing disturbances***, i.e., ***process variations***. These could be due to both random or repetitive disturbances, e.g.,

- Amplifier droop, i.e., a slow decrease of ΔV_c, which, for the klystrons, leads to a phase shift of the control loop according (4.2).
- Resonance frequency variations, i.e. variations of $\Delta\omega_a$, from mechanical deformation of the accelerator cavity caused by the Lorentz force[4] and due to microphonics[5].

These disturbances are slow/repetitive and can be almost perfectly canceled by the integral action of the feedback controller and the iterative learning control; from this perspective one could ignore them. However these disturbances give rise to parameter variations of the nominal process model (6.1), and thus either the controller needs to be robust to these variations in terms

---

[3] Frequently referred to as adaptive feedforward in the cavity field control literature.

[4] The electromagnetic field acting on the cavity walls.

[5] Unwanted mechanical vibrations, for example, from nearby traffic or the cryo plant.

of stability and performance, or the controller needs to be adaptive. Adaptive controllers are notoriously complex and troublesome, so we will focus on designing robust controllers.

The feedback from the controller $C(s)$ will attenuate both repetitive and non-predictable disturbances. However, since the repetitive disturbances such as the beam loading transient can be handled by iterative learning control, we will mainly consider non-predictable disturbances when we analyze closed-loop performance.

Remark: We classified high-frequency ripple on the amplifier supply voltage as a random disturbance, however as with droop it also leads to a change in the model. However, the amplitude of the high-frequency ripple is so small that it is a good approximation to consider it as a random load disturbance on the nominal model.

6.3 Closed-loop transfer functions

Without feedback, the disturbances will only be attenuated by the low-pass dynamics of the cavity (Figure 6.1). To attenuate slow disturbances, it is necessary to introduce a feedback controller.

The change in disturbance attenuation between no feedback, and feedback with the controller C, is, as discussed in Section 2.2, quantified by the so-called sensitivity function

$$S = 1/(1 + PC). \tag{6.3}$$

Using a controller with high gain at low frequencies, which is where we want to increase disturbance attenuation, we get a sensitivity function like the one in Figure 6.4; the transfer function examples throughout this chapter correspond to the nominal PI-controller design for a cold cavity from Chapter 7. In Figure 6.4 it is seen that disturbance attenuation is significantly improved at low frequencies, but worsened at around 10–100 kHz. This is unavoidable and known as the waterbed effect (see Section 2.2).

With feedback, the transfer function from disturbances to field errors becomes

$$G_{zd} = P_a/(1 + PC) = P_a S, \tag{6.4}$$

which is shown in Figure 6.5, together with the transfer function P_a of the accelerating cavity mode without feedback.

The transfer function

$$G_{un} = C/(1 + PC) = CS \tag{6.5}$$

quantifies how measurement noise n is amplified to the controlled signal z. High control signal activity tends to give poor control performance due

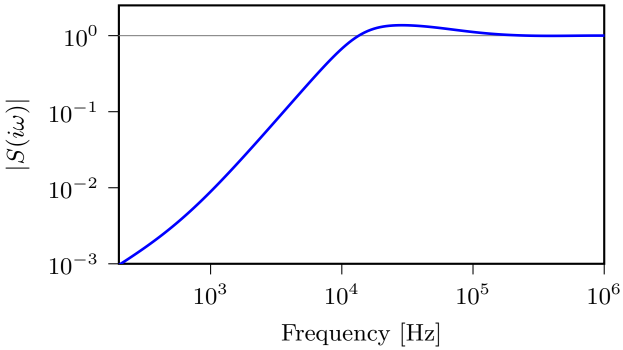


Figure 6.4 Sensitivity function $S = 1/(1 + PC)$ for a cold cavity controlled with a PI-controller, this, as well as the other examples in this chapter, correspond to the nominal design in Chapter 7.

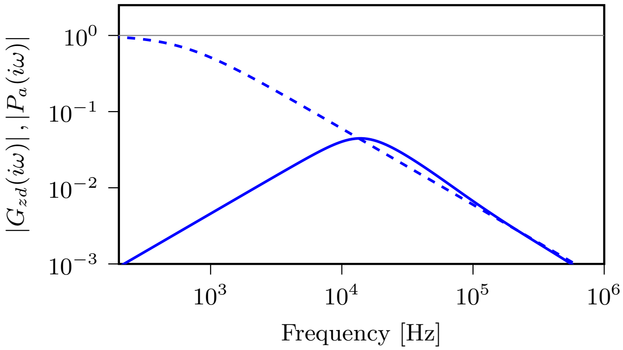


Figure 6.5 Transfer functions from disturbances d to field errors z: with feedback ($G_{zd} = P_a(1+PC)$, solid line), and without feedback (P_a, dashed line).

to excitation of unmodeled dynamics and an increased impact of amplifier nonlinearities. These effects are hard to quantify, so it is prudent to limit the control signal activity. Typically the controller gain needs to be large for the superconducting cavities, motivating the introduction of a low-pass filter in the controller in order to limit the amplification of high-frequency noise to the control signal.

Finally, we also mention the transfer function

$$G_{zn} = P_a C/(1 + PC) \tag{6.6}$$

which quantifies the impact of measurement noise on the controlled signal and also the response to set point changes; its typical characteristics are shown

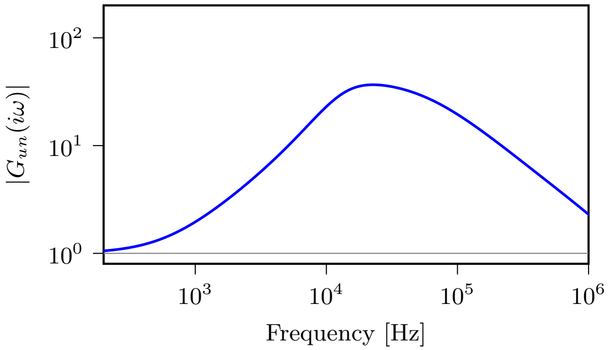


Figure 6.6 The transfer function $G_{un} = C/(1+PC)$ from measurement noise n to the control signal u.

in Figure 6.7. For ESS the impact of measurement noise on the field errors is small compared to the impact of disturbances from the amplifier and beam current variations. Also, during the pulses the set-point will be constant, and hence there are no demands to track fast set-points variations. Thus, we will primarily look at other closed-loop transfer functions than $G_{zn}(s)$ for our control analyses.

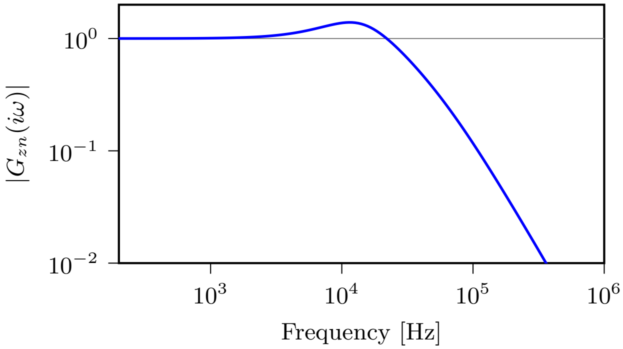


Figure 6.7 The transfer function $G_{zn} = P_aC/(1+PC)$ from measurement noise n to field errors z.

The transfer functions (6.3)-(6.6) are collectively known as the ”gang of four”, and capture the essentials of the closed-loop control performance. Typical gang-of-fours for a cold cavity and a warm cavity are shown in Figure 6.8.
.

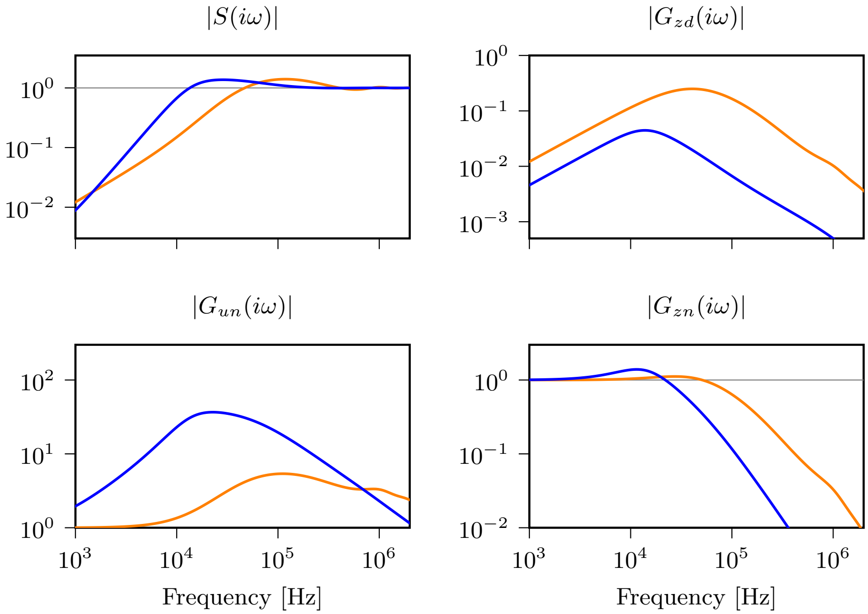


Figure 6.8 The set of transfer functions known as the "gang of four", for a warm cavity with a bandwidth of 12 kHz (orange lines), and a cold cavity with a bandwidth of 0.6 kHz (blue lines), with a PI-controller tuned for a maximum sensitivity of $S_{\max} = 1.6$.

Stability

It is not only sufficient to look at the magnitude of the closed-loop transfer function, but as we discussed in Section 2.2, it is important to consider stability when feedback is involved. For the applications in this thesis, closed-loop stability is conveniently verified by Nyquist's criterion, which we introduced in Section 2.2. The Nyquist curves and the Bode digram for the cold and warm cavities from our previous examples are shown in Figure 6.9.

6.4 Mathematical formulation of control specifications

In the previous section, we stated rather loosely that for good control performance it is desirable with high disturbance attenuation, sufficient robustness and limited control signal activity. Now, we will formulate these requirements mathematically, using th signal and system norms from Section 2.3.

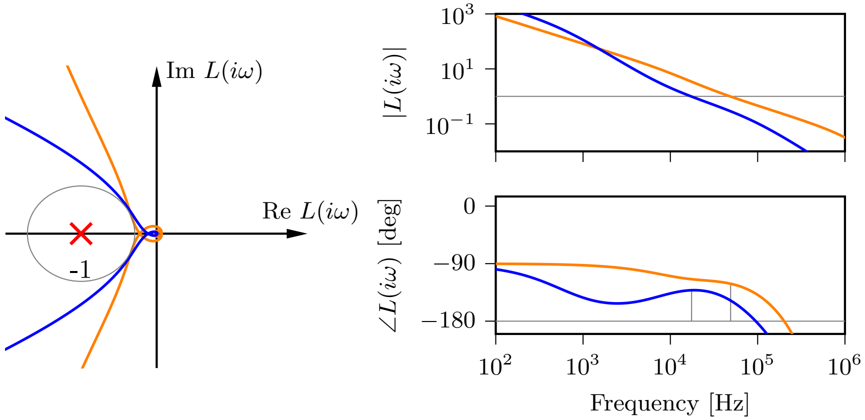


Figure 6.9 Nyquist and Bode diagrams for a warm and a cold cavity controlled by a PI-controller. The gray circle in the Nyquist diagram corresponds to a sensitivity constraint of $M_S \leq 1.6$, and the gray lines in the phase curve of the Bode diagram illustrate phase margins of about 45°; both robustness measures indicate good robustness of the closed-loop system. For more details on these robustness measures, see Section 2.2.

Cavity field error

By denoting the spectrum of the normalized load disturbances $\boldsymbol{d}$ by $D(s)$, we have, since $P_a(s)$ is normalized (Section 5.7), that the resulting (relative) rms error of the accelerating mode is

$$||G_{zd}D||_2 = \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} |G_{zd}(i\omega)|^2 \, |D(i\omega)|^2 \, d\omega}.$$

There is also a contribution $||G_{zn}N||_2$ due to measurement noise, however for the ESS accelerator this effect is not significant: assuming that the down-sampled 10 MHz signal has a signal-to-noise ratio of 70 dB, which should be achievable with the planned hardware, that the noise is white, and that the closed-loop bandwidth of the system is 100 kHz, then the effect on the cavity field is roughly

$$||G_{zn}N||_2 = (-70 - \log_{10}(100\,\text{kHz}/5\,\text{MHz}))\ \text{dBc} = -87\,\text{dBc} = 0.005\%,$$

which is small compared to the field error specification of 0.1 % and 0.1° rms[6].

---

[6] Note that for FEL accelerators, this error is of the same order as their typical control specification of 0.01 %

Robustness

We will quantify the robustness of the field control loop by the maximum value of the sensitivity function (6.3),

$$M_S = \sup_{\omega} |S(i\omega)| = ||S||_{\infty} .$$

This is, as we discussed in Section 2.2, a more general, and more convenient robustness measure than the traditionally used gain and phase margins. Typically the maximally allowed value for M_S is chosen between 1.4 and 2, with better robustness for smaller values.

Control signal activity

The control signal activity (rms), resulting from measurement noise with spectrum N, is given by $||G_{un}N||_2$. The peak-to-peak value is approximately 3 times the rms value, i.e. $\pm 3\, ||G_{un}N||_2$. At ESS the amplifier power overhead for regulation will be around 25 %, giving an amplitude overhead of 12 %, which we would like to stay well below.

Assuming that the measurement noise is white with $N(i\omega) = 1$ up to the Nyquist frequency ω_{Ny}, then the control signal activity due to measurement noise is given by

$$\frac{||G_{un}N||_2}{||N||_2} = \sqrt{\frac{1}{2\pi} \int_{-\omega_{\mathrm{Ny}}}^{\omega_{\mathrm{Ny}}} |G_{un}(i\omega)|^2 \, d\omega} \Big/ \sqrt{2\omega_{\mathrm{Ny}}/(2\pi)}.$$

Formal formulation of control problem

We are now in a position to formulate our control problem mathematically:

Control Problem

From a given class $\mathcal{C}$ of suitable controllers, we wish to find the controller C which solves the optimization problem:

$$\underset{C \in \mathcal{C}}{\text{minimize}} \quad \left|\left| \frac{P_a}{1 + PC} D \right|\right|_2 \tag{6.7a}$$

$$\text{subject to} \quad \left|\left| \frac{1}{1 + PC} \right|\right|_{\infty} \le S_{\max}. \tag{6.7b}$$

$$\left|\left| \frac{C}{1 + PC} N \right|\right|_2 \le B_{un}. \tag{6.7c}$$

where D is the spectrum of load disturbances, N is the spectrum of measurement noise, $S_{\max}$ is a limit on the maximal value of the sensitivity function, and B_{un} is a bound on the control signal activity due to noise.

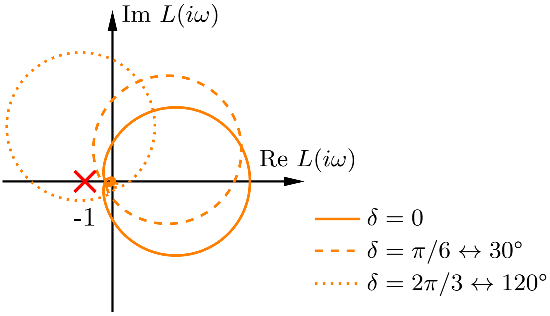


Figure 6.10 Illustration of the effect of loop phase variations for a warm cavity with a bandwidth of $\gamma = 12\,\text{kHz}$ which is controlled by a P-controller $C(s) = 5$.

6.5 Loop phase adjustment

We will now discuss how the loop phase adjustment angle θ_{adj} should be chosen. For this, we introduce $P_0(s) := P_{\text{cav}}(s)P_{\text{amp}}(s)\mathrm{e}^{-sL}$ so that

$$P(s) := P_0(s)\mathrm{e}^{-i\theta},$$

after which the open-loop transfer function can be written

$$L(s) = C(s)P(s) = C_0(s)\mathrm{e}^{i\theta_{\text{adj}}}P_0(s)\mathrm{e}^{-i\theta} = L_0(s)\mathrm{e}^{i\delta}$$

where $L_0(s) = C_0(s)P_0(s)$ and $\delta = \theta_{\text{adj}} - \theta$. From this we see that the Nyquist curve of $L(s)$ equals the Nyquist curve of $L_0(s)$ rotated by δ radians around the origin, see Figure 6.10.

If $L_0(s)$ has real-valued coefficients, for example if there is no detuning and parasitic modes are not modeled, then it is clear that choosing $\delta = 0$ gives the best robustness. In this case we also have that if $L_0(s)$ corresponds to a stable closed loop, then small values of δ correspond to a reduction of the phase margin by δ radians.

In the case where $L_0(s)$ has complex coefficients, the choice $\delta = 0$ may not necessarily give the best robustness, see Figure 6.11.

Methods for calibration of the loop phase adjustment

To calibrate the loop phase adjustment, the most straight-forward approach would be to do systems identification of $L(s)$ and then adjust ϕ_{adj} for optimal robustness. Realizing that adequate phase-margins are most critical around a system's bandwidth, the identification should be focused in this frequency region.

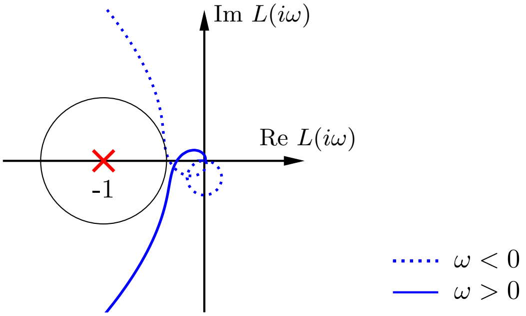


Figure 6.11 A situation where choosing θ_{adj} different from 0 gives improved robustness. The gray circle corresponds to $M_S = 1.6$. The Nyquist curve is for a cold cavity with a single parasitic mode, which is controlled by a PI-controller augmented with a second-order low-pass filter.

This approach essentially corresponds to the following model-free method for calibration of loop phase adjustment[7]:

1. Find an angle θ_a such that the closed-loop system is stable.
2. Increase θ_{adj} from θ_a until the loop starts to oscillate, i.e. is on the stability boundary, and denote the corresponding value of θ_{adj} by θ_b.
3. Similarly, decrease θ_{adj} from θ_a until oscillations start, and denote the angle θ_c.
4. Select $\theta_{\text{adj}} = (\theta_b + \theta_c)/2$.

This calibration should preferably be done with the controller and set-point that are intended to be used.

Calibration of the loop phase adjustment is also discussed in [Brandt, 2007, Sec. 4.3.3].

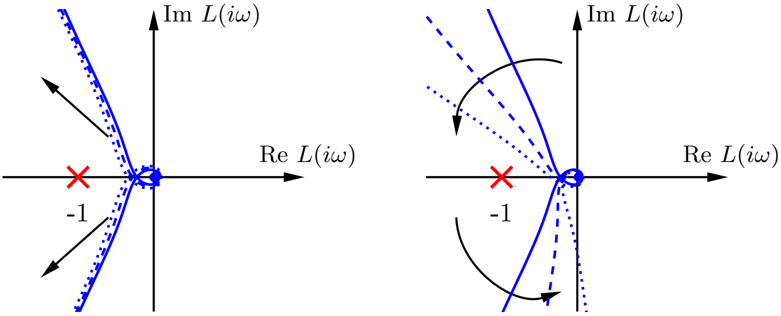
6.6 Process variations

An illustration of the effect is shown in Figure 6.12, where we see that gain variations of the plant reduce the disturbance rejection or reduce the robustness. Furthermore, phase variations give reduced robustness, and in particular, reduced phase margin.

The controller also needs to be robust against variations of the mode frequencies $\Delta\omega_a, \Delta\omega_1, \ldots, \Delta_N$.

---

[7] Oral tradition of the field control community. The author learned about it from Mark Crofford, SNS.



(a) Effect of gain variation.

(b) Effect of phase variation.

Figure 6.12 Effect of gain and phase variation on the closed-loop transfer function $L(s)$, which can be due to variations in the amplifier supply voltage or the electrical length of the wave-guide and cables.

7

Limits of Field Control Performance

7.1 Controller structure

An important aspect of control design is to decide on the controller structure—on the one hand it should be flexible enough to allow a small value of the objective function (6.7a), and on the other hand it should be as simple as possible, since this simplifies understanding, implementation, and leaves fewer things to go wrong.

For field control it has been proposed to use: PI(D)-controllers, observer-based state feedback, Smith predictor based controllers and model predictive controller. All these controllers, and almost all other controllers of practical use, are linear and time invariant (LTI), at least if there is no control signal saturation. Regardless of the controller however, there are *fundamental* limitations on the achievable control performance; when both the controller and plant are linear, one such fundamental limitation is given by Bode's Integral Formula (2.7).

In the LTI setting it is often possible to be more specific about the limits of control performance—by explicitly computing them! In this chapter we will do just that for the field control problem (6.7), and compare the performance of the optimal LTI controller with the performance of PI- and PID-controllers. It will turn out that the difference in performance is quite small, which indicates that there typically is little need to use more complicated controllers for field control. The work in this chapter has been inspired by [Garpinger, 2009], where similar investigations were made for low-order process models.

In Section 7.2 we provide the details for the comparison of the different controller types, and in Section 7.3 we compare the performance of the optimal LTI controller to the performance of optimal PI- and PID-controllers for different loop-delays and different levels of control signal activity.

In Appendix B.1 we give a short background on how to find optimal linear controllers and the specifics of the approach used for this thesis, and in Appendix B.2 we briefly discuss how optimal parameters of the PI- and PID-controllers were found.

7.2 Problem formulation

We will compare the optimal solutions to the control problem from the previous chapter,

$$\begin{array}{ll} \underset{C \in \mathcal{C}}{\text{minimize}} & \left\| \dfrac{P_a}{1+PC} D \right\|_2 \\ \text{subject to} & \left\| \dfrac{1}{1+PC} \right\|_\infty \leq S_{\max}. \\ & \left\| \dfrac{C}{1+PC} N \right\|_2 \leq B_{un} \end{array}$$

for the following classes $\mathcal{C}$ of controllers:

- The set of all linear time-invariant controllers.
- PI-controllers of the form

$$C_{\mathrm{PI}}(s) = K \left(1 + \frac{1}{sT_i}\right) \cdot \frac{1}{sT_f + 1}. \tag{7.1}$$

- PID-controllers of the form

$$C_{\mathrm{PID}}(s) = K \left(1 + \frac{1}{sT_i} + \frac{sT_d}{sT_d/N + 1}\right) \cdot \frac{1}{sT_f + 1}. \tag{7.2}$$

Below we introduce the assumed process models $P(s)$ and $P_a(s)$, the disturbance spectrum $D(s)$ and measurement noise spectrum $N(s)$.

