

Particle Acceleration and Detection

Stefan Simrock
Zheqiao Geng

Low-Level Radio Frequency Systems

Particle Acceleration and Detection

Series Editors

Alexander Chao, SLAC, Stanford University, Menlo Park, CA, USA

Katsunobu Oide, KEK, High Energy Accelerator Research Organization, Tsukuba, Japan

Werner Riegler, Detector group, CERN, Genéve, Switzerland

Vladimir Shiltsev, Accelerator Physics Center, Fermi National Accelerator Lab, Batavia, IL, USA

Frank Zimmermann, BE Department, ABP Group, CERN, Genéve, Switzerland

The series *Particle Acceleration and Detection* is devoted to monograph texts dealing with all aspects of particle acceleration and detection research and advanced teaching. The scope also includes topics such as beam physics and instrumentation as well as applications. Presentations should strongly emphasize the underlying physical and engineering sciences. Of particular interest are

- contributions which relate fundamental research to new applications beyond the immediate realm of the original field of research
- contributions which connect fundamental research in the aforementioned fields to fundamental research in related physical or engineering sciences
- concise accounts of newly emerging important topics that are embedded in a broader framework in order to provide quick but readable access of very new material to a larger audience.

The books forming this collection will be of importance to graduate students and active researchers alike

More information about this series at <https://link.springer.com/bookseries/5267>

Stefan Simrock • Zheqiao Geng

Low-Level Radio Frequency Systems



Springer

Stefan Simrock
ITER
St. Paul-lez-Durance, France

Zheqiao Geng
Paul Scherrer Institute
Villigen PSI, Aargau, Switzerland

ISSN 1611-1052 ISSN 2365-0877 (electronic)
Particle Acceleration and Detection
ISBN 978-3-030-94418-6 ISBN 978-3-030-94419-3 (eBook)
<https://doi.org/10.1007/978-3-030-94419-3>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*For Josef and Maria.
Stefan Simrock*

*For Yingxin and Xintian.
Zheqiao Geng*

Preface

Low-level radio frequency (LLRF) systems of particle accelerators become more critical because they become a limiting factor with the increasing RF field stability requirements demanded by the free-electron laser (FEL) machines and the fourth-generation synchrotron radiation sources. An LLRF system measures, regulates, and stabilizes the RF field used for beam acceleration and provides automation tools to set up, optimize, and automate the RF system operation.

Designing an LLRF system is challenging. An LLRF physicist/engineer must have a wide range of cross-domain knowledge, such as RF/digital circuits, software engineering, signals and systems, control theory, digital signal processing, and beam physics. The primary goal of this book is to provide an overview of the basic knowledge required for analyzing and designing an LLRF system. We focus on the basic principles and algorithms without giving many details about concrete implementations. Realizing the RF signal detection, actuation, and control algorithms strongly depends on the technology (e.g., in hardware, firmware, or software) chosen by the designer and is out of the scope of this book. The research and development of LLRF systems continue to evolve with the emerging new requirements of new accelerator projects. More advanced algorithms, benefiting from the modern control theory and the machine learning progress, are being developed. This book will not cover all these new development because of the limitation of space. However, the basic algorithms covered by this book are sufficient for most LLRF systems and are a basis for developing advanced algorithms.

This book is aimed at LLRF scientists and engineers at the postgraduate level. They should have basic knowledge of beam dynamics, college mathematics (calculus, linear algebra, and probability theory), signals and systems, control theory, and digital signal processing. The book is written with minimum math and provides many methods and results that can directly apply to engineering practices.

This book arose from the development of learning materials for the International Accelerator School of Linear Colliders (ISLC). Some examples in the book are based on the tests at the Free-Electron LASer in Hamburg (FLASH) when the author worked at DESY, Germany.

The book consists of nine chapters. Chapter 1 is an introduction to the accelerator RF systems and the context of LLRF systems. Chapter 2 provides an overview of the RF control strategies. Chapter 3 discusses the RF system models that are the basis of the remaining chapters. Chapter 4 collects the RF field and cavity resonance control algorithms, where the control theory is a solid background to understand the contents. Chapter 5 introduces the RF detection and actuation algorithms based mainly on the digital signal processing theory. In Chap. 6, we focus on the noise in LLRF systems. Typical literature on noise analysis has lots of mathematics and is not easy to follow. Therefore, we also include a brief introduction to the noise theory and the noise models of basic RF components, collecting the important concepts related to LLRF systems. Chapter 7 summarizes the nonlinearity issues and the mitigation methods for RF controls. Chapter 8 provides a conceptual overview of the timing and synchronization systems, highlighting the critical points in design and implementation. Finally, Chap. 9 discusses the high-level applications to set up, optimize, and automate the LLRF system operation.

We want to thank Enrico Rubiola (Femto-st) and Matthias Hoffmann (DESY) for the valuable discussion about noise analysis. We appreciate many colleagues and experts: Sven Reiche (PSI), Alexander Dietrich (PSI), Riccardo Zennaro (PSI), Dmitry Teytelman (Dimtel, Inc.), Enrico Rubiola (Femto-st), Rodolphe Boudot (Femto-st), and Valeri Ayvazyan (DESY), for providing some data and materials used in the book. Special thanks also to the RF & LLRF teams in DESY, SLAC, and PSI for the great support during my (Zheqiao Geng) career. We would also like to thank Alexander Wu Chao (SLAC/Stanford University) for proposing this book to Springer, which has greatly motivated us to finish this book much earlier than expected. Finally, we thank Springer staff, in particular Hisako Niko and Nagarajan Paramasivam, for their help and support.

St. Paul-lez-Durance, France

Stefan Simrock

Villigen PSI, Switzerland

Zheqiao Geng

Contents

1	Introduction	1
1.1	RF Systems of Particle Accelerators	1
1.2	Principles of Beam Acceleration	5
1.2.1	Acceleration in Standing-Wave Cavities	6
1.2.2	Acceleration in Traveling-Wave Structures	9
1.3	Disturbances to RF Fields	10
1.3.1	Electronic Noise	11
1.3.2	Temperature and Humidity	12
1.3.3	Mechanical Vibrations	14
1.3.4	Beam Loading	16
1.4	LLRF Systems Overview	16
1.4.1	Requirements and Architecture	17
1.4.2	Context in Particle Accelerators	19
1.4.3	A Brief History	21
1.5	Summary	23
References		23
2	RF Control Strategy	25
2.1	Feedback and Feedforward Control	25
2.2	Amplitude/Phase and In-Phase/Quadrature Control	26
2.3	RF Control Loop Architecture	28
2.3.1	Generator Driven Resonator	29
2.3.2	Self-Excited Loop	29
2.3.3	Phase-Locked Loop	30
2.4	Analog and Digital Control	31
2.5	Single-Cavity and Vector-Sum Control	33
2.6	Summary	34
References		34

3 RF System Models	35
3.1 General Assumptions	35
3.2 RF Modeling Method	37
3.2.1 RF Signal Description	37
3.2.2 Principle of RF Signal Detection	38
3.2.3 Phasor Laplace Transform	40
3.3 Single-Cell Cavity Model	43
3.3.1 Parallel RLC Circuit Model	43
3.3.2 Cavity Phasor Transfer Function	45
3.3.3 Cavity Step Response	46
3.3.4 Cavity Response to RF Power	48
3.3.5 Cavity Response to a Single Bunch	50
3.3.6 Cavity Response to a Bunch Train	51
3.3.7 Cavity Equation with Voltage Drives	53
3.3.8 Interaction Between Cavity Voltage and Beam	55
3.3.9 Forward and Reflected RF Power	59
3.3.10 Mechanical Model	62
3.4 Multi-cell Cavity Model	63
3.4.1 Coupled RLC Circuit Model	63
3.4.2 Multi-cell Cavity Phasor Equations	65
3.4.3 Passband Modes	67
3.4.4 Transient in Cavity Cells with RF Drive	70
3.5 Traveling-Wave Structure Model	71
3.5.1 Filling of Structures	72
3.5.2 Structure Phasor Transfer Function	74
3.6 Modeling of Important RF Devices	75
3.6.1 Transmission Line Model	75
3.6.2 RF Amplifier Model	76
3.6.3 RF Pulse Compressor Model	77
3.7 Application of RF System Models	78
3.8 Summary	80
References	80
4 RF Field Control	83
4.1 Requirements to RF Field Control	83
4.2 Generator Driven Resonator Control	85
4.2.1 Feedback Stability for Single-Cell Cavities	86
4.2.2 Feedback Stability for Multi-cell Cavities	89
4.2.3 Active Disturbance Rejection Control	90
4.2.4 Advanced Control Algorithms	94
4.3 Self-Excited Loop Control	98
4.3.1 Free-Running SEL	98
4.3.2 SEL with Amplitude Limiter	99
4.3.3 SEL with Feedback Control	101

4.4	Phase-Locked Loop Control	103
4.4.1	Introduction to PLL	104
4.4.2	Modeling of PLL for Cavity Control	105
4.4.3	Feedback Analysis of PLL	108
4.5	Adaptive Feedforward	113
4.5.1	Adapt Feedforward with Feedback Actuation	114
4.5.2	Iterative Learning Control	117
4.6	Cavity Resonance Control	122
4.6.1	Detuning Measurement	123
4.6.2	Cavity Tuners	129
4.6.3	Tuning Control Scheme	132
4.6.4	Ponderomotive Effects	142
4.7	Summary	146
	References	146
5	RF Detection and Actuation	149
5.1	RF Detection Schemes	149
5.1.1	Amplitude and Phase Detectors	150
5.1.2	Analog I/Q Demodulator	152
5.1.3	Down-Converter and IF Sampling	153
5.1.4	Direct RF Sampling	154
5.2	RF Detection Algorithms	155
5.2.1	I/Q Demodulation	155
5.2.2	Non-I/Q Demodulation	157
5.2.3	Digital Down-Conversion	166
5.2.4	Handling of Time-varying Frequency	167
5.2.5	RF Detection with Reference Tracking	168
5.3	RF Actuation Schemes	176
5.3.1	Direct Up-Conversion	177
5.3.2	Single Sideband Up-Conversion	178
5.3.3	IF Up-Conversion	180
5.4	Summary	181
	References	181
6	Noise in RF Systems	183
6.1	General Description of Noise	183
6.1.1	Basic Concepts of Noise	183
6.1.2	Estimation of PSD and SNR	186
6.1.3	Correlation of Noise	190
6.1.4	Additive Noise and Parametric Noise	195
6.1.5	White Noise and 1/f Noise	195
6.1.6	Noise Factor	196
6.1.7	Phase Noise and Amplitude Noise	197
6.1.8	Additive Noise and RF Jitter	201
6.1.9	Frequency-Domain Meaning of RMS Value	202
6.1.10	Drift and Jitter	202

6.2	Noise Model of Basic RF Components	204
6.2.1	Two-Port Passive RF Components	204
6.2.2	Power Splitter and Combiner	205
6.2.3	RF Amplifier	207
6.2.4	Mixer	209
6.2.5	Frequency Divider and Multiplier	211
6.2.6	Analog-to-Digital Converter	213
6.2.7	Digital-to-Analog Converter	221
6.3	Noise Transfer in RF Control Loops	226
6.3.1	Noise Transfer in Feedback Control	226
6.3.2	Noise Transfer in Pulse-to-Pulse Control	229
6.4	RF System Noise Specification	231
6.4.1	Noise Specification Strategy	231
6.4.2	Accelerator Global Noise Model	234
6.4.3	Specification of RF Reference Phase Noise	240
6.5	RF Station Noise Model	244
6.5.1	RF Station Noise Overview	244
6.5.2	RF Reference Phase Noise	245
6.5.3	RF Driving Chain Noise	246
6.5.4	RF Measurement Chain Noise	247
6.5.5	Estimation of RF Measurement Chain Noise	249
6.5.6	Estimation of RF Driving Chain Noise	251
6.5.7	Estimation of RF Field Noise	254
6.5.8	Validation of RF Station Noise Model	255
6.5.9	Specification of RF Component Noise	257
6.6	RF Detector Drift Correction	258
6.6.1	Reference Tracking	258
6.6.2	Drift Calibration	259
6.6.3	Beam-Based Feedback	261
6.7	Summary	262
	References	262
7	Nonlinearity in RF Systems	265
7.1	Basic Concepts	265
7.1.1	1-dB Compression Point	266
7.1.2	Third-Order Intercept Point	267
7.1.3	AM-PM Conversion	268
7.1.4	Nonlinearity Induced RF Detection Error	268
7.1.5	Nonlinearity Induced RF Driving Disturbances	269
7.2	Nonlinearity of RF Amplifiers	270
7.2.1	Look-Up Table Model	272
7.2.2	Analytical Model	274
7.2.3	Dynamical Model	276

7.3	Handling of Amplifier Nonlinearity in RF Control	277
7.3.1	RF Amplitude Control with High Voltage	278
7.3.2	Gain Scheduling	280
7.3.3	LUT-Based Linearization	281
7.3.4	Linearization Loop	282
7.4	Summary	285
	References	285
8	Timing and Synchronization	287
8.1	Overview	287
8.2	Master Oscillator	288
8.2.1	RF and Laser Oscillator	289
8.2.2	Synchronization of Two Oscillators	290
8.3	Timing System	292
8.3.1	Timing Fiducial Generation	293
8.3.2	Common Subharmonic	295
8.3.3	Client Trigger Generation	296
8.4	Synchronization System	297
8.4.1	Synchronization Signal Distribution	297
8.4.2	Phase Drift Mitigation	299
8.4.3	Client Synchronization	306
8.5	Robust Timing Relations	312
8.5.1	Timing Relation Highlight	312
8.5.2	Timing Relation Uncertainty	315
8.5.3	Strategies for Robust Timing Relations	317
8.6	Summary	323
	References	323
9	LLRF Applications	327
9.1	Overview	327
9.2	Parameter Optimization	328
9.2.1	RF Pulse Shaping	328
9.2.2	DAC Offset Correction	332
9.2.3	Parameter Scanning	333
9.3	RF Calibration	335
9.3.1	Beam-Induced Transient	336
9.3.2	Accelerating Voltage and Beam Phase Calibration	340
9.3.3	Cavity Input Power and Phase Adjustment	345
9.3.4	Vector-Sum Calibration	346
9.3.5	Cavity Forward and Reflected Signals Calibration	350
9.3.6	RF Signal Power Calibration	353
9.4	RF System Identification	354
9.4.1	Cavity Input Coupling Factor Identification	354
9.4.2	System Gain and System Phase Identification	358
9.4.3	Cavity Parameters Identification	362

9.5 Beam Loading Compensation	368
9.6 Summary	370
References	370
Index	373

Abbreviations

ADC	Analog-to-Digital Converter
ADRC	Active Disturbance Rejection Control
AFF	Adaptive Feedforward
AM	Amplitude Modulation
ANC	Adaptive Noise Cancellation
ASTRA	A Space Charge Tracking Algorithm
BAM	Bunch Arrival Time Monitor
BBFB	Beam-Based Feedback
BCM	Bunch Compression Monitor
BEPCII	Beijing Electron-Positron Collider II
BLC	Beam Loading Compensation
BOC	Barrel Open Cavity
BOM-PD	Balanced Optical to Microwave Phase Detector
BPF	Band-Pass Filter
BPM	Beam Position Monitor
CIC	Cascaded Integrator-Comb
CORDIC	Coordinate Rotation Digital Computer
CPSD	Cross Power Spectral Density
CW	Continuous Wave
DAC	Digital-to-Analog Converter
DC	Direct Current
DDC	Digital Down-Conversion
DDS	Direct Digital Synthesis
DFT	Discrete Fourier Transform
DOOCS	Distributed Object-Oriented Control System
DRO	Dielectric Resonator Oscillator
DSB	Double Sideband
DSP	Digital Signal Processor
EMI	Electromagnetic Interference
ENOB	Effective Number of Bits
EO	Electro-Optical

EPICS	Experimental Physics and Industrial Control System
EVG	Event Generator
EVR	Event Receiver
FEL	Free-Electron Laser
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FLASH	Free-Electron Laser in Hamburg
FPGA	Field Programmable Gate Array
FSM	Finite-State Machine
GA	Genetic Algorithm
GDR	Generator Driven Resonator
GPS	Global Positioning System
GPSD-Rb	GPS Disciplined Rubidium Oscillator
HLA	High-Level Application
HOM	Higher-Order Mode
HPRF	High-Power Radio Frequency
HVPS	High Voltage Power Supply
I/Q	In-phase/Quadrature
IF	Intermediate Frequency
IFFT	Inverse Fast Fourier Transform
IHEP	Institute of High Energy Physics
IIP3	Input Third-order Intercept Point
IIR	Infinite Impulse Response
ILC	Iterative Learning Control
IP3	Third-order Intercept Point
LCLS	Linac Coherent Light Source
LFD	Lorentz-Force Detuning
LHP	Left Half-Plane
LLRF	Low-Level Radio Frequency
LMO	Laser Master Oscillator
LMS	Least Mean Square
LO	Local Oscillator
LPF	Low-Pass Filter
LQG	Linear Quadratic Gaussian
LQR	Linear Quadratic Regulator
LUT	Look-Up Table
MIMO	Multiple Input Multiple Output
MO	Master Oscillator
MPC	Model Predictive Control
MPS	Machine Protection System
MSPS	Mega-Samples Per Second
NCO	Numerically Controlled Oscillator
NF	Noise Figure
OCXO	Oven-Controlled Crystal Oscillator

OIP3	Output Third-order Intercept Point
OPD	Operating Point Determination
PFD	Phase and Frequency Detector
PID	Proportional-Integral-Derivative. Architecture of feedback controllers, may be also configured as proportional (P), proportional-integral (PI), or proportional-derivative (PD) controller
PLL	Phase-Locked Loop
PM	Phase Modulation
PPS	Personnel Protection System
PRBS	Pseudo Random Binary Sequence
PSD	Power Spectral Density
PSO	Particle Swarm Optimization
RF	Radio Frequency
RHP	Right Half-Plane
RMO	RF Master Oscillator
RMS	Root Mean Square
SEL	Self-Excited Loop
SISO	Single Input Single Output
SLAC	Stanford Linear Accelerator Center
SLED	SLAC Energy Doubler
SNR	Signal-to-Noise Ratio
SSB	Single Sideband
SwissFEL	Swiss Free-Electron Laser. A FEL machine in Switzerland
TESLA	TeV-Energy Superconducting Linear Accelerator
VCO	Voltage-Controlled Oscillator
ZOH	Zero-Order Hold

Chapter 1

Introduction



The widely-used and best-understood technology for particle acceleration is the *radio frequency (RF)* standing-wave cavities or traveling-wave structures. High-power RF amplifiers produce the RF drive power to build up electromagnetic fields in the cavities or structures. We also denote the electromagnetic fields as *RF fields* or *cavity fields*, oscillating at the RF frequency. When a particle passes through the cavities or structures, it exchanges energy with the RF fields. Therefore, the RF field stability is critical to achieving a stable acceleration. A *low-level radio frequency (LLRF)* system, which is the main topic of this book, is typically employed to measure, control, and regulate the RF fields. The state-of-the-art accelerators, especially the FELs, require highly stable RF fields and complex operation scenarios, which impose critical requirements to the LLRF systems. In this chapter, the fundamental aspects of an LLRF system, including its functions, typical architecture, and its role in an accelerator, will be presented. We will also review the basic principles of beam acceleration and summarize the disturbances to the RF fields.

1.1 RF Systems of Particle Accelerators

Particle accelerators are widely used for high-energy physics study, synchrotron radiation, free-electron lasers (FEL), and so on. All these applications require particle beams (e.g., electrons, protons, or heavy ions) to be accelerated to energies ranging from kilo-electronvolt (KeV) to tera-electronvolt (TeV).

The beam acceleration is performed by an RF system. The RF system of a particle accelerator consists of two subsystems: LLRF and *high-power RF (HPRF)* systems. The HPRF system accepts a modulated RF signal produced by the RF controller and amplifies its power to a level up to tens of MW. The RF power drives the cavities or structures to establish RF fields for beam acceleration. Figures 1.1 (Phinney et al. 2007) and 1.2 (Loehl et al. 2013; Milne et al. 2017) illustrate two typical *RF stations* of the HPRF systems of two different accelerators. Both employ klystrons to amplify

the RF power and use a waveguide power distribution system to transmit the RF power to the cavities or structures. To increase the RF peak power for an RF station operating in the pulsed mode, a pulse compressor, such as the Barrel Open Cavity (BOC) in Fig. 1.2, is adopted. A high-power circulator is often installed before a standing-wave cavity to absorb the reflected RF power. High-power variable attenuators or phase shifters are often employed to adjust the power and phase of the inputs of multiple cavities or structures.

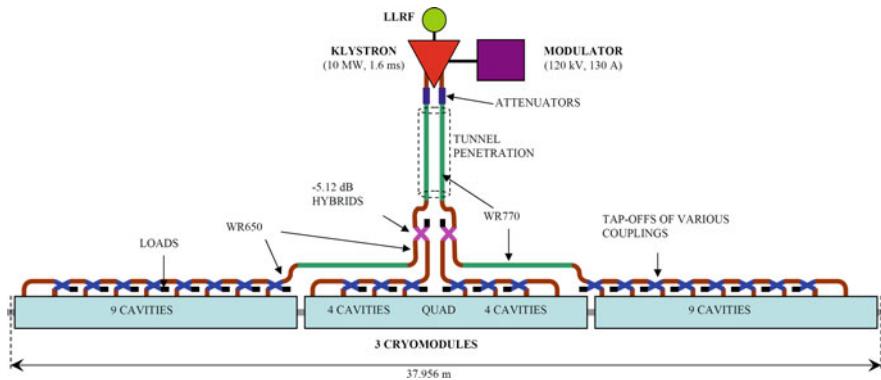


Fig. 1.1 An RF station of the Internaltional Linear Collider consisting of 26 standing-wave cavities (Courtesy of N. Phinney et al)

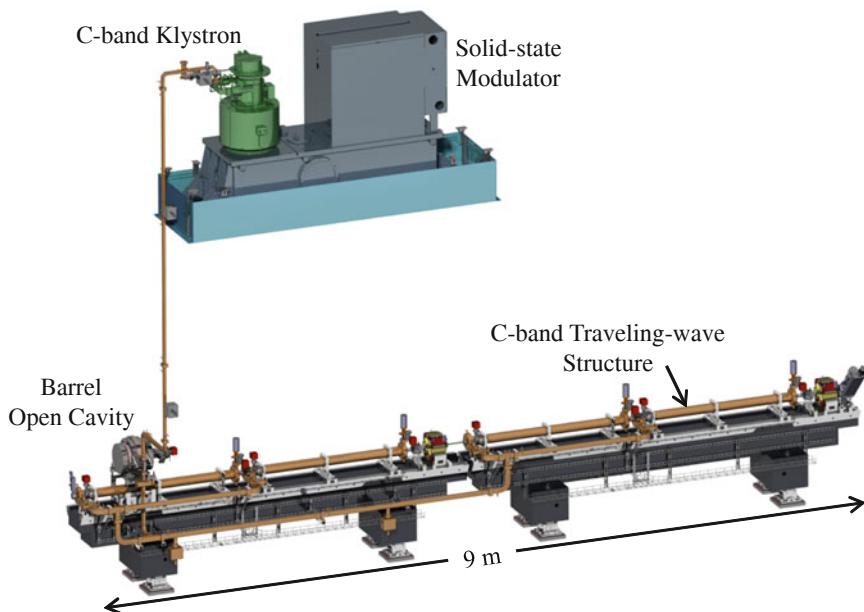


Fig. 1.2 A C-band (5712 MHz) RF station of the SwissFEL (Swiss X-ray Free-Electron Laser) consisting of 4 traveling-wave structures

An *LLRF node* is usually associated with an RF station to measure the RF signals at various locations (e.g., klystron input/output, cavity input, and cavity fields picked with a probe) and control the input to the HPRF amplifier. Since there are often multiple RF stations in a particle accelerator, multiple LLRF nodes are required. Therefore, the LLRF system is typically a distributed control system.

The RF station in Fig. 1.1 uses standing-wave cavities for beam acceleration, whereas the RF station in Fig. 1.2 adopts traveling-wave structures. These two types of RF structures (Miller 1986) have different properties that require different measurement and control strategies.

A *standing-wave cavity* has one input coupler to feed the RF drive power. The end walls of the cavity reflect the RF waves and form a standing wave in the cavity. The forward and backward waves propagate RF energy much faster (close to the velocity of light) than the energy propagation in a traveling-wave structure. On the cavity axis, the electric fields at different locations grow exponentially simultaneously. From the viewpoint of a particle passing through it, a standing-wave cavity provides resonating electric fields with the same phase along its axis. The input RF power, excluding the parts delivered to the beam and dissipated in the cavity walls, will be coupled back to the power transmission line through the input coupler. Therefore, a circulator is typically installed in the power transmission line to absorb the backward (reflected) RF power from the cavity. Figure 1.3 (Phinney et al. 2007; Marhauser et al. 2018; Ramberger et al. 2008; Ostroumov et al. 2012) shows some typical standing-wave cavities used in different accelerators.

A *traveling-wave structure* has one input coupler to feed the RF drive power and one output coupler to dump the transmitted RF power to a matched load. Compared to the standing-wave cavity, a well-tuned traveling-wave structure does not reflect much power from the input coupler, so a circulator is not needed. Of course, large reflection may happen when there are breakdowns in the structure or input/output couplers. It implies that a reflected power interlock is necessary to protect the RF power amplifier. In a traveling-wave structure, only a forward traveling wave is present. The energy propagation speed of the traveling wave is the *group velocity*, which is typically a few percent of the velocity of light. It allows building up the electric fields cell-by-cell towards the load of the structure. Such an RF field excitation procedure is also called *filling*, and the required time to establish the RF fields in the entire structure is denoted as the *filling time*. For example, the 2-m long C-band (5712 MHz) traveling-wave structure used in SwissFEL (Milne et al. 2017) has a filling time of 920 ns. The particle beam should be injected into the traveling-wave structure only after the entire structure is filled. From the viewpoint of a particle passing through it, a traveling-wave structure provides electric fields along its axis, traveling together with the particle. The *phase velocity* of the traveling wave is tuned to be the same as the particle velocity so that the particle experiences a constant RF phase along the axis, just like surfing on a wave. As an example, the SwissFEL C-band structures installed in the tunnel are depicted in Fig. 1.4.

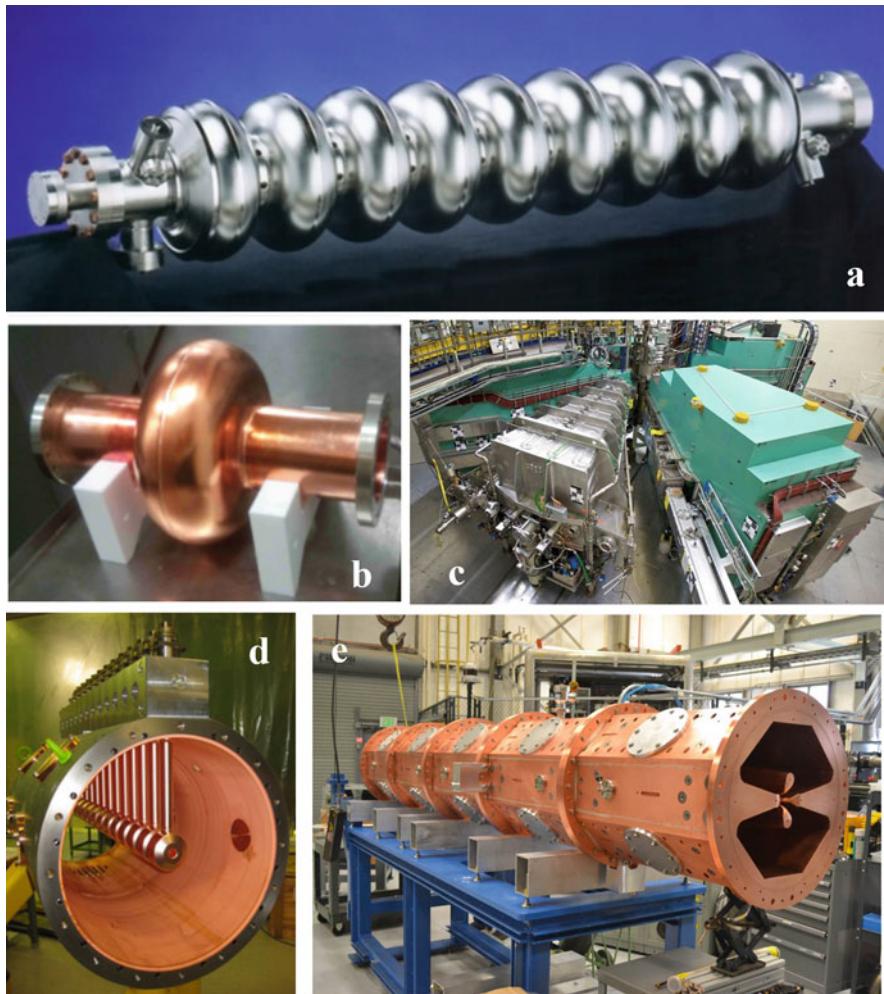


Fig. 1.3 Standing-wave cavities. (a) TESLA (TeV-Energy Superconducting Linear Accelerator) 9-cell superconducting cavity for electron beam acceleration (Courtesy of N. Phinney et al); (b) Single-cell normal-conducting cavity for electron beam acceleration (Courtesy of F. Marhauser et al); (c) The cavity of a proton cyclotron (<https://www.psi.ch/en/media/the-psi-proton-accelerator>); (d) Drift tube linear accelerator for proton beam acceleration (Courtesy of S. Ramberger et al); (e) Radiofrequency quadrupole for proton beam acceleration (Courtesy of P.N. Ostroumov et al)

Another critical component of an RF system is the RF power amplifier. Klystrons are still widely used in RF systems that require high peak power up to tens of MW. Nowadays, tens of kW power are achievable with solid-state amplifiers as commercial products for frequencies up to 1 GHz. Figure 1.5 shows examples of a klystron (Chin 2008) and a solid-state amplifier.



Fig. 1.4 C-band traveling-wave structures installed at SwissFEL (<https://www.psi.ch/en/swissfel>)

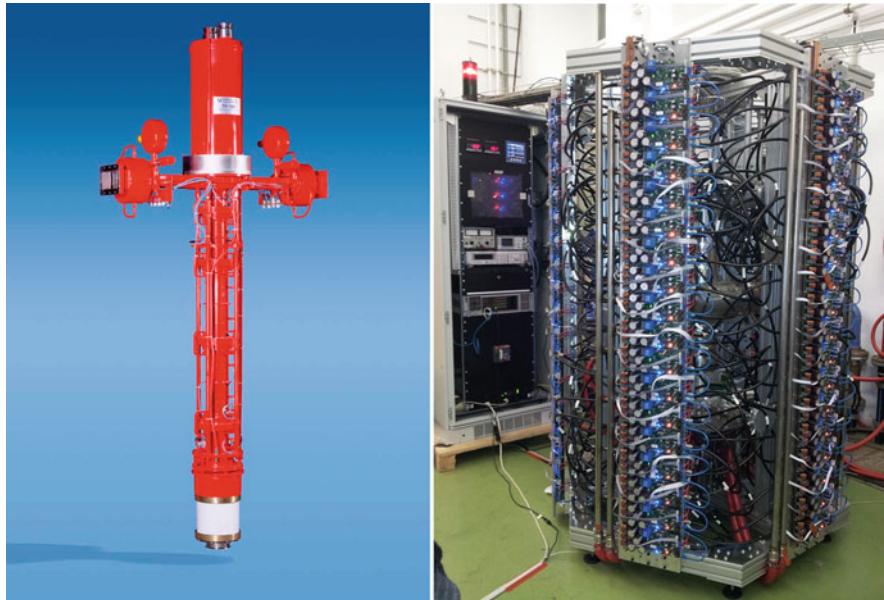


Fig. 1.5 RF power amplifiers. The THALES TH1801 klystron (left) (Courtesy of Y.H. Chin) and the 500 MHz solid-state amplifier developed at the Paul Scherrer Institut in Switzerland (right) (<https://www.psi.ch/en/amici/500mhz-rf-cavity-test-stand>)

In Chap. 3, we will study the HPRF system in more detail. Dynamical models will be built for standing-wave cavities and other HPRF components, with which we can explore the RF system behavior with both RF drive power and beam loading in the presence.

1.2 Principles of Beam Acceleration

In this section, we briefly review the principles of beam acceleration in standing-wave cavities and traveling-wave structures (Wangler 2008).

1.2.1 Acceleration in Standing-Wave Cavities

A standing-wave cavity may consist of multiple cells. We will start the discussion from a single-cell cavity and then expand the results to multi-cell cavities.

A cavity consists of multiple modes with different patterns of field distribution in space. Among these modes, the TM010 mode is used for beam acceleration because it provides an electric field along the beam trajectory. Figure 1.6 depicts the electric fields in a cylindrical cavity operating in the TM010 mode. It illustrates that at different locations on the cavity axis ($r = 0$), the electric fields are identical and oscillate in the same phase at the RF frequency. The spatial distribution of the electric field, which is a snapshot of the electric fields on the cavity axis at a time instant, is a constant at different locations (we neglect the electric field decay at both ends of the cavity to simplify the discussion) with a magnitude determined by the phase of the oscillation. The amplitude (maximum magnitude) of the electric field oscillation is denoted as E_0 .

We define the cavity length as L and the beam flight direction as the z -direction. For convenience, we define the z origin at the center of the cavity. When a particle passes through the cavity, the energy gain is determined by the integrated electric field that the particle experiences on its way. Note that the electric fields at different z locations oscillate in the same phase. Therefore, due to its finite flight velocity, the particle arrives at different locations in the cavity at different times, exposed to different electric field phases. The particle arrival time with respect to the RF field oscillation needs to be correct for proper acceleration, such as for maximum energy gain or optimal bunch compression. The relative timing between the RF field oscillation and the beam arrival time defines the accelerating phase.