Process model

We consider the process model (6.1),

$$P(s) = P_{\mathrm{cav}}(s) P_{\mathrm{amp}}(s) \mathrm{e}^{-sL} \mathrm{e}^{i\theta}.$$

To make the ideas come across clearly, we will neglect the impact of parasitic modes and only include the accelerating mode of the (normalized) cavity model,

$$P_{\mathrm{cav}}(s) = P_a(s) = \frac{\gamma_a}{s + \gamma_a} = \frac{1}{\tau_a s + 1}. \tag{7.3}$$

Table 7.1 Model parameters used in the numerical examples in Section 7.3.

Parameter	Unit	Value	
Amplifier bandwidth, $f_{\mathrm{amp}} = \omega_{\mathrm{amp}}/(2\pi)$	MHz	1.5	
Loop-delay, L	µs	1.0	
		Warm cavity	**Cold cavity**
Bandwidth of acc. mode, $f^a_{1/2} = \gamma_a/(2\pi)$	kHz	12	0.6
Time constant of acc. mode, τ_a	µs	13	265

where $\tau_a = 1/\gamma_a$ is the time constant of the accelerating mode, and we have assumed the detuning to be zero, to get a conjugate symmetric transfer function[1]. The assumption that the detuning is zero has little impact on field control performance, since the closed-loop bandwidth on $\approx 100\,\mathrm{kHz}$ is much greater than typical detuning. For the bandwidth γ_a we will consider two different values: 12 kHz, which corresponds to the (normal conducting) DTLs at ESS, and 0.6 kHz, which corresponds to the superconducting cavities at ESS.

We also assume that the loop phase is perfectly calibrated, i.e., $\theta_{\mathrm{adj}} = \theta$, and for simplicity, even that $\theta_{\mathrm{adj}} = \theta = 0$. The amplifier dynamics $P_{\mathrm{amp}}(s)$ is given by (4.1). The numerical parameter values that will be considered in the next section are given in Table 7.1, and an overview time delays in the field control loops at ESS is given in Table 7.2.

Disturbance spectrum

Since it has not been possible to obtain sufficient data on the disturbance spectrum $D(s)$ for the field control loops at ESS, we assume a simple, circular symmetric disturbance model, i.e., without the directionality discussed in Section 9, and with power spectral density

$$\Phi(\omega) = |D(i\omega)|^2 \quad \text{with} \quad D(s) = \frac{b(s+a)}{s(s+b)}, \tag{7.4}$$

where $a = 2\pi \cdot 3 \times 10^3$, and $b = 2\pi \cdot 3 \times 10^4$, see Figure 7.4; the large values of $|D(i\omega)|$ at low frequencies corresponds to droop of the amplifier power supply and pulse-to-pulse variations of the beam current, while the constant part up to the frequency b, corresponds to beam current ripple and fast switching in the amplifier power supply.

---

[1] Strictly speaking it still acts on complex-signals, but in the ideal case the dynamics for the real-part and the complex-parts are de-coupled.

Table 7.2 Estimated latencies in the field control loops at ESS. The RF-amplifier dynamics is modeled by a first-order system $P_{\mathrm{amp}}(s)$, instead of as a pure time delay; see Section 4.3.

Source of delay	**Delay [ns]**
Propagation Delays	
Amplifier to cavity (waveguide), $40\,\mathrm{m} \div 0.68c$	200
LLRF to amplifier (cable), $10\,\mathrm{m} \div 0.82c$	40
Cavity probe to LLRF (cable), $40\,\mathrm{m} \div 0.82c$	160
LLRF latencies	
ADC latency	130
30 FPGA clock cycles @100 MHz	300
DAC latency	90
Contingency	80
Total	1000

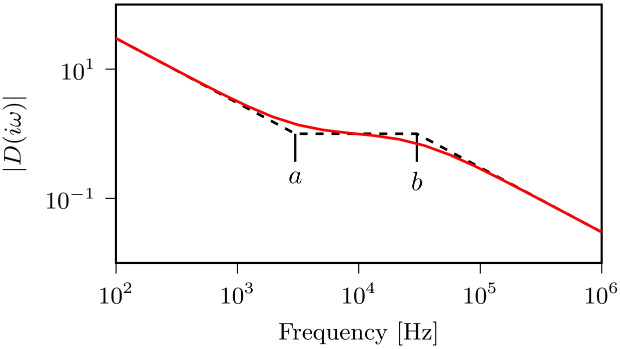


Figure 7.1 Assumed load disturbance spectrum for the numerical examples considered in this section. Since we will consider the performance of different control approaches relative to each other, the absolute scale of the disturbance spectrum is not important.

Measurement noise spectrum

We assume the measurement noise to be white, i.e., have constant magnitude up to the Nyquist frequency. Instead of specifying the spectrum $N(i\omega)$ corresponding to the actual noise level, we normalize so that $||N||_2 = 1$,

$$N(i\omega) = \begin{cases} \sqrt{2\pi/(2\omega_{\mathrm{Ny}})}, & \text{if } |\omega| \leq \omega_{\mathrm{Ny}} \\ 0, & \text{otherwise} \end{cases},$$

which makes it easier to compare different noise levels. We assume the Nyquist frequency $\omega_{\mathrm{Ny}} = 5\,\mathrm{MHz}$, which corresponds to a sampling rate of 10 MHz for the controller. The assumption on white noise will allow us to use a numerically preferable problem formulation to find the optimal LTI-controller.

7.3 Results

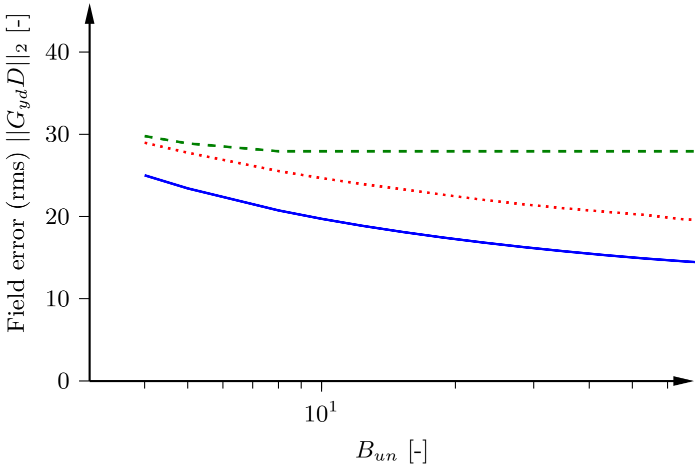
We consider the following nominal case: the time delay $L = 1\,\mu\mathrm{s}$, the maximum sensitivity $S_{\max} = 1.6$, and maximum control signal activity $B_{un} = 30$; for this nominal a case, a comparison of the open-loop frequency responses and the gang-of-fours, for the considered controller structures are shown in Figure 7.5 and Figure 7.4.

How the optimal control performance depends on the allowed control signal activity B_{un} is shown in Figure 7.2, and how the performance depends on the time delay L is shown in Figure 7.3.

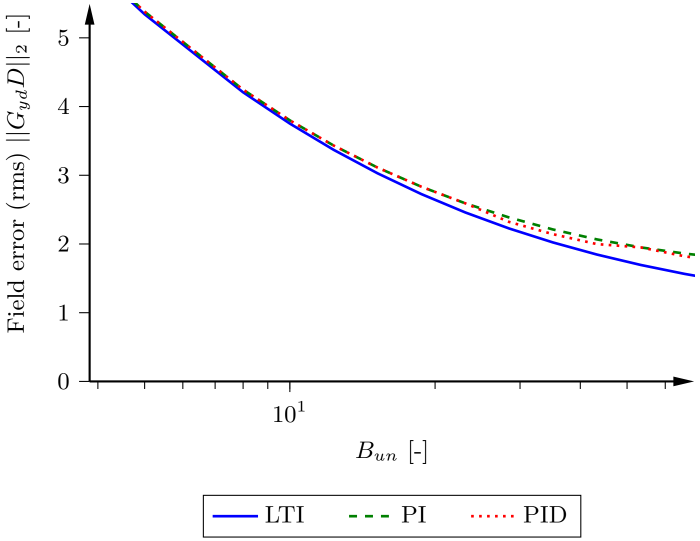
Remark

The sensitivity constraint of $M_S \leq 1.6$ is motivated by that $M_S \leq 1.4$ is commonly used for process control [Åström and Hägglund, 2006], where it is hard to obtain good process models, the controllers should function robustly even if there are significant process variations, and where there are thousands of control loops to tune and it does not make sense to push the performance of each and every one of them. The typical field control loop has rather simple dynamics (apart from the amplifier nonlinearity), and for accelerators that cost billions of euros, it is economically feasible, and justified, to spend some time on obtaining good process models. This makes it reasonable with a slightly larger values of M_S, which will enable better control performance.

The constraint on B_{un} can be motivated as follows. For ESS it has been estimated that the phase-adjusted baseband signal $\boldsymbol{y}$ will have a signal-to-noise ratio of 70 dB, which corresponds to relative rms variations on 0.03 %. The constraint $||G_{un}N||_2 \leq B_{un} = 30$ ensures that the control signal has relative rms variations of less than 1 %.

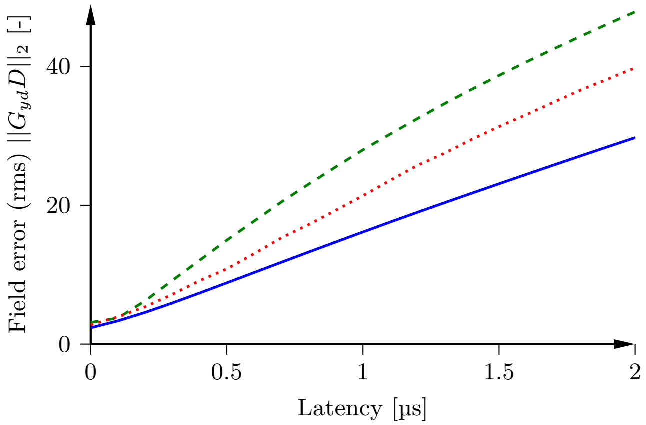


(a) Warm cavity.

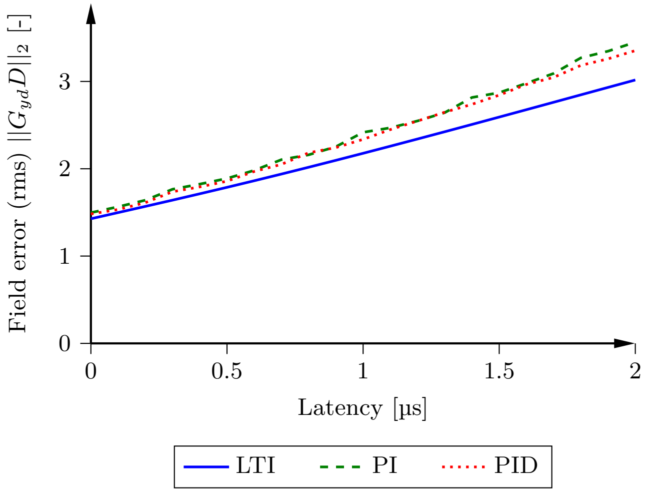


(b) Cold cavity.

Figure 7.2 Optimal control performance vs. allowed control signal activity B_{un}, for different controller types. For a small value of B_{un} it is not possible for the general LTI controller to use its flexibility to suitably shape the open loop system, since this requires a large control signal. With greater values of B_{un}, the performance advantage of the general LTI-controller over the PI- and PID-controllers increases.

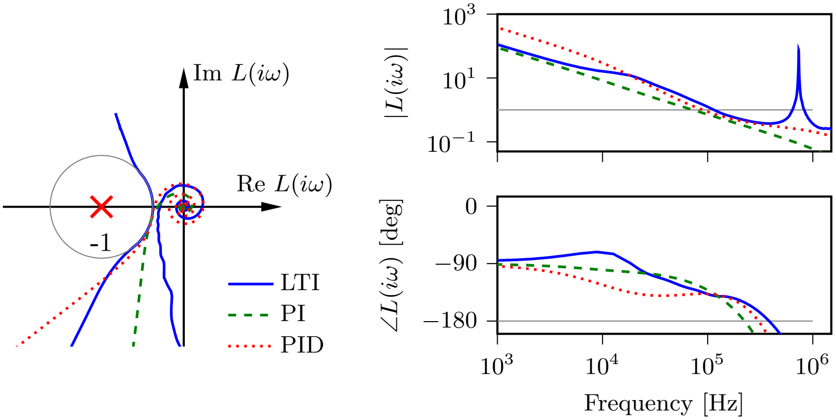


(a) Warm cavity.

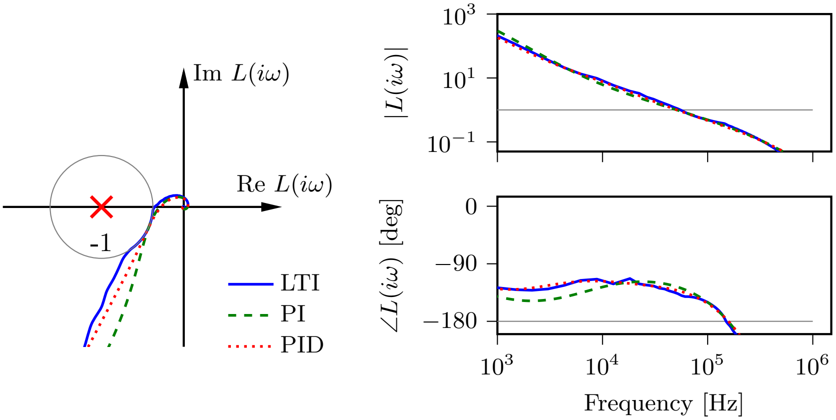


(b) Cold cavity.

Figure 7.3 Optimal control performance vs. time delay L in the field control loop, with the controller constraints given by $S_{\max} = 1.6$ and $B_{un} = 30$. It is seen that the achievable control performance, both for the warm and the cold cavity, can be significantly improved by reducing the time delay in the field control loop; the effect is most dramatic for the warm cavity, where the control performance is almost proportional to the time delay, this is because the main limit on the control performance of the warm cavity is the robustness constraint $M_S \leq 1.6$, which becomes much easier to handle for shorter time delays, for the cold cavity, the constraints on control signal activity does not allow a reduced time-delay to be fully exploited.

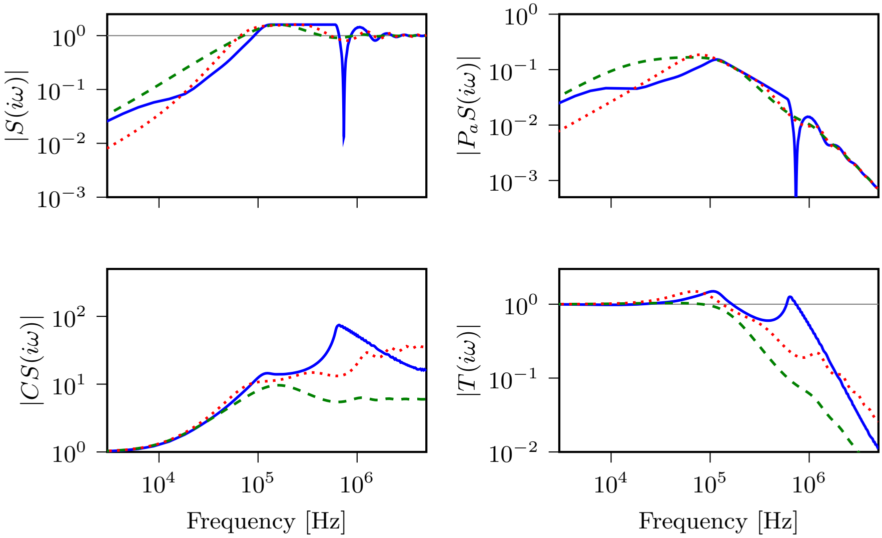


(a) Warm cavity. The Nyquist curve for the LTI-controller makes a big turn into the right half plane after it exits the top of the figure, before finally coming back and goes to zero for high frequency.

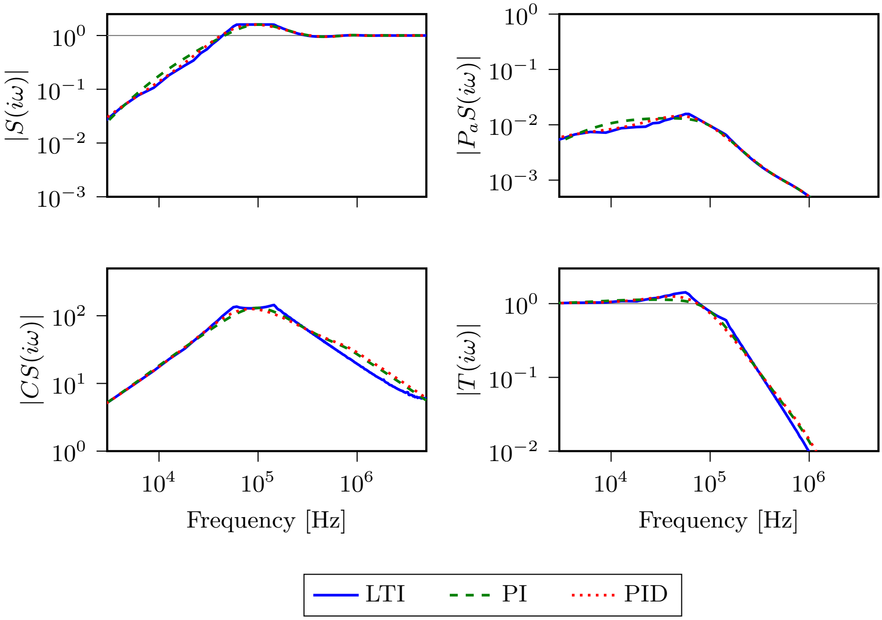


(b) Cold cavity.

Figure 7.4 Comparison between the open-loop frequency responses for three different controller types.



(a) Warm cavity.



(b) Cold cavity.

Figure 7.5 The gang of four for different optimal controllers subject to $S_{\max} = 1.6$ and $B_{un} = 30$.

7.4 Conclusions

From the Bode diagrams in Figure 7.4 we see that the open-loop transfer functions for all the different controllers are similar, and all have ”typically good” characteristics for open-loop transfer functions: high gain at low frequencies, and roll-off for higher-frequencies. However, the optimal LTI controllers do a better job of pushing up the loop gain at around 30 kHz, giving better disturbance rejection around that frequency (see also the sensitivity function in Figure 7.5); this is made possible by the high flexibility of the LTI controllers which allow them to push the sensitivity function all the way to the sensitivity constraint $M_S \leq 1.6$. With reference to the waterbed effect (Bode’s integral formula Section 2.2), one can say that the LTI controller corresponds to a softer waterbed than PI- and PID-controller. The same effect is seen in Figure 7.4 where the Nyquist curves for the LTI-controllers tightly follow the sensitivity circles.

In order to both push up the loop gain at around 30 kHz while still respecting the sensitivity constraint, the optimal LTI for the warm cavity makes a big turn into the right half-plane before finally going to zero for high frequencies; this big turn corresponds to the resonance peak in the Bode diagram. To implement such a controller in practice is very challenging, and it would be very sensitive to implementation errors and process variations.

Dependence on allowed control signal activity

From Figure 7.2 we see that control performance is strongly dependent on the allowed control signal activity B_{un}, especially for cold cavities. Cold cavities have low bandwidths so it is necessary to use a high-gain controller to increase the bandwidth of the closed-loop system; the high gain makes the controllers sensitive to measurement noise, and the constraint on control signal activity leaves little freedom for the general LTI controller to achieve better performance than the PI(D)-controllers.

With the given constraints on control signal activity there is little benefit from using a more complex controller than a PI-controller for controlling cold cavities. For warm cavities there are some benefits of using a PID-controller or a more complex controller over a PI-controller. Note that decreased levels of measurement noise directly leads to a corresponding allowed increase of B_{un}, and thus quite significant performance improvements.

Dependence on the loop delay

Figure 7.3 shows how the control performance depends on the time delay in the field control loop. It is seen that it is quite worthwhile to decrease it, in particular for warm cavities, for which the field error is almost proportional to the loop delay.

8

Field Control for Parasitic Cavity Modes

8.1 Introduction

We saw in Chapter 5 that RF cavities in addition to the accelerating mode, have infinitely many parasitic modes. The interaction of these modes with the beam and the field control loop is detrimental to beam quality and control performance.

Typically only one, or possibly two, of the parasitic modes are sufficiently close to RF frequency to require consideration in the field control design, see Figure 8.1. The beam interacts not only with the parasitic modes close to the RF frequency, but also with modes that are close to the *machine lines*, i.e., integral multiples of the bunch frequency. These interactions need consideration to ensure sufficient beam quality [Ainsworth and Molloy, 2012], but for the purpose of this thesis, we only need to consider the parasitic modes close to the RF frequency.

An early work which considered parasitic modes in the field control design is [Schilcher, 1998]. More details and discussions of different control strategies are provided in [Vogel, 2007]. For an example of how parasitic modes were handled in connection with robust control design, see [Pfeiffer et al., 2012; Schmidt et al., 2012], and for system identification of the parasitic mode frequencies, see [Pfeiffer, 2014].

In this chapter we will present and compare different approaches for dealing with parasitic modes in the control design, demonstrating how the complex-coefficient SISO representation of dynamics introduced in Appendix E, brings insight to the control design and analysis, avoiding the need to resort to the generalized Nyquist criterion [Schilcher, 1998], or to ignore the imaginary part of the cavity transfer function (which is acceptable for the accelerating mode, but it is not an adequate approximation for the parasitic modes) [Vogel, 2007].

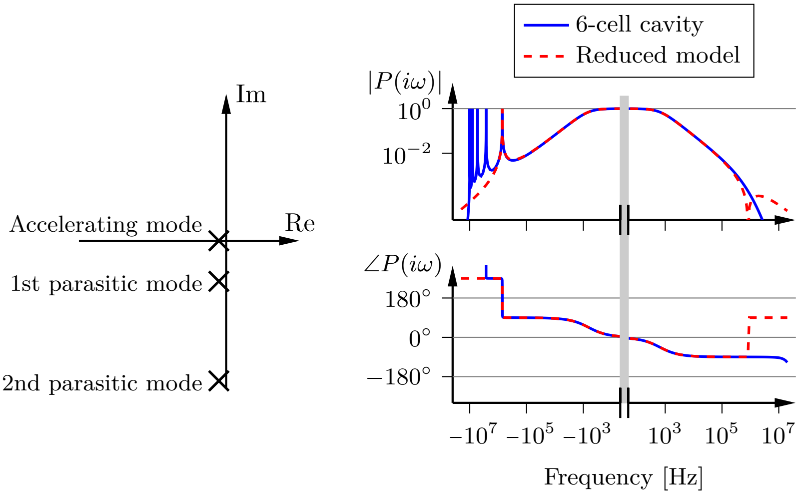


Figure 8.1 Pole-zero diagram and Bode diagram for the (baseband) transfer function from the power coupler to the measurement probe of a 6-cell elliptical cavity. Only the two parasitic modes closest to the accelerating mode are shown in the pole-zero diagram.

Illustration of the dangers of parasitic modes

Consider the process model $P(s)$ from Section 7.2, where the cavity model *does not* include parasitic modes, controlled by the PI-controller (7.1) with $K = 85$, $T_i = 24\,\mu\text{s}$, and $T_f = 0.48\,\mu\text{s}$; frequency domain plots for this process and controller combination are shown in Figure 8.2.

If the same controller $C_0(s)$ is used when the cavity model is taken to include parasitic modes (corresponding to the red curve of Figure 8.1), we get the Nyquist curve in Figure 8.3 that encircles the point -1, demonstrating that the closed-loop system is unstable. This simple example illustrates that parasitic modes need to be considered in the control design.

8.2 Process model

Throughout this chapter we will consider the RF system model

$$P(s) = P_{\text{cav}}(s)P_{\text{amp}}(s)\mathrm{e}^{-sL}\mathrm{e}^{-i\theta},$$

from (6.1), with the same parameters as in Section 7.2 ($L = 1\,\mu\text{s}$, $\theta = 0$), but include a *single* parasitic mode in the cavity model, i.e.,

$$P_{\text{cav}}(s) = \frac{\gamma_a}{s + \gamma_a - i\Delta\omega_a} - \frac{\gamma_1}{s + \gamma_1 - i\Delta\omega_1}. \tag{8.1}$$

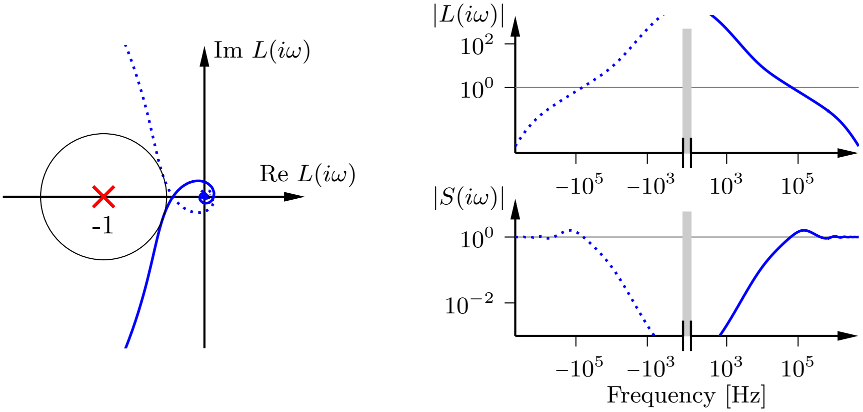


Figure 8.2 Frequency domain plots for a nominal controller design for a model of a cold cavity without parasitic modes.

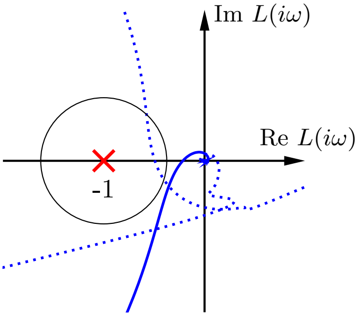


Figure 8.3 Nyquist curve, for when a controller designed for a cavity model *without* parasitic modes, is used for a cavity *with* parasitic modes. The bulge from the parasitic mode makes the Nyquist curve encircle -1 and the closed loop system is therefore unstable due to the effect of the parasitic mode. This illustrates that the parasitic modes needs to be considered in the controller design.

Table 8.1 Parameters used in the cavity model with one parasitic mode (8.1).

Parameter	**Unit**	**Value**
Accelerating mode		
Bandwidth, $f^a_{1/2} = \gamma_a/(2\pi)$	Hz	600
Detuning, $\Delta f_a = \Delta\omega_a/(2\pi)$	Hz	0
1st parasitic mode		
Bandwidth, $f^1_{1/2} = \gamma_1/(2\pi)$	Hz	1100
Detuning, $\Delta f_1 = \Delta\omega_1/(2\pi)$	kHz	−700

For our numerical examples we will use parameter values in Table 8.1, which are typical for ESS's medium-beta cavities[1]. The dashed, red lines of Figure 8.1 correspond to these parameter values. In practice it could be necessary to consider more than one parasitic mode, but to simplify the exposition, we limit ourselves to only one.

8.3 Control strategies for parasitic modes

To avoid instability from the parasitic mode, we want to ensure that the corresponding ”bulge” in the Nyquist curve $L(i\omega) = C(i\omega)P(i\omega)$, for ω close to $\Delta\omega_1$, does not encircle, or come close to the point -1. This can be achieved in two ways:

1. By sufficient *attenuation* of the parasitic mode so that the open-loop transfer function $L(s)$ is well below 1 around this frequency. Due to the normalization of $P_{\text{cav}}(s)$, this corresponds to that $|C(i\Delta\omega_1)|$ is well below 1 around $\Delta\omega_1$.

2. By using the controller $C(s)$ for *phase adjustment* of the open-loop transfer function $L(s)$ around $\Delta\omega_1$, so that the ”bulge” from the parasitic mode is directed away from the -1 point.

For option 1, attenuation, the phase of $L(i\omega)$ around $\Delta\omega_1$ is not important for stability, and thus the closed loop is robust to for example small variations of the loop delay. Option 2, on the other hand, where the bulge from the parasitic mode is kept large, is sensitive to phase variations around the parasitic mode frequency $\Delta\omega_1$.

---

[1] Except that we assume the detuning of the accelerating mode to be 0, since we did so in the previous section.

Attenuation

Below we discuss three different strategies to sufficiently attenuate the parasitic mode to ensure closed-loop stability: keeping the gain of the controller low, augmenting the controller with a high-order low-pass filter, and augmenting the controller with a notch filter.

Low gain Of course, one possibility would be to just use the PI-controller (7.1), and reduce the gain K and the bandwidth of the low-pass filter. While approach avoids instability, it would give poor control performance, since the low-frequency gain would be too low to suppress slow disturbances.

High-order low-pass filter Augmenting the controller with low-pass filters of the form

$$F_1(s) = \frac{1}{sT_f + 1}$$

and

$$F_2(s) = \frac{1}{s^2T_f^2 + sT_f + 1},$$

attenuates not only the closest parasitic mode with frequency $\Delta\omega_1$, but actually all parasitic modes. This strategy is not sensitive to the exact frequency of the parasitic mode.

Notch filter To specifically reduce the gain of $C(s)$ around $\Delta\omega_1$, the obvious choice would be to introduce a notch filter

$$\begin{aligned} F_n(s) &= \frac{s^2 + \omega_n^2}{s^2 + 2\zeta_n\omega_n s + \omega_n^2} \\ &= \frac{s - i\omega_n}{s + \zeta_n\omega_n - i\omega_n\sqrt{1-\zeta_n^2}} \cdot \frac{s + i\omega_n}{s + \zeta_n\omega_n + i\omega_n\sqrt{1-\zeta_n^2}}. \end{aligned} \tag{8.2}$$

in the controller, as in [Schmidt et al., 2012]. This notch filter has real coefficients and is thus conjugate symmetric, however from Figure 8.1 it is seen that the resonance from the parasitic mode is only present at negative baseband frequencies. Therefore it suffices to use a single-sided notch filter of the form

$$F_n(s) = \frac{s - i\omega_n}{s + \zeta_n\omega_n - i\omega_n}, \tag{8.3}$$

which requires slightly less FPGA resources for the implementation, and also gives slightly better control performance, since the phase reduction around cross-over could be somewhat reduced (at least after appropriate adjustment of θ_{adj}).

If the notch filter is designed to only reduce the magnitude of $C(\Delta\omega_1)$, the resulting design will be sensitive to the exact frequency of the parasitic mode.

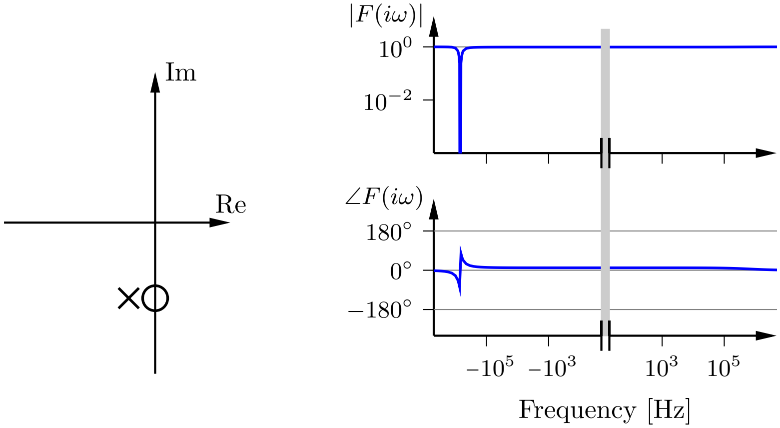


Figure 8.4 Pole-zero diagram for a single-sided notch filter F in continuous time.

A design criterion, which gives better robustness to this, is the following

$$|C(i\omega)| \leq 1 - \frac{1}{S_{\max}} \text{ for } |\omega| \in [\Delta\omega_1 - b, \Delta\omega_1 + b]\,, \tag{8.4}$$

where a low magnitude of C is enforced also in a neighborhood of $\Delta\omega_1$.

Remark on alternative notch filter implementation Another type of notch filter that is advantageous due it its low implementation complexity is [Vogel, 2007]

$$H(z) = \frac{1}{2}\left(1 + z^{-n}\right), \tag{8.5}$$

The filter (8.5) gives a notch at the frequency $1/(2nh)$, but there is also a version which allows adjustment of the notch frequency [Vogel, 2007]. A disadvantage of the filter (8.5) is that it does not allow adjustment of the *width* of the notch, like the notch filter in (8.3) does. Thus the notch filter (8.5) may give an unnecessarily large phase drop in the control loop; it is also not direct to obtain a one-sided version of 8.5

Phase adjustment

In order to adjust the exact phase of the parasitic mode and still keep a good behavior of the controller around the cross-over frequency and below, it is beneficial with a flexible controller structure.

8.4 Numerical comparison

We will consider the same control problem (6.7) as in the previous chapter, with the process model from Section 8.2, and with the disturbance and measurement noise spectra from Section 7.2. We will also consider the same nominal constraints as in Chapter 7, i.e., $S_{\max} = 1.6$ and $B_{un} = 30$.

For the cases when we want to attenuate the parasitic modes, we will augment the control problem with the constraint

$$|C(i\Delta\omega_1)| \leq 1 - \frac{1}{S_{\max}},$$

which guarantees that Nyquist curve stays outside of the M_S-circle regardless of its phase around the frequency $\Delta\omega_1$.

The approach to find the optimal controller parameters was the same as for the PI- and PID-controllers in the previous chapter, see Section B.2.

Controller structures

We will consider controllers of the form

$$C(s) = K\left(1 + \frac{1}{sT_i}\right) F(s) \mathrm{e}^{i\theta_{\mathrm{adj}}} \tag{8.6}$$

where the filter $F(s)$ is chosen as either:

1. A first-order low-pass filter $F_1(s) = \dfrac{1}{1 + sT_f}$.
2. A second-order low-pass filter, $F_2(s) = \dfrac{1}{1 + sT_f + s^2T_f^2}$.
3. A third-order low-pass filter $F(s) = F_3(s) := F_1(s)F_2(s)$.
4. A fourth-order low-pass filter $F(s) = F_4(s) := F_2(s)F_2(s)$.
5. A first-order low-pass filter + a notch filter $F(s) = F_1(s)F_n(s)$.
6. A second-order low-pass filter + a notch filter $F(s) = F_2(s)F_n(s)$.

The reason for including the factor $\mathrm{e}^{i\theta_{\mathrm{adj}}}$ in (8.6) is that the process transfer function $P(s)$ has complex coefficients when parasitic modes are included in the cavity model, and hence it is typically suboptimal to choose $\theta_{\mathrm{adj}} = 0$ as we did in the previous chapter.

Table 8.2 The performance for different approaches of dealing with parasitic modes in the control design. The table is normalized with respect to the performance for a cavity model without parasitic modes controlled by a PI-controller with a first-order low-pass filter.

Type	Controller Structure	Objective
Baseline, no PM	PI+1st order filter (Fig. 8.2)	1.00
PM attenuation	PI+1st order filter (Fig. 8.5)	5.94
	PI+2nd order filter (Fig. 8.6)	3.27
	PI+2nd order filter ($\theta_{\text{adj}} = 0$, Fig. 8.7)	3.31
	PI+3rd order filter	2.79
	PI+4th order filter	2.56
	PI + notch filter + 1st order filter	1.25
	PI + notch filter + 2nd order filter	1.25
	PI + double notch filter + 2nd order filter	1.30
	PI + wide notch filter + 1st order filter (Fig. 8.8)	2.11
PM phase adjust	PI+1st order filter	5.71
	PI+2nd order filter	2.70
	PI+3rd order filter ($\theta_{\text{adj}} = 0$)	1.96
	PI+3rd order filter (Fig. 8.9)	1.93

Results

For the case where we want to limit controller gain we consider all the filters 1–6 that are listed above, and for the case of phase adjustment we consider the filters 1–4.

In addition, we consider a few minor variations to the control structure: enforcing $\phi_{\text{adj}} = 0$, enforcing the wide-notch constraint (8.4), and consider a double sided notch filter instead of a single-sided one.

The field errors for the different designs, relative to a baseline situation without parasitic modes (i.e., the model in Chapter 7), are given in Table 8.2. For some controller designs we have provided illustrative frequency domain plots, see figures 8.5 to 8.9. A frequency domain comparison between four different control approaches is provided in Figure 8.10.

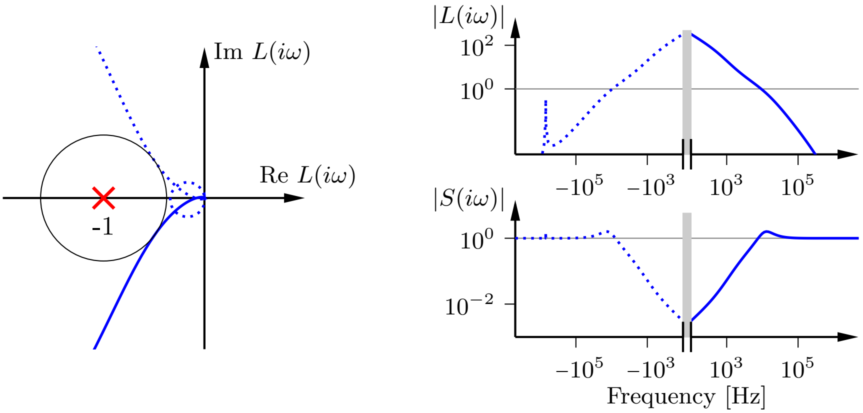


Figure 8.5 Attenuation of parasitic mode, PI-controller + 1st order filter.

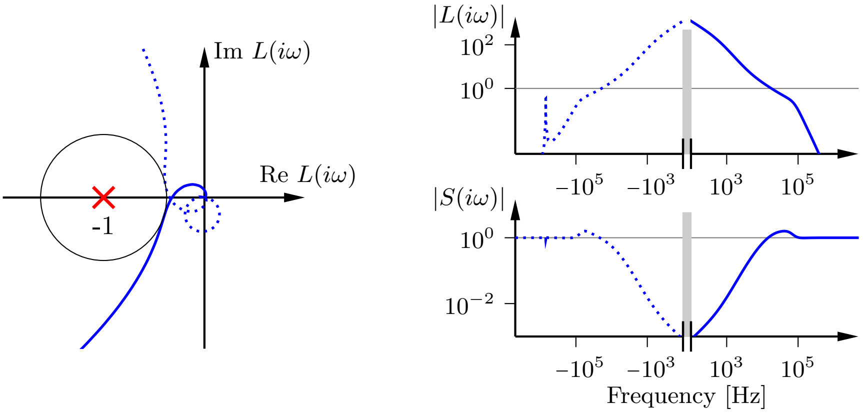


Figure 8.6 Attenuation of parasitic mode, PI-controller + 2nd order filter.

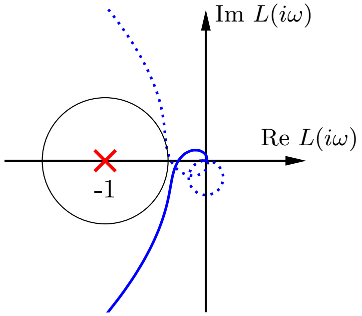


Figure 8.7 Attenuation of parasitic mode, PI-controller + 2nd order filter, with $\phi_{\mathrm{adj}} = 0$.

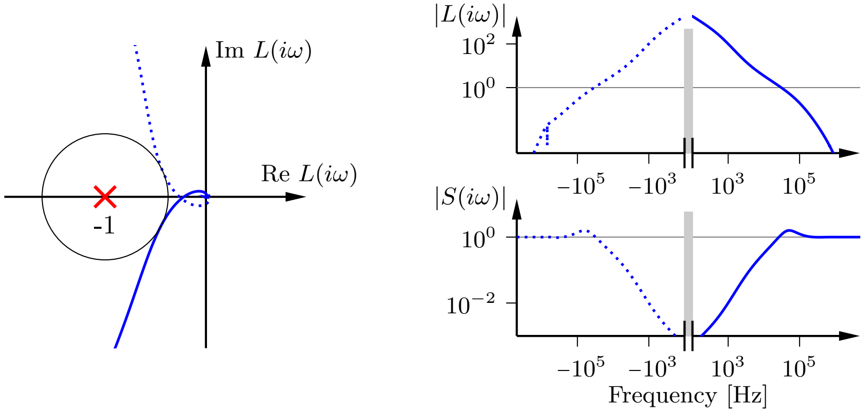


Figure 8.8 Attenuation of parasitic mode, PI-controller + wide notch filter.

8.5 Conclusions

Parasitic cavity modes that are close in frequency to the accelerating mode affects the control in that the low-frequency controller gain needs be reduced to maintain the desired robustness level. The performance degradation from parasitic mode depends strongly on the strategy chosen for dealing with them; see Table 8.2.

One simple approach to deal with parasitic modes is to design robust controller for a cavity model without parasitic mode, and then simply augment it with a notch filter. A perhaps better approach would be to include a notch filter with desired characteristics in the process model, after which standard design methods could be applied. A main drawback with a notch filter approach is that the frequency of the parasitic mode needs to be known

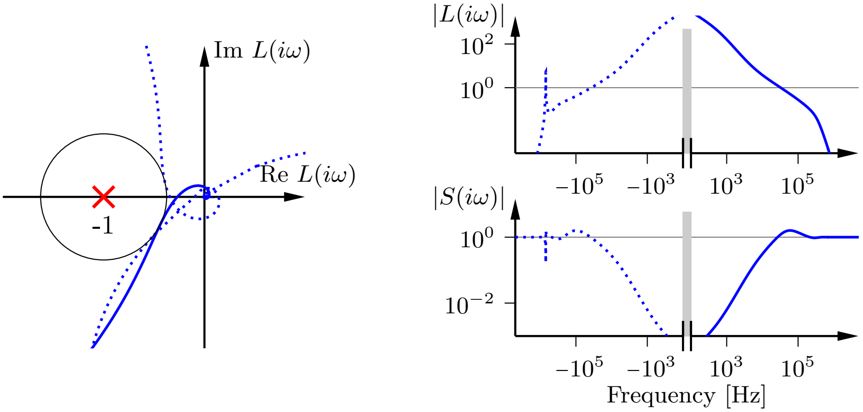


Figure 8.9 Phase adjustment of parasitic mode, PI-controller + 3rd order filter; note the big "bulge” in the Nyquist curve due to that the parasitic mode is not attenuated, if the phase at the parasitic mode frequency changes, the bulge might encircle the critical point -1, leading to instability.

with reasonable accuracy.

If a notch filter is not used, the consequences of the parasitic mode is that the bandwidth of the control loop must be reduced to maintain the specified robustness level, see Figure 8.10. This leads to reduced attenuation of low-frequency disturbances. If this is a problem or not depends on the disturbance spectrum. Also it could happen that the constraint on control signal activity also enforces a relative low bandwidth, in this case the penalty for choosing a different approach than the notch filter is not so great.

If the field error specification is easy to meet, the method of choice should be to use a high-order low-pass filter since this is both easier to implement and more robust than a notch filter.

This has been far from an exhaustive investigation, the conclusions are affected by many factors such as the considered controller structures, the assumed spectrum of load disturbances and the control specifications. We have also not considered time domain response of the system, this is important to consider since the convergence time of sharp notch filters could be long relative to the pulse length. However the results presented in this chapter gives the general ideas and an overview of different approaches to handle parasitic modes in the control design.

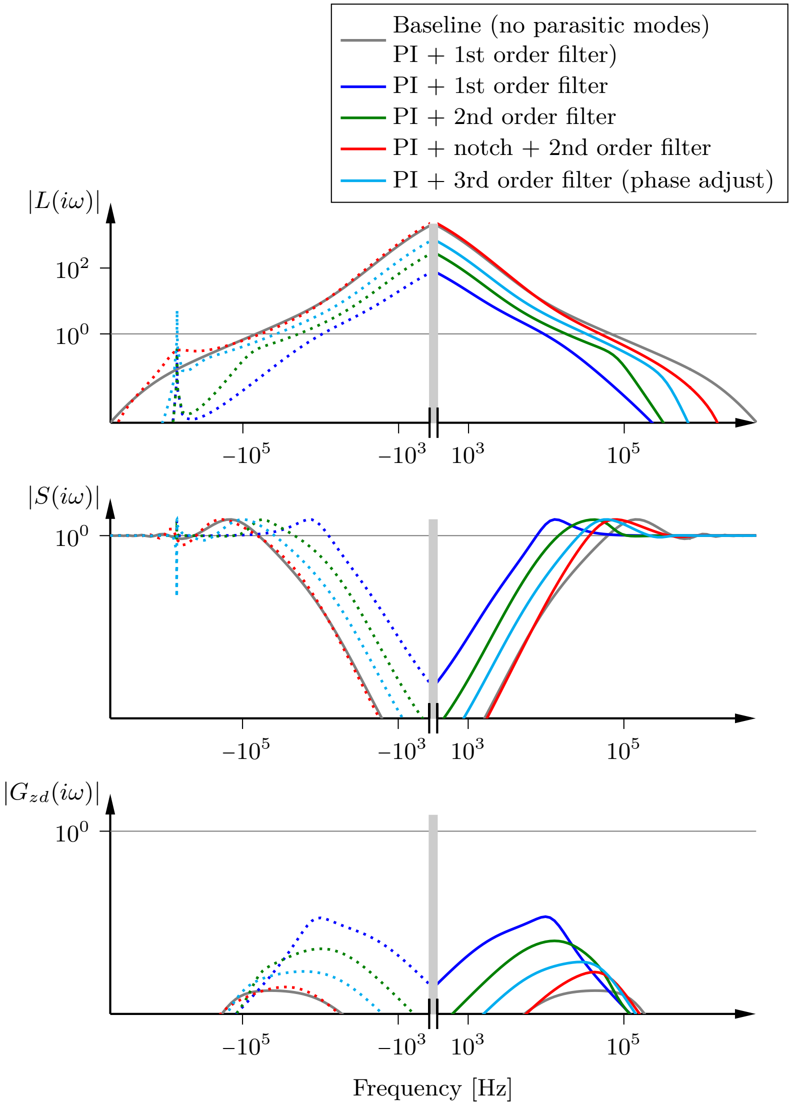


Figure 8.10 Comparison between different control strategies to avoid instability in the field control loop due to parasitic modes. It is seen that the achievable control bandwidth is dependent on the type of controller used, the controller with a notch filter achieves the highest bandwidth; whether a lower bandwidth is a problem or not, depends on the spectrum of the load disturbances.