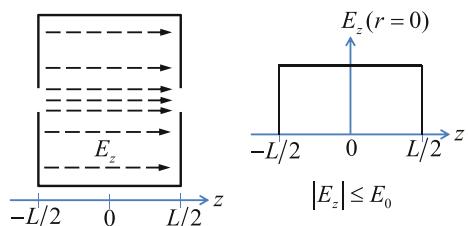
The on-axis electric field that the beam sees at a location z can be written as

$$E_{z, \text{axis}}(z) = E_0 \cos(\omega t(z) + \varphi_b), \quad (1.1)$$

where ω is the RF frequency and $t(z)$ is the time when the particle with a general velocity $v(z)$ arrives at z , which is defined as

$$t(z) := \int_0^z dz/v(z). \quad (1.2)$$

Fig. 1.6 A cylindrical cavity operating in the TM010 mode (left) and the electric fields on its axis snapshot at an arbitrary time instant (right)



We can see that at $z = 0, t = 0$. It means, at time zero, the particle arrives at the cavity center and experiences an electric field phase φ_b relative to the peak accelerating phase (i.e., on-crest acceleration).

The energy gain ΔW of a particle when it passes through the cavity can be then calculated as

$$\Delta W = \int_{-L/2}^{L/2} E_0 \cos(\omega t(z) + \varphi_b) dz =: V_0 T \cos \varphi_b. \quad (1.3)$$

Here V_0 is the integration of the on-axis electric field amplitude, and T is the *transit-time factor*, which is always smaller than 1. The transit-time factor is a reduction factor of energy gain due to the electric field phase variation when the particle passes through the cavity. The product of V_0 and T is defined as the *accelerating voltage* V_{acc} , which is the maximum energy gain of a particle accelerated on-crest ($\varphi_b = 0$):

$$V_{acc} := V_0 T, \text{ where } V_0 := \int_{-L/2}^{L/2} E_0 dz = E_0 L. \quad (1.4)$$

For the electric field distribution in Fig. 1.6, T can be calculated as

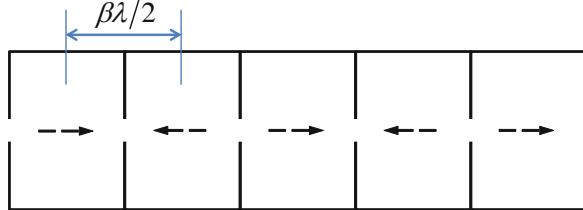
$$T = \sin \frac{\pi L}{\beta \lambda} / \frac{\pi L}{\beta \lambda}, \quad (1.5)$$

where $\beta = v/c$, λ is the RF wavelength in the vacuum space, and $\beta \lambda$ is the distance that the particle travels in an RF period. We notice that when $L \rightarrow 0$, $T \rightarrow 1$. In this case, the transit time of a particle in a short cavity is negligible compared to the RF period, resulting in less electric field phase variation when the particle passes through the cavity. Note that if the electric field distribution is not a simple square pattern as Fig. 1.6, the expression of T will be much more complicated than (1.5). In this book, the acceleration in a standing-wave cavity will be described with the *cavity voltage* \mathbf{v}_C (see Sect. 3.3), which is a complex quantity where its magnitude equals the accelerating voltage. That is,

$$|\mathbf{v}_C| = V_{acc}. \quad (1.6)$$

The phase φ_b is defined as the *beam phase*, representing the relative timing between the RF field oscillation and the particle arrival time. The relationship between the phase of \mathbf{v}_C and φ_b is discussed in Sect. 3.3.8. Note that with the definition here, the on-crest acceleration corresponds to $\varphi_b = 0$, which is typically defined in linear accelerators. For storage rings, the on-crest phase is defined as 90° , and the energy gain Eq. (1.3) should use the sine function instead of cosine. In this book, we will always use the beam phase definition given by (1.3).

Fig. 1.7 A multi-cell cavity operating in the π -mode



From the definition (1.4), V_{acc} depends not only on the electric field strength but also on the particle velocity. The acceleration efficiency of a cavity is described by the *effective shunt impedance* defined as

$$r = \frac{(V_0 T)^2}{P_{cav}} = \frac{V_{acc}^2}{P_{cav}}, \quad (1.7)$$

where P_{cav} is the RF power dissipated in the cavity wall. The value of r is expressed in megaohms ($M\Omega$). The physical meaning of r is the accelerating voltage produced by a unit of RF power dissipated in the cavity wall. Note that r depends on T , so it only describes what is experienced by the beam, but not the strength of the electric field. To describe the electric field strength produced by the RF power, we use the *shunt impedance* $r_s = V_0^2/P_{cav}$, a characteristic parameter of the cavity itself. Suppose a cavity is used to accelerate different particles (e.g., electron or proton). In that case, its r_s does not change, but r will be different due to different values of T . Another parameter, the *normalized effective shunt impedance* (r/Q), is defined as the ratio between r and the unloaded quality factor (Q_0 , defined in Sect. 3.3.1) of the cavity:

$$r/Q = \frac{r}{Q_0} = \frac{V_{acc}^2}{\omega_0 W}, \quad (1.8)$$

where ω_0 is the cavity resonance frequency, and W is the stored energy in the cavity. When studying the cavity model in Chap. 3, r/Q is an essential parameter in the differential equation of the cavity. Moreover, since the value of r/Q is dependent on the particle velocity, the cavity voltage calculated with the cavity differential equation is only meaningful for the specific particles. A multi-cell cavity consists of multiple cells that can be viewed as a chain of single-cell cavities. See Fig. 1.7. In each cell, the electric field follows the distribution as in Fig. 1.6. The relative amplitudes and phases of the electric fields in different cells have various patterns, denoted as *passband modes*. We will give more details about the passband modes in Sect. 3.4.3. The π -mode, in which the electric fields of adjacent cells have the same magnitude but 180° phase difference, is usually used for beam acceleration.

The length of a cell is selected to be $\beta\lambda/2$, the distance a particle travels in half the RF period. Therefore, when the particle travels from the center of one cell to that of the next, the electric field phase changes by 180° . If the particle is accelerated in the first cell with a beam phase φ_b , it will experience the same phase in the downstream cells, resulting in continuous acceleration.

The Eqs. (1.1) to (1.8) are also applicable to multi-cell cavities with the following definitions: L should be the total length of the cavity, V_{acc} is the accelerating voltage when the particle passes through all the cells, P_{cav} is the power dissipation in the walls of all the cells, and W is the stored energy in the entire cavity.

1.2.2 Acceleration in Traveling-Wave Structures

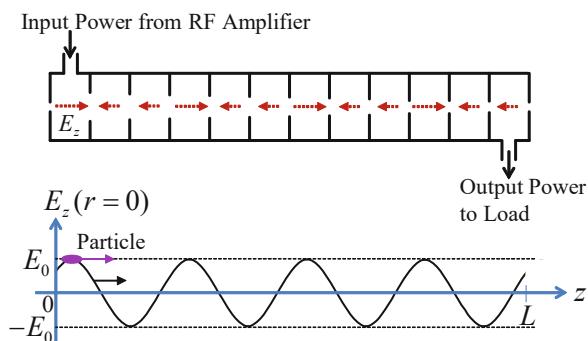
On the axis of a traveling-wave structure, the electric fields resonate at the RF frequency with different phases. Figure 1.8 is a snapshot of the electric fields along the axis of a structure at a particular time. The field pattern travels at the phase velocity v_p towards the load. Note that v_p is tuned to be the same as the velocity of the particles. Therefore, if a particle sees a phase φ_b of the electric field when it enters, it experiences the same phase when passing through the structure.

The on-axis electric field that the beam sees at a location z can be written as

$$E_{z, \text{axis}}(z) = E_0(z) \cos \varphi_b, \quad (1.9)$$

where E_0 is the amplitude of the accelerating electric field of the traveling wave. In a *constant-impedance traveling-wave structure*, E_0 decays exponentially along the axis, whereas in a *constant-gradient structure*, E_0 is a constant along the axis. Here we write E_0 as a function of z to include both cases. Then the energy gain of a particle passing through the structure can be calculated as

Fig. 1.8 A traveling-wave structure (constant-gradient) and the electric fields on its axis. The structure operates at the $2\pi/3$ mode, where the phase-lag per cell of the traveling wave is $2\pi/3$



$$\Delta W = \int_0^L E_0(z) \cos \varphi_b dz = V_0 \cos \varphi_b, \quad (1.10)$$

where $V_0 := \int_0^L E_0(z) dz = V_{acc}$.

Here L is the length of the structure, and we have chosen the z -axis origin at the beginning of the structure. Compare to (1.3) and (1.4), we notice that the transit-time factor of a traveling-wave structure is always 1. Similar to the standing-wave cavity case, the acceleration in a traveling-wave structure will be described with the *structure voltage* \mathbf{v}_{ACC} (see Sect. 3.5), which is a complex quantity where its magnitude equals the accelerating voltage:

$$|\mathbf{v}_{ACC}| = V_{acc}. \quad (1.11)$$

The beam phase φ_b has the same meaning as in the standing-wave cavities.

A traveling-wave structure is also characterized by shunt impedance, quality factor and r/Q , but they are defined per unit length. The *shunt impedance per unit length* (in $M\Omega/m$) is defined as

$$r_{s,tw} = \frac{E_0^2}{-dP_{tw}/dz}, \quad (1.12)$$

where P_{tw} is the traveling-wave power at a certain z location and $-dP_{tw}/dz$ is the power dissipation per unit length in the walls of the structure. The *quality factor* is given by

$$Q = \frac{\omega W_{tw}}{-dP_{tw}/dz}, \quad (1.13)$$

where ω is the RF frequency and W_{tw} is the stored energy per unit length in the structure. Then the *normalized shunt impedance per unit length* is written as

$$(r/Q)_{tw} = \frac{r_{s,tw}}{Q} = \frac{E_0^2}{\omega W_{tw}}. \quad (1.14)$$

We will discuss this in more detail in Sect. 3.5.

1.3 Disturbances to RF Fields

Achieving stable RF fields for beam acceleration is one of the primary goals of LLRF systems. In accelerators such as FELs, highly stable RF fields with amplitude jitter down to 1e-4 RMS and phase jitter down to 0.01° RMS are required. To obtain

the required RF stability, we must adopt low-noise RF components and suppress the disturbances coupled into the HPRF system. In general, a *disturbance* is any unwanted influence caused by the external environment of the RF system, such as temperature, humidity, or mechanical vibrations. In the control theory, a disturbance usually stands for an external input added to the input or output of the system. In contrast, an external influence that changes the dynamics of the system is called a *perturbation*. To simplify the discussion, we will not distinguish them in this book, and any unwanted external inputs are named as disturbances.

1.3.1 Electronic Noise

For highly stable RF fields, low noise RF components are essential, especially in the RF stations where fast feedbacks cannot be implemented, such as the normal-conducting RF stations operating in a short-pulsed mode. For such RF stations, the RF pulse is too short compared to the control loop delay, so intra-pulse feedback is not possible. The stability within a pulse is mainly determined by the noise in the master oscillator (MO), RF reference receiver, and the RF driving chain (see Fig. 1.13). If an RF feedback loop is closed, the noise in RF detectors will also affect the RF field stability. The noise transfer function in a feedback loop is discussed in Sect. 6.3 and the noise analysis for an RF station can be found in Sect. 6.5.

The master oscillator provides a common frequency source for the entire accelerator. The MO amplitude noise can be reduced by the RF actuator (with the MO input stage operating in saturation) and is not relevant to the RF field stability. However, the MO phase noise will be distributed to all RF stations, causing correlated phase fluctuations in the RF fields of different cavities or structures. We will discuss in more detail about the MO phase noise in Sect. 6.4.3.

The MO signal is distributed to each RF station by the synchronization system. At each RF station, the RF reference receiver produces an RF reference signal using the synchronization signal delivered by the synchronization system. The RF reference receiver introduces additional phase noise. The RF reference signal is then modulated in amplitude and phase by the RF actuator, which also adds additional amplitude and phase noise. As seen in Fig. 1.13, digital-to-analog converters (DACs) are often used to output the RF controller commands, modulating the RF reference signal using an I/Q modulator. The noise in RF mixers and DACs will be discussed in Sects. 6.2.4 and 6.2.7.

The RF amplifier noise, discussed in Sect. 6.2.3, increases further the amplitude and phase noise of the RF drive power. In particular, if a klystron is used to amplify the RF power, the modulator high voltage (HV) fluctuation is one of the significant noise sources in the klystron output. As an example, Fig. 1.9 illustrates the dependency of the klystron output amplitude and phase to the HV power supply (HVPS) settings of a modulator. The plots are measured at an S-band (2998.8 MHz) klystron of SwissFEL operating close to saturation with a constant RF input. The amplitude

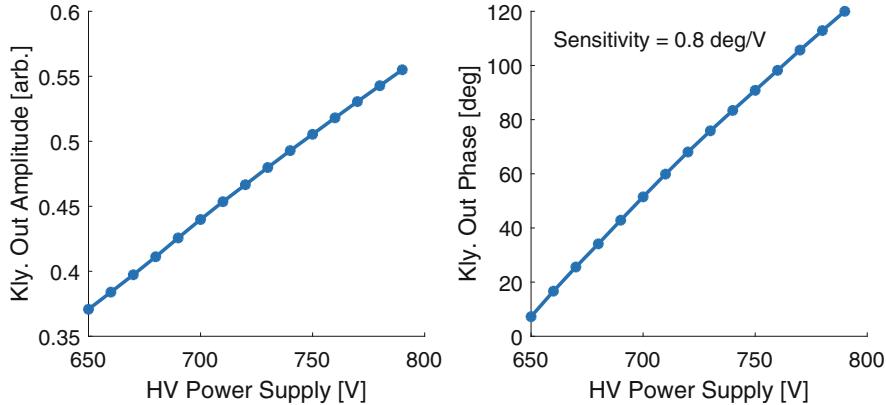


Fig. 1.9 Sensitivities of the klystron output amplitude and phase to the HV power supply

and phase of the klystron output change approximately linearly with the HVPS. Therefore, with an HV jitter ($\Delta V/V$)_{rms} = 100 ppm (i.e., 1e-4 relative jitter), the relative amplitude jitter ($\Delta A/A$)_{rms} \approx ($\Delta V/V$)_{rms} = 1e-4, and the phase jitter $\Delta\varphi_{rms} \approx 0.06^\circ$. Here we assume the HVPS operating point to be 750 V. The HV jitter causes pulse-to-pulse jitter in the klystron output, which is not controllable with the pulse-to-pulse feedback. Therefore, a stable klystron modulator is essential to reduce the RF field jitter. For example, the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory requires HV stability better than 40 ppm RMS (Decker et al. 2013) to achieve phase stability better than 0.1° RMS for the injector RF stations. With solid-state modulators, it has been demonstrated at SwissFEL that the HV jitter can be below 15 ppm (Epp et al. 2016), which significantly improves the RF field stability.

The RF detector noise induces amplitude and phase measurement errors, where the major noise sources are the mixers and analog-to-digital converters (ADCs). The ADC noise is discussed in detail in Sect. 6.2.6.

1.3.2 Temperature and Humidity

The temperature and humidity variations in the environment of an RF system are the main sources of the amplitude and phase drifts in the RF fields. These environmental factors cause slow variations in the gain (or attenuation) and phase shifts of the RF cables, waveguides, RF amplifiers and electronics. Furthermore, the resonance frequency of a normal-conducting cavity or structure will be affected by its temperature fluctuations, resulting in amplitude and phase errors in the RF fields.

An RF detector can only measure the relative phase between the RF fields and the RF reference signal. Hence, the phase drifts in the RF reference signal are not

detectable (then not controllable) by the LLRF system, and therefore, cause the same implications in the RF fields experienced by the beam.

An RF feedback controller can suppress the drifts in the RF driving chain, according to the discussion in Sect. 6.3. However, the drifts in the RF measurement chain, consisting of an RF signal pickup cable and an RF detector, cannot be compensated by the feedback. Therefore, the drifts in the RF measurement chain are critical to the long-term stability of the RF fields. For a long RF cable (e.g., a coaxial cable), both the temperature and humidity will affect its attenuation and phase shift. Figure 1.10 shows a long-term measurement at the X-band (11.995 GHz) RF station of SwissFEL. The gain and phase drifts in the RF measurement chain are evaluated by measuring the amplitude and phase of a stable RF reference signal. It shows that the gain drift is relatively small, whereas the phase drift is more sensitive to the humidity in the LLRF rack.

To reduce the RF field drifts, we must stabilize the temperature and humidity of the RF system's environment. For example, a rack with air conditioning significantly reduces the drifts in LLRF electronics, especially in RF detectors. We can also choose phase-stabilized RF cables that are less sensitive to temperature changes. Furthermore, routing a long cable in an environment with lower temperature and humidity variations (e.g., in the accelerator tunnel) can reduce the drifts. Moreover, techniques like reference tracking and drift calibration (see Sect. 6.6) have been developed to compensate for the RF measurement chain drifts. The drift compensation in the RF reference distribution (i.e., synchronization) system will be discussed in Chap. 8.

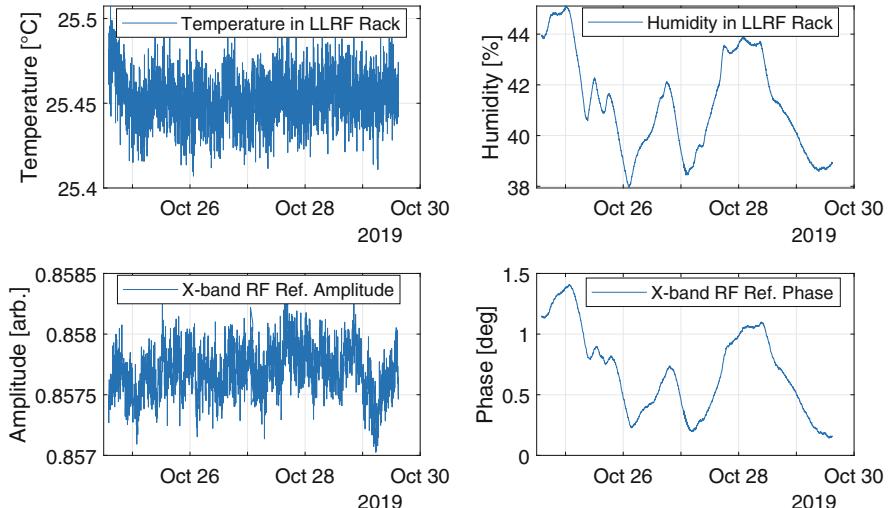


Fig. 1.10 Sensitivity of the RF measurement chain to temperature and humidity

1.3.3 Mechanical Vibrations

Mechanical vibrations in the environment where an RF cavity or structure is installed can stimulate mechanical oscillations in the cavity or structure. The mechanical oscillations will deform the shape of the cavity or structure, causing fluctuations in its resonance frequency, resulting in RF field errors. We have also observed that a cavity probe is disturbed by the mechanical vibrations and shows fluctuations in the cavity field measurement not actually experienced by the beam.

The mechanical vibrations may come from the following sources:

- (a) Pumps of vacuum or cooling systems,
- (b) Fans of air conditioning system,
- (c) Liquid flows (e.g., water or liquid helium) in cooling pipes,
- (d) Helium pressure fluctuations (for superconducting cavities) in cryomodules,
- (e) Ground motions caused by traffics or industry works.

The mechanical vibrations vary the cavity resonance frequency and modulate the electric fields in the cavity. This effect is denoted as *microphonics*. Superconducting cavities suffer more severely from microphonics because of the low rigidity due to thin walls. The change of cavity resonance frequency is called a *detuning* of the cavity. The microphonics-caused detuning can reach up to several tens of Hz (resonance frequency variation) with a changing rate up to several hundred Hz (changing rate of the time-varying detuning). Figure 1.11 (Neumann et al. 2010)

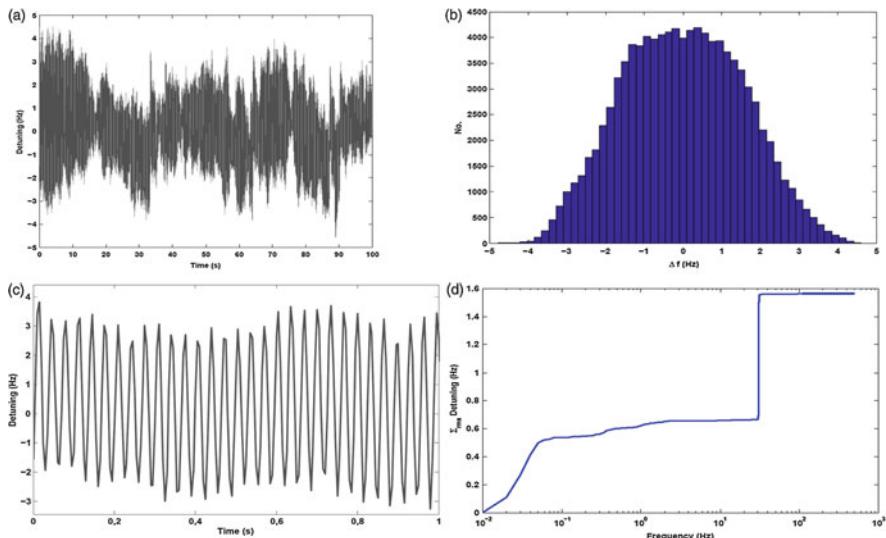


Fig. 1.11 Microphonics in a TESLA cavity (Courtesy of A. Neumann et al.). **(a)** Microphonics detuning measurement for a time period of 100 s; **(b)**. Statistics of the sampled detuning measurement. The RMS value is between 1.0 to 5.5 Hz; **(c)**. Zoom of the measurement data in 1 second showing an oscillation at around 30 Hz; **(d)**. Integration of the power spectral density of the detuning measurement from 0.01 Hz to different frequencies, illustrating the contribution to the RMS detuning value from different frequencies of microphonics

is an example of the effects of microphonics on a TESLA cavity. Zoom into the data shows a primary oscillation at around 30 Hz. This example mainly illustrates periodic microphonics. In many cases, wideband mechanical vibrations may exist, causing random fluctuations in the cavity resonance frequency.

For a superconducting cavity with a very high loaded quality factor (e.g., above 10^7), the microphonics-caused detuning may be comparable to or larger than the cavity bandwidth. In this case, the microphonics will cause significant cavity field errors and dramatically increase the required RF power to achieve the desired cavity voltage. Therefore, active controls of the cavity resonance frequency with a mechanical tuner are mandatory for operating such narrow-band cavities. The cavity tuning control will be discussed in Sect. 4.6.

Another source of mechanical deformation is the radiation pressure (also denoted as *Lorentz force*) induced by the electromagnetic fields in the cavity. See Fig. 1.12. In a cavity operating in the TM010 mode, the electric fields concentrate along the cavity axis while the magnetic fields are located near the cavity walls. The electric fields generate attracting forces at the cavity iris, whereas the magnetic fields produce repulsive forces at the cavity equator. The cavity shape is then deformed, reducing the cavity resonance frequency. The radiation pressure-caused detuning is denoted as the *Lorentz force detuning* (LFD). If the cavity voltage is in a steady state, the LFD has the following relationship with the electric field:

$$\Delta f_{LFD} = -K \cdot E_0^2, \quad (1.15)$$

where E_0 is the amplitude of the electric field on the cavity axis. The coefficient K is determined by the mechanical design of the cavity and can be measured experimentally. For a TESLA cavity, the typical value of K is 0.9 Hz/(MV/m) 2 (Schilcher 1998). Since the cavity fields and the LFD are coupled to each other, an instability may occur under certain conditions. This is the so-called *ponderomotive instability*, which will be discussed in Sect. 4.6.4. A cavity exhibits multiple mechanical oscillation modes. If the cavity operates in the pulsed mode, the time-varying Lorentz force will stimulate mechanical oscillations resulting in a time-varying detuning, as

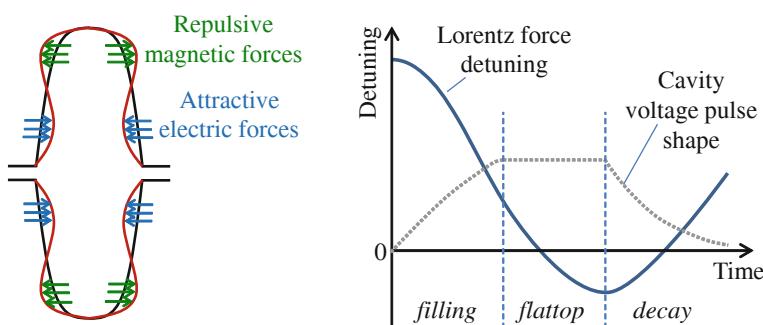


Fig. 1.12 The radiation pressure (Lorentz force) caused by the electromagnetic fields in the cavity (left) and the resulting detuning of a cavity operating in the pulsed mode (right)

shown in Fig. 1.12. The pulsed LFD is repetitive for successive RF pulses if the pulse parameters (e.g., pulse width, amplitude, and phase) are constant. Therefore, we may apply a feedforward signal in the RF drive power to compensate for the RF field errors caused by the pulsed LFD. Furthermore, with a pulsed drive to the cavity tuner, we can superimpose an additional mechanical oscillation that cancels the oscillations induced by the Lorentz force. This method is introduced in Sect. 4.6.3.

1.3.4 Beam Loading

A particle bunch passes through a cavity or structure and exchanges energy with the RF fields. The beam current, which is formed by the movement of the charged particles, also induces RF fields in the cavity or structure. We denote the beam-induced RF fields as *beam loading*. The overall RF fields in the cavity or structure are a superposition of the RF drive power and beam loading contributions. In many accelerators, we accelerate a bunch train consisting of multiple bunches. The time interval between adjacent bunches is much smaller than the time constant of the cavity or structure. Therefore, the beam loading of a single bunch disturbs the RF fields experienced by the following bunches. If with a constant bunch charge, the bunch train will generate a deterministic beam loading, which can be compensated for by a feedforward signal (see Sect. 9.5). However, if there are statistical fluctuations in the bunch charge or bunch arrival time, the beam loading effects will be random and can only be controlled by feedback. We will discuss in more detail about the feedback and feedforward in Sect. 2.1.

Another uncertainty caused by the beam loading is that the bunch train may be inhibited at a random time by the machine protection system. Such sudden changes in the beam will result in large transients in the RF fields in the cavities or structures. It may limit the RF field controllability with feedback or feedforward control.

1.4 LLRF Systems Overview

To perform desired beam acceleration, we must regulate the RF fields in the cavities or structures following even time-varying amplitude and phase set points. Furthermore, the effects of disturbances (see Sect. 1.3) should be reduced significantly to achieve stable beam parameters, such as beam energy, energy spread and bunch length. An LLRF system is crucial for monitoring and controlling the RF fields in the cavities or structures. The keyword “low level” indicates that the LLRF system manages low power RF signals (e.g., no larger than 1 W). The boundary between LLRF and HPRF is that the LLRF controls the drive power to the HPRF. The klystron drive RF amplifiers, which produce RF powers up to several hundred watts, may also be included in the scope of the LLRF system.

1.4.1 Requirements and Architecture

The primary functional requirements of an LLRF system include:

- (a) *Measure the RF fields used for beam acceleration accurately and precisely.* It implies two points. First, the amplitudes and phases of the RF signals should be measured precisely; second, the RF signal measurements should be interpreted accurately to estimate the accelerating voltage and beam phase.
- (b) *Control the RF fields to follow the set points and reduce the effects of disturbances.*
- (c) *Facilitate the RF system operation.* The LLRF system must support automation, including operational parameters optimization, RF measurements calibration, system status diagnostics, and RF system exceptions detection and handling. These functions are discussed in more detail in Chap. 9.

The performance of an LLRF system is quantified by non-functional requirements that are typically specified as:

- (a) *Amplitude and phase tolerances.* The RF field amplitude and phase stability are specified in terms of jitter and drift tolerances. The state-of-the-art FEL machines may require an RF stability as low as 1e-4 RMS and 0.01° RMS for amplitude and phase, respectively.
- (b) *Reliability.* A reliable LLRF system is essential to improve the availability of the particle beam. We use the *mean time between failures (MTBF)* to quantify the reliability, which illustrates how frequently a system fails.
- (c) *Reproducibility.* After rebooting an accelerator subsystem (e.g., timing, synchronization, Gun laser, LLRF, etc.), the LLRF system should restore the correct accelerating voltage and beam phase quickly. Good reproducibility is also essential for high beam availability.
- (d) *Maintainability.* To ease the maintenance of the LLRF system, we should consider modular design. The LLRF hardware should be remotely accessible, runtime accessible, or hot swappable.

The non-functional requirements may affect the design of an LLRF system significantly. For example, if very high reliability is required, such as for the LLRF system of a CW proton machine that drives a sub-critical nuclear reactor, redundancy in LLRF hardware may be needed. More terms of non-functional requirements can be found in the document ISO/IEC 9126.

A complete and detailed requirements analysis is necessary as inputs for the design of a large-scale system. From the requirements, we can propose an architecture of the LLRF system, which describes the decomposition and interfaces of the hardware, firmware (in embedded digital processors) and software components. Though it might be slightly different for different RF systems, a general (hardware) architecture of an LLRF node is depicted in Fig. 1.13.

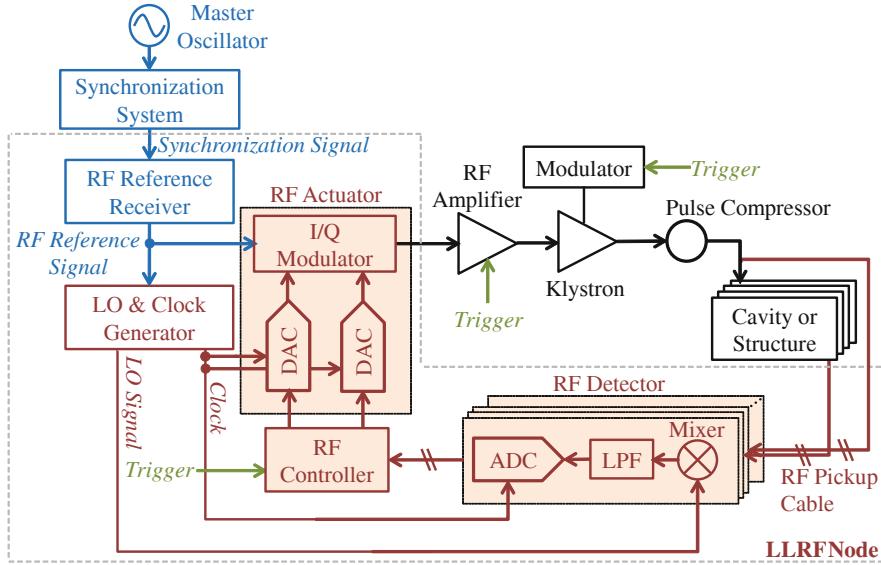


Fig. 1.13 A general architecture of an LLRF node controlling an RF station

An LLRF node provides control functions for an HPRF station. The cavity fields are sampled by a probe and delivered to the RF detector via an *RF pickup cable*. An *RF detector* measures the amplitude and phase (or the I/Q components) of the cavity probe signal. Figure 1.13 shows a typical implementation of an RF detector, which consists of a mixer, a low-pass filter (LPF) and an ADC. It is a widely used RF detector design and will be discussed in Sects. 5.1.3 and 5.2. An *RF controller* implements the feedback and feedforward control algorithms to regulate and stabilize the cavity fields. The RF controller outputs adjust the amplitude and phase of the cavity drive signal using an *RF actuator*. A typical RF actuator consists of one or two DACs and an I/Q modulator, which is discussed in more detail in Sect. 5.3. The frequency source of the RF station is provided by a *master oscillator (MO)*, which synchronizes all of the RF stations. The MO synchronization signal is distributed to different RF stations via a *synchronization system*. At each RF station, an *RF reference receiver* receives the MO synchronization signal and provides an *RF reference signal* for the RF station with the correct frequency and power. For the RF detector and actuator in Fig. 1.13, an *LO & clock generator* is required to generate the *local oscillator (LO)* and *clock* signals. The LO frequency is typically lower than the RF frequency and is used to convert the RF frequency signal to an *intermediate frequency (IF)* via a mixer. For convenience, we denote the combination of the RF actuator, RF amplifier, klystron, pulse compressor (for increasing the peak RF power, see Sect. 3.6.3) and cavities or structures as the *RF driving chain*. Similarly, the *RF measurement chain* includes the RF pickup cable and the RF detector. The concepts defined here will be used frequently in this book.

1.4.2 Context in Particle Accelerators

In a particle accelerator, the LLRF system interacts with the HPRF system and other systems depicted in Fig. 1.14. The figure is a context diagram of the LLRF system in a linear accelerator.

The *timing*, *synchronization* (including MO), *Gun laser* (generating electron bunches in a photocathode RF Gun) and LLRF systems are essential to define the time relationships between the beam and RF fields for proper acceleration. The timing system triggers different systems with a picosecond-level time resolution, corresponding to a fraction of the RF period. For example, in a pulsed machine, the timing system starts the macro RF pulses at proper times to establish the RF fields before the beam arrives. In comparison, the synchronization system provides femtosecond-level synchronization between the beam and RF fields. In Fig. 1.14, the phases of the Gun laser and the RF fields of different RF stations are all locked to the synchronization signals. Therefore, by introducing proper phase offsets in the RF stations, we guarantee that when a particle arrives at an RF station, it experiences the correct RF field phase. We will provide more details of the timing and synchronization systems in Chap. 8.