9

Directionality in the Control Problem

In the previous chapters, we modeled the dynamics of the cavity and RF system by a single-input single-output complex-coefficient system (under the assumption that imperfections in the vector-modulator , e.g., skew, are inverted by the controller)

$$P(s) = P_{\text{amp}}(s)P_{\text{cav}}(s)\mathrm{e}^{-sL}\mathrm{e}^{-i\theta}.$$

This representation implies that the dynamics of the plant is rotationally invariant in the baseband, i.e., a rotation of the input signal gives a corresponding rotation of the output signal,

$$P(i\omega)\left[\mathrm{e}^{i\rho}U(i\omega)\right] = \mathrm{e}^{i\rho} \cdot P(i\omega)U(i\omega).$$

this rotational invariance is natural and corresponds to that the original system is time invariant.

However, although the plant dynamics is rotationally invariant, the *true* objective function that quantify the beam losses (Ψ_{BL} from Section 3.4), and the disturbances acting on the system, are not; we will now discuss this directionality and its consequences.

Effect of directional input on complex-coefficient system

In this chapter we will consider how the output of a complex-coefficient SISO system[1] $G(s) = G_{\text{Re}}(s) + iG_{\text{Im}}(s)$, is affect by signals with a specific direction in the complex-plane, i.e., of the form

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 \cdot m(t) \qquad \text{where} \qquad m(t) \in \mathbb{R}, \text{ and } \boldsymbol{u}_0 \in \mathbb{C}.$$

---

[1] How to separate $G(s)$ in $G_{\text{Re}}(s) + iG_{\text{Im}}(s)$ is shown in footnote 3 on page 31.

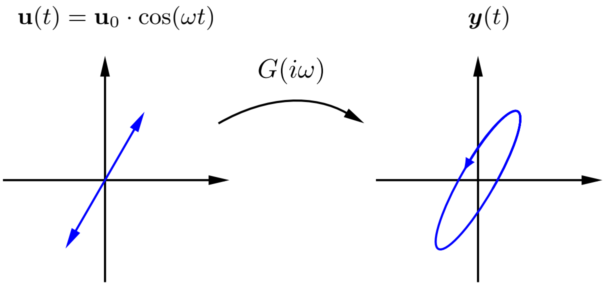


Figure 9.1 The output signal of a complex-coefficient system driven by a sinusoidal signal with a specific direction in the complex plane is given by a Lissajous oval in the complex plane, see (9.1).

To understand this it is helpful to consider the case when $m(t)$ is a real-valued sinusoidal signal,

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 \cdot \cos \omega t = \boldsymbol{u}_0 \cdot \frac{1}{2}(\mathrm{e}^{i\omega t} + \mathrm{e}^{-i\omega t}),$$

in which case, the output after transients becomes

$$\begin{aligned} \boldsymbol{y}(t) &= \boldsymbol{u}_0 \cdot \frac{1}{2}(G(i\omega)\mathrm{e}^{i\omega t} + G(-i\omega)\mathrm{e}^{-i\omega t}) \\ &= \boldsymbol{u}_0 \cdot \Big[A_{\mathrm{Re}} \cos(\omega t + \phi_{\mathrm{Re}}) + iA_{\mathrm{Im}} \cos(\omega t + \phi_{\mathrm{Im}})\Big], \qquad (9.1) \end{aligned}$$

where $A_{\mathrm{Re}} = |G_{\mathrm{Re}}(i\omega)|$, $\phi_{\mathrm{Re}} = \angle G_{\mathrm{Re}}(i\omega)$, $A_{\mathrm{Im}} = |G_{\mathrm{Im}}(i\omega)|$ and $\phi_{\mathrm{Re}} = \angle G_{\mathrm{Im}}(i\omega)$. From (9.1) we can now see that $\boldsymbol{y}(t)$ traces out a to a Lissajous oval in the complex plane; this is illustrated in Figure 9.1. In the special case of a transfer function with real coefficients, the direction of output signal is the same as for the input signal.

However, if $G(s)$ is given by a closed-loop transfer function of the field control loop, we typically have, at frequencies that are well greater than the detuning $\Delta\omega_a$, and well smaller than the frequency of the closest parasitic mode $\Delta\omega_1$, it holds that $G(i\omega)$ is quite close to being conjugate symmetric; thus, we have that $G_{\mathrm{Im}}(i\omega) = (G(i\omega) - G^*(-i\omega))/2i$ is rather small, making it a reasonable approximation that the effect of a disturbance on the cavity field is in the same direction as the disturbance.

Directionality of disturbances

In the case the RF amplifier is a klystron, we saw in equation (4.3) that the ripple from the AC-DC converter gives rise to a disturbance $\boldsymbol{K_g d_g} \approx \boldsymbol{K_g} i\xi\Delta_c$ on the cavity field; since the direction of $\boldsymbol{K_g}$ is the same as for $\mathbf{F^0_g}$, we have that the disturbance $\boldsymbol{K_g d_g}$ affects the cavity field in a direction perpendicular to $\mathbf{F_g^0}$.

Assuming that the only variations in $\mathbf{I_b}$ are from the beam current variations (due to ion source instabilities)[2], then only the amplitude of $\mathbf{I_b}$ changes, and thus $\boldsymbol{d_b}$ is real-valued, and the disturbance $\boldsymbol{K_b d_b}$ acts in the direction of $\mathbf{I_b^0}$.

To get insight in the directionality of the control problem it will will be helpful to consider the phasor diagrams that were introduced in Section 5.6; see Figure 9b for an illustration of the directions of the cavity field disturbances.

Directionality of the objective function

The requirements on cavity field errors are typically specified in terms of the allowed rms error of the cavity field amplitude $V_a = |\mathbf{V}_a|$ (in %) and the cavity phase $\phi_a = \angle \mathbf{V}_a$ (in °) [Altarelli, 2007; Peggs et al., 2013; Doolittle, 2015]. However, the actual quantities which ultimately affect the beam losses are: (1) the acceleration of the particle bunches, which is proportional to $V_a \cos(\phi_b - \phi_a)$; and (2) the longitudinal focusing of the particle bunches, which is proportional to $V_a \sin(\phi_b - \phi_a)$,

$$\begin{aligned}
&\text{Acceleration/} \\
&\text{Transferred Energy} \propto \quad V_a \cos(\phi_b - \phi_a) \propto \quad \operatorname{Re}\{\mathbf{V_a} \cdot \mathbf{I_b^*}\} && (9.2\text{a}) \\
&\text{Longitudinal Focusing} \propto \quad V_a \sin(\phi_b - \phi_a) \propto \quad \operatorname{Im}\{\mathbf{V_a} \cdot \mathbf{I_b^*}\} && (9.2\text{b})
\end{aligned}$$

The directionality of the control objectives (9.2) is illustrated in Figure 9a.

In Section 3.5 we discussed that the true objective function for high-intensity proton accelerators is a complicated nonlinear function $\Psi_{BL}(\Delta A_1, \Delta\phi_1, \Delta A_2, \Delta\phi_2, \ldots)$. This objective function is is too cumbersome to evaluate to be useful for control design; but specifying the errors on acceleration and the longitudinal focusing, should reflect the true objective Ψ_{BL} better, than if errors on amplitude and phase are considered, as is traditionally done for particle accelerators.

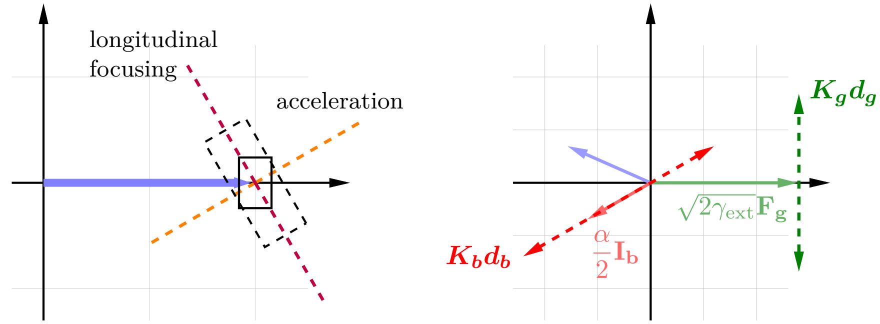
To allow more flexibility in the control design, one could consider an objective function of the form

$$J(\mathbf{e}) = \begin{bmatrix} \operatorname{Re} e & \operatorname{Im} e \end{bmatrix} Q_J \begin{bmatrix} \operatorname{Re} e \\ \operatorname{Im} e \end{bmatrix}$$

where Q_J is a symmetric, positive definite matrix, and $\mathbf{e} = \boldsymbol{z}_0 - \boldsymbol{z}$; this type of objective function would correspond to an ellipsoidal constraint in Figure 9a.

---

[2] There will of course also be variations in the arrival time of the particle bunches, which corresponds to phase variations, but these should be small if the field control in the upstream cavities is working satisfactorily.



(a) Directionality of the objective function. (b) Directionality of disturbances.

Figure 9.2 *Left:* The error directions that affect acceleration and focusing of the particle bunches. The solid rectangle correspond to a specficaition in terms of amplitude and phase errors. The dashed rectangle shows a specification in terms of acceleration and longitudinal focusing, which is seen to relax the error specification in the direction for longitudinal focusing. *Right:* Illustration of the directions that modulator ripple (green dashed line), and the beam current ripple (red dashed line) affect the cavity field in.

Why bother about this directionality?

Assume that the beam losses are significantly more dependent on the acceleration experienced by the particle bunches than the longitudinal focusing that they experience. In that case, we can from Figure 9 draw the conclusion that it is more worthwhile to improve the stability of the ion source, rather than to reduce the amplifier ripple.

Comment on error specifications of the form x%/x°

It should be pointed out that while a specification of the form $x\%/x^\circ$ is convenient to remember, it is actually not symmetric in the complex plane. Since $1\,\% = 0.01$ and $1^\circ = 0.017\,\text{rad}$ we notice that the phase errors are allowed to be significantly larger than the amplitude errors, see Figure 9.3. Little discussion of this can be found in the literature, and while $1\,\%$ and 1° are *roughly* of the same magnitude, there is indeed a factor 1.7 difference.

Without any prior knowledge of the effect of amplitude and phase disturbances, a symmetric error would be more natural. However, at least for free-electron-laser accelerators, only the acceleration is important [Mosnier and Tessier, 1994, Appendix C], and since ϕ_b is small for electron accelerators, it is much more important to keep the amplitude errors small. Thus for FEL accelerators, it not unreasonable to allow the phase errors to be larger than the amplitude errors. However, as we have already discussed, the situation is very different for high-intensity proton accelerators where the

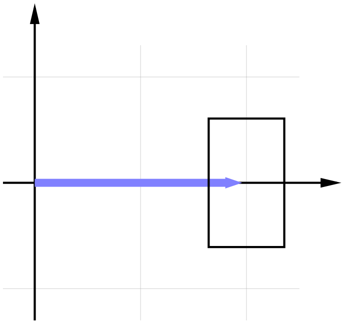


Figure 9.3 Illustration of a requirement on the form (x%, x°) in the IQ-plane. Note that the requirement is asymmetric.

longitudinal focusing is critical, and it is not clear that a specification of the form x% and x° is the most suitable.

Discussion

For a rotationally invariant, two-input two-output system

$$P(s) = \begin{bmatrix} P_{\mathrm{Re}}(s) & -P_{\mathrm{Im}}(s) \\ P_{\mathrm{Im}}(s) & P_{\mathrm{Re}}(s) \end{bmatrix},$$

with rotationlly invariant control objective and control constraints, it holds that also the optimal controller $C(s)$ is rotationally invariant[3].

A direct consequence of the directionality of the disturbances of the objective function for the field control loop is that the optimal field controller is *not* roationally invariant. For most practical applications however, the simplicity of the complex SISO-approach should outweigh the performance gains of a general two-input two-output controller.

---

[3] This can be shown by using the same idea as in the proof of Theorem 9 in [Bamieh et al., 2002]; note that the framework presented in the paper is not directly applicable to rotationally invariant systems.

10

Energy-Optimal Cavity Filling

This chapter is based on [Troeng and Bernhardsson, 2017].

10.1 Introduction

For pulsed particle accelerators, the energy required to *fill* the cavities (i.e., build up the electromagnetic fields), is significant but does not contribute to particle acceleration. For the ESS accelerator which has 2.86 ms long beam pulses at 14 Hz, it takes 150 µs to fill the superconducting cavities in the high-β section, which amounts to a yearly cost of 100 k€[1] [Peggs et al., 2013].

Currently, the most common RF amplifier for high-power particle accelerators is the klystron, which is somewhat inefficient since its power consumption is constant, regardless of output power. For the 84 cavities in the high-β section of the ESS accelerator, the industry has been encouraged to design an Inductive Output Tube (IOT), which is a newer type of RF amplifier that is usecommonlyd for television transmissions. IOTs have an almost constant efficiency (i.e. the power consumption is proportional to the output power) down to 30% of maximum output power, which will give substantial power savings.

For klystrons, which have constant power consumption, the energy-optimal filling strategy is simply to drive them at saturation. However, for IOTs, it is possible to reduce the filling energy with a more sophisticated strategy.

---

[1] The filling constitutes 0.15 ms/(2.86+0.15) ms = 5 % of the RF pulse, and the expected average electricity consumption of the RF amplifiers (inductive output tubes) is 6 MW, which gives the following cost estimate:

$5\,\% \times 6\,\mathrm{MW} \times 5000\,\mathrm{h/year} \times 0.07$ €/kWh = 100 k€/year.

As remarked in a recent article by [Bhattacharyya et al., 2015], little work has been done on energy-optimal filling of RF cavities. In their article, they derived analytically how to minimize the energy *reflected* from the cavity during filling, which corresponds to minimization of the wall-plug energy consumption for *ideal amplifiers*. In this contribution we show how to minimize the wall-plug energy for *arbitrary* amplifier characteristics.

Problem formulation

The baseband dynamics for the cavity voltage of the accelerating mode is given by (5.31), which we re-state for convenience

$$\frac{d\mathbf{V}}{dt} = (-\gamma + i\Delta\omega)\mathbf{V} + \alpha\sqrt{2\gamma_{\text{ext}}}\mathbf{F_g} + \frac{\alpha^2}{2}\mathbf{I_b}(t).$$

We will only do relative comparisons of the energy consumption for different filling approaches, so to make the exposition more clear, we normalize so that the nominal value of $\mathbf{V}$ is equal to 1, scale time so that $\gamma = 1$, and finally redefine $\mathbf{F_g}$ so that its coefficient equals 1 ($|\mathbf{F_g}|^2$ will still be proportional to the generator output power); thus we have the following normalized equation for the cavity voltage during filling (since $\mathbf{I_b} = 0$),

$$\frac{d\mathbf{V}}{dt} = (-1 + i\Delta\omega)\mathbf{V} + \mathbf{F_g}. \tag{10.1}$$

We want to determine how to fill the cavity, i.e., reach $\mathbf{V}(t_f) = 1$, while minimizing the energy consumption

$$W = \int_0^{t_f} P_{\text{amp}}(|\mathbf{F_g}|)\, dt, \tag{10.2}$$

where $P_{\text{amp}}(F_g) = F_g^2/\eta(F_g)$ is proportional to the wall-plug power drawn by the amplifier, and $\eta(F_g)$ is the amplifier efficiency as a function of output amplitude. The final time t_f is a free parameter and there is an upper limit $F_g^{\max}$ on $|\mathbf{F_g}|$, see Figure 10.1 for an illustration.

Formally stated, the problem is

$$\begin{aligned}
\underset{\mathbf{F_g},\, t_f}{\text{minimize}} \quad & \int_0^{t_f} P_{\text{amp}}(|\mathbf{F_g}|)\, dt && \text{(10.3a)}\\
\text{subject to} \quad & \dot{\mathbf{V}} = (-1 + i\Delta\omega(t))\mathbf{V} + \mathbf{F_g} && \text{(10.3b)}\\
& |\mathbf{F_g}| \le F_g^{\max} && \text{(10.3c)}\\
& \mathbf{V}(0) = 0 && \text{(10.3d)}\\
& \mathbf{V}(t_f) = 1. && \text{(10.3e)}
\end{aligned}$$

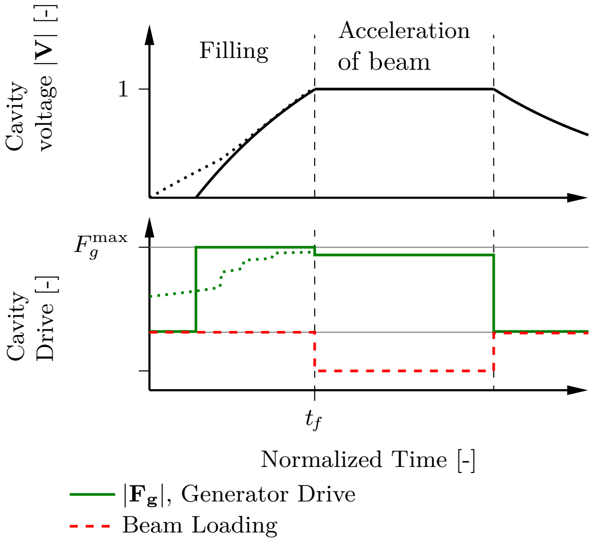


Figure 10.1 Normalized amplitudes of the cavity field and the generator current during an RF pulse, for minimum time filling, i.e., $|\mathbf{F_g}| = F_g^{\max}$ (solid lines), and energy-optimal filling for an IOT (dotted lines).

In the problem formulation (10.3), we assume the detuning $\Delta\omega(t)$ for $0 \leq t \leq t_f$ to be known in advance, that the amplifier has no dynamics, that there are no disturbances, and that all parameters for the cavity dynamics are perfectly known. These assumptions are approximations, but we believe that the results that will be computed under these assumpations are reasonably close to what can be achieved in practice.

Previous Work

An analytic solution for how to minimize the *reflected* energy for a fixed final time t_f for $\Delta\omega \equiv 0$ and no limit on $|\mathbf{F_g}|$, we derived in [Bhattacharyya et al., 2015]—the considered problem is similar to minimizing the energy consumption for an amplifier with constant efficiency. The optimal generator current profile was found to be

$$F_g^*(t) = \frac{\exp(t)}{\sinh(t_f)}, \tag{10.4}$$

with the corresponding cavity voltage

$$V(t) = \frac{\sinh(t)}{\sinh(t_f)}.$$

In case the amplifier saturation level is too low to implement (10.4), i.e., $\exp(t_f)/\sinh t_f > F_g^{\max}$, it was shown that the solution is given by $F_g^*(t) = \min(K\mathrm{e}^t, F_g^{\max})$, for a suitably chosen K. Bhattacharyya et al. also compared the energy consumption for amplifiers with different efficiency characteristics $\eta(F_g)$, with F_g given by (10.4). Note that F_g^* was computed for $\eta \equiv$ constant, and thus suboptimal for amplifiers whose efficiency depend on the output amplitude.

Outline of the chapter

In the next section we solve (10.3) for arbitrary efficiency characteristics $\eta(F_g)$ and time varying, but known, detuning $\Delta\omega(t)$. In Section 10.3 we compare the energy-optimal filling strategies for different efficiency characteristics $\eta(F_g)$ and different saturation levels $F_g^{\max}$. We conclude with a remark on cryogenic losses and a discussion of the results.

10.2 Energy Optimal Filling Profile

Optimal phase of $\mathbf{F_g}(t)$

In polar coordinates the cavity equation (10.3b) takes the form [Brandt, 2007],

$$V\dot{\phi} - \Delta\omega V = F_g \sin(\theta - \phi) \tag{10.5a}$$

$$\dot{V} + V = F_g \cos(\theta - \phi), \tag{10.5b}$$

where $F_g \geq 0$, $V \geq 0$, θ and ϕ are defined via

$$\mathbf{F_g}(t) = F_g(t)\mathrm{e}^{i\theta(t)} \tag{10.6}$$

$$\mathbf{V}(t) = V(t)\mathrm{e}^{i\phi(t)}. \tag{10.7}$$

By considering (10.5b), we see that choosing θ as

$$\theta^*(t) = \phi(t), \tag{10.8}$$

maximizes $\dot{V}$ for all $F_g \geq 0$. Since the cost (10.3a) is independent of ϕ, and we wish to minimize the cost for reaching $V(t_f) = 1$, it is clear that (10.8) is optimal. With this choice of $\theta(t)$, equation (10.5a) reduces to $\dot{\phi} = \Delta\omega$, and since we want $\phi(t_f) = 0$, we must have

$$\phi(t) = -\int_t^{t_f} \Delta\omega(t')\, dt'. \tag{10.9}$$

From (10.8) it follows that the optimal phase profile $\theta^*(t)$ of the generator current equals the right hand side of (10.9).

Remark: For superconducting cavities, the detuning $\Delta\omega(t)$ depends on the cavity field amplitude V via the Lorentz force detuning. However, since the the optimization of θ and F_g is decoupled, the optimal F_g and the corresponding V can be found first, before finding the optimal phase via (10.9).

Optimal amplitude of $\mathbf{F_g}(t)$

The optimal phase of the generator current is given by (10.8), so we only need to consider the amplitude dynamics given by (10.5b), and finding the optimal amplitude F_g reduces to the following problem

$$\begin{aligned}
\underset{F_g,\, t_f}{\text{minimize}} \quad & \int_0^{t_f} P_{\text{amp}}(F_g(t))\, dt && (10.10\text{a})\\
\text{subject to} \quad & \dot{V}(t) = -V(t) + F_g(t) && (10.10\text{b})\\
& |F_g(t)| \le F_g^{\max} && (10.10\text{c})\\
& V(0) = 0 && (10.10\text{d})\\
& V(t_f) = 1. && (10.10\text{e})
\end{aligned}$$

From (10.10), it seems reasonable that the optimal choice of F_g at each time instant maximizes the ratio between the increase of the cavity field and the power consumption, i.e.,

$$F_g^*(t) = \underset{F_g}{\text{argmax}} \frac{-V(t) + F_g}{P_{\text{amp}}(F_g)}. \tag{10.11}$$

That (10.11) indeed is optimal, follows from the following, slightly more general theorem.

Theorem 1

Consider the optimal control problem

$$\begin{aligned}
\underset{u,\, t_f}{\text{minimize}} \quad & \int_0^{t_f} r(u(t))\, dt && (10.12\text{a})\\
\text{subject to} \quad & \dot{x}(t) = f(x(t), u(t)) && (10.12\text{b})\\
& x(0) = 0 && (10.12\text{c})\\
& x(t_f) = 1 && (10.12\text{d})\\
& u(t) \in \mathcal{U}, && (10.12\text{e})
\end{aligned}$$

where $\mathcal{U}$ is a compact set, $r(u) > 0$ for all $u \in \mathcal{U}$, $f(x, u)$ and $r(u)$ are continuous functions of u, and

$$\forall x \in [0, 1] \;\; \exists u \in \mathcal{U} \text{ so that } f(x, u) \ge c > 0. \tag{10.13}$$

Define

$$u^*(x) := \operatorname*{argmax}_{u \in \mathcal{U}} \frac{f(x,u)}{r(u)}. \tag{10.14}$$

and assume that u^* is sufficiently well-behaved for $\dot{x} = f(x, u^*(t))$ to have a unique solution $x^*(t)$. Then the optimal control signal is given by $u(t) = u^*(x^*(t))$. □

Proof See Appendix D.2 for a proof based on the Hamilton-Jacobi-Bellman technique. □

Remark 1: The assumption (10.13) guarantees finite-time feasibility.

Remark 2: The maximum in (10.14) exists since a continuous function is optimized over a compact set. If several u maximize the expression, any can be chosen.

Remark 3: It is clear that the functions $r(u)$ and $f(x,u)$ considered in the optimal control problem (10.10) give a well-behaved u^*.

Remark 4: A constraint $t_f < t_{\max}$ on the final time can be handled by adding a constant term to $r(u)$, and doing a binary search over that constant.

Solution for constant efficiency For a constant efficiency $\eta \equiv \eta_0$ we have $P_{\text{amp}}(F_g) = F_g^2/\eta_0$, and (10.11) becomes,

$$F_g^* = \operatorname*{argmax}_{F_g} \frac{-V + F_g}{F_g^2/\eta_0} = 2V.$$

With this control signal V does not reach 1 in finite time, which is possible since the cost is not strictly greater than 0. If a fixed final time is imposed, the solution follows from the maximum principle, or as in [Bhattacharyya et al., 2015].

10.3 Results

In this section we compare the energy consumption for three filling strategies:

- *Minimum-time*, i.e., $F_g(t) = F_g^{\max}$
- *Minimum-reflection*, i.e., $F_g(t) = \min(\mathrm{e}^t / \sinh \hat{\tau}_i,\, F_g^{\max})$ with $\hat{\tau}_i = 2$ [Bhattacharyya et al., 2015]
- *Energy-optimal*, according to (10.11),

considering four amplifier efficiency characteristics $\eta(F_g)$, [Bhattacharyya et al., 2015], see Figure 10.2:

- Tetrode
- Solid-state amplifier (SSA)
- Inductive output tube (IOT)
- Constant efficiency (ideal) amplifier,

and two saturation levels, given in normalized units by:

- $F_g^{\max} = 1.5$
- $F_g^{\max} = 2.25$.

The levels correspond to a normal conducting cavity and a heavily beam-loaded superconducting cavity respectively.

The energy consumption for the different filling strategies and parameter combinations are shown in Figure 10.3. The relative energy consumption of energy-optimal filling compared to minimum time and minimum-reflection

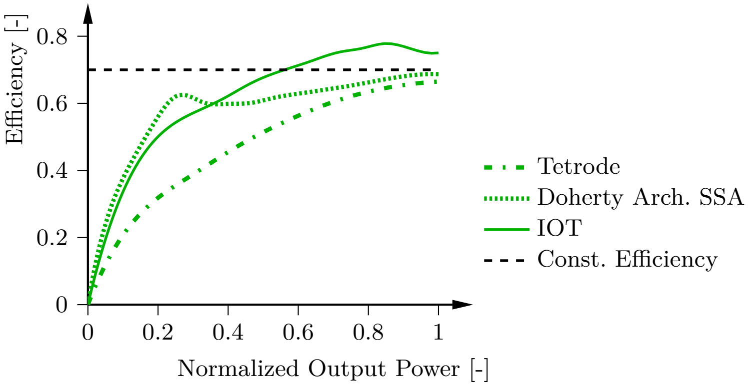
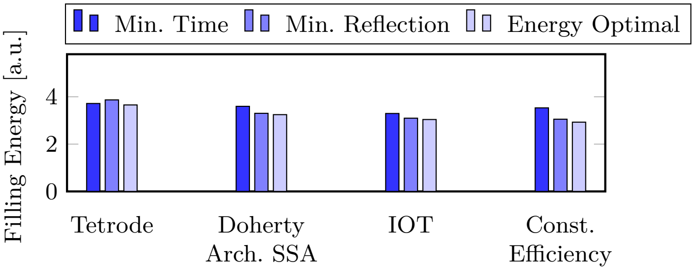


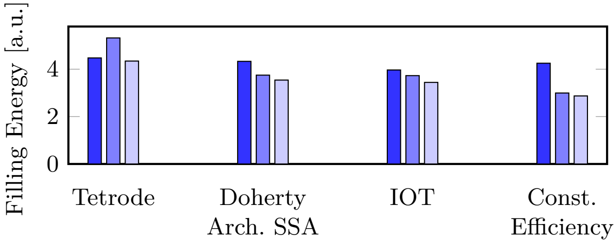
Figure 10.2 Efficiency as a function of normalized output power $(F_g/F_g^{\max})^2$ for the considered amplifiers types. The data is from [Bhattacharyya et al., 2015, Fig. 9], and has been slightly smoothed.

Table 10.1 Energy consumption of minimum-time (MT) and minimum-reflection (MR) filling relative energy-optimal (EO) filling.

Amplifier Type	$F_g^{\max} = 1.5$		$F_g^{\max} = 2.25$	
	$\frac{W_{\mathrm{EO}}}{W_{\mathrm{MT}}}$	$\frac{W_{\mathrm{EO}}}{W_{\mathrm{MR}}}$	$\frac{W_{\mathrm{EO}}}{W_{\mathrm{MT}}}$	$\frac{W_{\mathrm{EO}}}{W_{\mathrm{MR}}}$
Tetrode	98 %	94 %	97 %	82 %
Doherty Arch. SSA	90 %	98 %	82 %	94 %
IOT	92 %	98 %	87 %	92 %
Constant Efficiency	83 %	96 %	68 %	96 %



(a) $F_g^{\max} = 1.5$.

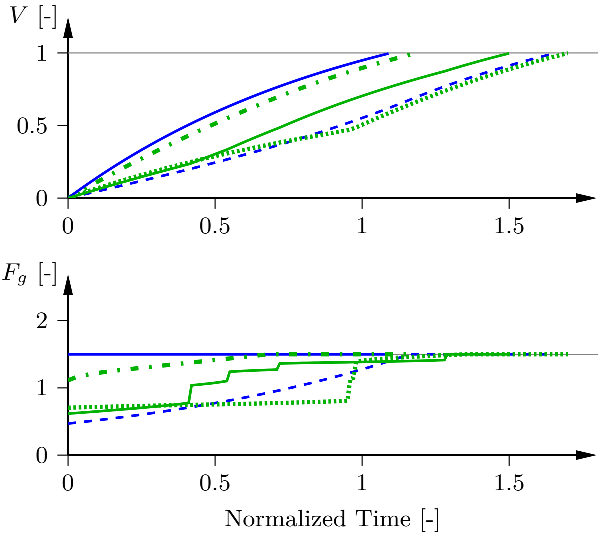


(b) $F_g^{\max} = 2.25$.

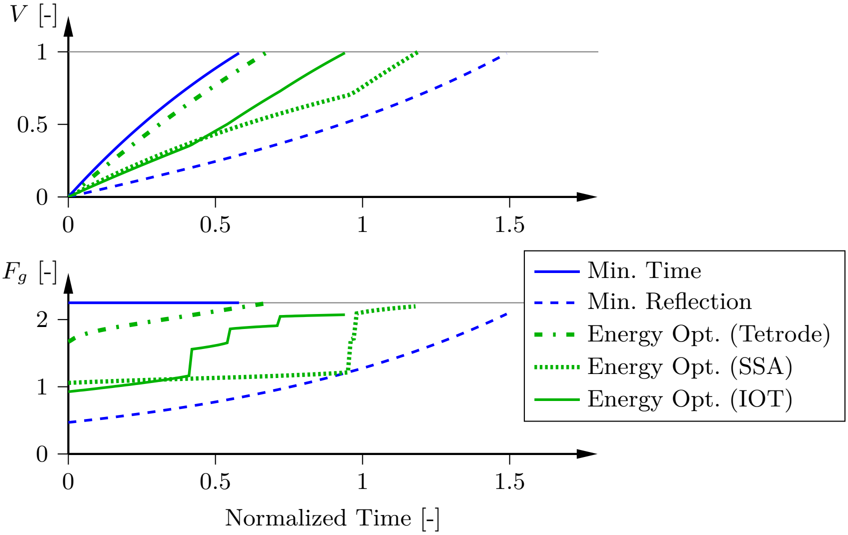
Figure 10.3 Comparison of the energy consumption for minimum-time filling (standard approach), minimum-reflection filling according to [Bhattacharyya et al., 2015], and energy-optimal filling according to (10.11). Four different amplifier types and two different saturation levels are considered.

filling is given in Table 10.1. The corresponding profiles for the cavity voltage and generator current are shown in Figure 10.4.

For the high-β section at ESS, where IOTs will be used, and with $F_g^{\max} = 2.25$, the energy reduction of using energy-optimal filling, relative to minimum time filling, is 13%, which corresponds to 13 k€/year.



(a) $F_g^{\max} = 1.5$.



(b) $F_g^{\max} = 2.25$.

Figure 10.4 Cavity voltage and generator current for different filling strategies: minimum time, minimum-reflection (↔ constant efficiency amplifier) and energy-optimal for the efficiency characteristics in Figure 10.2. Note that minimum time and minimum-reflection filling are independent of the amplifier characteristics. Two saturation levels are considered.

Remark on cryogenic losses

As seen in Figure 10.4, the energy-optimal filling profiles take about 50–100% longer time than minimum time filling. This implies increased RF heating of the cavities, and a higher load on the cryogenic system, estimated to a yearly cost of 1–2 k€[2]. This cost is clearly outweighed by the savings from using energy-optimal filling.

If desired, it is easy to incorporate cryogenic losses into the optimization problem by adding a term $\alpha_{\text{cryo}} V^2$ to the integrand in (10.10a). We performed some experiments with this, and found that the resulting energy-optimal trajectories were similar, but slightly faster, than the ones for which cryogenic losses were ignored.

10.4 Conclusions

We have shown how to minimize the energy required to build up the electromagnetic field in RF cavities. We proved that the amplitude and phase for the optimal generator current profile, for a normalized cavity, $\big(\dot{\mathbf{V}} = (-1 + i\Delta\omega)\mathbf{V} + \mathbf{F_g}\big)$, are given by:

$$F_g^*(t) = \operatorname*{argmax}_{F_g} \frac{-V(t) + F_g}{P_{\text{amp}}(F_g)}$$

$$\theta^*(t) = -\int_t^{t_f} \Delta\omega(t)\, dt.$$

We compared the energy savings for different amplifier characteristics and found that the energy consumption could be reduced by up to 30%. For the high-β section at the European Spallation Source we estimated the yearly savings to be about 10 k€. Since the energy-optimal filling profiles, or at least approximate versions thereof, are relatively easy to implement, they provide a straight-forward way to reduce the operating costs and environmental footprint of pulsed particle accelerators.

---

[2] The cryogenic load due to RF heating in the high-β section is 1.6 kW@4.5K. With a cooling efficiency of 250 W/W, the required wall-plug power is 400 kW [Peggs et al., 2013]. Assuming a linear increase of V during the filling, the average cryo load from the filling is proportional to $\int_0^T (V/T)^2 dV = T/3$, i.e., a third of the heating from T time units of flat-top operation. Thus the energy cost for the cryo load from minimum time filling, which take $\approx 5\,\%$ of the RF pulse length, is
$5\,\% \times 1/3 \times 400\,\text{kW} \times 5000\,\text{h/year} \times 0.07$ €/kWh = 2 k€/year.

11

Conclusions and Future Work

11.1 Conclusions

This thesis has addressed the problem of controlling the amplitude and phase of electromagnetic fields in accelerator cavities. The topics covered were selected based on their relevance to the high-intensity proton accelerator of ESS, but most of the material is relevant to other accelerator types as well.

Throughout, it has been demonstrated that a complex-coefficient representation of the cavity and RF system dynamics greatly facilitates control design and analysis. This representation was also helpful in the, somewhat non-standard, derivation of the cavity mode dynamics.

When the performance of simple PI(D)-controllers was compared to the optimal linear controller, it was found that there actually was not much difference, owing to the simple and well-behaved process dynamics of the field control loop (ignoring parasitic modes). Since PI(D)-controllers also are intuitive and easy to implement, they seem like the ideal controller choice for cavity field control.

To deal with parasitic modes in the control design, a number of different strategies were presented and discussed, each with their own pros and cons. Simply adding a second-order filter to the controller gives reasonable performance and avoids complications, but more complicated approaches allowed for better performance.

Just as the success story of complex-coefficient system representations was about to come to a happy end, it was shown that although the plant dynamics can be represented by complex-coefficient systems, and hence is rotationally invariant, this does not hold for the objective function or the disturbance distributions. Thus the optimal controller is not rotationally invariant, and cannot be represented by a complex-coefficient SISO system. However the complex-coefficient representation is good for insight, and for many cases of interest, the simplicity of a rotationally invariant controller probably outweighs the performance gains of a more complicated controller.

Lastly, we demonstrated how to reduce the energy required to build up the electromagnetic cavity fields, allowing a further improvement of the sustainability of ESS.

11.2 Future work

Improved computer tools for complex-coefficient control design

This thesis has illustrated the utility of complex-coefficient systems for analyzing the field control problem. For the control analysis, synthesis and testing of these systems it would be convenient with well-designed and well-tested computer software. The support provided by Matlab for complex-coefficient systems is however highly deficient, so it would be valuable to implement the necessary functionality either in Matlab, or Julia which is an up-and-coming language for technical computing.

Directionality in the control problem

As more data on the disturbances that will act on the ESS field control loops become available, it would be interesting to investigate the possible performance gains from using a general two-input two-output controller, rather than the rotationally invariant structure that was considered throughout this thesis. This study could be carried out using the Youla parametrization in a similar manner as it was used in Chapter 7.

Iterative Learning Control (ILC)

For the ESS accelerator it will be necessary to implement an ILC algorithm on top of the existing controller to update the table for predicted disturbances, $\mathbf{u}_k^{\mathrm{ILC}}$. The plan is to use the approach in [Norrlöf, 2000], where the update law is given by

$$\boldsymbol{u}_{k+1}^{\mathrm{ILC}} = Q(z)(1 - L(z)\mathbf{e}_k)\mathbf{u}_k^{\mathrm{ILC}},$$

and where $\mathbf{e}_k$ is the error vector from the previous pulse, $L(z)$ is a suitably chosen (possibly non-causal) filter, and $Q(z)$ is a zero-phase low-pass filter. This approach can be directly extended for the complex-coefficient representation used in this thesis, allowing a simple and elegant implementation of the algorithm.

Exploration of the objective function $\Psi_{\mathbf{BL}}$

It would be valuable to get a qualitatively better understanding of the structure of the objective function $\Psi_{\mathrm{BL}}(\Delta A_1, \Delta\phi_1, \Delta A_2, \Delta\phi_2, \ldots)$. This is a nontrivial task which would probably have to carried out by doing a very large number of numerically demanding computer code simulations in a joint effort with beam physicists.

Feedforward from beam current monitors

As the beam current of high-intensity proton accelerators keeps increasing, and the requirements on cavity field errors are made tighter, the impact of beam current variations on the cavity field becomes an issue.

An interesting possibility if the bandwidths of the field control loops does not allow sufficient attenuation of the beam current variations, is to use feedforward from measurements of the beam current that are made by the Beam Current Monitors (BCMs). This is however non-trivial since the propagation velocity of the beam and the electric BCM signals are of the same order; for a digital controller there are also processing delays. For these reasons, it is uncommon with beam feedforward for linear accelerators, and seems only to have been implemented with analog electronics [Jameson and Wallace, 1971].

Preliminary investigations for the ESS accelerator have indicated that BCM feedforward reduces the impact of beam current variations on the DTL fields by up to 30 %. This is a topic worthy of further investigation, since it would allow for higher beam currents, more ion source fluctuations, and tighter field control requirements.

Closing remark

In theory, there is no difference between practice and theory. In practice, there is. The question that will be answered, in one year form now, when the first field control loops of the ESS accelerator (Figure 11.1) will be commissioned, is *how* different the theory presented in this thesis is from practice.

Hopefully the basis for how to think about the field control problem for high-intensity accelerators that I have put forward in this thesis will be helpful in achieving the field control requirements of the ESS accelerator. I am looking forward to take part in the commissioning of its LLRF system, and to deal with the awaiting challenges.

Figure 11.1 The construction site of the European Spallation Source on 2017-04-17. The long gray building is the RF gallery that will house the RF systems of the ESS accelerator. Photo credit: ESS.

Bibliography

Ainsworth, R. and S. Molloy (2012). “Studies of parasitic cavity modes for proposed ESS linac lattices”. In: *Proceedings of LINAC2012.* (Tel-Aviv, Israel, Sept. 9–14, 2012).

Alberi, A. and M. Lacroix (2015). *RFQ mechanical design.* Presentation at CDR2 for the ESS RFQ, CEA Saclay, Paris.

Altarelli, M. et al. (2007). *The European X-Ray Free-Electron Laser, Technical design report.* Tech. rep. DESY 2006-097. Deutsches Elektronen-Synchrotron.

Arnaudon, L. et al. (2006). *Linac4 Technical Design Report.* Tech. rep. CERN-AB-2006-084. CERN, Geneva. URL: `http://cds.cern.ch/record/1004186`.

Åström, K. J. and T. Hägglund (2006). *Advanced PID Control.* The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC. ISBN: 978-1-55617-942-6.

Åstrom, K. J. and R. M. Murray (2010). *Feedback Systems: An Introduction for Scientists and Engineers.* Princeton University Press, Princeton, NJ.

Bamieh, B., F. Paganini, and M. A. Dahleh (2002). “Distributed control of spatially invariant systems”. *IEEE Transactions on Automatic Control* **47**:7, pp. 1091–1107. ISSN: 0018-9286.

Baudrenghien, P. (2000). “Low-level RF systems for synchrotrons”. In: *Proceedings of CERN Accelerator School 2000 — Radio Frequency Engineering.*

Baudrenghien, P., D. Valuch, D. Stellfeld, J. Galindo, J. Noirjean, and G. Hagmann (2014). “Commissioning of the linac4 low level RF and future plans”. In: *Proceedings of LINAC2014.* (Geneva, Switzerland, Aug. 31–Sept. 5, 2014).

Bhattacharyya, A. K., V. Ziemann, R. Ruber, and V. Goryashko (2015). “Minimization of power consumption during charging of superconducting accelerating cavities”. *Nuclear Instruments and Methods in Physics Research Section A* **801**, pp. 78–85.

Bousson, S. et al. (2014). “The ESS spoke cavity cryomodules”. In: *AIP Conference Proceedings*. Vol. 1573. 1. American Institute of Physics, pp. 665–672.

Boyd, S. and C. Barratt (1991). *Linear Controller Design: Limits of Performance*. Prentice-Hall, Englewood Cliffs, NJ. ISBN: 0-13-538687-X. URL: `http://stanford.edu/~boyd/lcdbook/`.

Brandt, A. (2007). *Development of a Finite State Machine for the Automated Operation of the LLRF Control at FLASH*. PhD thesis. Hamburg University.

Bristow, D. A., M. Tharayil, and A. G. Alleyne (2006). “A survey of iterative learning control”. *IEEE Control Systems* **26**:3, pp. 96–114. ISSN: 1066-033X.

Carter, R. G. (2010). “RF power generation”. In: *Proceedings of CERN Accelerator School 2010 — RF for Accelerators*.

Celona, L. et al. (2016). “Preliminary commissioning results of the proton source for ESS at INFN-LNS”. In: *Proceedings of the 7th International Particle Accelerator Conference*. (Busan, Korea, May 8–13, 2016), pp. 2628–2631.

CVX Research Inc (2012). *CVX: matlab software for disciplined convex programming, version 2.0*. `http://cvxr.com/cvx`.

Delayen, J. R. (1978). *Phase and Amplitude Stabilization of Superconducting Resonators*. PhD thesis. California Institute of Technology.

Doolittle, L. et al. (2015). “The LCLS-II LLRF system”. In: *Proceedings of the 6th International Particle Accelerator Conference*. (Richmond, VA, May 3–8, 2015).

Doolittle, L., H. Ma, and M. S. Champion (2006). “Digital low-level RF control using non-IQ sampling”. In: *Proceedings of LINAC2006*. (Knoxville, TN, Aug. 20–25, 2006).

Eshraqi, M. and Y. Levinsen (2016). *LLRF needs for Beam Physics*. Tech. rep. ESS-0049330. ESS internal document. ESS, Lund, Sweden.

ESSa (2017). *The unique capabilities of ESS*. Accessed 2017-04-12. URL: `https://europeanspallationsource.se/unique-capabilities-ess`.

ESSb (2017). *Elliptical SRF*. Accessed 2017-05-21. URL: `https://europeanspallationsource.se/elliptical-srf`.

Gallo, A. (2010). “Basics of RF electronics”. In: *Proceedings of CERN Accelerator School 2010 — RF for Accelerators*.

Garpinger, O. (2009). *Design of Robust PID Controllers with Constrained Control Signal Activity.* Licentiate thesis. Lund University, Sweden.

Gilmour, A. (2011). *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers, and Gyrotrons.* Artech House, Norwood, MA. ISBN: 978-1608071845.

Hara, M., T. Nakamura, and T. Ohshima (1998). “A ripple effect of a klystron power supply on synchrotron oscillation”. *Particle Accelerators* **59**, pp. 143–156.

Haus, H. A. (1983). *Waves and Fields in Optoelectronics.* Prentice-Hall, Englewood Cliffs, NJ. ISBN: 0-13-946053-5.

Hespanha, J. P. (2009). *Linear Systems Theory.* Princeton University Press, Princeton, NJ. ISBN: 978-0-691-14021-6.

Hoffmann, M. (2008). *Development of a multichannel RF field detector for the Low-Level RF control of the Free-Electron Laser at Hamburg.* PhD thesis. TU Hamburg-Harburg.