The *users* or the *physics applications* interact with the LLRF system via the *control system*. They tune the beam parameters by adjusting the amplitudes and phases of the RF stations. In particle accelerators, the following control systems are widely used: the Experimental Physics and Industrial Control System (EPICS) (<https://epics-controls.org/>), the Tango control system (<https://www.tango-controls.org/>), and the Distributed Object-Oriented Control System (DOOCS) (<https://doocs-web.desy.de/>). Protection systems, such as the *machine protection system* (MPS) or *personnel protection system* (PPS), may stop the power of an RF station for safety purpose.

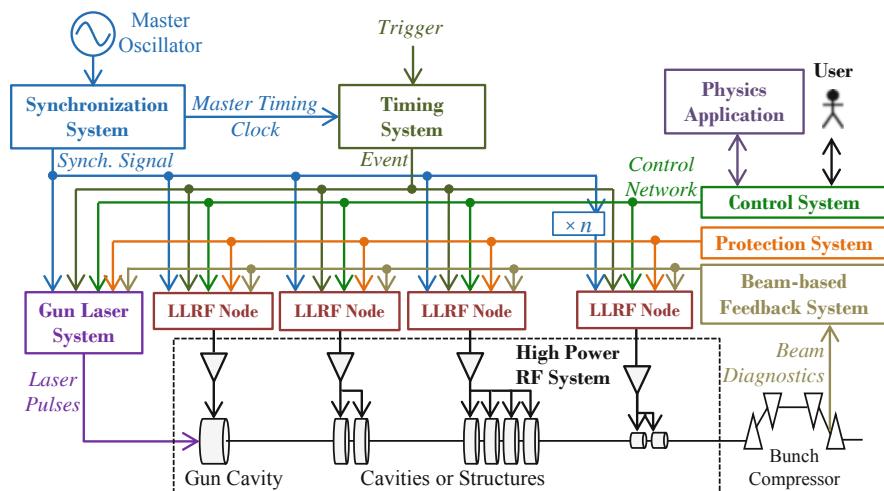


Fig. 1.14 Context of the LLRF system in a linear accelerator

Since the LLRF system only stabilizes the RF fields, the beam acceleration may still be affected by other factors, such as the beam arrival time and bunch charge fluctuations caused by the Gun laser timing jitter. Furthermore, the phase drifts in the synchronization system and RF pickup cables of LLRF are difficult to measure and control, resulting in residual errors even in the presence of LLRF controls. To stabilize the beam parameters directly, we adopt a *beam-based feedback system (BBFB)*. The BBFB system reads the beam parameters from the *beam diagnostics* and adjusts the amplitude or phase set points of the LLRF nodes. See Fig. 1.15. The BBFB loop and the LLRF feedback loop are cascaded loops that should be designed carefully to avoid instability. Furthermore, the transfer matrix between the RF amplitude and phase and the beam parameters may be ill conditioned, which introduces difficulties in designing the BBFB controller.

Figure 1.15 illustrates the beam control layers of a particle accelerator. Though the diagram shows the control of an FEL Linac, the concepts are also applicable to other accelerators like a storage ring. The first layer is the LLRF system regulating the RF fields, and the second layer is the BBFB that stabilizes the particle beam directly. They form the inner and outer loops of a cascaded loop. In many accelerators, the particle beam is not the final product. For example, in an FEL machine, the final product is the FEL photons. The quality of the photons, such as the photon energy in an FEL pulse, is optimized by adjusting the set points of the BBFB loops. The transfer model between the particle beam to the photon parameters might be complex and the linear control theory is no longer applicable. Therefore, modern optimization algorithms, such as the *genetic algorithm (GA)*, *particle swarm optimization (PSO)* algorithm, or *machine learning* algorithms, might be needed. These advanced optimization algorithms, as well as the BBFB algorithms, will be discussed in detail in a newly planned book *Intelligent Beam Control in Accelerators* that will come soon.

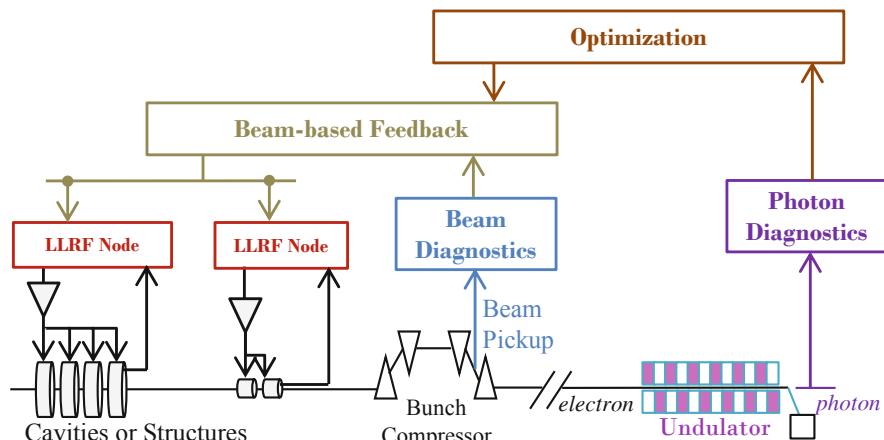


Fig. 1.15 A three-layer control strategy for a linear accelerator

1.4.3 A Brief History

Technologies for radio broadcasting became mature in the second half of 1920s and were first time applied to acceleration of charged particles in 1927 (Wideröe) at low frequencies (1 MHz) using drift tubes. In the 1930s the development advanced to higher frequencies and higher power levels paving the way to extensive use of radio frequency in particle accelerators. Control of amplitude, phase and frequency were initially controlled manually but with the increasing requirements on field stability the first analog control systems were introduced in the 1960s. This is the beginning of the LLRF control systems. In the following a chronological outline of the highlights of the LLRF development since 1960 is given.

Year	Highlights of LLRF development
1960s Various	Begin of the development of analog control systems for amplitude and phase. The main features were independent proportional-integral-derivative (PID) control loops using separate amplitude and phase detectors and actuators. Only manual control of set points and gain were supported. This is even today still often the preferred solution if only few cavities need to be controlled and the operating parameters are constant. This scheme is very reliable and cost efficient. Some “tweaking” is required to optimize controller performance. Resonance control loops have been added later.
1968 Perry Wilson, Alan Schwettman	The first demonstration of the control of superconducting cavity in electron Linac took place at Hansen Experimental Physics Lab in Stanford. (Wilson and Schwettman 1965; Smith et al. 1966). Only light beam loading was present and the very stiff standing wave structures supported excellent amplitude and phase stability.
1978 Jean Delayen	The Self-Excited Loop (SEL) has been introduced for the control of high-quality-factor resonators (Delayen 1978). The SEL utilizes the cavity as a resonant circuit determining the self-starting oscillator frequency. The main characteristics are an exponential start-up from noise and an amplitude limiter to avoid excessive gradients. A stable amplitude can be established in presence of large detuning without phase lock. Also, the control of ponderomotive instability is supported by this scheme. The SEL allows tracking of the resonance frequency during start-up which is a convenient method for finding the resonance frequency of cavity.
1980 Daniel Boussard	The modelling of high intensity beam loading has been introduced together with the concept of the optimal controller (Boussard 1985). The controller design became model based and allowed to control steady state and transient perturbations of the accelerating voltage by the beam current. Undesirable beam behaviour has been controlled through fast feedback and feedforward compensation techniques. The scheme is applicable to storage rings and to allows to suppress Robinson instability and other noise sources. Essential for the control performance are correct filtering and carefully selected delay lines.
1987 Roland Garoby	Semi-digital RF has been introduced to circular accelerators. The digital aspects were related to dynamic synchronization with variable frequencies. A master and slave direct digital synthesis (DDS) technology had been used to generate several revolution harmonics which were changing dynamically during the cycle. The technology facilitated RF gymnastics such as batch compression, bunch splitting and bunch merging. The time varying DDS frequency and phase offsets were programmed completely digitally (Pedersen 2005).

(continued)

Year	Highlights of LLRF development
1989 Stephen Jachim	The I/Q control scheme (Jachim 1990) has been introduced to allow field control in all 4 quadrants including zero field in presence of large beam currents. Key topics are the in-phase/quadrature (I/Q) field control, the application of six-port reflectometers for cavity instrumentation and resonance control, a technique for phase stabilization of critical RF cables, an application of state-variable feedback for field control and the direct integration of RF and computer-interfaced hardware. All these elements improved controllability of high-quality-factor cavities in presence of large microphonics significantly.
1990s (mid) Various	Standards are still the analog RF control systems which were still limited to simple PID controllers and separate control loops for amplitude, phase and resonance frequency. But meanwhile more often also I/Q control and direct RF feedback have been implemented. The RF control systems for circular machines started to become semi-digital control for the time-varying parameters mainly related to digital frequency synthesis. Typically, 1% amplitude and 1° phase stability have been achieved.
1994–1997 Flemming Pedersen	Analog field control loops are still dominant in circular accelerators with designs becoming more modular and including semi digital components for the frequency control aspects. Analog control signal processing has been added for beam phase, radial and synchro loops. These schemes have been used extensively at CERN in the PSB, PS, AAC until 1994–1997 (Pedersen 2005).
1996–2005+ Various	Starting in the mid 90s, digital technology with high-speed ADCs and high-performance processing in DSPs and later in FPGAs and CPUs allowed the continuous transition from analogue via semi-digital to full digital LLRF control systems. Digital technology was gradually applied to data acquisition and data processing, frequency synthesis, timing and synchronisation, frequency tuning with slow and fast actuators. Finally, with evolving technologies, fully digital field detection and actuation have been realized. Control algorithms started from SISO control of the various loops to full MIMO control of all field, frequency, and beam control loop. During this period the transition to digital technologies in all aspects of LLRF control occurred.
2001 LLRF community	A series of LLRF workshops started in 2001. The workshops were held every 2 years covering a large range of topics and types of accelerators. Initially the workshops were focused on hardware technologies for field control. Later software became more important and aspects like operability in the context of the accelerators including beam instrumentation and control have been added. Also, topics like system engineering methodologies for complex had been included. The website http://llrf.net provides an excellent source on information about all topics related to LLRF control.
2003 SNS LLRF team	Large scale digital LLRF controls were applied to the 96 Linac RF stations at the SNS with a mixture of different types of normal and superconducting cavities. The control system is challenging due to the RF interaction with non-relativistic beams in presence of heavy beam loading.
2006 Alba LLRF team	All digital RF control (FPGA) for circular accelerators have been deployed. They are based on digital IQ modulation/demodulation techniques implemented on commercial cPCI FPGA boards. The DLLRF controls the amplitude and phase of the RF cavity voltage within 0.1% and 0.1° respectively, as well as the resonance frequency of the cavity up to ±0.1°.
2008 Industry	Commercial LLRF control systems appear. Industry offers COTS integrated LLRF control solutions as well as custom solutions adapted to specific requirements. Also, a large number LLRF subsystems and components at chassis and board level are available. This includes master oscillators, up/

(continued)

Year	Highlights of LLRF development
	down-converters, amplitude and phase detectors, high speed data acquisition and high-performance data processing in FPGA and CPUs. Modular concepts are supported in various form factors including PXI-express, and MicroTCA.4.
2010+ DESY	Modular commercial LLRF control options in the MicroTCA.4 form factor have been developed at DESY for the XFEL project. Products and services are available from the MTCA technology lab (https://techlab.desy.de/) which is one of seven innovation labs funded by the Helmholtz Association. The goal is to foster the electronics standard MicroTCA in research and industry. The lab promotes the use of MicroTCA and support new and existing users of the technology with their products and services.
Future 2025 +	Forecasted is the application of artificial intelligence (AI) for RF parameter optimization with respect to performance, operability, and availability. In a later stage AI is expected to be deployed for exception detection and handling.

1.5 Summary

This chapter provides an overview of the HPRF and LLRF systems of particle accelerators. We reviewed the basic principles of beam acceleration in standing-wave cavities and traveling-wave structures. It provides the background knowledge of beam physics that is helpful to LLRF scientists and engineers. We also summarized the disturbances of RF systems and their effects on RF field stability. The strategies to overcome the disturbances are also briefly introduced, which will be studied in detail in this book. The context of the LLRF system illustrates its interfaces to other systems in the accelerator. We also introduced a layered structure for accelerator beam control, including LLRF, BBFB and global optimizations. It offers a guideline of what other topics can be studied beyond the LLRF system.

References

- Boussard D (1985) Control of cavities with high beam loading. *IEEE Trans Nucl Sci* 32(5): 1852–1856. <https://doi.org/10.1109/TNS.1985.4333745>
- Chin YH (2008) Design and performance of L-band and S-band multi beam klystrons. In: Proceedings of the XXIV linear accelerator conference, Victoria. 29 September – 3 October 2008
- Decker F-J, Colocco W, Huang Z et al (2013) Increased stability requirements for seeded beams at LCLS. In: Proceedings of the 35th international free-electron laser conference, New York, USA. 26–30 August 2013
- Delayen J R (1978) Phase and amplitude stabilization of superconducting resonators. Ph.D. thesis, California Institute of Technology
- Epp A, Frei M, Blokesch G et al (2016) Pulse to pulse stability of an Ampegon short pulse modulator. Poster presented in the IPAC2016 conference. https://ampegon.com/wp-content/uploads/2020/01/poster_ipac_2016_mf.pdf. Accessed 19 June 2021

- Jachim SP (1990) Some new methods of RF control. In: Proceedings of the 15th linear accelerator conference, Albuquerque, New Mexico. 10–14 September 1990
- Loehl F, Alex J, Blumer H et al (2013) Status of the SwissFEL C-band linear accelerator. In: Proceedings of the 35th international free-electron laser conference, New York. 26–30 August 2013
- Marhauser F, Castagnola S, Clemens WA et al (2018) 802 MHz ERL cavity design and development. In: Proceedings of the 9th international particle accelerator conference, Vancouver, 29 April – 4 May 2018. <https://doi.org/10.18429/JACoW-IPAC2018-THPAL146>
- Miller RH (1986) Comparison of standing-wave and traveling-wave structures. In: Proceedings of the 1986 international linac conference, Stanford. 2–6 June 1986
- Milne CJ, Schietinger T, Aiba M et al (2017) SwissFEL: the Swiss X-ray free electron laser. *Appl Sci* 7(7):720. <https://doi.org/10.3390/app7070720>
- Neumann A, Anders W, Kugeler O et al (2010) Analysis and active compensation of microphonics in continuous wave narrow-bandwidth superconducting cavities. *Phys Rev ST Accel Beams* 13(8):082001. <https://doi.org/10.1103/PhysRevSTAB.13.082001>
- Ostroumov PN, Mustapha B, Barcikowski A et al (2012) Development and beam test of a continuous wave radio frequency quadrupole accelerator. *Phys Rev ST Accel Beams* 15: 110101. <https://doi.org/10.1103/PhysRevSTAB.15.110101>
- Pedersen F (2005) LLRF developments at CERN: historic overview, highlights, future challenges. https://indico.cern.ch/event/421496/contributions/1901191/attachments/883929/1242355/LLRF_Developments_at_CERN.pdf. Accessed 14 August 2021
- Phinney N, Toge N, Walker N (2007) International linear collider reference design report, volume 3: accelerator. <http://cds.cern.ch/record/1061261>. Accessed 19 June 2021
- Ramberger S, Alharbi N, Bourquin P et al (2008) Drift tube linac design and prototyping for the CERN Linac4. In: Proceedings of the XXIV linear accelerator conference, Victoria. 29 September – 3 October 2008
- Schilcher T (1998) Vector sum control of pulsed accelerating fields in Lorentz force detuned superconducting cavities. Ph.D. thesis, Hamburg University
- Smith TI, Schwettman HA, Fairbank WM et al (1966) Stanford's superconducting accelerator program. In: Proceedings of the 1966 linear accelerator conference, Los Alamos. 3–7 October 1966
- Wangler TP (2008) RF linear accelerators, 2nded edn. Wiley-VCH, Weinheim
- Wilson PB, Schwettman HA (1965) Superconducting accelerators. *IEEE Trans Nucl Sci* 12(3): 1045–1052. <https://doi.org/10.1109/TNS.1965.4323780>

Chapter 2

RF Control Strategy



As more accelerators adopt superconducting cavities or require high RF field stability, the LLRF system becomes critical for meeting the accelerator performance objectives. Different control strategies have been developed to solve the RF control problems raised by different machines. In this chapter, we give an overview of the RF control strategies widely used in LLRF systems. We will briefly compare the feedback and feedforward control, the amplitude/phase and in-phase/quadrature (I/Q) control, the generator driven resonator (GDR), self-excited loop (SEL) and phase-locked loop (PLL) control, the analog and digital control, and the single-cavity and vector-sum control. The pros and cons of the control strategies will also be discussed.

2.1 Feedback and Feedforward Control

One of the major goals of an LLRF system is to stabilize the RF fields in standing-wave cavities or traveling-wave structures for beam acceleration. An RF controller provides typical functions like *command tracking* and *disturbance rejection*. For command tracking, the RF field amplitude and phase are regulated to follow the set points that may vary in time. For disturbance rejection, the errors caused by external disturbances, such as the thermal drifts, microphonics, and beam loading, must be reduced.

Feedback and *feedforward* are two basic methods for controlling the RF field. For feedback control, as shown in Fig. 2.1, the RF field is measured and compared with the set points. The error is then processed (e.g., scaled or integrated) to obtain a drive signal to the RF actuator. In a feedback loop, the system dynamics play an important role, and the closed-loop stability is of critical concern (Skogestad and Postlethwaite 2005). The RF feedback stability will be discussed in detail in Sect. 4.2.

Feedforward control is based on a pre-programmed drive signal to the RF actuator. It is often used to define the RF system's operating point, such as the RF

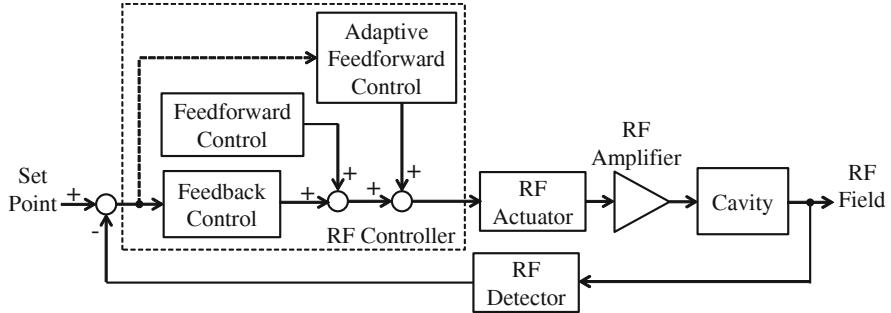


Fig. 2.1 Feedback and feedforward control of an RF station

pulse shape, cavity voltage and phase. For RF stations operating in the pulsed-mode, we may also adapt the feedforward signal to compensate for the pulse-to-pulse repetitive or slow time-varying RF field errors. This is the so-called *adaptive feedforward (AFF)* control, aiming to produce RF pulses with desired shapes or reduce the pulse-to-pulse fluctuations (Zhang et al. 1993). Typically, the AFF control refers to the RF field errors of the previous pulse and correct the feedforward signal for the next pulse. Therefore, it is a *pulse-to-pulse control* strategy. Note that in Fig. 2.1, the AFF control introduces a correction signal added to the nominal feedforward signal. However, in practice, the AFF control may also scale or rotate the nominal feedforward (Geng and Hong 2016). We will discuss the AFF algorithms in Sect. 4.5.

Feedforward control is a basic block of the RF controller, whereas AFF is typically implemented for RF stations operating in the pulsed mode. Feedback control is applied to RF stations operating in the continuous-wave (CW) mode or long-pulse mode. A long pulse means a pulse width much longer than the loop delay. If the RF pulse is too short (e.g., at SwissFEL, the pulse width is 1 μ s and the loop delay is over 300 ns), the intra-pulse feedback is not feasible. Still, the AFF control can be implemented to compensate for the slow pulse-to-pulse fluctuations (i.e., slower than 10% of the RF pulse repetition rate).

2.2 Amplitude/Phase and In-Phase/Quadrature Control

An RF signal is described by its amplitude, phase, and frequency. In Sect. 3.2.1, we will present the envelope of an RF signal as a complex signal called *phasor*. A phasor can be presented in terms of amplitude and phase (A/φ) or in-phase and quadrature (I/Q) components. Their relationship is given by

$$I = A \cos \varphi, Q = A \sin \varphi. \quad (2.1)$$

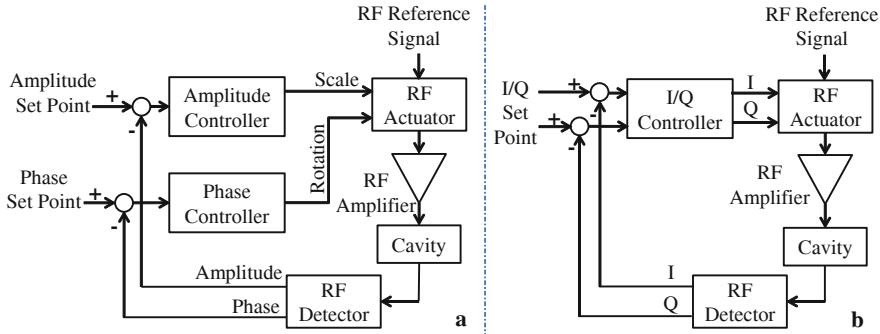


Fig. 2.2 RF controllers applied on A/φ (a) and I/Q (b)

An RF controller can stabilize either the A/φ or I/Q components of the RF field. If one pair of them are stabilized, the other pair will also be stable. However, due to the nonlinear relationship (2.1), feedback controllers applied to A/φ or I/Q behave differently, leading to different advantages and limitations. Figure 2.2 shows simplified block diagrams of the A/φ and I/Q control loops. In an A/φ control loop, the RF detector measures the amplitude and phase of the RF field, which are regulated by two separate *single-input single-output (SISO)* loops. Correspondingly, the RF actuator accepts inputs as scaling and rotation of the amplitude and phase of the RF reference signal. In contrast, in an I/Q control loop, the RF detector produces I/Q components of the RF field, and the RF actuator accepts inputs in the format of I/Q . The controller is a *multiple-input multiple-output (MIMO)* controller. However, as discussed in Chap. 3, if we view the I/Q set point, measurement and actuation as complex signals, the I/Q controller can also be viewed as a SISO controller.

For most LLRF systems, both A/φ and I/Q controls work well. However, there are some situations that one of them is more preferable.

With the A/φ control strategy, the amplitude and phase can be detected separately. It allows us to choose low-noise analog detectors. For example, a Schottky diode detector can measure the amplitude more precisely. Since A/φ control has two independent loops, we may operate them with different configurations (e.g., feedback gain) or only close one loop. One example is the RF stations with amplifiers or klystrons operating in saturation, e.g., the Linac Coherent Light Source (LCLS) (Akre et al. 2007) and the SwissFEL (Geng 2017). In this case, the amplitude control loop is not applicable, but the phase control loop can still be closed to stabilize the phase. For such RF stations with HPRF operating in saturation, the I/Q control strategy is often not applicable.

The implementation of the I/Q control and amplitude control loops are simple. However, some issues need to be considered for the correct implementation of the phase control loop. First, the phase measurement is noisy or jumping when the amplitude is close to zero, so we must disable the phase control when the amplitude is too small. Second, the phase is periodic with a period of 2π . Typically, in the phase control loop, the phase set point, measurement and error values should be converted

to the same range (e.g., $-\pi$ to π). The phase feedback gain should not be too large because the controller may amplify the phase error beyond 2π . Let us analyze a simple example. Assume that the phase error is 10° and the feedback gain is 36, then the resulting phase correction is 360° , which does not produce any phase corrections. This should be avoided.

The phase control loop may also cause problems when controlling a narrow-band cavity. The phase control often introduces fast phase transients, resulting in large RF frequency changes, which may exceed the cavity bandwidth and cause significant amplitude drops. This situation may cause strong oscillations due to the interaction between the amplitude and phase control loops. Furthermore, the RF amplifier may saturate if the amplitude control loop attempts to compensate for the amplitude drops. Due to these reasons, the I/Q control strategy is more suitable for the control of narrow-band cavities.

Another drawback of the A/ φ control is the latency when calculating the amplitude and phase with a digital processor. A state-of-the-art digital LLRF system often implements the RF controller in a Field Programmable Gate Array (FPGA). The RF signal is converted to an intermediate frequency (IF) and sampled by an ADC. In the FPGA, digital algorithms (see Sect. 5.2) are implemented to demodulate the ADC samples, and the results are typically in the format of I/Q. To calculate the amplitude and phase, we have to use a CORDIC (Coordinate Rotation Digital Computer) module that introduces significant latencies (Du et al. 2017). The loop latency will limit the reachable closed-loop bandwidth. Therefore, if a large closed-loop bandwidth is required, the I/Q control strategy is usually selected.

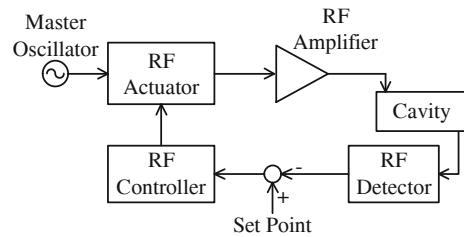
The I/Q control strategy is also preferable for the control of cavities with heavy beam loading. Usually, we implement a separate feedforward signal to compensate for the beam loading (see Sect. 9.5). With I/Q control, the beam loading compensation is more straightforward by adding a feedforward vector on the I/Q actuation signals. However, if we choose the I/Q control strategy, the system phase must be calibrated and corrected close to zero to avoid instability (Fong et al. 2008). This topic is discussed in detail in Sect. 4.2. Note that the A/ φ control does not have this constraint.

The analysis above provides some guidelines to choose the A/ φ or I/Q control strategies when designing an RF controller. In this book, the GDR (Sect. 4.2), SEL (Sect. 4.3), and adaptive feedforward controls (Sect. 4.5) will be described based on the I/Q control strategy; whereas the PLL (Sect. 4.4) and cavity resonance controls (Sect. 4.6) will adopt the A/ φ control strategy.

2.3 RF Control Loop Architecture

Multiple forms of control loops have been developed for the RF systems of particle accelerators (Schilcher 1998). The most widely used forms are the Generator Driven Resonator (GDR), Self-Excited Loop (SEL) and Phase-Locked Loop (PLL). They are used in various operating conditions with different control goals and are briefly introduced in this section. In Chap. 4, we will discuss their more in-depth principles.

Fig. 2.3 Architecture of GDR control



2.3.1 Generator Driven Resonator

In a GDR control loop, the RF field amplitude and phase are locked to an RF reference signal generated by a master oscillator (MO). See Fig. 2.3. Therefore, the RF system operating frequency is determined by the MO frequency. In a linear accelerator, the MO frequency is typically fixed, whereas, in a circular machine, the MO frequency may vary following the particle revolution frequency. Since the GDR control can generate stable RF fields with the desired frequency, amplitude and phase, it is adopted when the RF system is in beam operation. We notice that Figs. 2.1 and 2.2 are all GDR control loops.

When starting up a GDR control loop, the feedback is initially disabled, and the cavity is driven by a feedforward signal to establish an RF field near the desired operating point of the RF station. Then the feedback control is applied to obtain the desired RF field for beam acceleration. Since the MO determines the RF frequency, the cavity resonance frequency should be tuned close to the MO frequency. It implies that we may have difficulties starting up a narrow-band cavity with an unknown initial resonance frequency. If the cavity is detuned far from the MO frequency, the RF drive power will be fully reflected, and no RF fields can be built up in the cavity. Therefore, a procedure is required for the initial tuning of the narrow-band cavity when starting up a GDR control loop. To tune the cavity, we can measure the cavity resonance frequency by scanning the RF frequency with the method introduced in Sect. 4.6.1.1.

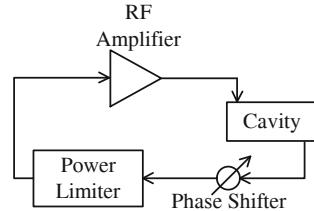
The algorithms and feedback stability issues of GDR control will be discussed in Sect. 4.2.

2.3.2 Self-Excited Loop

In an SEL, the master oscillator is not needed, and the cavity probe signal is directly driving the input of the high-power RF amplifier (Delayen 1978). A typical structure of the SEL is depicted in Fig. 2.4.

The SEL represents a positive feedback loop. If the loop phase is multiple of 2π and the loop gain is larger than one at a given frequency, the SEL will induce an exponential growth of the cavity voltage. As a result, the SEL operating frequency

Fig. 2.4 Architecture of SEL control



can be tuned using the phase shifter in Fig. 2.4. A limiter is necessary to control the maximum RF power produced by the RF amplifier.

One attractive feature of the SEL is that it can follow the cavity resonance frequency. The SEL operating frequency always satisfies the condition that the loop phase equals multiple of 2π . Therefore, if the loop is initially tuned to operate the cavity on resonance, the operating frequency will always track the cavity resonance frequency changes. This makes the SEL a perfect solution for cavity conditioning, for which the RF field stability requirements do not apply.

The SEL can also help to start up a narrow-band cavity with an unknown detuning. Assume that our final goal is to tune the cavity and control the cavity with a GDR loop for beam operation. By adjusting the loop phase, the SEL can be started at the cavity resonance frequency. We can compare the SEL operating frequency and the MO frequency and adjust the cavity tuner to reduce the detuning. As mentioned above, the SEL frequency will automatically track the cavity resonance frequency, which finally approximately matches the MO frequency. This procedure is used for the initial tuning of the cavity. Once the cavity is tuned close to the RF operating frequency, the SEL can be stopped, and the GDR control can be enabled to stabilize the RF fields for beam acceleration.

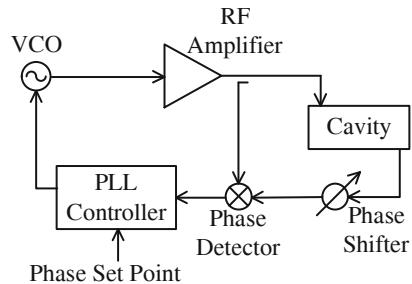
Note that the SEL may excite multiple modes of the cavity. This situation may happen if more than one passband modes or higher-order modes of the cavity satisfy the SEL starting up condition simultaneously. In practice, a band-pass filter can be installed to select the desired cavity mode.

The principle of SEL control will be discussed in Sect. 4.3. Particularly, in Sect. 4.3.3, the SEL in combination with a GDR feedback loop will be presented.

2.3.3 Phase-Locked Loop

The PLL, as in Fig. 2.5, is an alternative to SEL. They both can feed RF power into the cavity by tracking the RF frequency to the cavity resonance frequency. The PLL maintains the relative phase between the cavity input and output by adjusting the output frequency of a voltage-controlled oscillator (VCO). The loop phase shifter is used to tune the VCO frequency to operate the cavity on resonance. A PLL controller is used to lock the phase detector output to a set point. An example of the PLL-based RF controller design can be found in the reference article (Fu et al. 2020).

Fig. 2.5 Architecture of PLL control



After the PLL is initially tuned to operate the cavity on resonance, the VCO will follow the changes of the cavity resonance frequency. Assume that the output of the phase detector is zero (i.e., phase shift of the cavity without detuning) when the PLL is enabled. Then, if the cavity is detuned, the PLL maintains the output of the phase detector to be zero by changing the VCO frequency. Therefore, the cavity is always powered at the resonance frequency. With this feature, the PLL is also a commonly used solution to condition the cavity.

Note that an SEL starts up its operation spontaneously from noise, which needs time to ramp the RF drive power to the cavity. The SEL starting time depends on the noise level. One can add noise (or a seed signal) for a faster startup. However, in a PLL, the VCO always produces input power to the RF amplifier, so the cavity drive power always exists and does not need a ramping. It implies that a PLL may respond faster to the cavity resonance frequency changes than an SEL. This feature makes the PLL a candidate to test unloaded superconducting cavities with a very narrow bandwidth (e.g., several Hz) and disturbed by strong microphonics.

We will discuss more in detail about the PLL in Sect. 4.4. The design procedure of a feedback controller based on the traditional loop shaping method will also be demonstrated with the PLL.

2.4 Analog and Digital Control

When digital technology was not well developed in the early days, LLRF systems were implemented with analog circuits (Hassanzadegan et al. 2006). Nowadays, digital LLRF systems have become more and more popular due to the significant progress in digital technologies. In some existing circular accelerators operating in the CW mode, analog LLRF systems are still in operation. For increased flexibility, a digital system is often preferred in a newly developed or upgraded LLRF system.

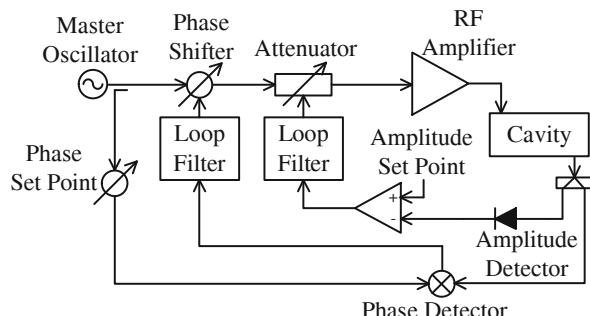
Figure 1.13 illustrates a typical digital LLRF system. It adopts ADCs to sample the RF or IF signals. The ADC samples are demodulated by digital algorithms to derive the RF field amplitude and phase. The RF controller is implemented in software or firmware of Digital Signal Processors (DSP) or FPGAs. The actuation signals from the RF controller are usually produced by DACs to drive the I/Q modulator.