Jameson, R. and J. Wallace (1971). “Feedforward control of accelerator RF fields”. *IEEE Transactions on Nuclear Science* **18**:3, pp. 598–600. ISSN: 0018-9499.

Jamieson, S. (2006). *Micro Telecommunications Computing Architecture Short Form Specification.* Tech. rep. PICMG MTCA.0. PICMG. URL: `https://www.picmg.org/wp-content/uploads/MicroTCA_Short_Form_Sept_2006.pdf`.

Levinsen, Y., M. Eshraqi, R. Miyamoto, M. Munoz, A. Ponton, and R. D. Prisco (2016). “Beam dynamics challenges in the ESS linac”. In: *Proceedings of 57th ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams.* (Malmö, Sweden, July 3–8, 2016).

Ma, H. et al. (2006). “Low-level RF control of Spallation Neutron Source: System and characterization”. *Physical Review Special Topics - Accelerators and Beams* **9** (3), p. 032001. URL: `https://link.aps.org/doi/10.1103/PhysRevSTAB.9.032001`.

MTCA.4 (2017). *Microtca.4 for industry and research.* DESY. URL: `http://mtca.desy.de/`.

Mokhov, N. V. and W. Chou, (Eds.) (1999). *Beam Halo and Scraping — Proceedings of the 7th ICFA Mini-Workshop on High-Intensity and High-Brightness Hadron Beams.* (Lake Como, Wisconsin, Sept. 13–15, 1999). Fermi National Accelerator Laboratory, Batavia, IL.

Montgomery, C. G., R. H. Dicke, and E. M. Purcell (1948). *Principles of Microwave Circuits.* McGraw-Hill, New York, NY.

Mosnier, A. and J. M. Tessier (1994). *Field stabilization study for TESLA.* Tech. rep. TESLA-1994-16. TESLA.

Neri, L. et al. (2014). “Improved design of proton source and low energy beam transport line for European Spallation Source”. *Review of Scientific Instruments* **85**:2, 02A723.

Norrlöf, M. (2000). *Iterative Learning Control: Analysis, Design, and Experiments.* PhD thesis. Linköping University, Sweden.

Omet, M. (2014). *Digital Low Level RF Control Techniques and Procedures Towards the International Linear Collider.* PhD thesis. The Graduate University for Advanced Studies, Japan.

Padamsee, H., J. Knobloch, and T. Hays (2008). *RF Superconductivity for Accelerators.* 2nd ed. Wiley-VCH, Weinheim.

Pedersen, F. (1975). “Beam loading effects in the CERN PS booster”. *IEEE Transactions on Nuclear Science* **22**:3, pp. 1906–1909. ISSN: 0018-9499.

Peggs, S. et al. (2013). *ESS Technical Design Report.* Tech. rep. ESS-doc-274. European Spallation Source.

Pfeiffer, S. (2014). *Symmetric Grey Box Identification and Distributed Beam-Based Controller Design for Free-Electron Lasers.* PhD thesis. TU Hamburg-Harburg.

Pfeiffer, S., G. Lichtenberg, C. Schmidt, H. Schlarb, and H. Werner (2012). “Design of an optimal and robust controller for a free-electron laser exploiting symmetries of the RF-system”. In: *Proceedings of the 51st IEEE Conference on Decision and Control.* (Maui, HI, Dec. 10–13, 2012). IEEE, pp. 4253–4258.

Rezaeizadeh, A. (2016). *Automatic Control Strategies for the Swiss Free Electron Laser.* PhD thesis. Swiss Federal Institute of Technology in Zurich.

Robinson, K. W. (1964). *Stability of Beam in Radiofrequency System.* Tech. rep. CEAL-1010. Cambridge Electron Accelerator, Cambridge, MA.

Rogers, E. et al. (2010). “Norm-optimal iterative learning control with application to problems in accelerator-based free electron lasers and rehabilitation robotics”. *European Journal of Control* **16**:5, pp. 497–522.

Russell, G. J. (1990). “Spallation physics — an overview”. In: *Proceedings of International Collaboration on Advanced Neutron Sources XI.* (Tsukuba, Japan, Oct. 22–26, 1990).

Scheinker, A. (2012). *Extremum seeking for stabilization.* PhD thesis. University of California, San Diego.

Schilcher, T. (1998). *Vector Sum Control of Pulsed Accelerating Fields in Lorentz Force Detuned Superconducting Cavities.* PhD thesis. Hamburg University.

Schilcher, T. (2007). “Digital signal processing in RF applications”. In: *Proceedings of CERN Accelerator School 2007 — Digital Signal Processing.*

Schmidt, C., S. Pfeiffer, J. Branlard, H. Schlarb, and W. Jalmuzna (2012). "Precision regulation of RF fields with MIMO controllers and cavity-based notch filters". In: *Proceedings of LINAC2012.* (Tel-Aviv, Israel, Sept. 9–14, 2012).

Schmidt, C. (2010). *RF System Modeling and Controller Design for the European XFEL.* PhD thesis. TU Hamburg-Harburg.

Skogestad, S. (2003). "Simple analytic rules for model reduction and PID controller tuning". *Journal of Process Control* **13**:4, pp. 291–309. ISSN: 0959-1524.

Skogestad, S. and I. Postlethwaite (2007). *Multivariable Feedback Control: Analysis and Design.* 2nd ed. Wiley New York.

Slater, J. C. (1950). *Microwave Electronics.* Van Nostrand, Princeton, NJ.

Stein, G. (2003). "Respect the unstable". *IEEE Control Systems* **23**:4, pp. 12–25. ISSN: 1066-033X.

Suh, W., Z. Wang, and S. Fan (2004). "Temporal coupled-mode theory and the presence of non-orthogonal modes in lossless multimode cavities". *IEEE Journal of Quantum Electronics* **40**:10, pp. 1511–1518.

Troeng, O. and B. Bernhardsson (2017). "Optimal excitation of radio-frequency cavity". In: *Proceedings of the 20th World Congress of the International Federation of Automatic Control.* (Toulouse, France, July 9–14, 2017), Accepted.

Tückmantel, J. (2011). *Cavity-Beam-Transmitter Interaction Formula Collection with Derivation.* Tech. rep. CERN-ATS-Note-2011-002 TECH. CERN.

Uriot, D. and N. Pichoff (2015). "Status of the Tracewin code". In: *Proceedings of the 6th International Particle Accelerator Conference.* (Richmond, VA, May 3–8, 2015).

UT-Battelle, L. (2006). *Spallation Neutron Source Project Completion Report.* Tech. rep. SNS 100000000-BL0005-R00. Oak Ridge National Laboratory, Oak Ridge, TN.

Vogel, E. (2007). "High gain proportional RF control stability at TESLA cavities". *Physical Review Special Topics - Accelerators and Beams* **10**:5, p. 052001.

Wangler, T. P. (2008). *RF Linear Accelerators.* 2nd ed. Wiley-VCH, Weinheim, Germany. ISBN: 978-3-527-40680-7.

Youla, D., H. Jabr, and J. Bongiorno (1976). "Modern Wiener-Hopf design of optimal controllers–Part II: the multivariable case". *IEEE Transactions on Automatic Control* **21**:3, pp. 319–338.

A

Dynamics of Digital Demodulation for Field Control

In this section we take a closer look at how the complex envelope of a signal (i.e. the equivalent baseband signal) is typically detected in low-level RF systems.

Although it was the availability of high-performance components developed for telecommunication systems that enabled digital LLRF systems, we find that the detection is done quite differently in LLRF and telecommunication systems. In telecommunication systems the signals are wideband—their amplitudes and phases are modulated rapidly in order to maximize data rates, and ideally only one sample should be spent to detect each symbol; for portable devices the design of hardware and firmware have been designed for low power consumption.

The signals that are measured by LLRF systems have complex envelopes that vary slowly in the time (for the cavity signal the envelope is ideally constant), so the challenge is not to track a rapidly varying envelope, but rather to detect it accurately. A typical implementation is shown in Figure A.1. First the RF signal is down-converted to an intermediate frequency (IF) signal in order to reduce the impact of LO and clock jitter, then it is sampled by a single high-speed ADC[1], and finally the sampled signal is digitally demodulated in the FPGA by a scheme known as Near-IQ or Non-IQ (NIQ) sampling [Doolittle et al., 2006; Schilcher, 2007; Hoffmann, 2008]. By carefully choosing the parameters for the NIQ sampling it is possible to reduce the effects of ADC nonlinearity and IF harmonics from the mixer.

---

[1] For telecom applications two mixers with LOs offset by 90° and two ADCs are typically used in order to achieve high data rates, however this introduces problems with phase alignment.

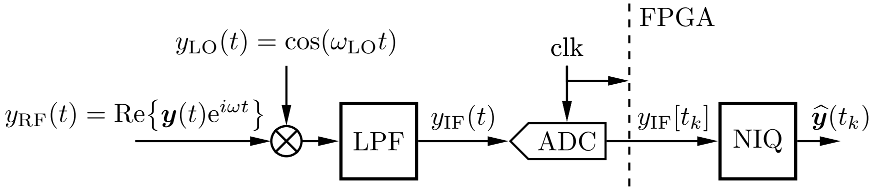


Figure A.1 Typical down-modulation scheme for LLRF systems, it produces an estimate $\widehat{\boldsymbol{y}}(t)$ for the complex envelope $\boldsymbol{y}(t)$ of the signal $y_{\mathrm{RF}}(t)$. The components shown are: down-conversion mixer, analog low-pass filter (LPF), analog-to-digital converter (ADC) and the digital down conversion scheme which is called Near-IQ (NIQ) sampling.

The near IQ complex baseband signal, *relative the master oscillator*, is computed from the IF signal as

$$\widehat{\boldsymbol{y}}[t_k] = \mathrm{e}^{-i\omega_{\mathrm{IF}} t_k} \frac{2}{N} \sum_{\ell=0}^{N-1} y_{\mathrm{IF}}[t_{k-\ell}] \mathrm{e}^{-2\pi i \ell M/N} \tag{A.1}$$

where $t_{k+1} - t_k = T_{\mathrm{adc}}$ is the sampling period of the ADC. M and N are integers chosen so that N samples are taken over M IF periods, i.e.

$$N T_{\mathrm{adc}} = M \frac{2\pi}{\omega_{\mathrm{IF}}}.$$

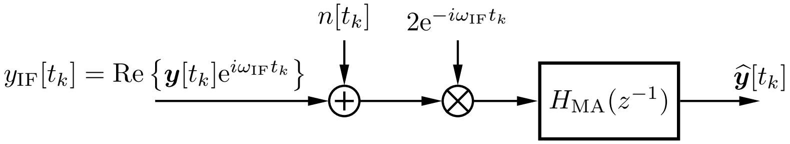
The first factor on the right hand side of (A.1) takes into account that the computed phasor should be relative to the phase of the master oscillator[2]. Notice how the second factor of (A.1) can be seen as a discrete Fourier transform, over a very short time interval, with only one bin, which is perfectly matched to the IF frequency.

To reduce aliasing of IF harmonics from the mixer, the number N needs to be chosen carefully, [Schilcher, 2007; Doolittle et al., 2006].

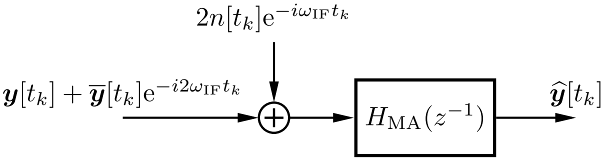
To save FPGA resources, one may compute only every Nth sample of (A.1). [Hoffmann, 2008, Figure 4.9] calls this approach *step window detection* and the original version *sliding window detection*. Step window detection is equivalent to directly decimating the signal in (A.1) by a factor N, and as is well known in signal processing, one should lowpass filter signals before decimating them in order to reduce aliasing. As we discuss in the chapters on control design, we wish to include a low-pass filter in the controller anyhow, so before the decimation could be a suitable option.

---

[2] Exactly how this is accomplished is implementation dependent, either the local oscillator and the clock signal are locked to the master oscillator, or both the signal of interest and the master oscillator signal are sampled after which the phase of the master oscillator is subtracted in the firmware.



(a) Illustration of the digital down conversion and how ADC noise enters the detection, $H_{\mathrm{MA}}(z^{-1})$ is defined in (A.3).



(b) Rearrangement of the figure above, using that $\mathrm{Re}\{A\} = (A + \bar{A})/2$.

Figure A.2 Detection of the complex envelope $\boldsymbol{y}$ of an IF signal $y_{\mathrm{IF}}(t) = \mathrm{Re}\left\{\boldsymbol{y}(t_k)\mathrm{e}^{i\omega_{\mathrm{IF}} t_k}\right\}$. The non-IQ sampling scheme essentially corresponds to a discrete-time Fourier transform with only one bin. We assume that the timing errors are negligible.

A.1 Baseband dynamics of non-IQ sampling

We will now study the dynamics this detection approach adds to the control loop and how the ADC noise enters; the situation is illustrated in Figure A.1.

Since the detection is done relative the master oscillator, the action of the non-IQ filter on the signal $y_{\mathrm{IF}}[t_k]$ is effectively time varying; this is illustrated in Figure A.2a. After a small rearrangement we get Figure A.2b, where we see that

$$\widehat{\boldsymbol{y}}[t_k] = \frac{1}{N}\sum_{\ell=0}^{N-1} \boldsymbol{y}[t_{k-\ell}] + \frac{1}{N}\sum_{\ell=0}^{N-1} \overline{\boldsymbol{y}}[t_{k-\ell}]\mathrm{e}^{-2i\omega_{\mathrm{IF}} t_{k-\ell}} + \frac{2}{N}\sum_{\ell=0}^{N-1} n[t_{k-\ell}]\mathrm{e}^{-2\pi i\ell M/N}. \quad \text{(A.2)}$$

The first term is the cavity field phasor filtered through an N sample moving average filter

$$H_{\mathrm{MA}}(z^{-1}) = \frac{1 + z^{-1} + \ldots z^{-N+1}}{N}, \quad \text{(A.3)}$$

the second term sums to 0 if $\boldsymbol{y}$ is constant, and is small otherwise, the final term is low-pass filtered noise from the ADC.

ADC quantization noise

In the case of step window detection, the variance of the noise added to the samples at times t_k, t_{k+N}, t_{k+2N}, …, are

$$\boldsymbol{n}[t_{Nm}] = \frac{2}{N}\sum_{\ell=0}^{N-1} n[t_{Nm-\ell}]\mathrm{e}^{-2\pi i\ell M/N}$$

It can be seen that $\mathrm{Re}\{\boldsymbol{n}\}$ and $\mathrm{Im}\{\boldsymbol{n}\}$ are white and uncorrelated with variance $2\sigma_n^2/N$. Thus $\boldsymbol{n}$ is white and circular symmetric with variance $4\sigma_n^2/N$

For sliding window detection the ADC noise gives rise to non-circular symmetric, cyclostationary noise on the baseband phasor.

B

Controller Optimization

B.1 Finding optimal linear controllers

The key idea that for finding (approximately) optimal, linear time-invariant controllers is the so-called Youla parametrization [Youla et al., 1976], whereby all internally stable closed-loop transfer functions are affinely parameterized in a stable transfer function $Q(s)$. From this parametrization it is possible to obtain a finite dimensional, convex problem which can be solved efficiently by standard solvers [Boyd and Barratt, 1991]; see [Hespanha, 2009] for a practical description of how to formulate the convex problem in the modeling system `cvx` [CVX Research Inc, 2012].

While the underlying ideas and the formulation of the optimization problem are relatively straight-forward, it has been my experience that numerous numerical problems are encountered for all but the smallest problems. The method presented in this section has grown out of my attempts to deal with these numerical problems.

The approach is entirely based on frequency domain computations, and only handles stable SISO plants, but this specialization allows the approach to be both transparent and easy to use. For example time-delays are effortlessly handled without approximation, and measured process frequency response, or disturbance spectra could be used without complications.

For the problem that we consider in Chapter 7 it is enough with the following special case of the Youla parametrization.

YOULA PARAMETRIZATION (SIMPLIFIED VERSION)
Assume that the SISO system $P(s)$ is stable, then

$$C(s) = \frac{Q(s)}{1 - P(s)Q(s)}$$

is an internally stabilizing controller for $P(s)$, if and only if $Q(s)$ is stable.□

With this parametrization, all closed loop transfer functions become affine in $Q(s)$, for example, the sensitivity function $S = (1-PQ)$, and we can re-write the control problem (6.7) as

$$\begin{aligned} \underset{Q}{\text{minimize}} \quad & ||P_a(1-PQ)D||_2 \\ \text{subject to} \quad & ||1-PQ||_\infty \le M_S. \\ & ||QN||_2 \le B_{un}. \end{aligned}$$

We see that the both the objective and the constraints are convex in Q.

Obtaining a finite-dimensional problem

To be able to solve the problem (B.1) numerically, it needs to be finite dimensional; we therefore let Q be approximated using a basis $\{Q_k\}_{k=1}^N$ of stable transfer functions,

$$Q = \sum_{k=1}^{N} \beta_k Q_k, \tag{B.2}$$

and we also select a finite number of frequencies $\Omega_1 = \{\omega_\ell\}_{\ell=1}^L$ where the sensitivity constraint is to be enforced. To allow for arbitrary disturbance spectra $D(i\omega)$, perhaps *empirically* measured, we express the 2-norm in the objective as a frequency-domain integral,

$$||G||_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(i\omega)|^2 \, d\omega \approx \frac{1}{2\pi}\sum_{m=1}^{M-1} |G(i\omega_m)|^2 \, \Delta\omega_m \tag{B.3}$$

where $\Omega_2 = \{\omega_m\}_{m=1}^M$ is a set of suitably chosen frequencies and $\Delta\omega_m = \omega_{m+1} - \omega_m$. In principle we could have $\Omega_1 = \Omega_2$, but having them different could allow for a more efficient frequencies selection. For example, the sensitivity constraint only needs to be enforced around the closed-loop system's cross-over frequency, so Ω_1 only needs to include frequencies in that range.

For the constraint on measurement-noise amplification of the control signal we can use the same type of integral expression as in (B.3).

Practical considerations

Basis functions The basis functions were taken as equal to 1 over some interval

$$Q_k(t) = \begin{cases} 1 & \text{if } t \in [t_k,\, t_{k+1}) \\ 0 & \text{otherwise} \end{cases} \tag{B.4}$$

with finite support in the time-domain (Figure B.1). This is similar to the choice in [Garpinger, 2009] where all intervals had the same length, $t_k = k\Delta t$.

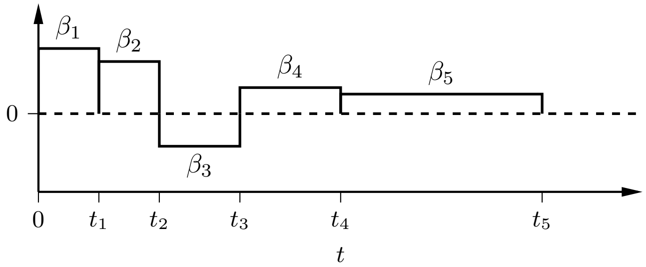


Figure B.1 Illustration of the impulse response of $Q = \sum \beta_k Q_k$ for when the basis functions are of the form (B.4).

Allowing the interval length to be adjusted makes it possible to increase the resolution of the impulse response of Q for small times t, where fast variations typically occur. Also note that the basis functions Q_k are orthogonal. In [Boyd and Barratt, 1991; Hespanha, 2009] a different choice of basis functions was used, $Q_k(s) = 1/(s+\alpha)^k$, however the parametrization (B.4) seems to have certain advantages, it is easy to specify the time-domain resolution, and their orthogonality allows simplification of the constraint (B.1), as we will see below.

Constraint on control signal activity If the measurement noise is white, i.e., $N(i\omega) = N(0)$, and if the basis functions Q_k are orthogonal, the constraint on control signal activity (B.1) can be expressed as

$$||QN||_2^2 = N(0)^2\, ||Q||_2^2 = N(0)^2 \sum_{n=1}^{N} |\beta_n|^2\, ||Q_n||_2^2 \leq B_{un}^2, \qquad \text{(B.5)}$$

which seemed numerically preferable to using a constraint based on the expression (B.3). One reason for the numerical advantages of (B.5) could be that it acts as a regularization constraint on the coefficient vector $\{\beta_n\}$.

B.2 Finding optimal low-order controllers

Some notes on the numerical optimization

It is tricky to hand-tune the controllers well enough for a fair comparison between them, so ”optimal” parameters for the different controller structures were found by local, gradient-based search performed with MATLAB’s `fmincon` with the methods `sqp` and `active set`. Note that the problem is non-convex, and hence it is not practically feasible to guarantee that the optimal parameters are found. The command `Multistart` was used for running multiple local searches from different starting points to increase the likelihood of finding the global optimum.

C

A practical matter: Anti-windup

When the actuator in a control loop is driven into saturation, the integral part of the PI(D) controllers causes a problematic effect known as integrator windup [Åström and Hägglund, 2006]. The problem is an effect of that a persistent non-zero control error is integrated during the time of saturation. When the control error later changes sign, it will take a long time for the control signal to counter-act the error, due to the large value of the integral part.

It is not expected that the saturation of the power amplifiers at ESS will be overly severe, our simulations show that for flat-top operation, even if the control signal saturates for short periods of time it does not lead to a significant reduction in control performance. Windup is however a notorious issue when working with controllers with integral action and it is considered good practice to always implement some means of anti-windup. In face of currently discussed reductions of the amplifier power overhead at ESS, which would increase the probability of saturation, it seems like a prudent decision to include anti-windup in the field controllers.

In the standard (real-valued) single-input single-output setting, there are several anti-windup strategies to choose from [Åström and Hägglund, 2006]. One simple option called clamping, is to stop increasing the I-part when the control signal saturates. Another approach is back-calculation, which is applicable to the complex signal setting if the saturation of a complex signal u is defined as

$$\boldsymbol{u}_{\text{sat}} = \begin{cases} \boldsymbol{u} & \text{if } |\boldsymbol{u}| \leq 1 \\ \boldsymbol{u}/\,|\boldsymbol{u}| & \text{if } |\boldsymbol{u}| \geq 1, \end{cases} \tag{C.1}$$

which corresponds to limiting the magnitude, but leaving the phase unchanged. Back-calculation in the complex-coefficient setting is illustrated in Figure C.1.

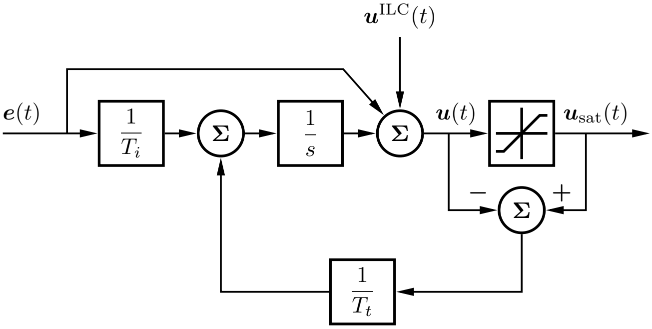


Figure C.1 Complex-valued anti-windup based on back-calculation. The nominal controller is given by $C(s) = (1 + 1/(sT_i))$. The signal $\boldsymbol{u}^{\mathrm{ILC}}(t)$ is the feedforward term that is updated by the ILC algorithm.