The digital systems have exhibited excellent performance and flexibility. With fast high-resolution (e.g., 400 MSPS, 14-bit) ADCs, the direct sampling of a 1.3 GHz signal has been demonstrated with an acceptable signal-to-noise ratio (SNR) (Geng and Simrock 2008). High-performance DSPs or FPGAs can implement complex RF control algorithms with latency down to several tens of nanoseconds. The keep-increasing computation power has enabled to implementation of control algorithms based on real-time optimizations (e.g., model predictive control). At the RF actuator side, fast DACs can directly generate RF signals with a frequency of up to several GHz. Since the RF control functions are implemented in software or firmware, they can be easily configured to support different operation scenarios of the RF system. Furthermore, the results or status of the system can be easily accessed by reading the variables or registers.

Though a digital LLRF system is preferred in most machines, an analog LLRF system may still be preferable in some situations. Compared to the digital system, an analog system can achieve a much smaller loop delay, a much lower noise level and much higher reliability. A lower latency allows a higher closed-loop bandwidth of the RF control loop. Typically, the ADC is a limiting factor for digital LLRF systems to reduce the RF measurement noise. A state-of-the-art fast ADC can achieve an SNR up to around 80 dB, often enough for most LLRF systems. However, if a higher SNR is required, we may have to adopt an analog system. Of course, we may also consider using multiple ADC channels to sample the same signal and averaging them to reduce the uncorrelated noise (see Sect. 6.1.3.2). The flexibility of a digital LLRF system is, on the other hand, a potential risk for errors. For example, after rebooting the software, it may stop working because of configuration errors (e.g., variable values not successfully restored). The analog system is typically much simpler, and therefore, more reliable. Of course, maturely designed software and firmware help to improve the reliability of digital LLRF systems. In summary, the analog solution is only considered if a digital system cannot fulfill the latency, noise level or reliability requirements.

As an example, Fig. 2.6 shows a typical analog LLRF system. The amplitude is detected with a diode detector, and the phase is detected with a mixer. The control knobs, such as the amplitude and phase set points, feedback gain and feedforward signal, are implemented with potentiometers. For diagnostics, we need additional test equipment like oscilloscopes.

Fig. 2.6 A typical analog LLRF system



2.5 Single-Cavity and Vector-Sum Control

For an RF system consisting of multiple cavities, there are two strategies to supply RF power. One choice is to power each cavity with a separate RF amplifier controlled by a separate LLRF node, as depicted in Fig. 2.3. We call this strategy *single-cavity control*. The alternative approach is to power multiple cavities with a high-power RF amplifier. See Fig. 2.7, which is denoted as *vector-sum control*.

With the single-cavity control strategy, the cost may be higher because more RF amplifiers and LLRF nodes are required. However, it allows us to adopt low-power solid-state amplifiers. Compared to a tube-based amplifier, such as a magnetron or klystron, the solid-state amplifier is more reliable with lower noise levels. From the LLRF system viewpoint, controlling a single cavity is easier without the need to calibrate the vector sum, which is necessary for the vector-sum control. Another benefit of single-cavity control is that each cavity can be adjusted independently, which offers more knobs for the parameters to tune the beam acceleration.

For vector-sum control, multiple cavities are driven by a single high-power RF amplifier, usually realized by a klystron. One primary benefit of the vector-sum control is the cost reduction on the high-power RF side. However, it brings several challenges to the RF system control and operation. First, the RF fields in different cavities are measured with different RF detector channels. They need to be calibrated to generate a vector sum representing the integrated RF field experienced by the beam when it passes through all cavities. The vector-sum calibration is challenging, and the calibration error will be a source of RF field errors in the presence of microphonics. We will cover this topic in Sect. 9.3.4. Furthermore, since multiple cavities are powered by a common amplifier, the cavity with the lowest breakdown or quench gradient will limit the achievable gradient of all other cavities. This problem needs to be overcome by tuning the input power to each cavity. Therefore, we need to install a high-power tunable attenuator (also a phase shifter) before each cavity, which potentially increases the cost. Particularly, if a cavity is defective, we have to detune it far from the RF operating frequency to bypass it.

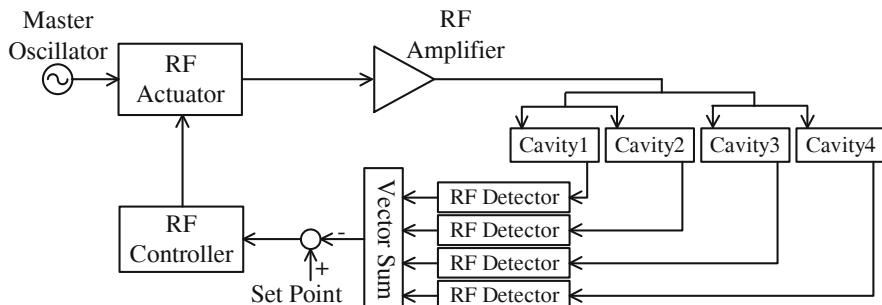


Fig. 2.7 Vector-sum control of 4 cavities

2.6 Summary

This chapter offers an overview of the strategies to control the RF systems of particle accelerators. The pros and cons of different strategies are compared. The discussion in this chapter allows determining the technology direction when starting the design of an LLRF system. The detailed discussions of the control strategies are distributed in the later chapters of this book. The referred chapters and sections have been given in the text.

References

- Akre R, Dowell D, Emma P et al (2007) LCLS LLRF upgrades to the SLAC linac. In: Proceedings of the 22nd IEEE particle accelerator conference, Albuquerque
- Delayen JR (1978) Phase and amplitude stabilization of superconducting resonators. Ph.D. thesis, California Institute of Technology
- Du BT, Lin HX, Liu W et al (2017) DSP frame and algorithm of LLRF of IR-FEL. In: Proceedings of the 8th international particle accelerator conference, Copenhagen
- Fong K, Laverty M, Zheng QW (2008) AM-PM conversion induced instability in an I/Q feedback control loop. In: Proceedings of the 24th linear accelerator conference, Victoria
- Fu X, Fong K, Yin Z et al (2020) A digital phase-locked loop based LLRF system. Nucl Instrum Methods Phys Res A962:163688. <https://doi.org/10.1016/j.nima.2020.163688>
- Geng Z (2017) RF control optimization and automation for normal conducting linear accelerators. IEEE Trans Nucl Sci 64(8):2361–2368. <https://doi.org/10.1109/TNS.2017.2723508>
- Geng Z, Hong B (2016) Design and calibration of an RF actuator for low-level RF systems. IEEE Trans Nucl Sci 63(1):281–287. <https://doi.org/10.1109/TNS.2015.2507204>
- Geng Z, Simrock S (2008) Evaluation of fast ADCs for direct sampling RF field detection for the European XFEL and ILC. In: Proceedings of the 24th linear accelerator conference, Victoria
- Hassanzadegan H, Salom A, Perez F (2006) Analog and digital low level RF for the ALBA synchrotron. In: Proceedings of the 10th European particle accelerator conference, Edinburgh
- Schilcher T (1998) Vector sum control of pulsed accelerating fields in Lorentz force detuned superconducting cavities. Ph.D. thesis, Hamburg University
- Skogestad S, Postlethwaite I (2005) Multivariable feedback control: analysis and design, 2nd edn. Wiley, New York
- Zhang R, Ben-Zvi I, Xie J (1993) A self-adaptive feedforward RF control system for Linacs. Nucl Instrum Methods Phys Res A324:421–428. [https://doi.org/10.1016/0168-9002\(93\)91045-O](https://doi.org/10.1016/0168-9002(93)91045-O)

Chapter 3

RF System Models



A (*mathematical*) *model* of the RF system is essential for the design and analysis of LLRF systems. The model here describes the (dynamical) relationship between the outputs and inputs of the RF system in the form of transfer functions. The transfer functions provide insights into the behavior and limitations of the RF system. For example, we may predict the RF system outputs in the presence of RF drive, beam loading and external disturbances (e.g., electromagnetic noise, thermal drifts or mechanical vibrations) using the transfer functions. Another useful information that the model can provide is the stability limits for the loop phase, loop gain and loop delay, which helps to determine these operational parameters of the RF controller. We will derive the RF system model through its physical principles, which can, therefore, also be used to optimize the operational parameters and to identify the characteristics of the RF system.

3.1 General Assumptions

The models discussed in this chapter will focus on the input-output relations of the RF system components (see Fig. 3.1), thereby allowing to design the RF system for control and operation. The RF cavity will be a primary component to be modeled. The cavity input and output are represented by the envelopes of the RF drive power and the RF field applied to the beam. The detailed spatial distribution of the RF field in the cavity will not be discussed since it is not relevant for the input-output relation. In this chapter, the RF system will be modeled as a linear system, and the nonlinear effects will be covered in Chap. 7. The linear model is simpler for the analysis and design of control systems. In practice, there exist non-linear components (e.g., RF amplifiers and klystron) in the RF systems of accelerators. However, if we consider a small range around the operating point of an RF system, a linear model is a good



Fig. 3.1 RF system model focusing on input-output relations

approximation. Furthermore, the drive terms of the cavity, including the RF drive power and the beam loading, will be treated as noiseless in the modeling. In Chap. 6, we will discuss the effects of the noisy RF drive power and beam loading on the cavity field.

The accelerator RF systems are typically narrowband operating at fixed or slowly time-varying frequencies. An RF reference signal provided by the synchronization system is used as a source of various frequencies in an RF station, such as the local oscillator (LO), clock and RF operating frequencies. We define the frequency of the RF reference signal as an *RF reference frequency* and use ω_{ref} to denote it. By default, we assume that the *RF operating frequency*, ω_{RF} , the actual working frequency of the RF station, equals ω_{ref} . Another concept we want to introduce is the *carrier frequency* ω_c , which is selected flexibly, and we often choose $\omega_c = \omega_{RF}$. For different RF signals, ω_c represents their common terms, which are omitted when describing the input-output relation of an RF component. If $\omega_c = \omega_{RF}$, we can obtain an RF system model in terms of the envelopes of its input and output.

When ω_{RF} is a constant, it is convenient to define ω_c the same as ω_{RF} , whereas for a time-varying ω_{RF} , we can treat ω_c in two alternative ways:

1. We can vary ω_c to track ω_{RF} . This method makes the changes of ω_{RF} transparent for the RF system model, which, therefore, only deals with the envelopes (i.e., baseband signals) of the RF signals.
2. We may choose a fixed ω_c . This causes a time-varying phase $\varphi(t) = [\omega_{RF}(t) - \omega_c]t + \varphi_{0i}$ (where φ_{0i} is a constant for each signal) in the signals handled by the RF system model. When using this method, the time-varying term $[\omega_{RF}(t) - \omega_c]t$ in the phases of all RF signals should be omitted to obtain the correct input-output relations of the RF components.

The first case above shifts the spectrum of an RF signal directly to around DC while the second case to around an intermediate frequency (IF). In this book, we will always assume that $\omega_c = \omega_{RF}$ to simplify the RF system modeling and the controller design.

Another assumption comes from the fact that the bandwidths of the RF systems of accelerators are much smaller than the RF operating frequency ($\omega_{1/2} \ll \omega_{RF}$, where $\omega_{1/2}$ is half of the full bandwidth). It implies that the RF signal envelopes change much slower than the RF frequency. Therefore, we can treat the RF signal envelopes approximately as constant within several periods of the RF frequency.

3.2 RF Modeling Method

The input-output relation of an RF component in Fig. 3.1 can be described as a *time-domain* function $g(t)$: $x(t) \rightarrow y(t)$. Here g is a scalar (or a matrix if x or y are vectors) for static systems and a (or a group of) differential equation(s) for dynamical systems. Using the Laplace transform, we can describe $g(t)$ as a transfer function $G(s)$ satisfying $Y(s) = G(s)X(s)$, where $X(s)$ and $Y(s)$ are the Laplace transforms of the time-domain input and output signals: $X(s) = \mathcal{L}\{x(t)\}$, $Y(s) = \mathcal{L}\{y(t)\}$. The operator $\mathcal{L}\{\cdot\}$ denotes the Laplace transform. Here we have assumed that $x(t) = 0$ and $y(t) = 0$ when $t \leq 0$. The variable s is a complex frequency defined as $s := \sigma + j\omega$. We will use the term *frequency domain (s domain)* to denote the collection of the Laplace transforms of signals and the transfer functions. The complex plane of s is referred to as the *s plane*. The Laplace transform is a powerful tool to model a linear system and forms the basis of classical control theories (Dorf and Bishop 2010). If the inputs and outputs of a system are *real signals*, which are time-domain scalars with real values, the transfer function of the system has real-valued coefficients, and its frequency response is symmetric for positive and negative frequencies.

As we have discussed, the input and output signals of an RF component contain the same carrier-frequency term, which does not provide useful information for the input-output relation. Therefore, we will only consider the transfer function obtained from the envelopes of the input and output signals. To do this, we remove the carrier-frequency terms and describe the RF signals as *phasors (complex signals)*. The conversions between phasors and RF signals match well the RF signal detection and actuation processes in an LLRF system. Applying Laplace transforms to the input and output phasors yields a transfer function with complex-valued coefficients and with asymmetric frequency responses (Novotny and Wouterse 1976; Martin 2004; Harnefors 2007; Brandt 2007; Troeng 2019). Such a transfer function derived from phasors has a primary advantage: we can reduce the order of the RF system model from 2 to 1. Later we will see that it simplifies the analysis and design of LLRF systems significantly.

3.2.1 RF Signal Description

An RF signal with a carrier frequency ω_c , a time-varying amplitude $x_0(t)$ and a time-varying phase $\varphi(t)$ can be described in the time domain as

$$x(t) = x_0(t) \cdot \cos(\omega_c t + \varphi(t)), \quad (3.1)$$

where the amplitude and phase contain all information relevant for the modeling of the dynamical behavior of the RF system. From Euler's formula $e^{j\theta} = \cos\theta + j\sin\theta$, we can define a complex time-domain signal by expanding the signal above with a sine term $jx_0(t)\sin(\omega_c t + \varphi(t))$ as the imaginary part, resulting in

$$\tilde{\mathbf{x}}(t) := x_0(t)e^{j(\omega_c t + \varphi(t))}. \quad (3.2)$$

The carrier-frequency term, $e^{j\omega_c t}$, is the same for all RF signals. It can be treated as a constant since $\omega_c = \omega_{RF}$ and ω_{RF} is a constant or varies slowly compared to the time constant of the RF system. By discarding $e^{j\omega_c t}$, we obtain a complex envelope of the RF signal with all the information of amplitude and phase kept:

$$\mathbf{x}(t) := x_0(t)e^{j\varphi(t)} = I(t) + jQ(t). \quad (3.3)$$

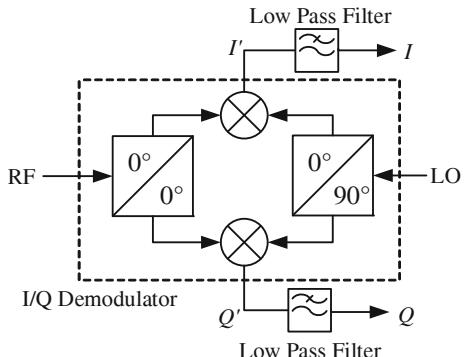
The complex envelope (3.3) is called a *phasor* standing for a “phase vector”. Its magnitude and phase correspond to the amplitude and phase of the signal (3.1), respectively. In this book, we will describe RF signals as phasors.

Note that the phasors are defined for a particular ω_c . When using a different carrier frequency ω'_c , we can rewrite (3.1) as $x(t) = x_0(t) \cdot \cos(\omega'_c t + (\omega_c - \omega'_c)t + \varphi(t))$, and the phasor becomes $\mathbf{x}'(t) = x_0(t)e^{j(\omega_c - \omega'_c)t + j\varphi(t)}$. We can see that there remains an oscillatory term in phase $\varphi'(t) = (\omega_c - \omega'_c)t + \varphi(t)$ if the carrier frequency differs from ω_{RF} ; that is, instead of converting the RF signal to baseband, we converted it to an intermediate frequency $\omega_{IF} = \omega_c - \omega'_c$.

3.2.2 Principle of RF Signal Detection

In an LLRF system, RF detectors convert RF signals to phasors, whereas RF actuators convert phasors back to RF signals. An RF controller processes the RF field phasors measured by the RF detectors and generates a drive phasor to the RF actuator. To understand the physical meaning of phasors, we take a look at the processes of RF detection and actuation. Since the RF actuation is an inverted process of RF detection, we only analyze the principle of RF detection, which is depicted using the I/Q demodulator in Fig. 3.2.

Fig. 3.2 A typical RF detector using an I/Q demodulator



To detect the phasor of an RF signal picked from the cavity, we introduce a local oscillator (LO) as a reference. The RF and LO signals are expressed as

$$x_{RF}(t) = x_{RF0} \cos(\omega_{RF}t + \varphi), x_{LO}(t) = x_{LO0} \cos(\omega_{LO}t),$$

where the LO phase is normalized to 0, and we assume the LO frequency $\omega_{LO} < \omega_{RF}$. The I/Q demodulator works as a multiplier of a real signal and a complex signal. The RF signal is split by an in-phase splitter and is equivalent to a real signal $x'_{RF}(t) = x_{RF}(t)/\sqrt{2}$, whereas the LO signal is split by a quadrature splitter and forms an equivalent complex signal

$$\tilde{x}'_{LO}(t) = \frac{x_{LO0}}{\sqrt{2}} [\cos(\omega_{LO}t) - j \sin(\omega_{LO}t)] = \frac{x_{LO0}}{\sqrt{2}} e^{-j\omega_{LO}t},$$

where the 0° and 90° branches are the real and imaginary parts, respectively. Note that the spectrum of x'_{RF} is symmetric for positive and negative frequencies, whereas the spectrum of \tilde{x}'_{LO} only appears at negative frequencies, see Fig. 3.3. The I/Q demodulator outputs are the products of x'_{RF} and \tilde{x}'_{LO} and are given by

$$I'(t) + jQ'(t) = \frac{x_{RF0}x_{LO0}}{4} \left(e^{-j[(\omega_{RF} + \omega_{LO})t + \varphi]} + e^{j[(\omega_{RF} - \omega_{LO})t + \varphi]} \right), \quad (3.4)$$

which contains two frequency components at $-(\omega_{RF} + \omega_{LO})$ and $(\omega_{RF} - \omega_{LO})$, respectively. In practice, we use low pass filters to filter out the frequency component at $-(\omega_{RF} + \omega_{LO})$, resulting in a time-varying phasor

$$x_{RF}(t) = I(t) + jQ(t) = \frac{x_{RF0}x_{LO0}}{4} e^{j[(\omega_{RF} - \omega_{LO})t + \varphi]}. \quad (3.5)$$

The frequency relations in Fig. 3.3 imply that we can derive the phasor spectrum from the RF signal spectrum in the following way: first, we shift the frequency origin to the LO frequency, then, we filter out the frequency components at $-(\omega_{RF} + \omega_{LO})$.

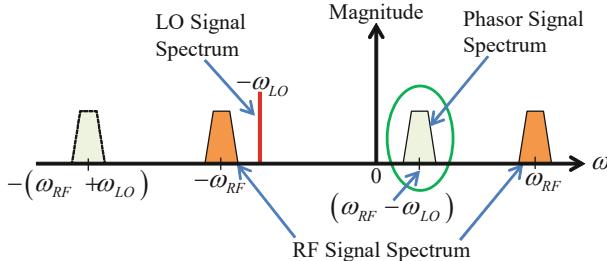
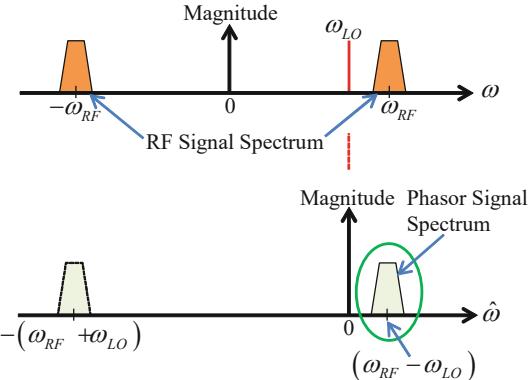


Fig. 3.3 Frequency domain conversion of RF detection. The RF detector output is a time-varying phasor of the RF signal with asymmetric spectrum. In the plot, the phasor spectrum only contains the circled part at positive frequencies

Fig. 3.4 Derive the phasor spectrum from the RF signal spectrum



See Fig. 3.4. This corresponds to defining a new frequency $\hat{\omega} = \omega - \omega_{LO}$ and only keeping the slowly changing envelope of the RF signal (i.e., the low-frequency components in the new frequency domain). Note that we will lose half of the magnitude of the input RF signal if we filter out one of the sidebands of its spectrum. This can be explained by the Fourier transform and the inverse Fourier transform of a sine signal. The spectrum of a sine signal with amplitude A and frequency ω_{RF} contains two components at $\pm\omega_{RF}$. These two spectrum lines are conjugate to each other with the same magnitude $A/2$ (or with a factor $\sqrt{2\pi}$ depending on the definition of the Fourier transform, but this factor will be compensated when making the inverse transform). With the inverse Fourier transform, we can reconstruct the original sine signal using these two spectrum lines. However, if we discard one of the spectrum lines and take the inverse Fourier transform, the amplitude of the reconstructed signal will be only half of the original signal. Based on this consideration, the magnitude of the phasor spectrum should be doubled to preserve the signal level if the high-frequency components are filtered out.

The frequency shifting can also be applied to the Laplace transforms of the RF signals to obtain the Laplace transforms of the corresponding phasors. The transfer functions of the RF system components will also be adjusted correspondingly to the shifted frequency.

3.2.3 Phasor Laplace Transform

The inputs and outputs of the RF systems of particle accelerators are RF signals. The model in Fig. 3.1 still contains the carrier-frequency term, but the input-output relation in terms of phasors (i.e., envelopes) is of more interest.

To derive the Laplace transform of the phasor (3.3), we apply Laplace transform to (3.2) and obtain

$$\tilde{\mathbf{X}}(s) = \mathcal{L}\{\tilde{\mathbf{x}}(t)\} = X(s) \left(1 + je^{-s\pi/(2\omega_c)} \right), \quad (3.6)$$

where $X(s)$ is the Laplace transform of the RF signal $x(t)$ in (3.1). Here we have used the approximation $x(t - \pi/(2\omega_c)) \approx x_0(t) \cdot \sin(\omega_c t + \varphi(t))$ when the variations of the amplitude and phase within a period of the RF signal is negligible. If both the input and output signals of an RF component are expressed as complex time-domain signals like (3.2), their Laplace transforms also satisfy $\tilde{\mathbf{Y}}(s) = G(s)\tilde{\mathbf{X}}(s)$, where $G(s)$ is the transfer function of the RF component. Phasors are complex envelopes of RF signals modulating the carrier frequency. According to Fig. 3.4, we can derive the spectrum of a phasor from the spectrum of the original RF signal by shifting the frequency origin by ω_c and removing the higher frequency components. Similarly, we may move the frequency axis of the s plane by ω_c to obtain a new form of Laplace transform named *phasor Laplace transform*. Then we can describe the inputs and outputs of an RF component in the form of phasor Laplace transforms, and model the RF component as a single-input single-output (SISO) *phasor transfer function*. The phasor Laplace transform defines a new complex frequency $\hat{s} := \sigma + j\hat{\omega}$, where $\hat{\omega} := \omega - \omega_c$ is the shifted frequency that is the frequency of the RF signal envelope. We have the relation $\hat{s} = s - j\omega_c$. We will use the term *offset frequency domain* (\hat{s} domain) to denote the collection of the phasor transfer functions and the phasor Laplace transforms of signals. The complex plane of \hat{s} will be denoted as the \hat{s} plane.

The relationship between the complex envelope (3.3) and the complex RF signal (3.2) can be written as $\mathbf{x}(t)e^{j\omega_c t} = \tilde{\mathbf{x}}(t)$. Take Laplace transform on both sides and we obtain $\mathbf{X}(s - j\omega_c) = \tilde{\mathbf{X}}(s)$, where $\mathbf{X}(s) = \mathcal{L}\{\mathbf{x}(t)\}$. Substitute the independent variable s with $s = \hat{s} + j\omega_c$, we obtain

$$\mathbf{X}(\hat{s}) = \tilde{\mathbf{X}}(\hat{s} + j\omega_c) = X(\hat{s} + j\omega_c) \left(1 + je^{-(\hat{s} + j\omega_c)\pi/(2\omega_c)} \right), \quad (3.7)$$

where the exponential term can be written as $e^{-\hat{s}\pi/(2\omega_c)}e^{-j\pi/2}$. If we only consider the low-frequency changes in the RF signal envelope compared to the carrier frequency, which means $|\hat{s}| \ll \omega_c$, we can get $e^{-\hat{s}\pi/(2\omega_c)} \approx 1$, and then the phasor Laplace transform of the RF signal in (3.7) can be simplified as

$$\mathbf{X}(\hat{s}) \approx 2X(s)|_{s=\hat{s}+j\omega_c}, \quad |\hat{s}| \ll \omega_c. \quad (3.8)$$

The factor 2 will compensate for the loss of half of the magnitude after shifting the frequency and discarding the higher frequency components, which is consistent with the discussions in Figs. 3.3 and 3.4. By converting the relation $\tilde{\mathbf{Y}}(s) = G(s)\tilde{\mathbf{X}}(s)$ into the \hat{s} domain, we get $\tilde{\mathbf{Y}}(\hat{s} + j\omega_c) = G(\hat{s} + j\omega_c)\tilde{\mathbf{X}}(\hat{s} + j\omega_c)$ and then $\mathbf{Y}(\hat{s}) = G(\hat{s} + j\omega_c)\mathbf{X}(\hat{s})$ according to (3.7). Hence, the phasor transfer function can be defined as

$$\mathbf{G}(\hat{s}) := G(\hat{s} + j\omega_c) = G(s)|_{s=\hat{s}+j\omega_c}, \quad |\hat{s}| \ll \omega_c. \quad (3.9)$$

Therefore, the relation between the phasor Laplace transforms of the input and output RF signals are given by

$$\mathbf{Y}(\hat{s}) = \mathbf{G}(\hat{s})\mathbf{X}(\hat{s}). \quad (3.10)$$

Note that $\mathbf{G}(\hat{s})$ is usually applied under the assumption $|\hat{s}| \ll \omega_c$, which means, we are more interested in the low-frequency (in the \hat{s} domain) dynamics of $\mathbf{G}(\hat{s})$ and neglect its high-frequency modes. Equations (3.8, 3.9, and 3.10) are the foundation of the RF system modeling method based on phasor Laplace transforms. With this method, we describe the RF signal envelopes with phasor Laplace transforms and model the RF system behaviors using phasor transfer functions. As depicted in Fig. 3.5, we can obtain the \hat{s} plane by shifting the real axis of the s plane to $j\omega_c$. Such a treatment will change the dynamical properties of a system. For example, suppose we have two conjugate poles p_1 and p_2 on the s plane; that is, we have a second-order system in the s domain. After converting to the \hat{s} plane, the two poles, \hat{p}_1 and \hat{p}_2 , are no longer symmetric to the real axis, resulting in two different dynamical modes. The pole \hat{p}_1 is at a lower frequency ($\omega - \omega_c$) representing the behavior of the input-output envelopes, whereas \hat{p}_2 is at a higher frequency ($\omega + \omega_c$). We will discard \hat{p}_2 in the phasor transfer function because we are not interested in the fast time-varying terms but only the envelopes. It implies that we can use a first-order phasor transfer function to replace the s domain second-order transfer function.

If either the RF system or the RF signal is narrowband compared to the carrier frequency, the following procedure can be used to model the RF system:

- Derive a transfer function from the RF system's differential equations, then replace s with $\hat{s} + j\omega_c$ to obtain the phasor transfer function. The dynamics in the phasor transfer function at frequencies much higher than the system bandwidth should be neglected.

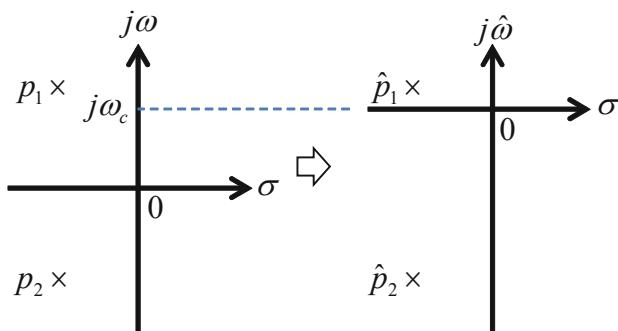


Fig. 3.5 Transfer from s plane to \hat{s} plane

2. Derive the Laplace transform of the input RF signal, then replace s with $\hat{s} + j\omega_c$ and multiply by 2 to obtain its phasor Laplace transform. The frequency components of the RF signal's phasor Laplace transform at frequencies much higher than the system bandwidth should be neglected.
3. Use (3.10) to calculate the phasor Laplace transform of the output RF signal. With the inverse Laplace transform, the output phasor in the time domain can be constructed.

Note that the phasor transfer function $\mathbf{G}(\hat{s})$ may contain complex-valued coefficients that might not be well supported by the existing software tools like Matlab. Nevertheless, it can be easily converted to a 2-by-2 multiple-input multiple-output (MIMO) system by separating the real and imaginary parts, which fits better to the software tools. In this case, the input and output signals should be both described as I/Q vectors representing the real and imaginary parts of the phasors. The following form of the system transfer function is equivalent to (3.10):

$$\begin{bmatrix} Y_I(\hat{s}) \\ Y_Q(\hat{s}) \end{bmatrix} = \begin{bmatrix} G_{11}(\hat{s}) & G_{12}(\hat{s}) \\ G_{21}(\hat{s}) & G_{22}(\hat{s}) \end{bmatrix} \begin{bmatrix} X_I(\hat{s}) \\ X_Q(\hat{s}) \end{bmatrix},$$

where G_{11} , G_{12} , G_{21} and G_{22} contain only real coefficients, $\mathbf{Y}(\hat{s}) = Y_I(\hat{s}) + jY_Q(\hat{s})$ and $\mathbf{X}(\hat{s}) = X_I(\hat{s}) + jX_Q(\hat{s})$. The MIMO form is suitable for numerical simulations, while (3.10) is more potent for the analytical study of the RF system's behaviors and characteristics. To simplify our presentation, we will omit the independent variables s , \hat{s} and t in the expressions when the meaning is clear.

3.3 Single-Cell Cavity Model

Single-cell standing-wave cavities are widely used in storage rings. In this section, we will model both the electrical and mechanical properties of single-cell cavities. The cavity behaviors in the presence of RF drive power, beam loading and external disturbances will be studied. The single-cell cavity model is also a basis for analyzing multi-cell cavities. A multi-cell cavity has multiple passband modes, and each mode can be described by a similar transfer function to that of a single-cell cavity. The multi-cell cavities will be discussed in the next section.

3.3.1 Parallel RLC Circuit Model

A cavity driven by an RF power source can be modeled as a parallel RLC circuit (Schilcher 1998). See Fig. 3.6. The RF power source can be modeled as a current source, and the beam in the cavity appears as the second current source driving the cavity. The input coupler, which couples RF power into the cavity, can be modeled

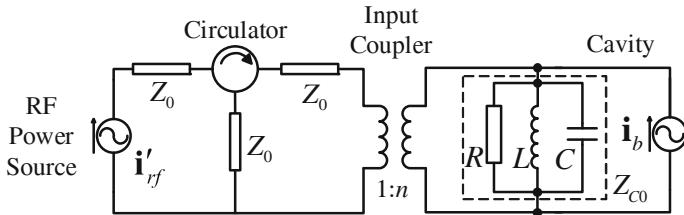
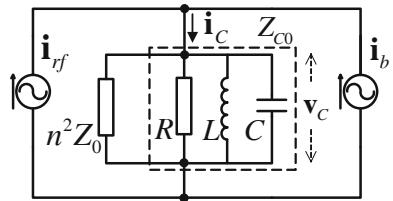


Fig. 3.6 Circuit model of a single-cell cavity driven by an RF power source and beam loading

Fig. 3.7 Circuit model of a single-cell cavity converted to the cavity side of transformer



as a lossless transformer with a voltage ratio of $1:n$. The characteristic impedance of the transmission line is a pure resistance Z_0 , and a circulator absorbs the RF power reflected from the cavity. The cavity's phasor transfer function describes the cavity behaviors with the input as a drive current phasor and the output as a cavity voltage phasor.

If we convert the RF drive current and the transmission line impedance to the cavity side of the transformer, Fig. 3.6 changes to Fig. 3.7. Based on the transformer's principle, the transmission line impedance Z_0 introduces the term $n^2 Z_0$ parallel to the cavity impedance at the cavity side. The RF power-induced drive current is denoted as a phasor i_{rf} , which is a superposition of the forward and reflected RF-induced currents at the input coupler. The beam-induced drive current is i_b . The overall drive current and the resulting cavity voltage are denoted as phasors i_C and v_C , where i_C is calculated as

$$i_C = i_{rf} + i_b. \quad (3.11)$$

A cavity is typically characterized by several measurable parameters like resonance frequency, coupling factor and loaded quality factor. To derive a meaningful cavity model from the equivalent circuit, we use the relations between the cavity parameters and the parallel RLC circuit parameters. See Table 3.1.

In the table above, W is the stored electromagnetic energy in the cavity; P_{cav} is the dissipated RF power in the cavity wall, and P_{ext} is the dissipated RF power on the external impedance, which corresponds to the emitted RF power out of the cavity through the input coupler.

Table 3.1 Relations between the cavity parameters and the parallel RLC circuit parameters

Cavity parameter	Notation	Relation with circuit parameters
Resonance frequency	ω_0	$\omega_0 = 1/\sqrt{LC}$
Unloaded quality factor	Q_0	$Q_0 = \frac{\omega_0 W}{P_{cav}} = \omega_0 RC = \frac{R}{L\omega_0}$
External quality factor	Q_{ext}	$Q_{ext} = \frac{\omega_0 W}{P_{ext}} = \omega_0 n^2 Z_0 C$
Coupling factor	β	$\beta = \frac{P_{ext}}{P_{cav}} = \frac{Q_0}{Q_{ext}} = \frac{R}{n^2 Z_0}$
Loaded quality factor	Q_L	$Q_L = \frac{Q_0}{1+\beta}$
Normalized shunt impedance	r/Q	$R = \frac{1}{2}(r/Q)Q_0$
Loaded resistance	R_L	$R_L = R/(1+\beta) = \frac{1}{2}(r/Q)Q_L$
Half-bandwidth	$\omega_{1/2}$	$\omega_{1/2} = \frac{\omega_0}{2Q_L}$

Table 3.2 Complex impedances of basic circuit elements

Element	Notation	s Domain impedance	\hat{s} Domain impedance
Resistance	R	R	R
Capacitance	C	$\frac{1}{sC}$	$\frac{1}{(\hat{s}+j\omega_c)C}$
Inductance	L	sL	$(\hat{s}+j\omega_c)L$

3.3.2 Cavity Phasor Transfer Function

Circuit models can be analyzed with the phasor Laplace transform method. The \hat{s} domain complex impedances of the basic circuit elements are defined in Table 3.2.

According to Fig. 3.7, if the cavity input phasor \mathbf{i}_C and the output phasor \mathbf{v}_C are presented as phasor Laplace transforms $\mathbf{I}_C(\hat{s})$ and $\mathbf{V}_C(\hat{s})$, the phasor transfer function of the cavity $\mathbf{G}_C(\hat{s})$ can be written as

$$\mathbf{G}_C(\hat{s}) = \frac{\mathbf{V}_C(\hat{s})}{\mathbf{I}_C(\hat{s})} = \frac{(\hat{s} + j\omega_c)/C}{(\hat{s} + j\omega_c)^2 + (\hat{s} + j\omega_c)/(R_L C) + 1/(LC)}.$$

With the relations in Table 3.1, it can be rewritten with the cavity parameters:

$$\mathbf{G}_C(\hat{s}) = \frac{2\omega_{1/2}R_L(\hat{s} + j\omega_c)}{(\hat{s} + j\omega_c)^2 + 2\omega_{1/2}(\hat{s} + j\omega_c) + \omega_0^2}. \quad (3.12)$$

The two poles of (3.12) are $\hat{s}_{p1,2} = -\omega_{1/2} \pm j\sqrt{\omega_0^2 - \omega_{1/2}^2} - j\omega_c$. The cavity bandwidth is typically much smaller than its resonance frequency, that means $\omega_{1/2} \ll \omega_0$, then the poles can be approximated to $\hat{s}_{p1} \approx -\omega_{1/2} + j(\omega_0 - \omega_c)$ and $\hat{s}_{p2} \approx -\omega_{1/2} - j(\omega_0 + \omega_c)$. The two poles define two damped resonances at frequencies $\omega_0 - \omega_c$ and $\omega_0 + \omega_c$ respectively. If we define the *cavity detuning* as

$$\Delta\omega = \omega_0 - \omega_c, \quad (3.13)$$

and assume $\Delta\omega \ll \omega_0$, Eq. (3.12) can be simplified as

$$\mathbf{G}_C(\hat{s}) = \frac{\omega_{1/2}R_L}{\hat{s} + \omega_{1/2} - j\Delta\omega}. \quad (3.14)$$

Here only the slow changes in the RF signal envelopes are considered, and we have assumed $|\hat{s}| \ll \omega_c \approx \omega_0$. The passband at around $-2\omega_c$ of (3.12) was neglected. See the discussion of Fig. 3.5. As seen in (3.14), the \hat{s} -domain cavity model is simplified to a first-order system with a single complex pole, resulting in asymmetric frequency responses at positive and negative frequencies. The differential equation for the time-domain cavity input/output phasors can be derived from (3.14), which is useful for the time-domain analysis of the cavity behaviors:

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = \omega_{1/2}R_L\mathbf{i}_C. \quad (3.15)$$

By separating the real and imaginary parts of the phasors, the cavity state-space equation can be written as

$$\begin{bmatrix} \dot{v}_{CI} \\ \dot{v}_{CQ} \end{bmatrix} = \begin{bmatrix} -\omega_{1/2} & -\Delta\omega \\ \Delta\omega & -\omega_{1/2} \end{bmatrix} \begin{bmatrix} v_{CI} \\ v_{CQ} \end{bmatrix} + \omega_{1/2}R_L \begin{bmatrix} i_{CI} \\ i_{CQ} \end{bmatrix}, \quad (3.16)$$

where $\mathbf{v}_C = v_{CI} + jv_{CQ}$ and $\mathbf{i}_C = i_{CI} + ji_{CQ}$. The state-space equation matches well to the modern control theory. It is suitable for implementing discrete algorithms derived from the cavity model.

3.3.3 Cavity Step Response

The cavity response to a step change in the input RF power is helpful to illustrate the dynamical behavior of the cavity. Assume the cavity drive current is turned on at $t = 0$ with a constant amplitude i_{C0} and a phase φ_0 . Its time-domain expression is $i_C = i_{C0}\cos(\omega_c t + \varphi_0) \cdot u(t)$ where $u(t)$ is the unit step function. The Laplace transform of i_C is $I_C(s) = i_{C0}(s \cos \varphi_0 - \omega_c \sin \varphi_0)/(s^2 + \omega_c^2)$. From (3.8), the phasor Laplace transform of i_C can be written as

$$\mathbf{i}_C(\hat{s}) = 2I_C(\hat{s} + j\omega_c) \approx i_{C0}(\cos \varphi_0 + j \sin \varphi_0)/\hat{s} = \mathbf{i}_C/\hat{s}, \quad (3.17)$$

where $\mathbf{i}_C = i_{C0}e^{j\varphi_0}$ is a constant phasor of the cavity drive current. When making the approximation in (3.17), the higher-frequency components in the \hat{s} domain have been neglected. If we apply (3.17) to the phasor transfer function of the cavity (3.14), we can derive the time-domain phasor of the cavity voltage

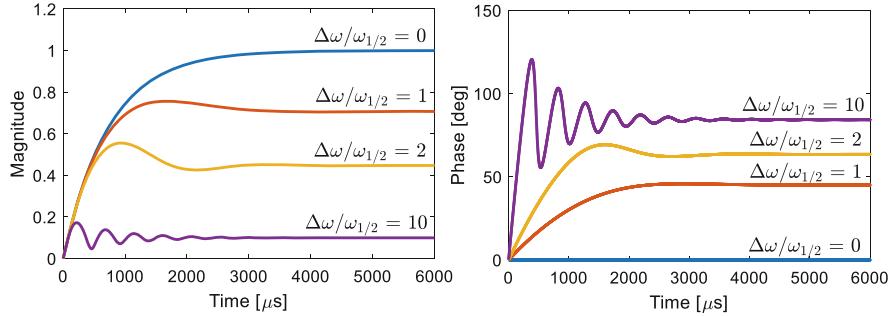


Fig. 3.8 Amplitude and phase step responses of a cavity with different detuning. The maximum magnitude of \mathbf{v}_C has been normalized to 1. When the cavity is detuned, overshoots and oscillations appear in the transient response, where the oscillation frequency equals to $\Delta\omega$

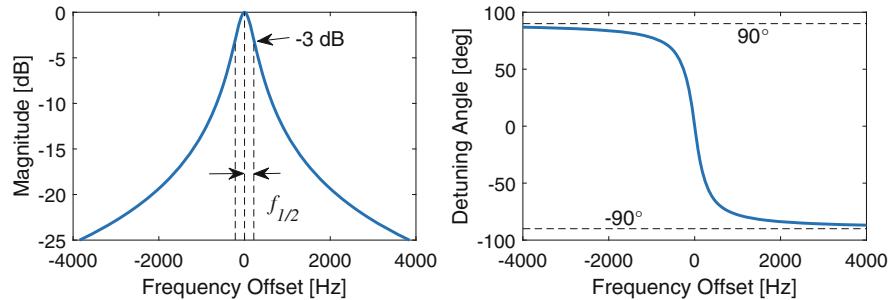


Fig. 3.9 Resonance curves of cavity voltage with different detuning

$$\mathbf{v}_C = \frac{\omega_{1/2} R_L \mathbf{i}_C}{\omega_{1/2} - j\Delta\omega} \left(1 - e^{-(\omega_{1/2} - j\Delta\omega)t} \right). \quad (3.18)$$

Here we have assumed the initial condition of the cavity voltage is zero; that is, $\mathbf{v}_C = 0$ when $t = 0$. The amplitude and phase of the cavity voltage in response to a step change in the drive current are depicted in Fig. 3.8. In (3.18), when $t \rightarrow \infty$, \mathbf{v}_C approaches the steady state with a vanishing exponential term. In the steady state, the relative angle ψ between \mathbf{v}_C and \mathbf{i}_C is determined by the ratio of $\Delta\omega$ and $\omega_{1/2}$ of the cavity. We define ψ as a *detuning angle* given by

$$\psi = \tan^{-1}(\Delta\omega/\omega_{1/2}). \quad (3.19)$$

The steady-state amplitude and phase of \mathbf{v}_C for $\mathbf{i}_C = 1$ are plotted as functions of the frequency offset between the RF operating frequency and the cavity resonance frequency. They are known as the *resonance curves*. See Fig. 3.9. The curves can also be derived from the frequency responses of the phasor transfer function (3.14) as the magnitude and phase of $\mathbf{G}_C(j\hat{\omega})$.

Note that the *steady state* indicates a state where the cavity voltage is a constant; that is, $\dot{\mathbf{v}}_C = 0$. In the steady state, the differential Eq. (3.15) becomes an algebraic equation $(\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = \omega_{1/2}R_L\mathbf{i}_C$. If the cavity drive has a step change in its envelope, the cavity reaches the steady state if we wait for a long time (e.g., longer than several times the cavity time constant). Alternatively, we may also achieve a constant cavity voltage by manipulating \mathbf{i}_C , which allows us to create a steady state even in a short RF pulse.

3.3.4 Cavity Response to RF Power

The cavity drive term \mathbf{i}_C in (3.15) consists of equivalent drive currents driven from both the RF drive power (\mathbf{i}_{rf}) and beam (\mathbf{i}_b). Let us first find out the relation between the RF drive power and the corresponding \mathbf{i}_{rf} in the cavity equation.

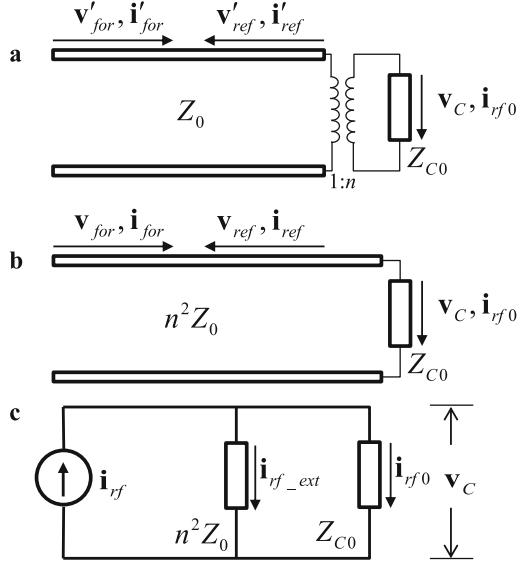
In the cavity circuit model of Fig. 3.7, Z_{C0} is the *unloaded impedance* of the cavity, which denotes the parallel impedance of the unloaded resistance R , capacitance C and inductance L . We use $//$ to indicate the parallel connection, so $Z_{C0} = R//L//C$. Connecting the cavity to a coupler and a transmission line corresponds to introducing an *external impedance* $Z_{ext} = n^2Z_0$. The *loaded impedance* of the cavity is then given by $Z_{CL} = Z_{ext}//Z_{C0}$. From the parallel circuits principle, the same voltage \mathbf{v}_C applies to all the circuit components Z_{ext} , R , L and C , whereas they obtain different currents depending on their impedances. The sum of these currents is \mathbf{i}_C . Note that R and Z_{ext} consume RF power. The power consumption in R corresponds to the power dissipation in the cavity wall. The consumption in Z_{ext} is the RF power emitted out of the cavity through the input coupler.

When there is no beam, the drive current \mathbf{i}_{rf} applies to Z_{CL} , as in Fig. 3.10c. Here \mathbf{i}_{rf0} and \mathbf{i}_{rf_ext} are the currents passing through Z_{C0} and Z_{ext} , respectively. To obtain the relation between \mathbf{i}_{rf} and the RF drive power P_{for} , we can simplify Fig. 3.6 into Fig. 3.10a. The phasors of the forward voltage and current at the transmission line side are denoted as \mathbf{v}'_{for} and \mathbf{i}'_{for} , while the reflected phasors are \mathbf{v}'_{ref} and \mathbf{i}'_{ref} . In Fig. 3.10b, all the quantities are transformed to the cavity side of the transformer, then Z_0 becomes n^2Z_0 and the forward and reflected phasors are labeled without'. The principle of transformers tells $\mathbf{v}_{for} = n\mathbf{v}'_{for}$, $\mathbf{v}_{ref} = n\mathbf{v}'_{ref}$, $\mathbf{i}_{for} = \mathbf{i}'_{for}/n$, $\mathbf{i}_{ref} = \mathbf{i}'_{ref}/n$. Then, the RF drive power can be calculated as

$$P_{for} = |\mathbf{i}'_{for}|^2 Z_0/2 = |\mathbf{i}_{for}|^2 n^2 Z_0/2. \quad (3.20)$$

From the transmission line theory (Pozar 2011), the following relations exist at the end of the transmission line in Fig. 3.10b:

Fig. 3.10 Simplified model of cavity with transmission line



$$\begin{aligned} \mathbf{v}_{for} &= n^2 Z_0 \mathbf{i}_{for}, \mathbf{v}_{ref} = n^2 Z_0 \mathbf{i}_{ref}, \mathbf{v}_C = Z_{C0} \mathbf{i}_{rf0} \\ \mathbf{v}_C &= \mathbf{v}_{for} + \mathbf{v}_{ref}, \mathbf{i}_{rf0} = \mathbf{i}_{for} - \mathbf{i}_{ref}. \end{aligned} \quad (3.21)$$

Then we obtain $\mathbf{i}_{rf0} = 2\mathbf{i}_{for}n^2Z_0/(Z_{C0} + n^2Z_0)$. This is the RF power-induced drive current passing through Z_{C0} . Then the current passing through Z_{ext} can be calculated from Fig. 3.10c as $\mathbf{i}_{rf_ext} = \mathbf{i}_{rf0}Z_{C0}/(n^2Z_0)$, and finally

$$\mathbf{i}_{rf} = \mathbf{i}_{rf0} + \mathbf{i}_{rf_ext} = 2\mathbf{i}_{for}. \quad (3.22)$$

The magnitude of the forward drive current \mathbf{i}_{for} of the forward RF power can be calculated from (3.20) and Table 3.1 as

$$|\mathbf{i}_{for}| = \sqrt{\frac{2P_{for}}{n^2Z_0}} = \sqrt{\frac{2\beta P_{for}}{R}} = \sqrt{\frac{2\beta}{\beta + 1} \cdot \frac{P_{for}}{R_L}}. \quad (3.23)$$

Using the two above equations, we can calculate $|\mathbf{i}_{rf}|$ from P_{for} if β and R_L are given. Let us use the 9-cell TESLA cavity as an example. If we only consider the π -mode, it can be viewed as an equivalent single-cell cavity. The resonance frequency of the cavity is $f_0 = 1.3$ GHz, the r/Q for the π -mode is 1036Ω , and the loaded quality factor is $Q_L = 3 \times 10^6$. The loaded resistance can be calculated as $R_L = 1554 \text{ M}\Omega$. Since the TESLA cavity has a large coupling factor ($\beta > 3000$), we approximate $\beta/(\beta + 1) \approx 1$. Then a drive power of 200 kW corresponds to a forward current $|\mathbf{i}_{for}| \approx 16 \text{ mA}$. So the RF-induced drive current is $|\mathbf{i}_{rf}| = 2|\mathbf{i}_{for}| \approx 32 \text{ mA}$.

3.3.5 Cavity Response to a Single Bunch

In an RF cavity, charged particles are typically accelerated in bunches. When a single bunch passes through a cavity, it will induce a transient cavity voltage. The transient voltage can be measured if its level is larger than the noise of the RF detectors, and the information can be used for beam diagnostics (e.g., bunch arrival time) and beam-based calibrations (e.g., beam phase calibration) (Pawlik et al. 2005).

From the phasor transfer function (3.14), we may calculate the cavity response to a single bunch. Assume we have a bunch with a charge Q_b . The bunch can be viewed as a short Gaussian-shaped current pulse when it flies through the cavity (see Fig. 3.11a), and the area under the current pulse equals the bunch charge. When the bunch length Δt is much smaller than the time constant of the cavity, the single bunch current pulse in the time domain can be simplified to a Dirac delta function as $i_{bs} = Q_b\delta(t)$ (Koontz et al. 1977). See Fig. 3.11b. The subscript “ bs ” stands for the cavity drive term of a single bunch. Here we assume that the bunch is injected at $t = 0$.

The Laplace transform of $\delta(t)$ is 1, so from (3.8), the phasor Laplace transform of i_{bs} is

$$\mathbf{I}_{bs}(\hat{s}) = 2Q_b. \quad (3.24)$$

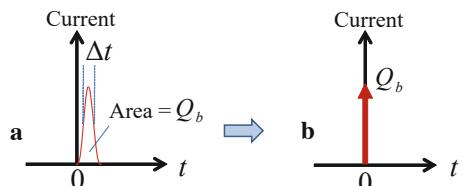
Then we can calculate the phasor Laplace transform of the induced cavity voltage $\mathbf{V}_{Cbs} = \mathbf{G}_C \mathbf{I}_{bs}$. Using the inverse Laplace transform, we obtain the time-domain phasors of the single-bunch drive current and the induced cavity voltage

$$\mathbf{i}_{bs} = 2Q_b\delta(t), \mathbf{v}_{Cbs} = 2\omega_{1/2}R_L Q_b e^{-(\omega_{1/2}-j\Delta\omega)t}. \quad (3.25)$$

In the calculations above, we have assumed $\mathbf{v}_{Cbs}(t = 0) = 0$. The amplitude of \mathbf{v}_{Cbs} decays with the cavity time constant $\tau = 1/\omega_{1/2}$, and its phase has a slope of $\Delta\omega$. The figure below shows the simulation of the cavity voltage induced by a single bunch in a detuned TESLA cavity.

A bunch induces a voltage in the cavity and loses energy. From the energy conservation theory, the energy in the signal \mathbf{v}_{Cbs} equals to the energy loss of the bunch in the cavity. The cavity voltage dissipates power in the equivalent resistive load of the cavity, and the total energy consumption can be calculated as

Fig. 3.11 Time domain current of a single bunch



$$\Delta E_{Cbs} = \int_0^\infty |\mathbf{v}_{Cbs}|^2 / (2R_L) dt = \omega_{1/2} R_L Q_b^2.$$

This is the energy loss of a single bunch passing through the cavity. Compared to (3.25), the bunch experiences an equivalent decelerating voltage equal to half of the maximum voltage induced by it. The energy loss of a bunch in a cavity is independent of the cavity detuning. The frequency of the cavity voltage induced by a single bunch equals the resonance frequency of the cavity. The phase slope in Fig. 3.12 comes from the fixed carrier frequency defining the phasors.

3.3.6 Cavity Response to a Bunch Train

An accelerator uses the cavities to achieve the required acceleration of a bunch train. Suppose that a bunch train is injected into the cavity at $t = 0$ with a bunch charge Q_b and a bunch spacing time T_b . To simplify the discussion, we assume an infinite bunch train as shown in Fig. 3.13. Here we have neglected the bunch length and present every single bunch as a δ -function.

The beam current of the bunch train can be described in the time domain as $i_{bt} = Q_b \sum_{n=0}^{\infty} \delta(t - nT_b)$ (Kramer and Wang 1999), where the subscript “ bt ” stands for the cavity drive term induced by the bunch train. Its Laplace transform is $I_{bt}(s) = Q_b \sum_{n=0}^{\infty} e^{-nT_b s}$. From (3.8), the phasor Laplace transform of i_{bt} is $\mathbf{I}_{bt}(\hat{s}) = 2Q_b \sum_{n=0}^{\infty} e^{-nT_b (\hat{s} + j\omega_c)}$. For most accelerators, the bunch repetition rate is a

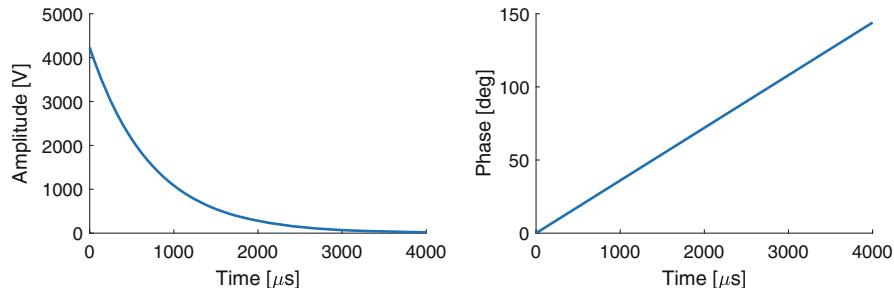
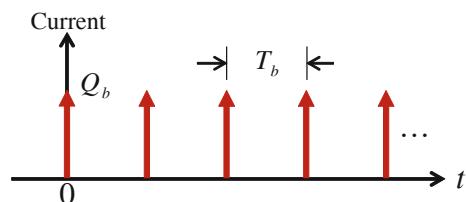


Fig. 3.12 Single-bunch induced cavity voltage in a TESLA cavity. The cavity parameters are the same as the example in Sect. 3.3.4. In addition, the bunch charge is $Q_b = 1$ nC and the cavity detuning is $\Delta\omega = 2\pi \times 100$ rad/s

Fig. 3.13 Time domain current of a bunch train



sub-harmonic of the RF operating frequency ω_{RF} (we assume the carrier frequency $\omega_c = \omega_{RF}$), so that all bunches experience the same RF phase. Therefore, $nT_b\omega_c$ is integer times of 2π and $e^{-jnT_b\omega_c} = 1$. Finally, we obtain

$$\mathbf{I}_{bt}(\hat{s}) = 2Q_b \sum_{n=0}^{\infty} e^{-nT_b\hat{s}}. \quad (3.26)$$

Using the phasor transfer function of the cavity \mathbf{G}_C , we can calculate the resulting cavity voltage as $\mathbf{V}_{Cbt} = \mathbf{G}_C \mathbf{I}_{bt}$ in the \hat{s} domain. Then the time-domain phasor of the cavity voltage induced by the bunch train can be calculated as

$$\mathbf{v}_{Cbt} = 2\omega_{1/2}R_L Q_b \sum_{n=0}^{\infty} e^{-(\omega_{1/2}-j\Delta\omega)(t-nT_b)} u(t - nT_b). \quad (3.27)$$

Here $u(t - nT_b)$ is the unit step function at time nT_b . The bunch train-induced cavity voltage (3.27) is the superposition of the cavity voltages induced by the single bunches at nT_b . The infinite sum in the drive term (3.26) can be simplified by applying the geometric series theory with $\sum_{n=0}^{\infty} e^{-nT_b\hat{s}} = 1/(1 - e^{-T_b\hat{s}})$, resulting in $\mathbf{I}_{bt}(\hat{s}) = 2Q_b/(1 - e^{-T_b\hat{s}})$. The delay term $e^{-T_b\hat{s}}$ can be further simplified using the first-order Padé approximation $e^{-T_b\hat{s}} \approx (-T_b\hat{s} + 2)/(T_b\hat{s} + 2)$ and we get

$$\mathbf{I}_{bt}(\hat{s}) \approx Q_b(T_b\hat{s} + 2)/(T_b\hat{s}) = Q_b + 2I_{b0}/\hat{s}. \quad (3.28)$$

Here we define the average beam current $I_{b0} = Q_b/T_b$. With (3.28), we can calculate the time-domain phasors of the beam current (\mathbf{i}_{bt}) and cavity voltage (\mathbf{v}_{Cbt}) as

$$\begin{aligned} \mathbf{i}_{bt} &\approx 2\mathbf{i}_{b0} + Q_b\delta(t)e^{j\angle\mathbf{i}_{b0}} \\ \mathbf{v}_{Cbt} &\approx \frac{2\omega_{1/2}R_L\mathbf{i}_{b0}}{\omega_{1/2} - j\Delta\omega} \left(1 - e^{-(\omega_{1/2}-j\Delta\omega)t}\right) + \omega_{1/2}R_L Q_b e^{j\angle\mathbf{i}_{b0}} e^{-(\omega_{1/2}-j\Delta\omega)t}. \end{aligned} \quad (3.29)$$

The phasor \mathbf{i}_{b0} is the average beam current with its magnitude equals I_{b0} and its phase denoted as $\angle\mathbf{i}_{b0}$. If \mathbf{i}_{b0} is time varying (e.g., due to fluctuations in bunch charge), we should calculate the cavity response numerically using the differential equation (3.15). In a bunch train, a later bunch will be affected by the cavity voltage induced by the previous bunches. Furthermore, as discussed in the last section, a bunch also experiences an equivalent decelerating voltage induced by itself, which is approximately modeled by the term $Q_b\delta(t)$ in the overall beam current \mathbf{i}_{bt} . Note that in this approximation, we only included the decelerating voltage induced by the first bunch (at $t = 0$), and its effect vanishes after several time-constants of the cavity, see the second term of \mathbf{v}_{Cbt} . In most cases, the magnitude of \mathbf{i}_{b0} is much larger than Q_b , then the term $Q_b\delta(t)$ in (3.29) is typically neglected for further simplification.

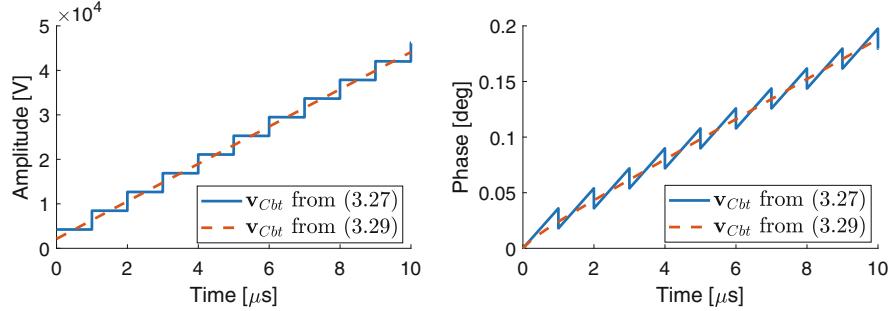


Fig. 3.14 Bunch train induced cavity voltage in a TESLA cavity (zoomed in for the first 10 bunches). The cavity and bunch parameters are the same as the example in Fig. 3.12 and the bunch repetition rate is assumed to be 1 MHz, which means, $T_b = 1 \mu\text{s}$

Figure 3.14 shows an example of the response to a bunch train of a TESLA cavity. The envelope of the beam induced cavity voltage is plotted according to (3.27) and its approximation (3.29).

In principle, (3.27) offers a more accurate estimate of the cavity voltage induced by the bunch train. As seen in Fig. 3.14, the single-bunch induced transients are visible in the cavity voltage. In practice, if the RF detector has enough signal-to-noise-ratio (SNR), the single-bunch induced transients may be detectable from the measurement of the cavity voltage. For example, if the cavity voltage is 25 MV, to detect the single-bunch transients in Fig. 3.12, we require the SNR of the RF detector to be better than 75 dB.

The analytical results above are only valid for a regular bunch train with uniform Q_b and T_b . If either Q_b or T_b is time varying, it is not easy to get an analytical description of the resulting cavity voltage. In this case, the cavity voltage may be calculated by solving the differential equation (3.15) numerically using the time-varying \mathbf{i}_{bt} . Nevertheless, the analysis in this section is still very useful to understand the essential behavior of the cavity when a bunch train passes through.

3.3.7 Cavity Equation with Voltage Drives

Suppose that we use an RF cavity driven by RF power to accelerate a bunch train. The drive terms induced by the RF power (\mathbf{i}_{rf}) and by the bunch train (\mathbf{i}_{bt}) are given by (3.22) and (3.29), respectively. We will neglect the minor term $Q_b\delta(t)$ in \mathbf{i}_{bt} in the following discussions; that is, $\mathbf{i}_{bt} \approx 2\mathbf{i}_0$. Then the cavity differential equation (3.15) can be rewritten as

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = 2\omega_{1/2}R_L(\mathbf{i}_{for} + \mathbf{i}_{b0}). \quad (3.30)$$

We assume that both \mathbf{i}_{for} and \mathbf{i}_0 are turned on at $t = 0$. The drive terms of the cavity equation may be transformed into voltages if the input-output relations in terms of voltages are of interest. First, we define the beam drive voltage as $\mathbf{v}_{b0} := R_L \mathbf{i}_{b0}$. With the formulas in (3.21) and Table 3.1, we obtain the relation between \mathbf{i}_{for} and the RF drive voltage \mathbf{v}_{for} as $\mathbf{i}_{for} = \beta/(\beta + 1) \cdot \mathbf{v}_{for}/R_L$. Then (3.30) becomes

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = 2\omega_{1/2} \left(\frac{\beta}{\beta + 1} \mathbf{v}_{for} + \mathbf{v}_{b0} \right). \quad (3.31)$$

For superconducting cavities, the input coupling factor $\beta \gg 1$ because the RF power dissipation in the cavity wall is much smaller than the RF power emission out of the cavity. For example, the TESLA cavity operates with $\beta > 3000$, then (3.31) can be simplified by neglecting the term $\beta/(\beta + 1)$.

When there is no beam loading, the steady-state cavity drive \mathbf{v}_{for} , reflection \mathbf{v}_{ref} and voltage \mathbf{v}_C satisfy the following relations according to (3.21) and (3.31):

$$\begin{aligned} \mathbf{g}_C &:= \frac{\mathbf{v}_C}{\mathbf{v}_{for}} = \frac{2\omega_{1/2}}{\omega_{1/2} - j\Delta\omega} \frac{\beta}{\beta + 1} \\ \boldsymbol{\gamma}_C &:= \frac{\mathbf{v}_{ref}}{\mathbf{v}_{for}} = \frac{\mathbf{v}_C - \mathbf{v}_{for}}{\mathbf{v}_{for}} = \mathbf{g}_C - 1. \end{aligned} \quad (3.32)$$

The complex number \mathbf{g}_C is the gain between the steady-state \mathbf{v}_C and the corresponding \mathbf{v}_{for} . We define $\boldsymbol{\gamma}_C$ as a *complex reflection factor* of the cavity, where the cavity is treated as a reactive load of the RF power transmission line. Both \mathbf{g}_C and $\boldsymbol{\gamma}_C$ appear as circles in the complex plane as functions of β and $\Delta\omega$. See Fig. 3.15. Here we only show \mathbf{g}_C , and $\boldsymbol{\gamma}_C$ can be derived by shifting the circles of \mathbf{g}_C to the left by 1. Note that when $\Delta\omega$ approaches infinity, the \mathbf{g}_C circles always cross the complex plane's origin, whereas the $\boldsymbol{\gamma}_C$ circles always cross the $(-1, 0)$ point. It implies that with a large detuning, all the forward RF power is reflected from the cavity input coupler, and the phase of the reflected signal is offset by 180° from the phase of the forward signal.

Fig. 3.15 Complex gain \mathbf{g}_C for different coupling factors and detuning

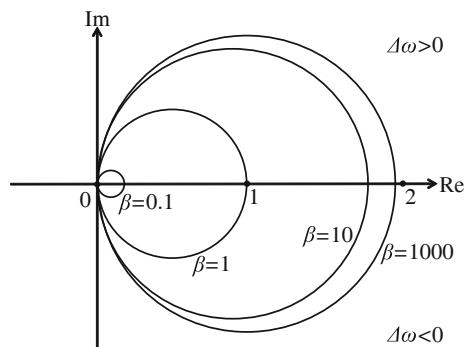
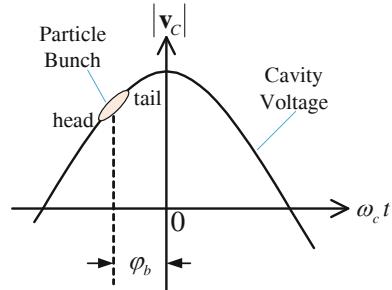


Fig. 3.16 Time dependent cavity voltage and interaction with a particle bunch



3.3.8 Interaction Between Cavity Voltage and Beam

The ultimate purpose of a cavity is to accelerate particle beams using the RF field. The RF field experienced by the beam is a superposition of the contributions from the RF power and the beam loading. If we inject a bunch, which contains many particles, into the cavity, different particles will experience different accelerations, depending on the location of the particle in the bunch and the time when the bunch arrives. The cavity voltage oscillates at a frequency the same as the RF frequency. The time-dependent cavity voltage is shown in Fig. 3.16.