A nice feature of back-calculation in the complex setting, is that the phase of the resulting control signal $u_{\mathrm{sat}}(t)$ equals that of the nominal control signal $u(t)$. This seems tricky to achieve with a clamping-based approach.

D

Proofs and Calculations

D.1 Evaluation of beam-interaction term

We wish to compute the beam-interaction term

$$\Gamma_a(t) := \frac{1}{\omega_a}\frac{\partial}{\partial t} \underbrace{\iiint_V \boldsymbol{J}(\boldsymbol{r},t)\cdot \boldsymbol{E}_a(\boldsymbol{r})\, dV}_{=:\, g_a(t)} . \tag{D.1}$$

for the accelerating mode in (5.7).

The current of particle bunches, with time-separation T_{bunch}, is well described by point charges traveling along the z-axis through the center of the cavity. Thus, we model as a train of Dirac pulses, $\text{III}_T(t) = \sum_{\ell=-\infty}^{\infty} \delta(t-\ell T)$, with slowly varying DC-level,

$$\begin{aligned} I(z,t) &= I_{\text{DC}}(t)\frac{1}{f_{\text{bunch}}}\,\text{III}_{T_{\text{bunch}}}\left[\frac{z}{v} - \left(t + \frac{\phi_b(t)}{\omega_{\text{RF}}} - \frac{\phi_0}{\omega_{\text{RF}}}\right)\right] \\ &= I_{\text{DC}}(t)\sum_{\ell=-\infty}^{\infty} \exp\left[i\ell\omega_{\text{bunch}}\left[\frac{z}{v} - \left(t + \frac{\phi_b(t)}{\omega_{\text{RF}}} - \frac{\phi_0}{\omega_{\text{RF}}}\right)\right]\right], \end{aligned} \tag{D.2}$$

where the second equality follows from the Fourier expansion of the impulse train[1], and where the phase angles ϕ_0 and $\phi_b(t)$ are those discussed in Section 5.3; note that we via the time-dependence of ϕ_b allow for variations in the arrival time/phase of the bunches.

---

[1]

$$\text{III}_T(t) = \frac{1}{T}\sum_{\ell=-\infty}^{\infty} \mathrm{e}^{i2\pi\ell t/T}$$

Now, by inserting the expression (D.2) for the beam current into the integral in (D.1), we have

$$\begin{aligned} g_a(t) &= \int_0^L I(z,t)E_z(z)\,dz \\ &= I_{\mathrm{DC}}(t)\int_0^L \sum_{\ell=-\infty}^{\infty} \exp\left[i\ell\omega_{\mathrm{bunch}}\left[\frac{z}{v}-\left(t+\frac{\phi_b(t)}{\omega_{\mathrm{RF}}}-\frac{\phi_0}{\omega_{\mathrm{RF}}}\right)\right]\right] E_z(z)\,dz \\ &= I_{\mathrm{DC}}(t)\sum_{\ell=-\infty}^{\infty}\exp(-i\ell\omega_{\mathrm{bunch}}t) \\ &\qquad\times\int_0^L \exp\left[i\ell\omega_{\mathrm{bunch}}\left(\frac{z}{v}-\frac{\phi_b(t)}{\omega_{\mathrm{RF}}}+\frac{\phi_0}{\omega_{\mathrm{RF}}}\right)\right] E_z(z)\,dz \end{aligned} \tag{D.3}$$

Only the term where $-\ell\omega_{\mathrm{bunch}} = \omega_{\mathrm{RF}}$ will[2] be able to significantly excite the mode amplitude $\mathbf{a}$ in (5.7), so neglecting the other terms we have that

$$\begin{aligned} g_a(t) &\approx I_{\mathrm{DC}}(t)\mathrm{e}^{i\omega_{\mathrm{RF}}t}\int_0^L \exp\left[-i\omega_{\mathrm{RF}}\left(\frac{z}{v}-\frac{\phi_b(t)}{\omega_{\mathrm{RF}}}+\frac{\phi_0}{\omega_{\mathrm{RF}}}\right)\right] E_z(z)\,dz \\ &= I_{\mathrm{DC}}(t)\mathrm{e}^{-i\phi_b(t)}\int_0^L \mathrm{e}^{-i(\omega_{\mathrm{RF}}z/v+\phi_0)}E_z(z)\,dz\cdot\mathrm{e}^{i\omega_{\mathrm{RF}}t} = -\alpha\mathbf{I_b}\mathrm{e}^{i\omega_{\mathrm{RF}}t} \end{aligned} \tag{D.4}$$

where we introduced the phasor for the beam current

$$\mathbf{I_b} = -I_{\mathrm{DC}}(t)\mathrm{e}^{-i\phi_b(t)} = -I_{\mathrm{DC}}(t)\mathrm{e}^{i(\pi-\phi_b(t))}, \tag{D.5}$$

(the extra minus-sign will be convenient later on), and recognized the definition of α from (5.25) with the help of [Wangler, 2008, Exercise 2.12].

Now we are able to compute

$$\Gamma_a(t) \approx \frac{1}{\omega_a}\frac{d}{dt}g_a(t) = \frac{1}{\omega_a}\frac{d}{dt}\left(-\alpha\mathbf{I_b}\mathrm{e}^{i\omega_{\mathrm{RF}}t}\right) \approx \alpha\frac{\omega_{\mathrm{RF}}}{\omega_a}\mathbf{I_b} \approx i\alpha\mathbf{I_b}, \tag{D.6}$$

where we for the first approximation used that $\mathbf{I_b}$ varies slowly relative to $\mathrm{e}^{i\omega_{\mathrm{RF}}t}$, and in the second inequality we used that the modes that we consider have resonance frequencies close to the RF frequency.

---

[2] In cases of practical interest, the bunch period is a multiple of the RF period, so there is always one such term.

D.2 Proof of Theorem 1

Define the optimal cost-to-go function

$$V(x) := \int_x^1 \frac{r(u^*(x'))}{f(x', u^*(x'))}\, dx'.$$

Let u be an arbitrary control signal with $u(t) \in \mathcal{U}$, such that the corresponding state trajectory x satisfies (10.12b)–(10.12d). It then holds that

$$\begin{aligned}
r(u) + \frac{d}{dt} V(x(t)) &= r(u) + \frac{dV}{dx} f(x, u) \\
&= r(u) - \frac{r(u^*(x))}{f(x, u^*(x))} f(x, u) = \\
&= \frac{r(u) r(u^*(x))}{f(x, u^*(x))} \left(\frac{f(x, u^*(x))}{r(u^*(x))} - \frac{f(x, u)}{r(u)} \right) \geq 0, \qquad \text{(D.7)}
\end{aligned}$$

where the inequality follows from $f(x, u^*(x)) > 0$, $r(u) > 0$ and the definition (10.14) of u^*. Equality holds for $u = u^*$.

Integration of (D.7) gives, since $x(0) = 0$ and $x(t_f) = 1$,

$$\int_0^{t_f} r(u(t))\, dt \geq V(x(0)) - V(x(t_f)) = V(x(0)),$$

with equality for $u = u^*$. This proves optimality of u^*. □

E

Complex-Coefficient Systems in Control

This appendix contains the paper

Troeng, O., B. Bernhardsson, and C. Rivetta* (2017). "Complex-coefficient systems in control". In: *Proceedings of the 2017 American Control Conference.* (Seattle, WA, May 24–26, 2017).

*C. Rivetta is with Stanford Linear Accelerator, Stanford University.

Abstract

Complex-valued dynamics can be used for modeling rotationally invariant two-input two-output systems and bandpass systems when they are considered in the baseband. In a few instances, control design has been done in the complex domain, which facilitated analysis and synthesis. While previous work has been application specific, we will discuss more generally how complex valued dynamics arise, basic properties of these systems, revisit some classic control theoretic results in the complex setting, and discuss two novel examples of control design in the complex domain—accelerator cavity field control and feedback linearization of RF amplifiers.

E.1 Introduction

Certain systems are most conveniently modeled by complex-coefficient differential equations [Novotny and Wouterse, 1976; Harnefors, 2007; Dòria-Cerezo and Bodson, 2016; Byun and C.-W. Lee, 1988; Ren et al., 2013; Pippard, 2007; Martin, 2004; Beauchard et al., 2007; James and Kosut, 2010; Dong and Petersen, 2010]. In two cases, also control design has been done in the complex domain: regulation of electric machines [Novotny and Wouterse, 1976; Harnefors, 2007; Dòria-Cerezo and Bodson, 2016] and active vibration damping of rotating machinery [Byun and C.-W. Lee, 1988; Ren et al., 2013]. For these applications it was found that the complex formulation facilitated design and analysis compared to previous real-valued formulations.

The utility of complex-coefficient representations has also become apparent in the authors' work on accelerator cavity field control at the European Spallation Source [Peggs et al., 2013] and the SLAC National Accelerator Laboratory [Arthur, 2002]. When complex-coefficient transfer functions are analyzed in existing cavity field control literature, either the coefficients are assumed to be real, or an equivalent, real-coefficient, two-input two-output representation is considered, which complicates analysis and synthesis. With complex-coefficient systems, the standard Nyquist criterion can be used, rather than the less intuitive MIMO Nyquist criterion as in [Schilcher, 1998, p. 85].

As remarked in [Dòria-Cerezo and Bodson, 2016], little has been written on complex-coefficient systems in the control literature, some noteworthy examples are: linear systems theory [Lancaster and Rodman, 1995], Routh-Hurwitz's stability criterion [Frank, 1946], Kharitonov's theorem [Bose and Shi, 1987], the Nyquist stability criterion [Gataric and Garrigan, 1999] and root locus [Dòria-Cerezo and Bodson, 2016].

In the next section we look at how complex-coefficient dynamics arise in real-world applications, in Section E.3 we discuss basic properties of complex systems, in Sections E.4 and E.5 we consider some classic control theoretic results in the complex setting, and in the final two sections we discuss two novel applications of complex-coefficient systems for control analysis: cavity field control and Cartesian feedback linearization of RF amplifiers. In the Appendix we mention some pitfalls when analyzing complex systems with Matlab.

E.2 Origin of complex-valued dynamics

Rotationally invariant TITO systems

A two-input two-output (TITO) system

$$\mathbf{G}(s) = \begin{bmatrix} G_1(s) & -G_2(s) \\ G_2(s) & G_1(s) \end{bmatrix} \tag{E.1}$$

acting on signals $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ can be compactly represented by the complex SISO system

$$G(s) = G_1(s) + iG_2(s) \tag{E.2}$$

acting on signals $x_1 + ix_2$.

For example, the dynamics of the Foucault pendulum in the xy-plane, can, subject to small angle approximation, be represented by the complex differential equation

$$\ddot{z} + 2i\Omega\dot{z}\sin\phi + \omega^2 z = 0$$

where $z = x+iy$, ω is the natural frequency of the pendulum, Ω the rotational frequency of the Earth and ϕ is the latitude where the pendulum is located. See [Pippard, 2007] for similar examples.

Two other examples are the dynamics of balanced three-phase electric machines, which take the form (E.1) after application of an $\alpha\beta$-transformation [Novotny and Wouterse, 1976; Harnefors, 2007; Dòria-Cerezo and Bodson, 2016], and vibrations in rotating machines [Byun and C.-W. Lee, 1988], where the states x_1 and x_2 correspond to the x- and y-positions of the rotating shaft.

Bandpass systems

In applications such as telecommunications, where the signals of interest are narrowband around some frequency ω_c, it is convenient to consider the complex envelopes of the signals [Crochiere and Rabiner, 1983; Martin, 2004; Schreier and Scharf, 2010].

If the physical signal is given by

$$\begin{aligned} x_{\mathrm{c}}(t) &= A(t)\cos\left(\omega_c t + \phi(t)\right) \\ &= \mathrm{Re}\left\{A(t)\mathrm{e}^{i\phi(t)}\mathrm{e}^{i\omega_c t}\right\}, \end{aligned} \tag{E.3}$$

where the modulation, i.e. $A(t)$ and $\phi(t)$, varies slowly, then the complex envelope, or the equivalent baseband signal, is given by

$$\begin{aligned} \mathbf{x}_{\mathrm{BB}}(t) &:= A(t)\mathrm{e}^{i\phi(t)} \\ &= x_{\mathrm{Re}}(t) + ix_{\mathrm{Im}}(t), \end{aligned} \tag{E.4}$$

where $x_{\mathrm{Re}}(t)$ and $x_{\mathrm{Im}}(t)$ are real-valued.

An input-output-relation

$$Y_c(s) = G_c(s)U_c(s)$$

in the Laplace domain, is conveniently transformed to the base-band via the transformation $s \mapsto s + i\omega_c$, which gives

$$Y_c(s + i\omega_c) = G_c(s + i\omega_c)U_c(s + i\omega_c).$$

$Y_c(s + i\omega_c)$ and $U_c(s + i\omega_c)$ are equivalent baseband signals and thus the equivalent baseband model of $G_c(s)$ can be identified as

$$G_{\mathrm{BB}}(s) = G_c(s + i\omega_c). \tag{E.5}$$

If the signals of interest are have narrow support around ω_{RF}, high-frequency dynamics of $G_{\mathrm{BB}}(i\omega)$ can be neglected. Typically the resulting $G_{\mathrm{BB}}(s)$ has complex coefficients [Martin, 2004].

Example, baseband model of complex pole pair: The second order resonant system

$$\frac{2\zeta_0\omega_0 s}{s^2 + 2\zeta_0\omega_0 s + \omega_0^2}$$

has a baseband model given by

$$\frac{2\zeta_0\omega_0(s + i\omega_c)}{(s + \zeta\omega_0 + i\omega_0 + i\omega_c)(s + \zeta\omega_0 - i\omega_0 + i\omega_c)}.$$

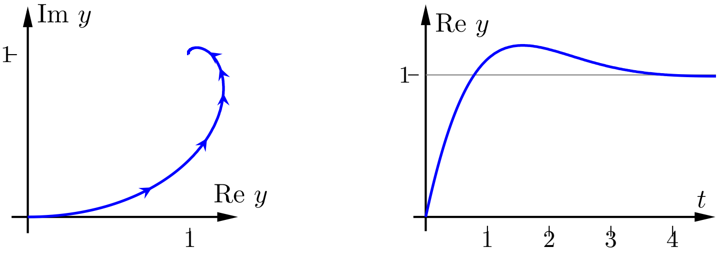
If $\omega_0 \approx \omega_c$ and the damping factor ζ is small, then for small s the first term in the denominator is $\approx 2i\omega_c$ and the following first-order approximation holds,

$$G_{\mathrm{BB}}(s) \approx \frac{\zeta_0\omega_0}{s + \zeta\omega_0 + i(\omega_c - \omega_0)}.$$

Example, baseband model of time delay: The baseband model of a time delay e^{-sT} is $\mathrm{e}^{-(s+i\omega_c)T} = \mathrm{e}^{-sT}\mathrm{e}^{-i\omega_c T}$. If ω_c is large, the phase of baseband model is sensitive to variations in T.

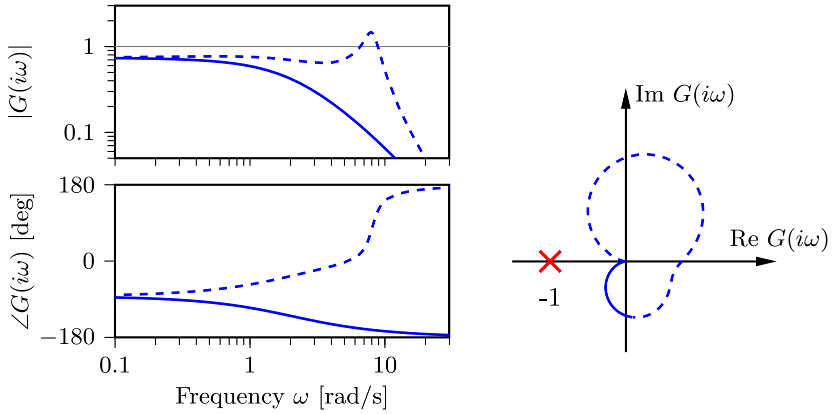
Quantum systems

Linear stochastic quantum systems are naturally described by complex, quantum stochastic differential equations—see [Beauchard et al., 2007; Dong and Petersen, 2010; James and Kosut, 2010] for control design for these systems.



(a) $y(t)$ in $\mathbb{C}$. (b) $\operatorname{Re} y(t)$.

Figure E.1 Step response $y(t)$ of first-order system $2/(s+1-i)$.



(a) Bode plot. (b) Nyquist curve.

Figure E.2 Frequency response of $G(s) = 12/(s+1+8i)/(s+2)$. Solid lines correspond to $\omega > 0$ and dashed lines to $\omega < 0$.

E.3 Complex signals and systems

In the previous section we motivated the study of systems of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{E.6}$$

where the signals and matrices are complex. The complex setting gives rise to some peculiarities not seen for real systems. In Fig. E.1 it is seen that a first-order complex system may exhibit an oscillatory step response and in Fig. E.2 it is seen that the frequency response is not necessarily conjugate symmetric with respect to positive and negative frequencies.

To better understand the structure of the system (E.6) we split its impulse

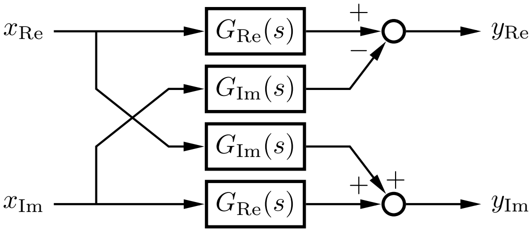


Figure E.3 Illustration of how a complex transfer function $G(s) = G_{\text{Re}}(s) + iG_{\text{Im}}(s)$ acts on a signal $x = x_{\text{Re}} + ix_{\text{Im}}$ to produce a signal $y(t) = y_{\text{Re}} + iy_{\text{Im}}$.

response into its real and imaginary parts

$$g(t) = g_{\text{Re}}(t) + ig_{\text{Im}}(t).$$

Denoting the Laplace transform of g_{Re} and g_{Im} by $G_{\text{Re}}(s)$ and $G_{\text{Im}}(s)$ respectively, it is seen that the transfer function for (E.6) is given by

$$G(s) = G_{\text{Re}}(s) + iG_{\text{Im}}(s). \tag{E.7}$$

Note that $G_{\text{Re}}(s)$ and $G_{\text{Im}}(s)$ are not the real and imaginary parts of $G(s)$, but that the subscripts are motivated by their relative contribution to the impulse response. Since $g_{\text{Re}}(t)$ and $g_{\text{Im}}(t)$ are real it follows that $g^*(t) = g_{\text{Re}}(t) - ig_{\text{Im}}(t)$ and

$$G^*(\bar{s}) = G_{\text{Re}}(s) - iG_{\text{Im}}(s),$$

due to conjugate symmetry of $G_{\text{Re}}(s)$ and $G_{\text{Im}}(s)$. Thus the decomposition (E.7) can be recovered from $G(s)$ via

$$G_{\text{Re}}(s) = \frac{G(s) + G^*(\bar{s})}{2}, \; G_{\text{Im}}(s) = \frac{G(s) - G^*(\bar{s})}{2i}. \tag{E.8}$$

The action of the complex coefficient transfer function (E.7) on a signal $x(t) = x_{\text{Re}}(t) + ix_{\text{Im}}(t)$ is illustrated in Fig. E.3.

Correspondence to real-valued representation

In some applications, [Schilcher, 1998; Dawson and T. H. Lee, 2004], the complex transfer function (E.7) is represented as a real, two-input two-output (TITO) system of the form

$$\mathbf{G}(s) = \begin{bmatrix} G_{\text{Re}}(s) & -G_{\text{Im}}(s) \\ G_{\text{Im}}(s) & G_{\text{Re}}(s) \end{bmatrix}, \tag{E.9}$$

acting on real-valued vector signals $\begin{bmatrix} x_{\text{Re}} & x_{\text{Im}} \end{bmatrix}^T$.

To better understand the relationship between the real system representation (E.9) and the complex representation (E.7), we consider the eigenvalue factorization of (E.9),

$$\mathbf{G}(i\omega) = \mathbf{S}^* \begin{bmatrix} G(i\omega) & 0 \\ 0 & G(-i\omega) \end{bmatrix} \mathbf{S}, \quad \mathbf{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}, \tag{E.10}$$

from which we see that the eigenvectors are independent of frequency, and that the eigenvector $\begin{bmatrix} 1 & -i \end{bmatrix}^T \leftrightarrow \begin{bmatrix} \cos\omega t & \sin\omega t \end{bmatrix}^T \leftrightarrow \mathrm{e}^{i\omega t}$, and similarly for $\begin{bmatrix} 1 & i \end{bmatrix}^T$.

While the real-coefficient representation (E.9) is necessary for physical implementation of complex transfer functions, it contains redundant information, and from (E.10) we see that the frequency responses of $G(i\omega)$ for positive and negative frequencies are intertwined, complicating analysis.

The eigenvectors of $\mathbf{G}(i\omega)$ are orthogonal, so the singular values of $\mathbf{G}(i\omega)$ are the modulus of the eigenvalues, thus

$$||\mathbf{G}||_\infty = ||G||_\infty \tag{E.11}$$

$$||\mathbf{G}||_2 = \sqrt{2}\, ||G||_2 \,. \tag{E.12}$$

Response to signal with specific direction

Even if the dynamics of a system is rotationally invariant, and hence can be represented as a complex SISO system, disturbances may have a specific direction. Consider for example phase noise in radio-frequency applications.

To illustrate the general behavior, consider without loss of generality, the output of (E.6) when subjected to a purely real signal $u(t) = \cos(\omega t)$,

$$\begin{aligned} y(t) = &\, |G_{\mathrm{Re}}(i\omega)| \cos\left(\omega t + \angle G_{\mathrm{Re}}(i\omega)\right) \\ &+ i\, |G_{\mathrm{Im}}(i\omega)| \cos\left(\omega t + \angle G_{\mathrm{Im}}(i\omega)\right). \end{aligned} \tag{E.13}$$

The signal (E.13) corresponds to Lissajous ovals in the complex plane, see Fig. E.4.

E.4 Frequency domain analysis

When analyzing complex systems in the frequency domain it is necessary to consider both positive and negative frequencies, as illustrated in Fig. E.2. For example a factor $\mathrm{e}^{i\epsilon}$ gives the impression of an improved phase margin if only positive frequencies are considered.