If we inject a bunch into the cavity when the cavity voltage is ramping up (at the left part of the curve in Fig. 3.16), the head of the bunch will experience less acceleration than the tail. We define a *beam phase*, denoted as φ_b , to represent the relative time between the on-crest cavity voltage and the bunch arrival time. In this book, we define $\varphi_b = 0$ when the center of the bunch obtains maximum acceleration, which is called “on-crest” acceleration. In the particular case in Fig. 3.16, $\varphi_b < 0$. We can calculate φ_b using the phases of the beam current and cavity voltage

$$\varphi_b = \pi + (\angle v_C - \angle i_{b0}). \quad (3.33)$$

Then the energy gain of the particle at the center of the bunch can be calculated as $\Delta E = |v_C| \cos \varphi_b$. Note that the definition of (3.33) is typically used in linear accelerators, whereas for storage rings, the on-crest phase is often defined as 90°. First, let us study the steady-state cavity voltage when both the RF drive power and the beam loading are present. From the cavity Eq. (3.30), we can present the steady-state cavity voltage as

$$v_C = 2R_L(i_{for} + i_{b0}) \cos \psi \cdot e^{j\psi}, \quad (3.34)$$

where ψ is the detuning angle defined in (3.19). The relations between the cavity input/output phasors are depicted in Fig. 3.17. It can be seen that the beam phase is determined by the relative phase between the overall cavity voltage and the beam drive current, as calculated by (3.33). Equation (3.34) shows that to accelerate a specific CW beam i_{b0} with a desired steady-state voltage v_C , the power and phase of

Fig. 3.17 Relations between cavity drive and cavity voltage phasors in the steady state

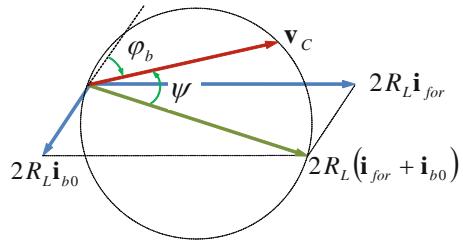
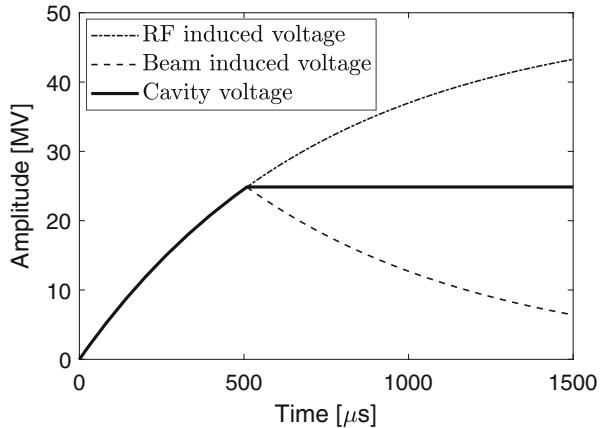


Fig. 3.18 Flattop in RF pulse generated by the constant RF drive power and beam loading



the required RF drive \mathbf{i}_{for} depend on the bandwidth $\omega_{1/2}$ and detuning $\Delta\omega$ of the cavity. Besides, we may choose proper values of the loaded quality factor Q_L and $\Delta\omega$ to minimize the required RF power for achieving the desired cavity voltage (see next section).

For a cavity operating in the pulsed mode, the beam may be injected before the cavity enters a steady state. This happens quite often for superconducting cavities with a large time-constant. Before the beam is injected, the cavity voltage increases exponentially with a constant RF drive power turned on at $t = 0$. Assume that at $t = T_f$, we inject the beam into the cavity, then the cavity voltage is also affected by the beam loading. With a specific T_f , the increase of the cavity voltage generated by the RF drive power can be compensated by the voltage induced by the beam, resulting in a flattop in the RF pulse (see Fig. 3.18).

To determine the value of T_f to achieve a flattop, we write the cavity voltage after injecting the beam ($t > T_f$) as

$$\mathbf{v}_C|_{t>T_f} = \frac{2\omega_{1/2}R_L}{\omega_{1/2} - j\Delta\omega} \left[\mathbf{i}_{for} \left(1 - e^{-(\omega_{1/2} - j\Delta\omega)t} \right) + \mathbf{i}_{b0} \left(1 - e^{-(\omega_{1/2} - j\Delta\omega)(t-T_f)} \right) \right].$$

As discussed before, \mathbf{v}_C should be a constant for a flattop after the beam injection, which requires the derivative of \mathbf{v}_C to be zero, demanding $\mathbf{i}_{for}e^{-(\omega_{1/2} - j\Delta\omega)T_f} = -\mathbf{i}_{b0}$. Therefore, T_f should satisfy

$$T_f = \ln(|\mathbf{i}_{for}| / |\mathbf{i}_{b0}|) / \omega_{1/2}, \quad (3.35)$$

and the cavity voltage at the flattop can be written as

$$\mathbf{v}_{C,FT} = \frac{2\omega_{1/2}R_L(\mathbf{i}_{for} + \mathbf{i}_{b0})}{\omega_{1/2} - j\Delta\omega},$$

which is consistent to the steady-state solution of the cavity Eq. (3.30). It can be seen that the cavity voltage jumps from a transient state immediately to a steady state (flattop) due to the beam injection. As an example, to accelerate a beam $i_{b0} = |\mathbf{i}_{b0}| = 8$ mA with a flattop voltage $v_{C0} = |\mathbf{v}_{C,FT}| = 25$ MV, a TESLA cavity requires an RF drive power $P_{for} = 200$ kW. According to (3.23), we estimate that $|\mathbf{i}_{for}| = 16$ mA. The cavity has a $Q_L = 3 \times 10^6$ and operates at a frequency $f_0 = 1.3$ GHz, then we have $\omega_{1/2} = 1361.4$ rad/s. Using (3.35), we can estimate the optimum beam injection time as $T_f \approx 510$ μ s to obtain a flattop in the RF pulse while keeping the RF drive power constant. Note that it is not always possible to achieve a flattop only with the beam loading (e.g., in case of small beam currents) without introducing additional steps in the RF drive power.

A larger T_f implies that we have more time to fill the cavity, which reduces the required peak RF drive power. However, in some cases, the largest RF pulse width is limited by the RF amplifier, and we need to fill the cavity in a shorter period with higher peak power. After filling the cavity to the desired voltage, the flattop for beam acceleration needs to be generated by both the beam loading and a step in the RF drive power. Let us use an example to demonstrate the flattop generation for a cavity operating in the pulsed mode.

Example 3.1 RF Feedforward for Flattop Generation

Consider a TESLA-type superconducting cavity with the following parameters: $f_0 = 1.3$ GHz, $r/Q = 1036$ Ω and $Q_L = 3 \times 10^6$. The cavity operates in the pulsed mode and the desired cavity voltage at the flattop is $v_{C0} = 25$ MV. We want to accelerate a beam with an average current $i_{b0} = 3$ mA and a beam phase $\varphi_b = -30^\circ$ during the flattop. The cavity is detuned by $\Delta\omega = 2\pi \times 100$ rad/s and the beam is injected at $T_f = 510$ μ s after the RF power is turned on. What are the required RF drives during the filling stage and the flattop stage?

From the formulas in Table 3.1, the half bandwidth and the loaded resistance of the cavity can be calculated as $\omega_{1/2} = 1361.3$ rad/s and $R_L = 1554$ M Ω , respectively. During the filling stage, the cavity voltage is built up with the RF drive power. To reach v_{C0} before injecting the beam at T_f , the RF drive current $\mathbf{i}_{for,fill}$ should satisfy

$$v_{C0} = \frac{2\omega_{1/2}R_L\mathbf{i}_{for,fill}}{\omega_{1/2} - j\Delta\omega} \left(1 - e^{-(\omega_{1/2} - j\Delta\omega)T_f} \right).$$

Here we have normalized the phase of the cavity voltage to zero. With the parameters above, we can calculate the amplitude and phase of the RF drive current as $|i_{for,fill}| = 16.1 \text{ mA}$ and $\angle i_{for,fill} = -8.1^\circ$, respectively. This corresponds to 202.3 kW RF drive power according to (3.23). Since the cavity voltage phase is normalized to zero, we obtain $\varphi_b = \pi - \angle i_{b0}$ from (3.33), and the beam current phasor can be presented as

$$i_{b0} = i_{b0} e^{j\angle i_{b0}} = i_{b0} e^{j(\pi - \varphi_b)}.$$

From the discussions when deriving (3.35), if we want to create a flattop only by the beam loading while keeping the RF drive current constant over the entire RF pulse, i.e., $|i_{for,flat}| = |i_{for,fill}|$, we need a beam current i'_{b0} that must satisfy

$$i'_{b0} = -i_{for,fill} e^{-(\omega_{1/2} - j\Delta\omega)T_f}.$$

With the given $T_f = 510 \mu\text{s}$ and the $i_{for,fill}$ calculated above, we require $|i'_{b0}| = 8 \text{ mA}$. But the actual beam current is $|i_{b0}| = 3 \text{ mA}$, so the RF drive current at the flattop $|i_{for,flat}|$ must be reduced by about 5 mA compared to $|i_{for,fill}|$. The difference between $i_{for,flat}$ and $i_{for,fill}$ is equivalent to a virtual beam Δi_{b0} satisfying $i'_{b0} = i_{b0} + \Delta i_{b0}$. Therefore, the required RF drive current during the flattop can be calculated as

$$i_{for,flat} = i_{for,fill} + \Delta i_{b0} = i_{for,fill} + i'_{b0} - i_{b0}.$$

Then we can calculate the amplitude and phase of the RF drive current at the flattop as $|i_{for,flat}| = 10.9 \text{ mA}$ and $\angle i_{for,flat} = -11.7^\circ$, respectively. This corresponds to 91.8 kW RF drive power. With the results above, we plot the amplitude waveforms of the cavity voltage, forward and reflected signals, as well as the waveform of the beam phase in Fig. 3.19. Here we have illustrated the forward and reflected signals as voltages referring to the same reference plane as the cavity voltage.

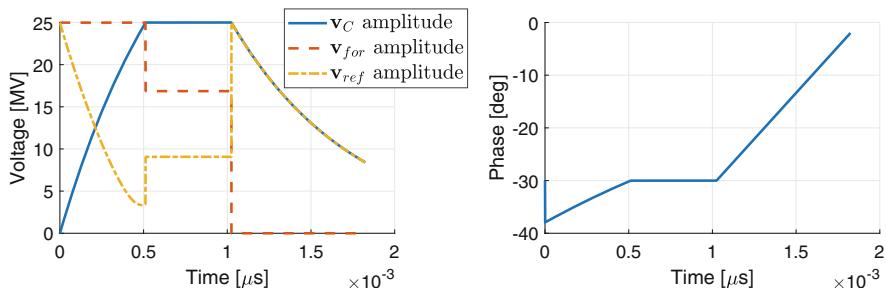


Fig. 3.19 Waveforms of the cavity voltage, forward and reflected signals (left) and the beam phase during the RF pulse (right)

3.3.9 Forward and Reflected RF Power

For RF amplifiers or klystrons, the required RF drive power and the reflected RF power from the cavity are the most relevant parameters.

When a cavity operates in the CW mode or the pulsed mode but at the flattops of RF pulses, the cavity voltage is in the steady state, and the forward and reflected RF powers can be derived from (3.34). Suppose that the steady-state cavity voltage is v_{C0} and a beam i_{b0} is accelerated at a beam phase φ_b . Using the cavity voltage as a reference with its phase normalized to zero, that is, $\mathbf{v}_C = v_{C0}$, we can simplify the beam current phasor as $\mathbf{i}_{b0} = -i_{b0}e^{-j\varphi_b}$ according to (3.33). Then from the steady-state cavity voltage in (3.34), the required RF drive current will be

$$\mathbf{i}_{for} = \frac{v_{C0}}{2R_L \cos \psi \cdot e^{j\psi}} + i_{b0}e^{-j\varphi_b}. \quad (3.36)$$

With the relation between \mathbf{v}_{for} and \mathbf{i}_{for} as in (3.31), the forward RF power (i.e., RF drive power of cavity) is calculated as $P_{for} = \mathbf{v}_{for}\mathbf{i}_{for}^*/2 = R_L(\beta+1)|\mathbf{i}_{for}|^2/(2\beta)$. Substitute \mathbf{i}_{for} with (3.36), we can rewrite the steady-state forward RF power as

$$P_{for} = \frac{\beta+1}{\beta} \frac{v_{C0}^2}{8R_L} \left[\left(1 + \frac{2R_L i_{b0} \cos \varphi_b}{v_{C0}} \right)^2 + \left(\tan \psi + \frac{2R_L i_{b0} \sin \varphi_b}{v_{C0}} \right)^2 \right]. \quad (3.37)$$

Similarly, the reflected RF power in the steady state can be calculated as

$$P_{ref} = \frac{\beta+1}{\beta} \frac{v_{C0}^2}{8R_L} \times \left[\left(\frac{\beta-1}{\beta+1} - \frac{2R_L i_{b0} \cos \varphi_b}{v_{C0}} \right)^2 + \left(\tan \psi + \frac{2R_L i_{b0} \sin \varphi_b}{v_{C0}} \right)^2 \right]. \quad (3.38)$$

As mentioned before, we can optimize the cavity parameters (e.g., input coupling factor β , loaded quality factor Q_L and detuning $\Delta\omega$) to minimize P_{ref} (and the required P_{for}). When $P_{ref} = 0$ in the steady state, the cavity appears as a matched load of the RF transmission line. To achieve this *matched case*, the two square terms in (3.38) should be zero:

$$\frac{\beta-1}{\beta+1} - \frac{2R_L i_{b0} \cos \varphi_b}{v_{C0}} = 0, \tan \psi + \frac{2R_L i_{b0} \sin \varphi_b}{v_{C0}} = 0. \quad (3.39)$$

The optimal $\Delta\omega$ for the matched cased can be derived from the second equality and the formulas in Table 3.1 as

$$\Delta\omega_{opt} = -\frac{\omega_0(r/Q)i_{b0}\sin\varphi_b}{2v_{C0}}. \quad (3.40)$$

It implies that the cavity should be detuned if the beam is accelerated off-crest if aiming for minimum reflected RF power. To obtain the optimal Q_L , we will deal with (3.39) differently for superconducting cavities and normal-conducting cavities:

1. *Superconducting cavity*: β is much larger than 1, and the first equality in (3.39) can be approximated to $v_{C0} = 2R_L i_{b0} \cos\varphi_b$. Referring to Table 3.1, we can obtain the optimal Q_L as

$$Q_{L,opt} = \frac{v_{C0}}{(r/Q)i_{b0} \cos\varphi_b}. \quad (3.41)$$

2. *Normal-conducting cavity*: β is close to 1, and we can determine the optimal β first and then the optimal Q_L :

$$\begin{aligned} \beta_{opt} &= (r/Q)Q_0 i_{b0} \cos\varphi_b / v_{C0} + 1 \\ Q_{L,opt} &= Q_0 / (\beta_{opt} + 1). \end{aligned} \quad (3.42)$$

With the optimal parameters, the matched case can be achieved, in which all the forward RF power is consumed by the beam and the cavity wall resistance. Note that the matched case cannot always be achieved due to the limitations in the input coupler and frequency tuner of the cavity. Tradeoffs should typically be made between the cavity parameters and the allowed reflected RF power. For example, when i_{b0} is very small, the resulting $Q_{L,opt}$ calculated by (3.41) might be too large to be realized by adjusting the cavity input coupler. Therefore we have to choose a lower Q_L and allow some reflected RF power.

In general, the difference between the forward and reflected RF powers is the power dissipation in the cavity

$$P_{diss} = P_{for} - P_{ref} = \frac{v_{C0}^2}{2R} + v_{C0} i_{b0} \cos\varphi_b.$$

Here the relation $R = R_L(\beta + 1)$ has been applied. The first term is the power dissipation in the cavity wall and the second term the power delivered to the beam. For a superconducting cavity, the resistance R is usually very large ($>10^{12}$), so the power dissipation in the cavity wall is very small, and almost all of the RF power dissipated in the cavity is delivered to the beam. This also means that if there is no beam in the cavity, nearly all of the forward RF power will be reflected. When a cavity operates in the pulsed mode, a *filling stage* appears in the RF pulse to build up the cavity voltage before injecting the beam. The forward and reflected RF powers within an RF pulse have the typical shapes in Fig. 3.20.

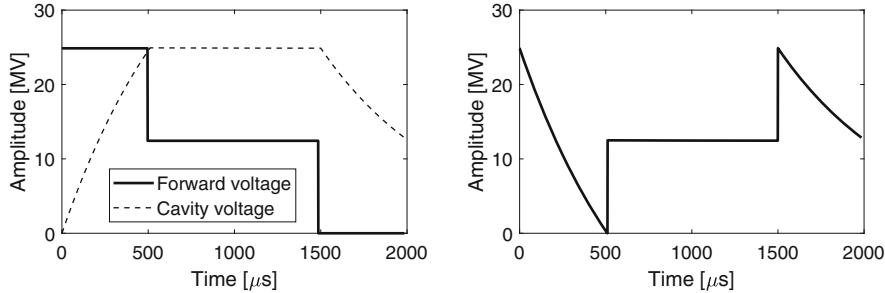


Fig. 3.20 Forward (left) and reflected (right) RF powers within an RF pulse

The relation between the forward voltage v_{for} , the reflected voltage v_{ref} and the cavity voltage v_C is given by (3.21). At the starting time ($t = 0$) of an RF pulse, all of the forward RF power is reflected back because $v_C(0) = 0$ and we have $v_{ref}(0) = -v_{for}(0)$ since $v_C = v_{for} + v_{ref}$. With the increasing v_C , v_{ref} decreases because $v_{ref} = v_C - v_{for}$, and if the cavity is not detuned, the phases of v_C and v_{for} are the same. We may explain the reduction of v_{ref} in another way. The overall v_{ref} consists of two parts, $v_{ref} = v_{ref1} + v_{ref2}$, where v_{ref1} is the full reflection of the forward voltage with an opposite phase ($v_{ref1} = -v_{for}$), whereas v_{ref2} corresponds to the RF power emitted from the cavity back into the transmission line. Since v_{ref2} is proportional to v_C that has the same phase as v_{for} , v_{ref2} will cancel part of v_{ref1} , resulting in a decreasing v_{ref} . During the *flat-top stage* of the RF pulse, the cavity is in a steady state, so (3.37) and (3.38) can be used to calculate the forward and reflected RF powers. After the forward RF power is turned off, the cavity voltage decays following the cavity time constant, and the reflected RF power only contains the RF power emitted from the cavity. This part of the RF pulse, after turning off the forward RF power, is denoted as the *decay stage*.

During the filling stage, the required forward RF power to fill the cavity to a desired voltage v_{C0} within a filling time T_f can be calculated as

$$P_{fill} = \frac{\beta + 1}{\beta} \frac{v_{C0}^2}{8R_L} \frac{1}{\cos^2 \psi [1 + e^{-2\omega_{1/2} T_f} - 2e^{-\omega_{1/2} T_f} \cos(\Delta\omega T_f)]}. \quad (3.43)$$

It shows that the required filling power is determined by the desired cavity voltage v_{C0} , filling time T_f , loaded quality factor Q_L and detuning $\Delta\omega$. If we choose proper values of T_f and Q_L , we may reduce the required P_{fill} . Reducing the required peak power helps avoid saturating the amplifiers or klystrons and is critical for a practical RF system. The determination of T_f and Q_L usually needs a tradeoff between the output power and the RF pulse width available from the RF amplifiers or klystrons. For example, if we need a smaller P_{fill} for some reasons (e.g., the power limit of the cavity input coupler), we may increase T_f and select a proper Q_L . Figure 3.21 shows an example of the required P_{fill} for different Q_L and T_f , from which we can determine the T_f and Q_L values taking into account the available RF power and pulse width.

Fig. 3.21 Filling power for different loaded quality factors and filling times

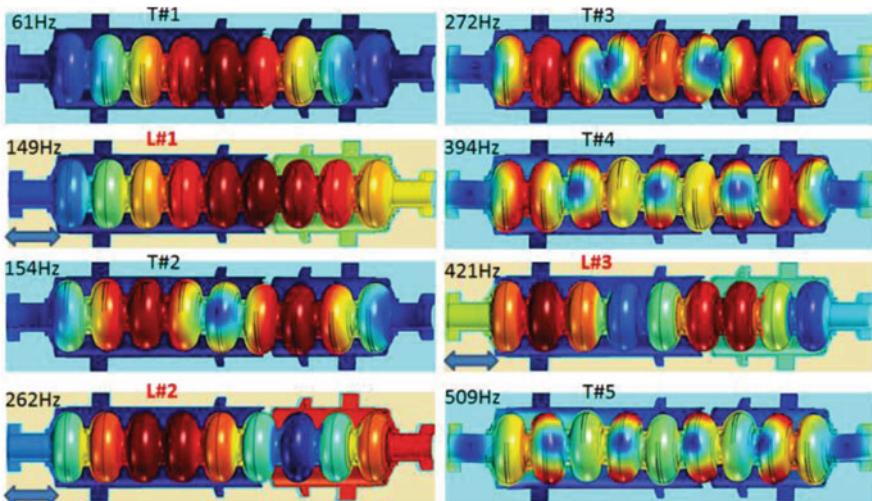
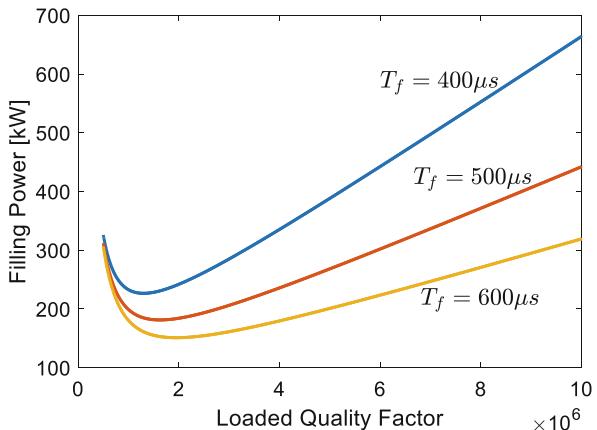


Fig. 3.22 Frequencies and patterns of the eight lowest mechanical modes of a TESLA cavity. (Courtesy of N. Solyak et al.)

3.3.10 Mechanical Model

The electrical and mechanical aspects of a cavity are coupled through the Lorenz force applied to the cavity wall induced by the RF field. Superconducting cavities suffer more severely from mechanical oscillations due to the thinner walls than normal-conducting cavities. The mechanical oscillations will change the cavity geometry and shift its electrical resonance frequency.

A cavity has multiple mechanical modes at different frequencies with various patterns of oscillation. As an example, Fig. 3.22 shows the patterns of the eight lowest mechanical modes of a TESLA cavity (Solyak et al. 2013).

The cavity mechanical model describes the cavity voltage as an input and the resulting detuning as the output. The differential equation of the m th mechanical mode can be written as

$$\frac{d^2\Delta\omega_m}{dt^2} + \frac{\omega_m}{Q_m} \frac{d\Delta\omega_m}{dt} + \omega_m^2 \Delta\omega_m = -K_m \omega_m^2 v_C^2, \quad (3.44)$$

where $v_C = |\mathbf{v}_C|$ is the magnitude of the cavity voltage and $\Delta\omega_m$ is the time-varying detuning caused by the m th mechanical mode. A mechanical mode is characterized by the parameters ω_m , Q_m and K_m that are the frequency, quality factor and Lorenz force detuning constant, respectively. In the steady state, we have $\Delta\omega_m = -K_m v_C^2$, which is consistent with the Lorenz force Eq. (1.15) introduced in Chap. 1. Note that the mechanical model is driven by the square of the cavity voltage, so it is a nonlinear model. Taking into account all the mechanical modes, we can calculate the overall time-varying detuning caused by the mechanical oscillations as

$$\Delta\omega_{mech}(t) = \sum_{m=1}^M \Delta\omega_m(t). \quad (3.45)$$

Here we have included M most significant mechanical modes under the working condition of the cavity.

3.4 Multi-cell Cavity Model

The multi-cell cavity model can be derived in a similar approach as for a single-cell cavity. It helps us understand the behavior of multi-cell cavities and is useful for feedback stability analysis. As shown later, a multi-cell cavity has multiple eigenmodes (passband modes), and each of them can be described with a phasor transfer function similar to that of a single-cell cavity. In particular, we may equivalent a multi-cell cavity to a single-cell cavity if we only consider one passband mode (e.g., the π -mode of a TESLA cavity for beam acceleration). As an example of multi-cell cavities, a TESLA type 9-cell cavity is depicted in Fig. 3.23 (Aune et al. 2000).

3.4.1 Coupled RLC Circuit Model

Similar to single-cell cavities, a multi-cell cavity can be modeled as coupled RLC circuits (Sekutowicz 1994; Liepe 2001; Vogel 2007; Hall 2011). See Fig. 3.24. Here we have used serial RLC circuits, where the RF power source is represented by a voltage source and the input coupler by a lossless transformer. Each cell of the

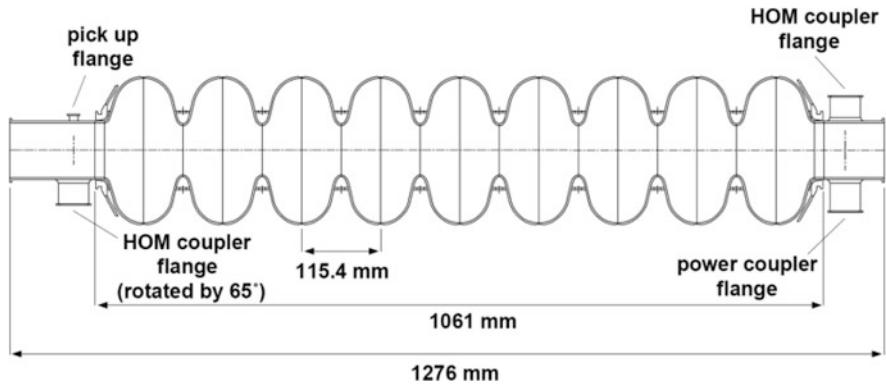


Fig. 3.23 Side view of a 9-cell TESLA cavity with a main power coupler and two higher order mode (HOM) couplers. (Courtesy of B. Aune et al.)

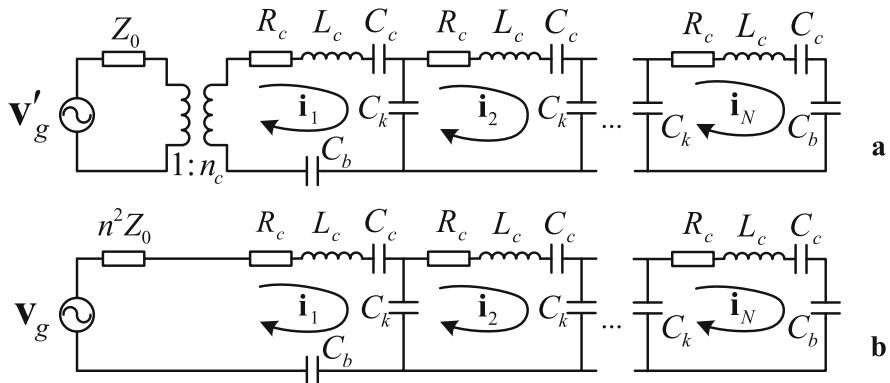


Fig. 3.24 Coupled circuit model for multi-cell cavities. Compared to the plot **a**, the plot **b** has converted the RF power source and the transmission line impedance to the cavity side

cavity is modeled as a serial RLC circuit coupled via a capacitor C_k . We assume all the cells are identical to simplify the discussion. For the two end cells, we use a capacitance C_b to include the effects of the beam pipes.

Similar to Table 3.1, the relations between the multi-cell cavity parameters and the circuit parameters (labeled with foot script “*c*” for cell) are summarized in Table 3.3.

In the table above, W_c is the stored electromagnetic energy in a cell, P_{cell} is the dissipated RF power in the cavity wall in a cell, and P_{ext} is the dissipated RF power in the external impedance, which corresponds to the RF power emitted out of the cavity through the input coupler.

Table 3.3 Relations between the multi-cell cavity parameters and the parameters of the serial RLC circuits

Cavity parameter	Notation	Relation with circuit parameters
Resonance frequency of single cell	ω_{0c}	$\omega_{0c} = 1/\sqrt{L_c C_c}$
Unloaded quality factor of single cell	Q_{0c}	$Q_{0c} = \frac{\omega_{0c} W_c}{P_{cell}} = \frac{\omega_{0c} L_c}{R_c} = \frac{1}{\omega_{0c} R_c C_c}$
External quality factor of the first cell	$Q_{ext,c}$	$Q_{ext,c} = \frac{\omega_{0c} W_c}{P_{ext}} = \frac{\omega_{0c} L_c}{n^2 Z_0}$
Input coupling factor of the first cell	β_c	$\beta_c = \frac{P_{ext}}{P_{cell}} = \frac{Q_{0c}}{Q_{ext,c}} = \frac{n_c^2 Z_0}{R_c}$
Loaded quality factor of the first cell	Q_{Lc}	$Q_{Lc} = \frac{Q_{0c}}{1+\beta_c}$
Loaded resistance of the first cell	R_{Lc}	$R_{Lc} = n_c^2 Z_0 + R_c = (1 + \beta_c) R_c$
Cell-to-cell coupling factor	κ	$\kappa = 2C_c/C_k$
End cell-beam tube coupling factor	γ	$\gamma = C_c/C_b$

3.4.2 Multi-cell Cavity Phasor Equations

The phasor Laplace transform method can be used to analyze the coupled circuit in Fig. 3.24b. When applying Kirchhoff's voltage rule to each current loop, we can write the s -domain circuit equations as (3.46), where $V_g(s)$ is the Laplace transform of the drive voltage, and $I_n(s)$ is the Laplace transform of the current in the n th cell ($n = 1, \dots, N$). The three equations are for the first cell equipped with an input coupler, the middle cells, and the last cell, respectively.

$$\begin{aligned} & \left((\beta_c + 1)R_c + sL_c + \frac{1}{sC_c} \right) I_1 + \frac{1}{sC_k} (I_1 - I_2) + \frac{1}{sC_b} I_1 = V_g, \\ & \frac{1}{sC_k} (I_n - I_{n-1}) + \left(R_c + sL_c + \frac{1}{sC_c} \right) I_n + \frac{1}{sC_k} (I_n - I_{n+1}) = 0, \quad 1 < n < N \\ & \frac{1}{sC_k} (I_N - I_{N-1}) + \left(R_c + sL_c + \frac{1}{sC_c} \right) I_N + \frac{1}{sC_b} I_N = 0. \end{aligned} \tag{3.46}$$

Apply the relations in Table 3.3 and the equations can be rewritten as

$$\begin{aligned} & (s^2 + 2\omega_{1/2,L}s + \omega_{e0}^2) I_1 - \frac{\kappa\omega_{0c}^2}{2} I_2 = 2s\omega_{1/2,L}V_g/R_{Lc}, \\ & (s^2 + 2\omega_{1/2,0}s + \omega_{m0}^2) I_n - \frac{\kappa\omega_{0c}^2}{2} I_{n-1} - \frac{\kappa\omega_{0c}^2}{2} I_{n+1} = 0, \quad 1 < n < N \\ & (s^2 + 2\omega_{1/2,0}s + \omega_{e0}^2) I_N - \frac{\kappa\omega_{0c}^2}{2} I_{N-1} = 0. \end{aligned} \tag{3.47}$$

Here the following parameters have been defined:

1. Effective resonance frequency of the end-cells (the first and the last cell):
 $\omega_{e0} := \omega_{0c}\sqrt{1 + \kappa/2 + \gamma}.$
2. Effective resonance frequency of the middle cells: $\omega_{m0} := \omega_{0c}\sqrt{1 + \kappa}.$
3. Half-bandwidth of the first cell with an input coupler: $\omega_{1/2,L} := \omega_{0c}/(2Q_{Lc}).$
4. Half-bandwidth of other cells: $\omega_{1/2,0} := \omega_{0c}/(2Q_{0c}).$

From (3.47), we can draw an initial conclusion that the multi-cell cavity has multiple passbands, each corresponding to a resonance mode at frequencies around ω_{e0} and ω_{m0} . Since typically $\kappa \ll 1$ and $\gamma \ll 1$, we have $\omega_{e0} \approx \omega_{m0} \approx \omega_{0c}$. Furthermore, the quality factors Q_{Lc} and Q_{0c} are much larger than 1, so we assume that $\omega_{1/2,L} \ll \omega_{0c}$ and $\omega_{1/2,0} \ll \omega_{0c}$ in our following discussions.

We will derive the phasor transfer functions of the multi-cell cavity by converting (3.47) into the \hat{s} domain, for which we choose a carrier frequency $\omega_c = \omega_{RF} \approx \omega_{0c}$ to define the phasors. From (3.8), the phasor Laplace transforms of the input V_g and outputs I_n ($n = 1, \dots, N$) of the multi-cell cavity can be written as

$$\mathbf{V}_g(\hat{s}) = 2V_g(\hat{s} + j\omega_c), \quad \mathbf{I}_n(\hat{s}) = 2I_n(\hat{s} + j\omega_c), \quad n = 1, \dots, N..$$

To derive the phasor transfer functions of the multi-cell cavity, we substitute the variable s in (3.47) with $\hat{s} + j\omega_c$, resulting in a group of functions of \hat{s} . In the \hat{s} domain, the cavity passbands are shifted to two different ranges of frequencies, one is around 0 ($\omega_{e0} - \omega_c \approx 0$ and $\omega_{m0} - \omega_c \approx 0$) and the other is around $2\omega_{0c}$ ($\omega_{e0} + \omega_c \approx 2\omega_{0c}$ and $\omega_{m0} + \omega_c \approx 2\omega_{0c}$). The relations between the envelopes of the inputs and outputs of the multi-cell cavity are more interested; that is, we focus on the frequencies with $|\hat{s}| \ll \omega_{0c}$. Therefore, we neglect the passbands at higher frequencies around $2\omega_{0c}$ and obtain the simplified equations of the multi-cell cavity:

$$\begin{aligned} (\hat{s} + \omega_{1/2,L} - j\Delta\omega_e) \cdot \mathbf{I}_1 - m\mathbf{I}_2 &= \omega_{1/2,L}\mathbf{V}_g/R_L, \\ (\hat{s} + \omega_{1/2,0} - j\Delta\omega_m) \cdot \mathbf{I}_n - m\mathbf{I}_{n-1} - m\mathbf{I}_{n+1} &= 0, \quad 1 < n < N, \\ (\hat{s} + \omega_{1/2,0} - j\Delta\omega_e) \cdot \mathbf{I}_N - m\mathbf{I}_{N-1} &= 0. \end{aligned} \quad (3.48)$$

Here we have defined $\Delta\omega_e := \omega_{e0} - \omega_c$, $\Delta\omega_m := \omega_{m0} - \omega_c$ and $m := -jk\omega_{0c}^2/(4\omega_c)$. Equation (3.48) can be rewritten as a time-domain phasor state-space equation

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}. \quad (3.49)$$

By defining $a := -\omega_{1/2,L} + j\Delta\omega_e$, $b := -\omega_{1/2,0} + j\Delta\omega_m$, $c := -\omega_{1/2,0} + j\Delta\omega_e$ and \mathbf{I}_N a N -by- N unit matrix, it can be elaborated as

$$\mathbf{x} = [\mathbf{i}_1 \quad \mathbf{i}_2 \quad \cdots \quad \mathbf{i}_n \quad \cdots \quad \mathbf{i}_{N-1} \quad \mathbf{i}_N]^T,$$

$$\mathbf{u} = [\mathbf{v}_g \quad 0 \quad \cdots \quad 0 \quad \cdots \quad 0 \quad 0]^T,$$

$$\mathbf{A} = \begin{bmatrix} a & m & & & & \\ m & b & m & & & \\ \ddots & \ddots & \ddots & & & \\ & m & b & m & & \\ & & \ddots & \ddots & \ddots & \\ & & & m & b & m \\ & & & m & c & \end{bmatrix}, \mathbf{B} = \frac{\omega_{1/2,L}}{R_L} \mathbf{I}_N. \quad (3.50)$$

Note that (3.50) is for the cavities with an input coupler installed in the first cell. If the input coupler is installed in one of the middle cells, similar techniques can be used to derive the phasor equation with the same form but a different matrix \mathbf{A} .

3.4.3 Passband Modes

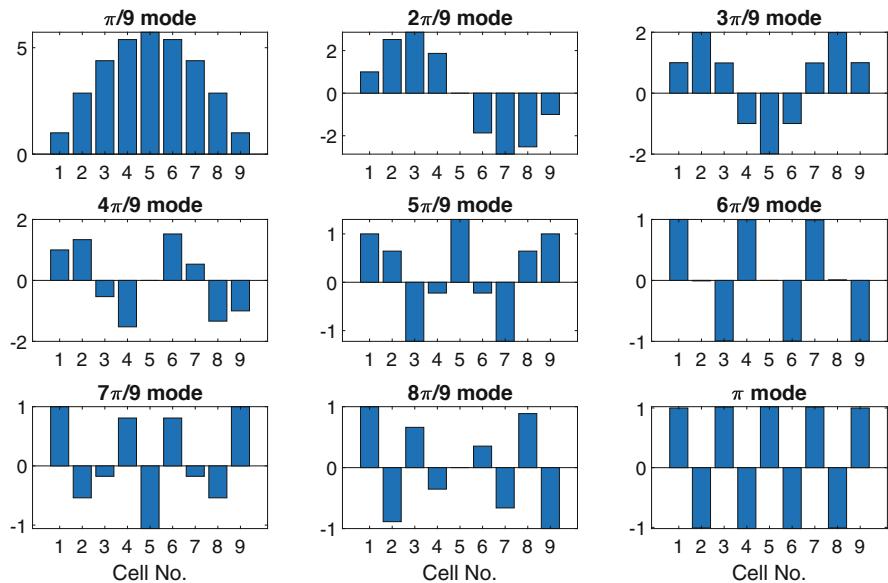
The frequency response of a multi-cell cavity consists of multiple passband modes at various frequencies and with different cell-to-cell field distributions. We can calculate the passband modes as the poles of the phasor state-space equation (3.49). The eigenvalues of the matrix \mathbf{A} have the following form representing the poles of the state-space equation

$$\lambda_k = -\omega_{1/2,k} + j\Delta\omega_k, k = 1, \dots, N. \quad (3.51)$$

The real part of the eigenvalue represents the half-bandwidth of the corresponding passband mode, whereas the imaginary part represents its frequency offset compared to the carrier frequency ω_c . For a particular multi-cell cavity, we can formulate the matrix \mathbf{A} and calculate its eigenvalues and eigenvectors numerically. As an example (Liepe 2001; Vogel 2007), the TESLA 9-cell superconducting cavity has the following parameters: the single-cell resonance frequency $\omega_{0c}/(2\pi) = 1275$ MHz, the single-cell unloaded quality factor $Q_{0c} = 10^{10}$, the cell-to-cell coupling factor $\kappa = 1.98\%$, and the input coupling factor of the first cell $\beta_c = 3e4$. With these parameters, the matrix \mathbf{A} can be calculated with the carrier frequency ω_c selected to be the same as ω_{0c} . The eigenvalues of \mathbf{A} are listed in Table 3.4, together with the frequency $\omega_{0,k}$ and loaded quality factor $Q_{L,k}$ of each passband mode. Here $\omega_{0,k} = \Delta\omega_k + \omega_c$ and $Q_{L,k} = \omega_{0,k}/(2\omega_{1/2,k})$, which are derived from the eigenvalue λ_k ($k = 1, \dots, N$).

Table 3.4 Passband modes of the TESLA cavity

Passband Mode	Eigen value λ_k	$\omega_{0,k}/(2\pi)$ (MHz)	$Q_{L,k}$
$\pi/9$	$-80.9 + j4.39e6$	1275.7	49.52e6
$2\pi/9$	$-312.8 + j18.14e6$	1277.9	12.84e6
$3\pi/9$	$-667.9 + j39.21e6$	1281.2	6.03e6
$4\pi/9$	$-1103.3 + j65.06e6$	1285.4	3.66e6
$5\pi/9$	$-1566.0 + j92.57e6$	1289.7	2.59e6
$6\pi/9$	$-1999.1 + j118.41e6$	1293.8	2.03e6
$7\pi/9$	$-2346.1 + j139.49e6$	1297.2	1.74e6
$8\pi/9$	$-2534.7 + j153.24e6$	1299.4	1.61e6
π	$-1289.3 + j158.12e6$	1300.2	3.17e6

**Fig. 3.25** Eigenvectors of the passband modes of the TESLA cavity

The eigenvector of a passband mode describes the relative field distribution (i.e., the magnitude and direction of the electric field) in different cells of the cavity. The eigenvectors of the passband modes of a TESLA cavity are shown in Fig. 3.25. Here we have normalized the magnitude of the electric field in the first cell to 1 for all of the passband modes.

We calculate the frequency response of a multi-cell cavity using the phasor transfer function derived from (3.49), which can be written as

$$\mathbf{X}(\hat{s}) = (\hat{s}\mathbf{I}_N - \mathbf{A})^{-1} \mathbf{B} \mathbf{U}(\hat{s}),$$

where \mathbf{X} and \mathbf{U} are the phasor Laplace transforms of the time-domain phasor-vectors \mathbf{x} and \mathbf{u} , respectively. To calculate the frequency response, we substitute the variable

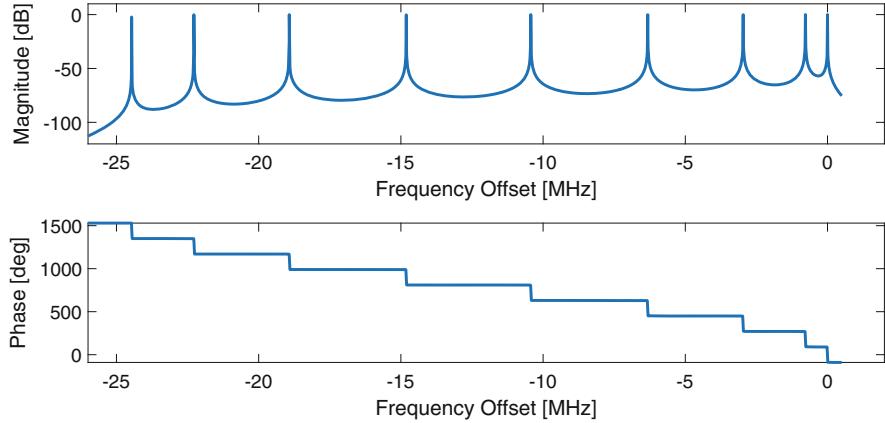


Fig. 3.26 Frequency response of the RF field in the ninth cell of a TESLA cavity. The π -mode has the highest resonance frequency normalized to zero

\hat{s} with $j\hat{\omega}$, where $\hat{\omega}$ is the frequency offset compared to the carrier frequency ω_c . For example, the frequency responses of the amplitude and phase in the ninth cell (where the probe is installed) of a TESLA cavity are shown in Fig. 3.26. The RF field in a multi-cell cavity is typically measured with a probe installed in a specific location in the cavity, like at the ninth cell of the TESLA cavity. In an RF control loop, we use the probe signal to represent the RF field experienced by the beam. This is an approximation because the beam is accelerated by the integrated electric field from all cavity cells. In practice, the relative field distribution in different cells is constant for the passband mode used for beam acceleration. Therefore, it is relatively accurate to represent the integrated field with the probe signal. In a feedback loop, a multi-cell cavity appears as a two-port device; that is, the controller manipulates the cavity input and measures the probe signal as output. The transfer function between the cavity input signal and the probe signal is more interested in designing a feedback control system. This is because, for feedback controller design, the multi-cell internal dynamics are not relevant, and the controller relies on the measurement of the probe signal. Such a transfer function is the basis to analyze (e.g., stability analysis) and design RF controllers.

According to the frequency response in Fig. 3.26, the phasor transfer function of a multi-cell cavity has multiple poles, each corresponding to a passband mode. From the phasor transfer function of a single-cell cavity (3.14), the phasor transfer function of a multi-cell cavity, including all passband modes, can be written as

$$\mathbf{G}_{C,multi-cell}(\hat{s}) = \sum_{k=1}^N \frac{\omega_{1/2,k} G_k}{\hat{s} + \omega_{1/2,k} - j\Delta\omega_k}. \quad (3.52)$$

Here $\omega_{1/2,k}$ and $\Delta\omega_k$ are from the eigenvalue described in (3.51) for the k th passband mode. The scale factor G_k reflects the relative magnitude and direction of the electric fields in the pickup cell for different passband modes. When modeling a multi-cell

cavity, the passband mode used for beam acceleration is usually studied to determine its equivalent single-cell cavity model. Then, the transfer functions of other passband modes can be derived with the values of G_k in the pickup cell, as well as the values of $\omega_{1/2,k}$ and $\Delta\omega_k$ ($k = 1, \dots, N$). Note that G_k is typically given relative to the gain of the transfer function of the passband mode used for beam acceleration. For example, the π -mode of a TESLA cavity is used for beam acceleration, and its equivalent single-cell cavity model is described by (3.14). From the results in Fig. 3.25, different passband modes in the ninth cell have the same magnitudes but opposite phases compared to their neighbor modes. Therefore, we can obtain the following phasor transfer function for the TESLA cavity with a probe installed in the ninth cell:

$$\mathbf{G}_{TESLA}(\hat{s}) = \frac{\omega_{1/2,(\pi)}R_L}{\hat{s} + \omega_{1/2,(\pi)} - j\Delta\omega_{(\pi)}} - \frac{\omega_{1/2,(8\pi/9)}R_L}{\hat{s} + \omega_{1/2,(8\pi/9)} - j\Delta\omega_{(8\pi/9)}} + \\ \frac{\omega_{1/2,(7\pi/9)}R_L}{\hat{s} + \omega_{1/2,(7\pi/9)} - j\Delta\omega_{(7\pi/9)}} - \dots$$

Note that R_L is defined for the π -mode. It is the integrated loaded resistance experienced by the beam when it passes through all of the nine cells. From the viewpoint of the beam, other passband modes contribute zero acceleration. Nevertheless, if we only consider the electric field in the ninth cell, the magnitudes of different passband modes are the same for a constant RF drive power at the corresponding frequencies. Therefore, we use the same R_L value in the transfer functions of other passband modes to describe their relative magnitudes compared to the π -mode, but the physical meaning of R_L should be kept in mind.

In Fig. 3.26, the resonance frequencies of other passband modes are smaller than that of the π -mode; that is, the frequency response is asymmetric around the π -mode resonance frequency. When designing a controller regulating the π -mode, we often draw a Bode plot of the cavity, taking into account the other passband modes. Note that the Bode plot will convert the other passbands to positive offset frequencies. It does not reflect the exact physics. If a control actuation affects both sidebands in the frequency domain, the actuation may interact with the other passband modes. In some cases, the control actuation may only affect one sideband (e.g., actuation via single sideband modulation). If the actuation is at higher frequencies than the π -mode, it will not affect other passband modes, even if the actuation frequency overlap with the resonance frequencies of the other passband modes in the Bode plot.

3.4.4 Transient in Cavity Cells with RF Drive

The phasor equation of multi-cell cavities (3.49) can be used to estimate the transient response in each cell of the cavity (Sekutowicz 1994, 2008). For that, we select a carrier frequency $\omega_c = \omega_{RF}$ and calculate the matrix \mathbf{A} , \mathbf{B} , and the cavity drive phasor \mathbf{u} . By solving the differential equation numerically, we obtain the response of the cavity in each cell as phasors referring to ω_c .

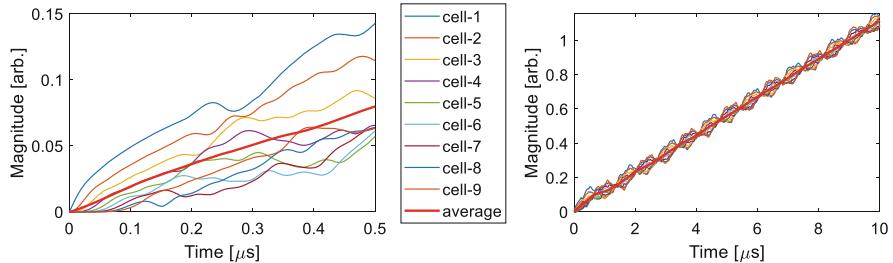


Fig. 3.27 Transient response in the cells of a TESLA cavity during the first 0.5 μs (left) and 10 μs (right) after the RF power is turned on. The average magnitude of all cells corresponding to the field felt by the beam is also plotted

Figure 3.27 shows the transient response of a TESLA cavity to a step change in \mathbf{u} turned on at $t = 0$. The cavity parameters are taken from Sect. 3.4.3 (i.e., the example to derive Table 3.4), and the RF operating frequency ω_{RF} is chosen to be the resonance frequency of the π -mode. The results show RF field fluctuations in the cavity cells caused by the passband modes. We also observe a group delay of about 100 ns from the first cell to the last cell, which is the delay of RF energy flow. The delay of energy flow will limit the maximum bunch rate that the cavity can accelerate. The RF energy taken away by the previous bunch should be filled by the RF drive before the following bunch arrives. Therefore, a group delay of 100 ns will limit the bunch rate to be no more than 10 MHz.

3.5 Traveling-Wave Structure Model

In a standing-wave cavity, the accelerating voltage experienced by the beam can be fully described by the RF measurement via a probe installed in the cavity. That is, the accelerating voltage is directly measurable.

When a particle passes through a traveling-wave structure, it experiences an integrated electric field along the beam path (Holtkamp and Joestingmeier 1996; Rezaeizadeh 2016). Unlike the standing-wave cavity, there is no direct measurement of the accelerating voltage in a traveling-wave structure. This is because the electric field has a spatial distribution in the structure, and there are no deterministic relations for the electric fields at different locations of the structure. Therefore, we cannot represent the integrated field using the measurement from a single probe installed somewhere in the structure. As it is not practical to measure the electric field at each point in the structure, we need an alternative method to estimate the accelerating voltage from the measurement of the input RF signal.

The input of a traveling-wave structure model is the forward RF drive voltage \mathbf{v}_{for} and the output is the accelerating voltage \mathbf{v}_{ACC} experienced by the beam. An overview of the RF power filling procedure is helpful to derive the phasor transfer function of a traveling-wave structure.

3.5.1 Filling of Structures

Due to the small group velocity v_g , the RF drive power of a traveling-wave structure launches a wave that builds the RF field cell-by-cell (Wangler 2008). This is the so-called filling of the structure. The *filling time* T_f , in which the structure is fully filled, is typically in the range of hundreds of nanoseconds to several microseconds. After the structure is filled, the beam is injected and accelerated by the electric field distributed on the axis of the structure. The phase velocity v_p of the traveling wave is the same as the beam flight velocity so that the beam experiences a constant accelerating phase in the structure. As the phase velocity is much higher than the group velocity, the RF field distribution in the structure can be viewed as time-invariant when the beam passes by.

A structure with a constant group velocity at different locations along its axis is called a *constant impedance structure*, which has all the cells with identical dimensions. Since there is power dissipation in the structure wall, the electric field has a maximum strength at the input coupler and decays exponentially in space towards the load location. Practically, we prefer the *constant gradient structure*, which has a space-varying group velocity, resulting in a constant gradient at different locations of the structure. The constant gradient structure has several advantages. For example, the constant gradient implies a uniform power dissipation in the structure wall that can simplify the design of the water cooling system. Furthermore, a constant gradient structure has a lower peak gradient than a constant impedance structure if we want to achieve the same accelerating voltage. Lowering the peak gradient is critical to avoid breakdowns in the structure. A constant gradient structure is constructed by varying the dimensions of its cells to decrease the group velocity when the drive wave travels toward the load. The result is that the drive wave is compressed in space to compensate for the power loss in the structure wall, resulting in a constant gradient. See Fig. 3.28. It shows that if we inject a short RF pulse into the constant gradient structure, the RF power travels along the structure axis at the z direction (beam flight direction). The short pulse of RF power will build up electric fields in the structure. The tail of the pulse travels faster than the head due to the location-dependent group velocity. Therefore, the electric field distribution will be compressed in space, as seen from the electric fields in t_1 , t_2 , and t_3 in Fig. 3.28. As a result, the electric field gradient is kept constant by shortening the pulse against the RF power loss in the structure walls.

First, let us summarize the basic principles of a constant gradient structure. In Chap. 1, we have introduced the characteristics of a traveling-wave structure. From the quality factor (1.13), we obtain a differential equation for the traveling-wave power P_{tw}

$$\frac{dP_{tw}}{dz} = -\frac{\omega W_{tw}}{Q}. \quad (3.53)$$

Here W_{tw} is the stored energy per unit length that can be calculated as $W_{tw} = P_{tw}/v_g$. Therefore, we can rewrite the equation above as

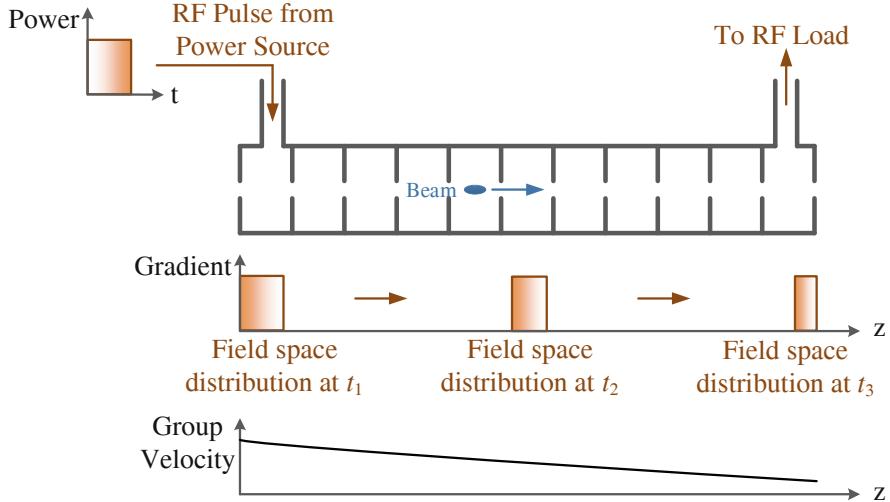


Fig. 3.28 Filling of a constant gradient traveling-wave structure

$$\frac{dP_{tw}}{dz} = -\frac{\omega P_{tw}}{v_g Q} := -2\alpha P_{tw}, \quad (3.54)$$

where α is the *attenuation length* representing the exponential decrease of P_{tw}

$$\alpha := \frac{\omega}{2v_g Q}. \quad (3.55)$$

Then from the shunt impedance per unit length (1.12), the gradient E_0 is given by

$$E_0^2 = -r_{s,tw} dP_{tw}/dz = 2\alpha P_{tw} r_{s,tw}. \quad (3.56)$$

In a constant-gradient structure, E_0 is constant at different locations (z from 0 to L , where L is the structure length) on the axis of the structure. Since $r_{s,tw}$ is approximately a constant, we require that $\alpha P_{tw} = \text{constant}$. And because P_{tw} decays along the z axis, α should also be a function of z . Therefore, from (3.54), we can write

$$\int_{P_0}^{P_L} dP_{tw}/P_{tw} = -2 \int_0^L \alpha(z) dz, \quad (3.57)$$

where P_0 and P_L are the traveling-wave powers at the entry and end of the structure, respectively. For convenience, we define $\tau_{tw} := \int_0^L \alpha(z) dz$ as the total power attenuation factor. Therefore, we can calculate the traveling-wave power at different z locations as

$$P_{tw}(z) = P_0 [1 - (1 - e^{-2\tau_{tw}})z/L]. \quad (3.58)$$

Note that for a constant gradient, the traveling-wave power decreases linearly with z . Substituting (3.54), we can work out the expression of v_g as

$$v_g(z) = \frac{\omega L [1 - (1 - e^{-2\tau_{tw}})z/L]}{Q(1 - e^{-2\tau_{tw}})}, \quad (3.59)$$

which also decreases linearly with z . Then the gradient can be calculated from (3.55) and (3.56) as

$$E_0 = \sqrt{P_0 r_{s,tw} (1 - e^{-2\tau_{tw}})/L}, \quad (3.60)$$

and the filling time of the structure can be calculated as

$$T_f = \int_0^L dz/v_g(z) = 2Q\tau_{tw}/\omega. \quad (3.61)$$

Equations (3.60) and (3.61) can be used to estimate the gradient for an input power P_0 with the shunt impedance per unit length and the structure filling time given.

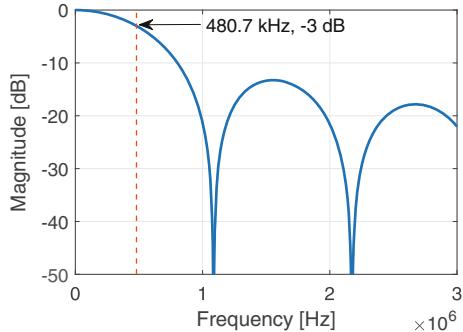
3.5.2 Structure Phasor Transfer Function

If we fill the structure with an RF pulse presenting between $t - T_f/2$ and $t + T_f/2$, we can estimate the beam accelerating voltage, which is also called the *structure voltage*, as

$$\mathbf{v}_{ACC}(t) = \frac{1}{T_f} \int_{t-T_f/2}^{t+T_f/2} w(\tau) \mathbf{v}_{for}(\tau) d\tau.$$

Here $w(\tau)$ is a weight function applied to \mathbf{v}_{for} representing the contributions of different parts of the RF drive pulse to \mathbf{v}_{ACC} . It is not easy to obtain an analytical expression of $w(\tau)$. Typically, we obtain $w(\tau)$ as a numerical look-up table using the physical model of the traveling-wave structure. In this book, we will not discuss deep into the calculation of $w(\tau)$. It is often adequate if we assume $w(\tau) = 1$ when the purpose of the structure transfer model is to analyze and design RF controllers. Therefore, the phasor transfer function of a traveling-wave structure can be derived from the integral equation above as

Fig. 3.29 Frequency response of the structure with a filling time of 920 ns



$$\mathbf{G}_{ACC}(\hat{s}) = \frac{e^{\hat{s}T_f/2} (1 - e^{-\hat{s}T_f})}{\hat{s}T_f}. \quad (3.62)$$

It implies that the half-bandwidth of the structure experienced by the beam is mainly determined by T_f and can be calculated as

$$f_{1/2,ACC} = 0.443/T_f \text{ [Hz]}. \quad (3.63)$$

For example, the half-bandwidth of a structure with a filling time $T_f = 920$ ns is about 480.7 kHz, as seen in the frequency response of (3.62) in Fig. 3.29. The bandwidth discussed here ($f_{1/2,ACC}$ [Hz]) is the bandwidth in which the RF field interacts with the beam. That is, the fluctuations in the RF field faster than $f_{1/2,ACC}$ do not have significant effects on the beam acceleration, even if they are still visible in the measurement of the RF field. From the viewpoint of RF (e.g., scatter parameter S21 of the structure), the bandwidth of the structure could be much larger than the bandwidth experienced by the beam.

3.6 Modeling of Important RF Devices

3.6.1 Transmission Line Model

Transmission lines (e.g., waveguides, coaxial cables or micro-strip lines) are essential for an RF system to transfer RF power and deliver RF signals. In RF system models, a transmission line mainly contributes an attenuation in magnitude, a phase shift at the carrier frequency, and a group delay of the input RF signal envelope.

The phasor transfer function of a transmission line can be written as

$$\mathbf{G}_T(\hat{s}) = g_T \cdot e^{-j\omega_c \Delta t_p} e^{-\Delta t_g \hat{s}}, \quad (3.64)$$

where g_T is the gain (or attenuation since $g_T < 1$) of the transmission line. We have defined two delay terms: Δt_p is the RF phase propagation delay resulting in a phase

lag at ω_c , whereas Δt_g is the propagation delay of the RF envelopes. They are determined by the transmission line length L , phase velocity v_p and group velocity v_g :

$$\Delta t_p = L/v_p, \quad \Delta t_g = L/v_g.$$

Note that v_p and v_g might be different for a transmission line with dispersion and are therefore frequency dependent. If with dispersion, the delay terms Δt_p and Δt_g should be described as functions of \hat{s} and (3.64) becomes complicated. Fortunately, the dispersion of the transmission lines used in accelerator RF systems can often be neglected within the system bandwidths. Therefore, we can assume a frequency-independent v_p that implies $v_g = v_p$. In this case, Δt_p and Δt_g will be constant and identical. Then a transmission line only introduces an envelope delay and a phase shift in the input RF signal without changing its shape.

3.6.2 RF Amplifier Model

RF amplifiers produce the RF power required by the downstream RF components. For low-power or medium-power amplifiers (up to hundreds of kW), solid-state technology is more and more adopted, while for high-power amplifiers, electron tubes (e.g., klystrons) are still widely in use.

The bandwidths of RF amplifiers are typically much larger than the cavity bandwidth. We often neglect the dynamics of amplifiers when analyzing and designing RF control systems. If the amplifier dynamics are not negligible, we may model it as a first-order low-pass filter in the \hat{s} domain. The phasor transfer function of an RF amplifier can be written as

$$\mathbf{G}_A(\hat{s}) = g_A \cdot e^{-j\varphi_A} e^{-\Delta t_A \hat{s}} \frac{\omega_{1/2,A}}{\hat{s} + \omega_{1/2,A}}, \quad (3.65)$$

where g_A , φ_A and Δt_A are the gain, phase-lag and group delay of the amplifier. A real pole at the half-bandwidth of the amplifier $-\omega_{1/2,A}$ models the dynamics of the amplifier. An RF amplifier may be non-linear with its gain and phase-lag as functions of the input signal level. However, the linear model above is still valid if we only consider the behavior of the amplifier in a small region around an operating point. Several algorithms have been developed to linearize the RF amplifiers over a broad operating range that will be discussed in Chap. 7.

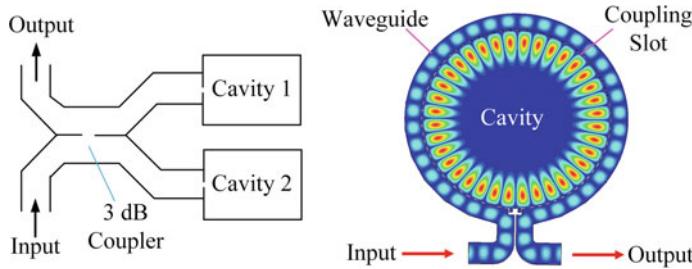


Fig. 3.30 SLED (left) and BOC (right): two typical RF pulse compressors. (Courtesy of R. Zennaro)

3.6.3 RF Pulse Compressor Model

RF pulse compressors are often equipped in RF stations where high peak powers are needed within short RF pulses. Due to the limitations of amplifiers or klystrons, the available peak power might not be sufficient to produce the desired accelerating gradient. However, we may insert a passive pulse compressor after the amplifier or klystron to achieve the required peak power. Figure 3.30 shows two types of RF pulse compressors widely used in various machines: the SLAC Energy Doubler (SLED) (Farkas et al. 1974) and the Barrel Open Cavity (BOC) (Shu et al. 2015).

Both the SLED and BOC use standing-wave cavities with a coupling factor larger than one to store the RF energy injected from the first part of the RF pulse. See Fig. 3.31. After an instant phase inversion in the pulse, the RF power emitted from the cavity sums in-phase with the second part of the pulse, increasing the peak power. If we use voltages to represent the input and output of the pulse compressor, they can be written below referring to the cavity Eq. (3.31):

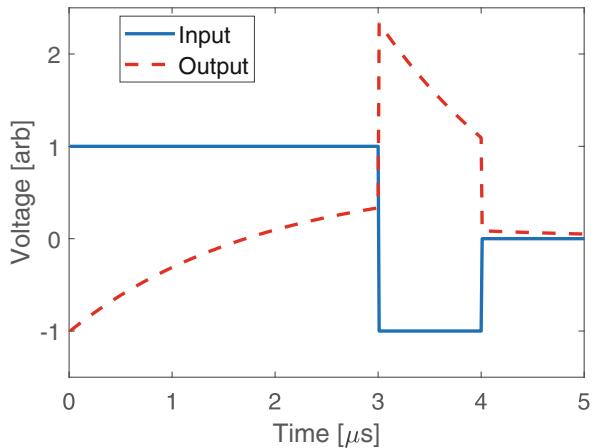
$$\mathbf{v}_{PC,IN} = \mathbf{v}_{for}, \mathbf{v}_{PC,OUT} = \mathbf{v}_C - \mathbf{v}_{for},$$

where the subscript “ $_{PC}$ ” stands for “pulse compressor”. It can be seen that the pulse compressor output is equivalent to the RF power reflected from the cavity. Therefore, the phasor transfer function of the RF pulse compressor can be written as

$$\mathbf{G}_{PC}(\hat{s}) = \frac{2\beta}{\beta + 1} \frac{\omega_{1/2}}{\hat{s} + \omega_{1/2} - j\Delta\omega} - 1. \quad (3.66)$$

Here β is the input coupling factor of the cavity, and $\omega_{1/2}$ and $\Delta\omega$ are the half-bandwidth and detuning of the cavity, respectively. Typically, the cavity is well-tuned for normal operation, so $\Delta\omega = 0$. In some special cases, the cavity might be detuned to shape the output pulse or bypass the RF pulse compressor completely. As an example, Fig. 3.31 shows the input and output voltage waveforms of the SLED used in the BEPCII Linac at the Institute of High Energy Physics (IHEP) in China.

Fig. 3.31 Waveforms of the SLED input and output voltages. The amplitude of the input voltage has been normalized to 1 and the detuning of the cavity has been assumed to zero



The RF frequency is 2856 MHz, the unloaded quality factor of the SLED cavity is 10^5 and the input coupling factor is 5. In the input RF pulse, the phase is shifted by 180° instantly at 3 μs after the RF pulse starts. The SLED output (from 3 to 4 μs) is used to fill the traveling-wave structures, increasing the accelerating voltage by a factor of about 1.7 compared to that achieved without the SLED.

The enhancement of the accelerating voltage experienced by the beam is denoted as the *energy multiplication factor*, which is typically smaller than 2 for the existing RF pulse compressors.

3.7 Application of RF System Models

RF system models are widely used in the design and analysis of LLRF systems. The contents of this chapter are the basis of the following chapters. In this section, we briefly introduce the applications of RF system models.

The most straightforward application of RF system models is an RF system simulator, with which the RF fields under different RF drives, beam loadings and disturbances can be simulated (Grassellino et al. 2007; Pucyk et al. 2007; Geng et al. 2008; Orel and Mavric 2013; Qiu et al. 2018). In addition to software-based simulations, we may implement the RF system simulator in an FPGA equipped with ADCs and DACs, as shown by the example in Fig. 3.32 (Qiu et al. 2018). Such a hardware-based simulator can represent the RF system to be controlled physically and can be used for testing a physical LLRF controller. This is very helpful when designing LLRF controllers, especially when the high power RF system is not available.