Nyquist's stability criterion

The assumptions and standard proof of the Nyquist stability criterion require no change as the argument principle is valid for any meromorphic function [Gataric and Garrigan, 1999].

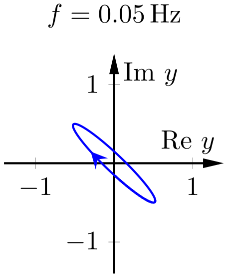

$f = 0.05\,\mathrm{Hz}$

Figure E.4 Lissajous ovals in the complex plane generated by excitation of $1/(s+1+i)$ by $u = \sin(2\pi f t)$.

Bode's sensitivity integral

Bode's sensitivity integral is typically considered over only positive frequencies [Freudenberg and Looze, 1985], however double-sided integration is necessary for complex coefficient transfer functions,

$$\int_{-\infty}^{\infty} \log |S(i\omega)| \, d\omega = 2\pi \sum_{k=1}^{N_p} \operatorname{Re} p_k, \tag{E.14}$$

where $\{p_k\}$ are the RHP poles of G. Also, unlike the real case, it is crucial to take the real part of the poles in (E.14). The proof is the same [Freudenberg and Looze, 1985].

That the single-sided version of (E.14) fails to hold in the complex case, is seen from that $G(s + i\delta)$ would correspond to different lower limits of integration for different δ.

Bode's complementary sensitivity integral

The relationship for the complementary sensitivity function [Middleton and Goodwin, 1990] needs the same modifications as in (E.14) to cover complex coefficient transfer functions,

$$\int_{-\infty}^{\infty} \log |T(i\omega)| \frac{d\omega}{\omega^2} = -\pi K_v^{-1} + \pi\tau + 2\pi \sum_{k=1}^{N_z} \operatorname{Re} \frac{1}{z_k}, \tag{E.15}$$

where τ is the time delay of the system, $\{z_k\}$ are the RHP zeros of $G(s)$ and $K_v = \lim_{s\to 0} sL(s)$. The result follows, with minor modifications, from the proof in [Middleton and Goodwin, 1990].

Bode's gain-phase relationship

Bode's gain-phase relationship which relates the phase of a real, minimum phase system $G(s)$, to the slope of its gain curve in logarithmic scale, does not hold for complex $G(s)$. The less intuitive relationship given by the double-sided version of the Kramers-Kronig relations [Kronig, 1926],

$$\angle G(i\omega_0) = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \frac{\log|G(i\omega)|}{\omega - \omega_0}\, d\omega,$$

where $\mathcal{P}$ denotes the Cauchy principle value, still holds for complex, minimum phase systems.

E.5 State-space analysis

Notions such as controllability, stability, etc. are analogous to the real case [Lancaster and Rodman, 1995]. Below, some special results are discussed in more detail.

$\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms

The H_2-norm can be calculated using the same formulas as in the real case, i.e. $\|C(sI - A)^{-1}B\|_2^2 = \operatorname{trace}(B^*YB) = \operatorname{trace}(CXC^*)$, where $X = X^*$ and $Y = Y^*$ are solutions to the complex Lyapunov equations $XA + A^*X + B^*B = 0$ and $AY + YA^* + CC^* = 0$ respectively.

The linear matrix inequalities for calculating the H_∞-norm also carry over, given that Hermitian transposition is used.

Remark: MATLAB's functions for $\mathcal{H}_\infty$-synthesis does not handle complex systems correctly.

LQR

In [Lancaster and Rodman, 1995] the optimal feedback for a complex linear system with respect to a cost functional

$$J = \int_0^\infty \begin{bmatrix} x(t) & u(t) \end{bmatrix}^* \begin{bmatrix} Q & N \\ N^* & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt$$

is derived. The optimal feedback is given by $u = -Kx$ with

$$K = R^{-1}(N^* + B^*X),$$

where the Hermitian matrix $X \geq 0$ satisfies the complex Riccati equation

$$A^*X + X^*A - (N + XB)R^{-1}(N^* + B^*X) + Q = 0.$$

For a first order system it seen that when $N = 0$, BK is real and positive, and the linear optimal regulator moves the closed-loop pole, $A - BK$, parallel to the real axis, further into the LHP.

The amazing robustness properties of LQR (for $N = 0$), that hold in the real case (infinite gain margin and $\geq 60^\circ$ phase margin), hold also in the complex case, even if the Nyquist curve is not symmetric with respect to the real axis.

Kalman filter for complex-valued normally distributed noise

It is natural to allow the state and measurement noise to be complex-valued. A complex-valued normally distributed variable Z with zero mean is determined by the matrices $E(ZZ^*)$ and $E(ZZ^T)$. If the latter is zero, one says that Z is circular-symmetric, which means that the distribution function is rotationally invariant in the complex plane. The paper [Dini and Mandic, 2012] discusses the general problem and introduces the concept of "widely linear state space models" to describe the optimal estimator in an aesthetic form.

E.6 Example I: Amplifier linearization

Cartesian feedback linearization of power amplifiers was actively studied 10–20 years ago as a means to reduce power consumption and adjacent channel interference in telecommunications [Johansson, 1991; Briffa and Faulkner, 1996; Dawson, 2003]. To avoid instability and performance degradation, the phase shift ϕ between up- and down-conversion needs to be properly compensated by an adjustment phase $\hat{\phi}$, see Fig. E.5.

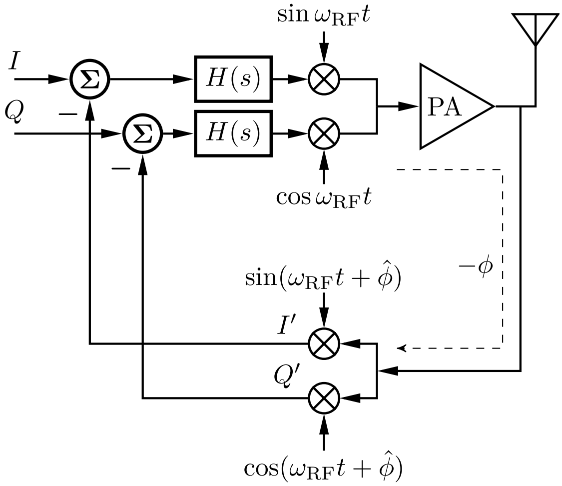


Figure E.5 Amplifier linearization by Cartesian feedback [Dawson, 2003], loop filters, up- and down-conversion mixers are shown.

If the amplifier is operating in an almost linear region, the open-loop system is well approximated by

$$G(s) = H(s)P(s)\mathrm{e}^{-\tau s}\mathrm{e}^{i\delta}, \tag{E.16}$$

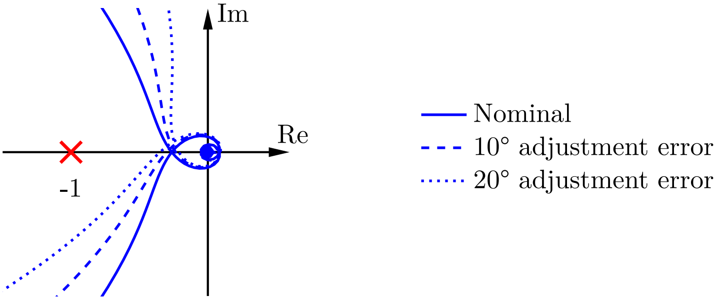


Figure E.6 Nyquist curves for Cartesian feedback loop with different phase adjustment errors. The nominal curve is from [Johansson, 1991, Sec. 4.2].

where $H(s)$ is the loop filter, $P(s)$ is a baseband model of the mixer and amplifier dynamics, τ is the loop delay and $\delta := (\hat{\phi} - \phi)$ is the phase adjustment error.

Although [Briffa, 1996] simulated (E.16) as a complex systems, the stability properties were analyzed using the equivalent TITO form (E.1). After algebraic computations and a clever observation it was shown that an adjustment error δ translates directly to a corresponding reduction in phase margin.

In the complex setting the same conclusion follows directly from the Nyquist criterion (Sec. E.4), by noting that the factor $\mathrm{e}^{i\delta}$ corresponds to a rotation of the Nyquist curve $H(i\omega)P(i\omega)\mathrm{e}^{-i\tau\omega}$ by δ radians, see Fig. E.6 for an illustration.

E.7 Example II: Cavity field control

In radio-frequency accelerators, particle bunches are accelerated by electromagnetic fields confined in RF cavities. The amplitude of the fields, and their phase relative the particle bunches, need to be precisely controlled [Wangler, 2008].

To see how complex transfer functions play a role in this, we first derive the baseband equations for the cavity and the RF system (Fig. E.7), and then design a complex $\mathcal{H}_\infty$-controller.

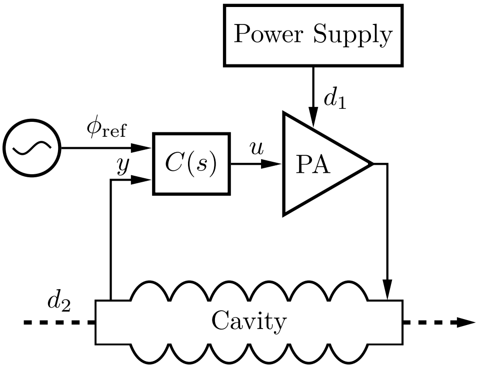


Figure E.7 Block diagram of RF system for one cavity. Variations in the voltage supply to the power amplifier (d_1) and ripple on the accelerated current (d_2), affect the system as load disturbances on the cavity input.

Model of cavity dynamics

From Maxwell's equations it follows that the electric field in the cavity can be expressed as a linear combination of eigenmodes,

$$\mathcal{E}(r,t) = \sum_{k=0}^{\infty} v_k(t)\mathbf{E}_k(r),$$

where the mode amplitudes e_k satisfy [Wangler, 2008, Ch 5, 10]

$$\frac{d^2}{dt^2}v_k + 2\gamma_k \frac{d}{dt}v_k + \omega_k^2 v_k = 2\kappa_k \frac{d}{dt}i_g + 2\alpha_k \frac{d}{dt}i_b, \tag{E.17}$$

where ω_k is the resonance frequency and γ_k the half bandwidth of mode k. κ_k and α_k quantify how the output of the power amplifier, modeled as a current i_g, and the accelerated particle current i_b, couple to the cavity field. The amplifier output i_g, can considered as the control signal. Variations in i_b enter as load disturbances.

The distribution of modes in the fundamental passband for an elliptical cavity is shown in Fig. E.8. The mode that is used for particle acceleration is typically the π-mode, and the purpose of the RF system is to excite the π-mode and control its phase and amplitude.

After both Laplace and baseband transformations of (E.17), we get

$$V_k(s) := \frac{1}{s + \gamma_k + i\Delta\omega_k}\left(\kappa_k I_g(s) + \alpha_k I_b(s)\right), \tag{E.18}$$

where $\Delta\omega_k = \omega_{\mathrm{RF}} - \omega_k$.

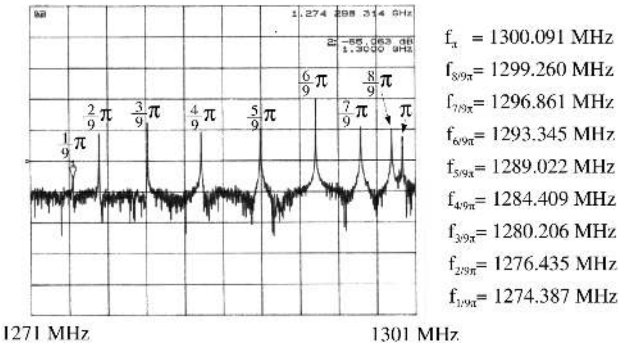


Figure E.8 Fundamental passband modes in the nine-cell TESLA cavity [Schilcher, 1998], that is used in many large accelerators, e.g. LCLS II. The π-mode is used for particle acceleration, while all other modes are detrimental to both particle acceleration and RF system stability.

A baseband model of the RF system in Fig. E.7, including the accelerating π-mode and one parasitic mode, now takes the form

$$P(s) = P_{\text{amp}}(s) e^{-i\omega_{\text{RF}}\tau} e^{-\tau s} \times \left[\frac{c_\pi \kappa_\pi / 2}{s + \gamma_\pi + i\Delta\omega_\pi} + \frac{c_1 \kappa_1 / 2}{s + \gamma_1 + i\Delta\omega_1} \right], \quad \text{(E.19)}$$

where $P_{\text{amp}}(s)$ is the dynamics of the power amplifier, τ is the system time delay, c_k quantify the coupling of mode k to the measurement probe, and $e^{-i\omega_{\text{RF}}\tau}$ is an additional factor resulting from the baseband transformation of the loop delay (cf. E.2). Complex quantities in (E.19) have been highlighted.

$\mathcal{H}_\infty$-synthesis example

As we demonstrate in the Appendix, the MATLAB functions for $\mathcal{H}_\infty$-synthesis does not work for complex coefficient systems, instead we used the TITO representation (E.1), which resulted in a controller that also had structure (E.1), from which we recovered a complex controller.

Specifications The main requirements for cavity field control is to suppress load disturbances while maintaining good robustness and avoiding excessive control signal activity. These requirements correspond to the following weights for mixed sensitivity synthesis,

$$\begin{aligned} W_S(s) &= 1 \\ W_{PS}(s) &= k_1 \cdot \frac{1}{s + \epsilon} \\ W_{KS}(s) &= k_2 \cdot \frac{s + \omega_{\text{bw}}}{s + N\omega_{\text{bw}}}. \end{aligned}$$

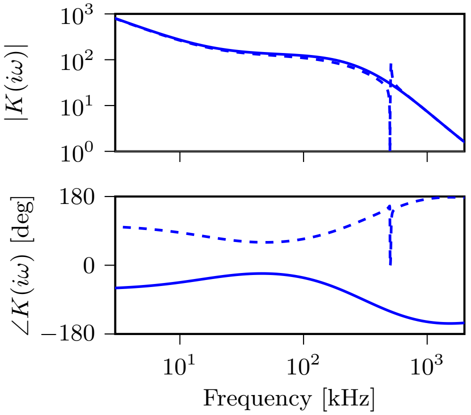


Figure E.9 Bode plot of $\mathcal{H}_\infty$-controller K, solid (dashed) lines correspond to positive (negative) frequencies.

By tuning how the parameters of the weighting functions, we arrived at a reasonable controller design.

Results The frequency response of the controller is shown in Fig. E.9, note the asymmetry with respect to positive and negative frequencies, which imply that the controller has complex-coefficients. It can also be seen that a notch has been by the $\mathcal{H}_\infty$-design, which allow a high control bandwidth for negative frequencies, without introducing positive feedback via the parasitic mode. The gang of four for the design is shown in Fig. E.10.

E.8 Conclusions

We have described some applications where system dynamics are conveniently modeled by complex-coefficient systems. Most control theoretic results developed for the real-valued case carry to the complex setting with minor changes, such as ensuring that transposition is Hermitian and that both negative and positive frequencies are considered. Instances where complex systems are erroneously handled by Matlab were pointed out.

A design example for cavity field control was presented in some detail and it was demonstrated that it was possible to synthesize a complex $\mathcal{H}_\infty$ controller. Due to the special structure of the plant, with resonance peaks only occurring at negative frequencies, the optimal controller had a markedly different frequency response for positive and negative frequencies.

We believe that there are many applications where a complex approach could bring increased insight and that there are related control theoretical

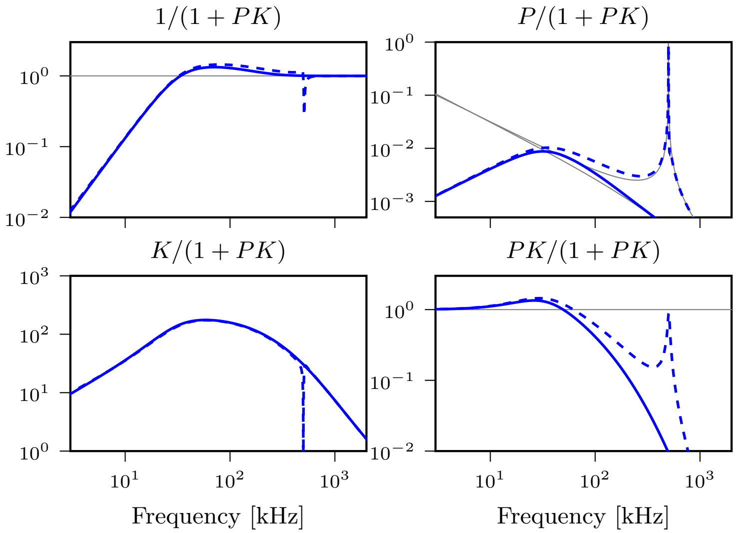


Figure E.10 Gang of four for the $\mathcal{H}_\infty$-design. The frequency response is not conjugate symmetric, so both positive frequencies (solid) and negative frequencies (dashed) are shown. The gray curves in the upper right plot is the frequency response of the plant P.

questions worthy of further investigation.

References

Arthur, J. et al. (2002). *Linac Coherent Light Source II, Conceptual Design Report.* Tech. rep. SLAC-R-593. Stanford Linear Accelerator.

Beauchard, K., J. M. Coron, M. Mirrahimi, and P. Rouchon (2007). “Implicit Lyapunov control of finite dimensional Schrödinger equations”. *Syst. & Control Lett.* **56**:5, pp. 388–395.

Bose, N. and Y. Shi (1987). “A simple general proof of Kharitonov’s generalized stability criterion”. *IEEE Trans. Circuits Syst.* **34**:10, pp. 1233–1237. ISSN: 0098-4094.

Briffa, M. A. (1996). *Linearization of RF Power Amplifiers.* PhD thesis. Victoria University of Technology.

Briffa, M. A. and M. Faulkner (1996). “Stability analysis of Cartesian feedback linearisation for amplifiers with weak nonlinearities”. *IEE Proceedings - Communications* **143**:4, pp. 212–218. ISSN: 1350-2425.

Byun, S.-W. and C.-W. Lee (1988). “Pole assignment in rotating disk vibration control using complex modal state feedback”. *Mech. Syst. and Signal Process.* **2**:3, pp. 225–241.

Crochiere, R. E. and L. R. Rabiner (1983). *Multirate Digital Signal Pocessing.* Prentice Hall, Englewood Cliffs, NJ.

Dawson, J. L. and T. H. Lee (2004). “Cartesian feedback for RF power amplifier linearization”. In: *Proc. 2004 Amer. Control Conf.* Vol. 1, pp. 361–366.

Dawson, J. L. (2003). *Feedback Linearization of RF Power Amplifiers.* PhD thesis. Stanford University.

Dini, D. H. and D. P. Mandic (2012). “Class of widely linear complex Kalman filters”. *IEEE Trans. Neural Netw.* **23**:5, pp. 775–786. ISSN: 2162-237X.

Dong, D. and I. R. Petersen (2010). “Quantum control theory and applications: a survey”. *IET Control Theory & Applicat.* **4**:12, pp. 2651–2671.

Dòria-Cerezo, A. and M. Bodson (2016). “Design of controllers for electrical power systems using a complex root locus method”. *IEEE Trans. Ind. Electron.* **63**:6, pp. 3706–3716. ISSN: 0278-0046.

Frank, E. (1946). “On the zeros of polynomials with complex coefficients”. *Bulletin of the Amer. Math. Soc.* **52**:2, pp. 144–157.

Freudenberg, J. S. and D. P. Looze (1985). “Right half plane poles and zeros and design trade-offs in feedback systems”. *IEEE Autom. Control* **30**:6, pp. 555–565.

Gataric, S. and N. R. Garrigan (1999). “Modeling and design of three-phase systems using complex transfer functions”. In: *30th Annu. IEEE Power Electron. Specialists Conf.* Vol. 2, pp. 691–697.

Harnefors, L. (2007). “Modeling of three-phase dynamic systems using complex transfer functions and transfer matrices”. *IEEE Trans. Ind. Electron.* **54**:4, pp. 2239–2248. ISSN: 0278-0046.

James, M. R. and R. L. Kosut (2010). “Quantum estimation and control”. In: Levine, W. S. (Ed.). *The Control Systems Handbook: Control System Applications.* Chap. 31.

Johansson, M. (1991). *Linearization of RF power amplifiers using Cartesian feedback.* Lic. thesis. Lund University.

Kronig, R. (1926). “On the theory of dispersion of X-rays”. *J. Opt. Soc. Am.* **12**:6, pp. 547–557.

Lancaster, P. and L. Rodman (1995). *Algebraic Riccati Equations.* Oxford University Press, Oxford, UK.

Martin, K. W. (2004). “Complex signal processing is not complex”. *IEEE Transactions on Circuits and Systems I* **51**:9, pp. 1823–1836.

Middleton, R. and G. Goodwin (1990). *Digital Control and Estimation: A Unified Approach.* Prentice Hall, Englewood Cliffs, NJ, pp. 423–433.

Novotny, D. W. and J. H. Wouterse (1976). “Induction machine transfer functions and dynamic response by means of complex time variables”. *IEEE Trans. Power App. Syst.* **95**:4, pp. 1325–1335. ISSN: 0018-9510.

Peggs, S. et al. (2013). *ESS Technical Design Report.* Tech. rep. ESS-doc-274. European Spallation Source.

Pippard, A. B. (2007). *The Physics of Vibration.* Cambridge University Press, New York, NY. ISBN: 9780521033336.

Ren, Y., D. Su, and J. Fang (2013). “Whirling modes stability criterion for a magnetically suspended flywheel rotor with significant gyroscopic effects and bending modes”. *IEEE Trans. Power Electron.* **28**:12, pp. 5890–5901. ISSN: 0885-8993.

Schilcher, T. (1998). *Vector Sum Control of Pulsed Accelerating Fields in Lorentz Force Detuned Superconducting Cavities.* PhD thesis. Hamburg University.

Schreier, P. J. and L. L. Scharf (2010). *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals.* Cambridge University Press, Cambridge, UK. ISBN: 978-0-511-67772-4.

Wangler, T. P. (2008). *RF Linear Accelerators.* 2nd ed. Wiley-VCH, Weinheim, Germany. ISBN: 978-3-527-40680-7.

E.9 Appendix: Complex-coefficient systems in Matlab

MATLAB handles complex-coefficient systems incorrectly. We detected the following issues with version R2016a (Linux).

In the `nyquist` plot, the frequency response for negative frequencies equals that at positive frequencies, which is incorrect for complex coefficients system. `hinfnorm` only considers positive frequencies, while `minreal` does not support complex data at all.

The first example of $\mathcal{H}_\infty$-synthesis in the MATLAB documentation,

```
G = (s-1)/(s+1)^2;
W1 = 0.1*(s+100)/(100*s+1);
W2 = 0.1;
[~,~,GAM] = mixsyn(G,W1,W2,[])
```

gives `GAM=0.23`. Multiplying the plant `G` by a complex factor `exp(0.4i)` should not affect the resulting value, as the factor could be canceled by the controller. However the result in this case is `GAM=0.40`, thus demonstrating incorrectness.