The RF system model is also a basis for determining the RF system parameters for the desired operating point. An operating point is typically represented by the required RF pulse shape, the desired cavity voltage, and the given beam current

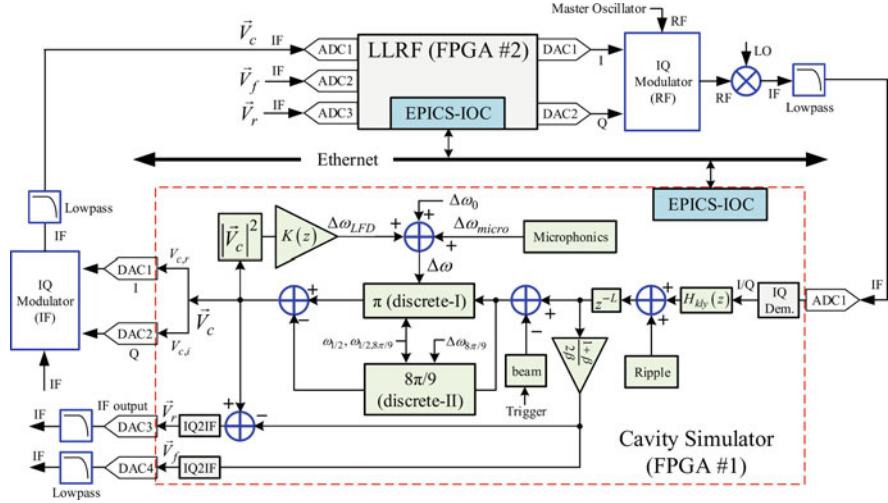


Fig. 3.32 An RF system simulator implemented in an FPGA and the test bench for the design of LLRF controllers. In the RF system simulator, the klystron model, the cavity electrical model including the π -mode and $8\pi/9$ -mode, and the cavity mechanical model are implemented. (Courtesy of F. Qiu et al.)

and beam phase. Example 3.1 is a good example to determine the feedforward drive for the desired accelerating voltage and beam loading. Another example is that the loaded quality factor and detuning of the cavity can be optimized to minimize the required RF drive power, as discussed in Sect. 3.9.

When designing an RF controller, which typically consists of feedback and feedforward controls, the RF system dynamical model in the form of phasor transfer functions plays an important role. The underlying architecture and parameters of the feedback controller can be initially determined using the RF system model via simulations. Some essential characteristics of the feedback loop, such as the gain margin and phase margin, can already be evaluated with simulations. Of course, we may identify the RF system model using the empirical system identification methods, yielding a more accurate model for an RF system in practice. However, the theoretical model of the RF system could still be useful for the system identification process. It helps determine the model structure and define the experiments for system identification, such as the input signals (e.g., pseudorandom binary sequence (PRBS) or Gaussian random inputs), the sampling rate, the data length, and the identification algorithms.

When operating an RF system, the knowledge of the system characteristics, such as the system gain, system phase, input coupling factor, loaded quality factor and detuning of the cavity, is useful to optimize the RF system parameters and diagnose exceptional conditions. Based on the RF system models and the RF signal measurements at different locations in the system, such as the cavity probe, forward and reflected signals, the RF system characteristics can be calculated. The RF system

model is also the basis to perform necessary calibrations in the RF measurements, such as the calibration of the vector sum, the cavity voltage, beam phase, and the forward and reflected RF powers. We will discuss the RF system parameter identification and the RF system calibration in detail in Chap. 9.

3.8 Summary

This chapter studies the models of the key RF components of accelerator RF systems. We emphasized the RF cavity and structure models that are critical for determining the RF system operational parameters and designing RF controllers. The models focus on the input-output dynamics of the RF components instead of the detailed RF field distributions. The knowledge introduced in this chapter is the basis of understanding other chapters, especially Chaps. 4, 6, and 9.

References

- Aune B, Bandelmann R, Bloess D et al (2000) Superconducting TESLA cavities. *Phys Rev Spec Top Accel Beams* 3(9):092001. <https://doi.org/10.1103/PhysRevSTAB.3.092001>
- Brandt A (2007) Development of a finite state machine for the automated operation of the LLRF control at FLASH. Ph.D thesis, Hamburg University
- Dorf RC, Bishop RH (2010) Modern control systems, 12th edn. Pearson Education, London
- Farkas Z D, Hoag H A, Loew G A et al (1974) SLED: a method of doubling SLAC's energy. In: Proceedings of the 9th international conference on the high-energy accelerators, Stanford, CA, USA, 2–7 May 1974
- Geng ZQ, Hou M, Pei GX (2008) Modeling of RF system for the normal conducting linac. *Chin Phys C* 32(1):56–59. <https://doi.org/10.1088/1674-1137/32/1/013>
- Grassellino A, Keung J K, Newcomer M et al (2007) FPGA based ILC cavity simulator. In: Proceedings of the 22nd IEEE particle accelerator conference, Albuquerque, NM, USA, 25–29 June 2007
- Hall D (2011) A review of the equivalent circuit model of RF accelerating cavities in the context of an analysis of the 1.3 and 3.9 GHz cavities at FLASH. <http://www.hep.manchester.ac.uk/u/danhall/publications.html>. Accessed 26 May 2021
- Harnefors L (2007) Modeling of three-phase dynamic systems using complex transfer functions and transfer matrices. *IEEE Trans Ind Electron* 54(4):2239–2248. <https://doi.org/10.1109/TIE.2007.894769>
- Holtkamp N, Joestingmeier A (1996) Transient beam loading compensation in traveling wave linear accelerators. <https://cds.cern.ch/record/300739/files/SCAN-9604090.pdf>. Accessed 25 May 2021
- Koontz RF, Loew GA, Miller RH et al (1977) Single bunch beam loading on the SLAC three-kilometer accelerator. *IEEE Trans Nucl Sci* 24(4):1493–1495. <https://doi.org/10.1109/TNS.1977.4328989>
- Kramer SL, Wang JM (1999) Beam induced RF cavity transient voltage. *Nucl Instrum Methods Phys Res A Accel Spectrom Detect Assoc Equip* 423(2–3):260–264. [https://doi.org/10.1016/S0168-9002\(98\)01369-2](https://doi.org/10.1016/S0168-9002(98)01369-2)
- Liepe MU (2001) Superconducting multicell cavities for linear colliders. Ph.D. thesis, Hamburg University

- Martin KW (2004) Complex signal processing is not complex. *IEEE Trans Circuits Syst I Regul Pap* 51(9):1823–1836. <https://doi.org/10.1109/TCSI.2004.834522>
- Novotny DW, Wouterse JH (1976) Induction machine transfer function and dynamic response by means of complex time variables. *IEEE Trans Power App Syst* 95(4):1325–1335. <https://doi.org/10.1109/T-PAS.1976.32227>
- Orel P, Mavric U (2013) Analog cavity simulator. *Nucl Instrum Methods Phys Res A* 729(11): 241–246. <https://doi.org/10.1016/j.nima.2013.06.059>
- Pawlak P, Grecki M, Simrock S (2005) Single bunch transient detection for the beam phase measurement in superconducting accelerators. In: Proceedings of the 7th European workshop on beam diagnostics and instrumentation for particle accelerators, Lyon, France, 6–8 June 2005
- Pozar DM (2011) Microwave engineering, 4th edn. Wiley, New York
- Pucyk P, Jalmuzna W, Simrock S (2007) Real time cavity simulator for European XFEL. In: Proceedings of the 15th IEEE-NPSS real-time conference, Batavia, IL, USA, 29 April-4 May 2007
- Qiu F, Michizono S, Miura T et al (2018) Real-time cavity simulator-based low-level radio-frequency test bench and applications for accelerators. *Phys Rev Accel Beams* 21(3):032003. <https://doi.org/10.1103/PhysRevAccelBeams.21.032003>
- Rezaeizadeh A (2016) Automatic control strategies for the Swiss free electron laser. Ph.D. thesis, Eidgenössische Technische Hochschule Zürich
- Schilcher T (1998) Vector sum control of pulsed accelerating fields in Lorentz force detuned superconducting cavities. Ph.D. thesis, Hamburg University
- Sekutowicz J (1994) Transient state in standing wave accelerating structures. Part Accel 45:47–58
- Sekutowicz J (2008) Multi-cell superconducting structures for high energy e+e- colliders and free electron laser linacs EuCARD editorial series on accelerator science, vol 01
- Shu G, Zhao FL, He X (2015) RF study of a C-band barrel open cavity pulse compressor. *Chin Phys C* 39(5):057005. <https://doi.org/10.1088/1674-1137/39/5/057005>
- Solyak N, Awida M, Gonin I et al (2013) Simulations of mechanical resonances in SRF cavities in low beam current CW operation. In: Proceedings of the 16th international conference on RF superconductivity, Paris, France, 23–27 September 2013
- Troeng O (2019) Cavity field control for linear particle accelerators. Ph.D. thesis, Lund University
- Vogel E (2007) High gain proportional RF control stability at TESLA cavities. *Phys Rev Spec Top Accel Beams* 10(5):052001. <https://doi.org/10.1103/PhysRevSTAB.10.052001>
- Wangler TP (2008) RF linear accelerators, 2ed edn. Wiley, Weinheim

Chapter 4

RF Field Control



One of the LLRF system's core functions is controlling the RF field for particle beam acceleration. An RF controller implements feedback or feedforward algorithms, regulating the RF field to follow set points or stabilizing it against disturbances. This chapter will discuss the principles of GDR, SEL, and PLL control strategies. Several widely used feedback and feedforward algorithms will be introduced. These algorithms can be applied to practical LLRF systems directly. We will also discuss the cavity resonance control, which keeps the cavity on-resonance and is essential to enhance the performance of RF field controls.

4.1 Requirements to RF Field Control

Deeply understanding the requirements for RF field control is essential to design an RF controller successfully. The requirements should be collected from LLRF system users, such as the RF system experts, beam physicists and machine operators. The RF field control requirements are usually given as follows:

- (a) *Tolerances of beam accelerating voltage and phase.* A parameter's tolerance describes its maximum allowed deviation from its desired value. We also use the term *(RF) amplitude* to denote the accelerating voltage. The tolerances quantify the RF controller performance for both command tracking and disturbance rejection.

We may define different tolerances for the low- and high-frequency fluctuations (denoted as *drift* and *jitter*, respectively). For example, the RF amplitude and phase jitter tolerances for the European XFEL are 0.01% RMS and 0.01° RMS, respectively. However, as the drift is compensable by beam-based feedbacks, the drift tolerances can be larger (e.g., 10 times larger).

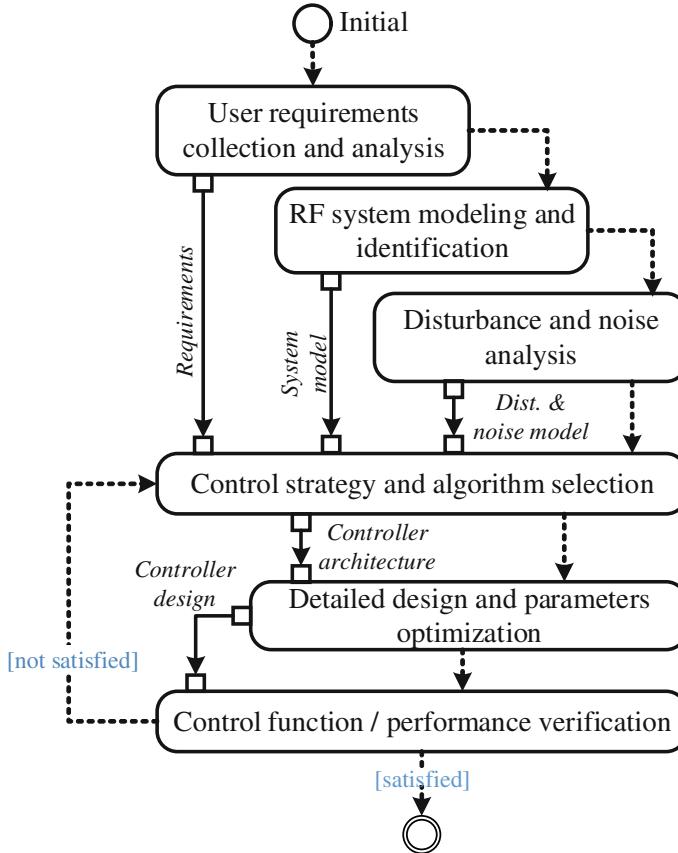


Fig. 4.1 RF controller design procedure. Here only the main steps are depicted. The rounded boxes are action steps, the dotted arrows are control flows and the solid arrows are object flows

- (b) *RF system operation scenarios.* An RF controller is used for testing or conditioning RF cavities, accelerating particle beams with different bunch charges and rates, and operating cavities in the pulsed or CW mode. These requirements determine the control strategies (GDR, SEL, or PLL) and control algorithms (feedback or feedforward) of the RF controller.

The user requirements provide a starting point to analyze and design an RF controller. We typically follow a systematic approach as shown in Fig. 4.1.

Comprehensive knowledge of RF systems is essential for designing an RF controller. In Chap. 3, we have derived the physical models of the RF components based on their physical principles (e.g., differential equations). The RF system parameters, such as the klystron bandwidth, cavity resonance frequency and quality factor, define the model parameters. The physical models provide insights into the RF system dynamics, which are often used to study the RF system behavior, determine the operational parameters, and perform the preliminary controller design

before the RF system is constructed. For a practical RF system, we may identify its model with empirical methods: fitting a dynamical model using the inputs and outputs measured from a system in operation. The empirical model fits better to the response of a real system, but its parameters' physical meanings are not obvious. The empirical model identification can benefit from the physical model study. The physical model helps define the empirical model structure, such as the model order or the number of states. By exercising the system identification process on a physical model, we can define the experiment, such as the inputs to be used, the frequency and length of outputs to be measured, etc.

The disturbances and noise affecting the control performance should also be analyzed. In superconducting cavities, Lorentz force detuning, microphonics, and beam loading are well-known disturbances that should be dealt with by RF controllers. Whereas for normal-conducting accelerators, klystron modulator high voltage jitter and cooling water temperature fluctuations are primary concerns. The analysis should provide information like the magnitude and frequency of the disturbances and noise.

We can then determine the control strategies and algorithms based on the user requirements, system model, disturbances, and noise. For example, an SEL or PLL loop might be required if we want to test or condition the RF cavities. Adaptive feedforward control should be implemented to deal with the repetitive disturbances like the Lorentz force detuning and beam loading. Considering the time-varying set points to be tracked or the disturbance and noise bandwidths, we can estimate the required closed-loop bandwidth of the feedback controller. Furthermore, the feedback loop limits, such as the gain margin, phase margin, and loop delay, can be derived. All the information leads to a selection of feedback algorithms, such as the PID control, optimal control, active disturbance rejection control, or robust control.

The RF controller parameters should be determined considering all constraints, such as the available klystron power and the maximum cavity gradient. We can perform closed-loop simulations using the RF system model to verify the controller's functionality and performance. Finally, we obtain a controller design with certain architecture and parameters. Of course, the process above should be executed iteratively. If the design cannot satisfy the requirements, a revision should be made by either updating the control algorithm or changing its parameters.

We will discuss the noise model of RF systems in Chap. 6. In this chapter, we will introduce several widely used RF control strategies and algorithms. The readers should have some basic knowledge of feedback control theory to understand the contents of this chapter.

4.2 Generator Driven Resonator Control

In Chap. 2, we have briefly introduced the GDR control strategy. Since the GDR control provides a stable RF field in both amplitude and phase, it is suitable for regular beam operation (Schilcher 1998; Fafara et al. 2007; Michizono et al. 2009).

Fig. 4.2 Simplified block diagram of a GDR control loop. Here \mathbf{r} is the set point phasor, \mathbf{y} is the cavity voltage phasor

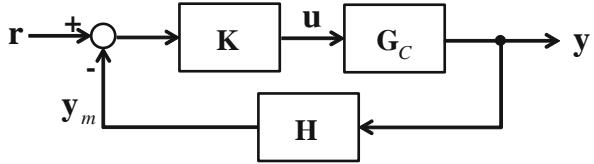


Figure 4.2 is a simplified block diagram of a GDR control loop, which adopts the I/Q control strategy where both the input \mathbf{r} and output \mathbf{y} are phasors. To simplify the discussion, we only included the phasor transfer function of the cavity (\mathbf{G}_C) in the plant to be controlled. Such a simplification is often valid because the cavity bandwidth is much smaller than the bandwidths of other RF components, and then dominates the system dynamics. Therefore, we neglect the dynamics of the RF driving chain components, such as the I/Q modulator, pre-amplifier, and klystron, when analyzing and designing an RF controller.

In Fig. 4.2, \mathbf{G}_C is the phasor transfer function of the cavity, and \mathbf{H} models the loop delay and the phase shift of the RF components. A typical controller \mathbf{K} employs the proportional-integral (PI) control algorithm. These phasor transfer functions are written as

$$\begin{aligned}\mathbf{K}(\hat{s}) &= K_P + K_I/\hat{s}, \quad \mathbf{G}_C(\hat{s}) = \frac{\omega_{1/2}}{\hat{s} + \omega_{1/2} - j\Delta\omega}, \\ \mathbf{H}(\hat{s}) &= g_{sys} e^{-j\varphi_{sys}} e^{-\hat{s}\Delta t},\end{aligned}\quad (4.1)$$

where K_P is the proportional gain, K_I is the integral gain, and φ_{sys} , g_{sys} and Δt are the *system phase*, *system gain* and *loop delay*, respectively. The *system phase* (φ_{sys}) and *system gain* (g_{sys}) are the phase shift and gain between the RF detector output \mathbf{y}_m and the RF controller output \mathbf{u} , excluding the cavity dynamics (i.e., let $\mathbf{G}_C = 1$). That is, φ_{sys} and g_{sys} are determined by the phase shifts and gains of the static RF components in the loop, such as the waveguides, RF cables, RF actuator, power amplifier and RF detector with their dynamics neglected. Note that φ_{sys} is defined at the RF operating frequency, which is typically the same as the carrier frequency defining $\Delta\omega$ and the phasor transfer functions (see Chap. 3).

In this section, we will define $g_{sys} = 1$ to simplify the discussion, i.e., use $\mathbf{H}(\hat{s}) = e^{-j\varphi_{sys}} e^{-\hat{s}\Delta t}$ in analysis. The algorithm to calibrate the system phase and system gain will be discussed in Sect. 9.4.2.

4.2.1 Feedback Stability for Single-Cell Cavities

The input and output of the loop in Fig. 4.2 are both phasors. The system phase is critical for the stability of such an I/Q control loop, which will be discussed in detail in this section. The open-loop transfer function can be written as

$$\mathbf{L}(\hat{s}) = \mathbf{H}(\hat{s})\mathbf{G}_C(\hat{s})\mathbf{K}(\hat{s}). \quad (4.2)$$

The loop is stable only when the closed-loop poles are located at the open left half-plane (LHP) of the \hat{s} complex plane, which requires the real part of the roots of $1 + \mathbf{L}(\hat{s}) = 0$ to be negative. Here we will assume that the controller contains only a proportional part to highlight the RF cavity's essential behaviors. With $\mathbf{K}(\hat{s}) = K_P$, the closed-loop pole can be calculated as

$$\mathbf{p}_{CL} = -\omega_{1/2} + j\Delta\omega - K_P\omega_{1/2}e^{-j\varphi_{sys}},$$

where the loop delay is neglected (i.e., $\Delta t = 0$) in the analysis, which still results in a good approximation of the closed-loop pole if Δt is much smaller than the cavity time constant. The loop stability implies that $-1 - K_P \cos \varphi_{sys} < 0$, leading to

$$\cos \varphi_{sys} > -1/K_P, \quad (4.3)$$

which is the system phase criterion for the loop stability. Let us make some detailed analysis of the stable regions of φ_{sys} for typical K_P values:

1. If $K_P \leq 1$, there are no limits for φ_{sys} . However, a feedback loop with a loop gain smaller than 1 is not useful for improving the RF system performance.
2. If $K_P \gg 1$, the loop stability requires that $-90^\circ < \varphi_{sys} < 90^\circ$.
3. For other K_P values ($K_P > 1$ but not much larger than 1), the stable region of the system phase is $-\arccos(-1/K_P) < \varphi_{sys} < \arccos(-1/K_P)$.

In the Bode and Nyquist plots (see Fig. 4.3) of the open-loop transfer function \mathbf{L} , φ_{sys} introduces an offset to the Bode plot phase response and rotates the entire Nyquist plot. The instability caused by an improper φ_{sys} happens when the Bode plot phase response reaches $\pm 180^\circ$ with the open-loop gain above 1 (i.e., 0 dB), or when the Nyquist plot is rotated to encircle the -1 point (Skogestad and Postlethwaite 2005). A larger system phase makes the Bode and Nyquist plots “closer” to the instability boundary, resulting in lower phase and gain margins and reduced control performance.

The sensitivity function \mathbf{S} and the complementary sensitivity function \mathbf{T} of the closed-loop system are calculated from the open-loop transfer function as

$$\mathbf{S}(\hat{s}) = \frac{1}{1 + \mathbf{L}(\hat{s})}, \quad \mathbf{T}(\hat{s}) = \frac{\mathbf{L}(\hat{s})}{1 + \mathbf{L}(\hat{s})}.$$

Note that $\mathbf{S} + \mathbf{T} = 1$ and both are complex-valued functions of frequency. With the same model parameters as Fig. 4.3, the frequency responses of \mathbf{S} and \mathbf{T} are shown in Fig. 4.4, illustrating the feedback performance for different system phases. The magnitude of \mathbf{S} (denoted as $|\mathbf{S}(j\hat{\omega})|$ at frequency $\hat{\omega}$) should be small at all frequencies where good performance is required in terms of command tracking and disturbance rejection. Whereas $|\mathbf{T}(j\hat{\omega})|$ should be small at higher offset

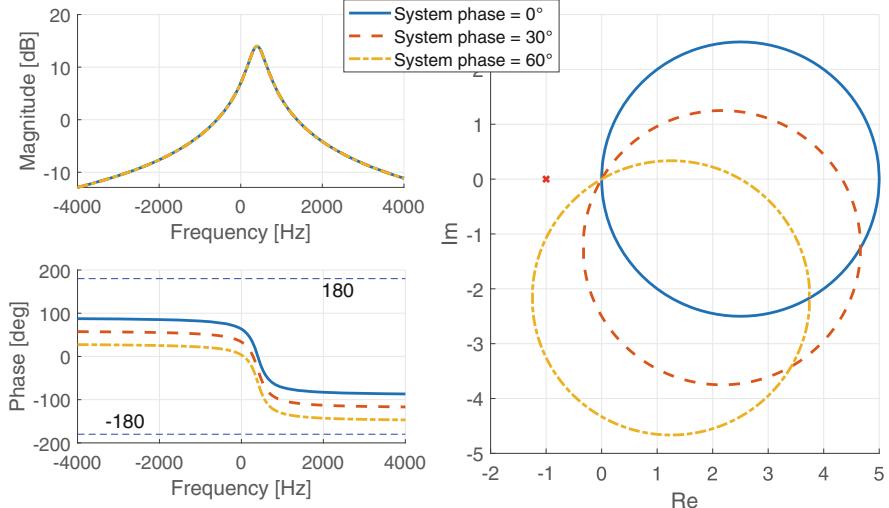


Fig. 4.3 Bode (left) and Nyquist (right) plots of \mathbf{L} for different system phases. The model parameters are $\omega_{1/2} = 2\pi \times 200$ rad/s, $\Delta\omega = 2\pi \times 400$ rad/s, $K_P = 5$ and $\Delta t = 0$. The Bode plot frequency axis is in linear scale to display the asymmetric frequency responses

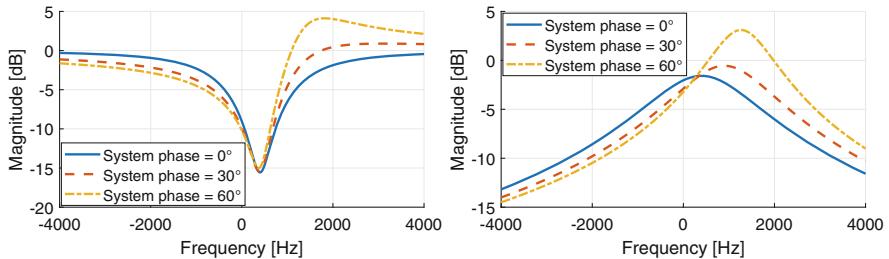


Fig. 4.4 Magnitudes of \mathbf{S} and \mathbf{T} for different φ_{sys} . Left: \mathbf{S} illustrates disturbance suppression and command tracking errors; Right: \mathbf{T} depicts transferring of measurement noise into the cavity

frequencies to avoid transferring high-frequency RF detector noise into the cavity and to improve the feedback loop robustness against RF system model uncertainties. The peak values of $|\mathbf{S}(j\hat{\omega})|$ and $|\mathbf{T}(j\hat{\omega})|$ should be small (e.g., less than 2). A large peak of $|\mathbf{S}(j\hat{\omega})|$ implies a small phase margin, resulting in significant overshoots in the closed-loop step responses.

Note that the stability condition (4.3) is for proportional control, where \mathbf{K} does not contribute to the phase response of \mathbf{L} . If the controller introduces additional phase-lag or phase-lead, the stability condition for φ_{sys} could be different. It can be explained by the Bode plot in Fig. 4.3. If \mathbf{K} introduces additional phases, the $\pm 180^\circ$ lines may be touched even if φ_{sys} is within the region determined by (4.3).

At the beginning of Sect. 4.2, we introduced the concepts of system phase and system gain. They are constants for a particular RF system and are defined at the RF operating frequency. In control theory, we have learned the *loop phase* and *loop gain* concepts. They are the phase shift and gain of $\mathbf{L}(j\hat{\omega})$ and are frequency-dependent, as shown in the Bode plot.

If the loop delay $\Delta t > 0$, it will introduce an additional phase-lag at $\hat{\omega} > 0$ and an additional phase-lead at $\hat{\omega} < 0$ to the Bode plot phase response. The phase lead/lag introduced by Δt is given by $\Delta\varphi_{\text{delay}} = -\hat{\omega}\Delta t$. As a result, the loop delay will reduce the stable region of φ_{sys} compared to the one determined by (4.3).

4.2.2 Feedback Stability for Multi-cell Cavities

For multi-cell cavities, passband modes are another factor that should be considered in the stability analysis (Vogel 2007). Typically, one specific passband mode (e.g., the π -mode) is selected to accelerate charged particles. The system phase of this passband mode should be adjusted close to zero so that the stability margins of the I/Q feedback loop are maximized.

If the cavity field measurement contains signals of other passband modes, the feedback loop may become unstable even if the system phase is well tuned. We use the TESLA cavity as an example. Except for the π -mode (used for beam acceleration) frequency, the $8\pi/9$ -mode frequency often falls in the RF detector bandwidth. We assume that all other passband-mode frequencies are filtered out by the RF detector and do not appear in the measurement results. In this case, the phasor transfer function of the cavity can be written as

$$\mathbf{G}_{C,\text{TESLA}}(\hat{s}) = \frac{\omega_{1/2,(\pi)}}{\hat{s} + \omega_{1/2,(\pi)} - j\Delta\omega_{(\pi)}} - \frac{\omega_{1/2,(8\pi/9)}}{\hat{s} + \omega_{1/2,(8\pi/9)} - j\Delta\omega_{(8\pi/9)}}, \quad (4.4)$$

where the magnitudes of different modes are normalized to 1. The half-bandwidth of the π -mode is $\omega_{1/2,(\pi)} = 2\pi \times 200$ rad/s, whereas the $8\pi/9$ -mode is -800 kHz away from the π -mode frequency with a half-bandwidth about $\omega_{1/2,(8\pi/9)} = 2\pi \times 400$ rad/s. The $8\pi/9$ -mode term is negative because its phase is 180° shifted w.r.t the π -mode in the cavity's 9th cell, where the pickup probe is installed. If the cavity is well tuned so that the π -mode frequency is the same as the RF operating frequency, we have $\Delta\omega_{(\pi)} = 0$ and $\Delta\omega_{(8\pi/9)} = -2\pi \times 800$ rad/s. The Bode and Nyquist plots of the open loop \mathbf{L} for a TESLA cavity considering these two passband modes are illustrated in Fig. 4.5. Here we have replaced \mathbf{G}_C in \mathbf{L} with $\mathbf{G}_{C,\text{TESLA}}$. The feedback loop becomes unstable due to the $8\pi/9$ -mode when the loop delay $\Delta t = 0$. In the Bode plot of Fig. 4.5, the phase response (solid curve) crosses the -180° line at the $8\pi/9$ -mode frequency, where the loop gain is larger than 1 due to the resonance peak of the $8\pi/9$ -mode. In the Nyquist plot, the -1 point is encircled by the contribution from the $8\pi/9$ -mode. Both indicate an unstable closed loop.

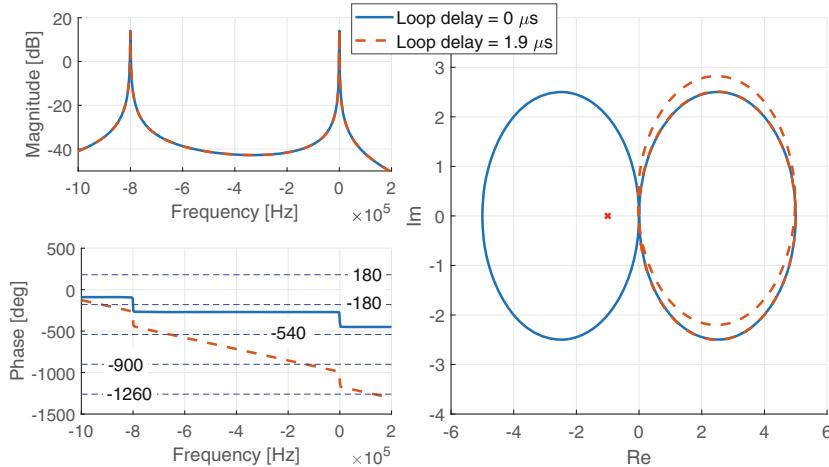


Fig. 4.5 Bode (left) and Nyquist (right) plots of \mathbf{L} for a TESLA cavity considering the π -mode and $8\pi/9$ -mode for different loop delays. The system phase $\varphi_{\text{sys}} = 0$

One method to overcome the influence of other passband modes is to add notch filters at the passband frequencies in the RF field measurement. Notch filters can reduce the loop gain at these frequencies to below 1 so that the closed-loop is stable. Another possibility to mitigate the instability caused by the passband modes is to introduce an extra delay in the RF feedback loop. As shown in the Bode plot, a non-zero loop delay will tilt the phase response curve and may avoid instability. For example, with $\Delta t = 1.9 \mu\text{s}$, the phase response crosses -180° (or $-180^\circ + k \times 360^\circ$, k is an integer) at frequencies where the loop gain is smaller than one. Therefore, the closed loop becomes stable. In the Nyquist plot, the loop delay will rotate the $8\pi/9$ -mode away from the -1 point, which also implies a stable closed loop.

In a real RF system with multi-cell cavities, typically, only a few passband modes fall into the RF detector bandwidth and other passband modes can be neglected.

4.2.3 Active Disturbance Rejection Control

PI controller is widely used in LLRF systems. It can be implemented without system models but still perform quite well. As an extension to PI control, the Active Disturbance Rejection Control (ADRC) (Han 2009; Vincent et al. 2011; Herbst 2013; Geng 2017a, b) is also widely used in RF control systems. It estimates the disturbances in the RF system and rejects them by acting on the system inputs. The disturbance observers often respond faster than the feedback loop so that the disturbances can be rejected in advance before they generate significant errors in the cavity. ADRC controllers demonstrate excellent performance for disturbance

rejection and robustness against system model uncertainties. In this section, we will introduce a disturbance observer based on the cavity model. The observer will be combined with a PI controller to enhance the disturbance rejection capability of the RF controller.

We start from a cavity model described by the equation

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega) \mathbf{v}_C = \omega_{1/2} \mathbf{g}_{sys} \mathbf{v}_{act}. \quad (4.5)$$

Here the cavity Eq. (3.31) with voltage drives has been used. We will treat the beam drive term as a disturbance and remove it from the cavity equation. The input of (4.5), \mathbf{v}_{act} , is generated by the RF controller. It produces the cavity drive voltage via a complex system gain $\mathbf{g}_{sys} = g_{sys} e^{-j\varphi_{sys}}$. Note that \mathbf{g}_{sys} should be calibrated using the measurements of the RF forward signal and the cavity probe signal (see Chap. 9). The cavity model above is characterized by its half-bandwidth $\omega_{1/2}$ and detuning $\Delta\omega$. These parameters are given either by the physical model or by the system identification experiments.

RF controllers are often designed based on the cavity model (4.5). A robust RF controller should still perform satisfactorily even if the model has uncertainties compared to the real system. Model uncertainties come from either an improper model structure (e.g., missing of hidden dynamics) or incorrect parameters. Notably, the cavity detuning variations caused by the Lorentz force or microphonics also vary the cavity dynamics. External disturbances, such as the beam loadings and RF amplifier noise, should also be suppressed by the controller. Both the model uncertainties and the external disturbances are defined as a uniformed “disturbance” here.

We define the terms

$$\mathbf{a} = -\omega_{1/2} + j\Delta\omega, \mathbf{b} = \omega_{1/2} \mathbf{g}_{sys}, \quad (4.6)$$

then (4.5) can be rewritten as $\dot{\mathbf{v}}_C = \mathbf{a}\mathbf{v}_C + \mathbf{b}\mathbf{v}_{act}$. Since \mathbf{a} and \mathbf{b} contain model uncertainties or calibration errors, we define the cavity disturbance model as

$$\begin{aligned} \dot{\mathbf{v}}_C &= (\mathbf{a}\mathbf{v}_C + \delta\mathbf{b}\mathbf{v}_{act} + \Delta) + \mathbf{b}\mathbf{v}_{act} \\ &= \mathbf{f} + \mathbf{b}\mathbf{v}_{act}, \end{aligned} \quad (4.7)$$

where \mathbf{f} is the general disturbance term which will be estimated by the ADRC observer. In (4.7), $\delta\mathbf{b}$ captures the errors in \mathbf{b} coming from the uncertainties in $\omega_{1/2}$ and \mathbf{g}_{sys} ; that is, $\mathbf{b} + \delta\mathbf{b}$ is the true value. The term Δ models all other disturbances like the beam loading and the RF amplifier noise. The estimates of the cavity voltage and the general disturbance term can be defined as $\hat{\mathbf{v}}_C$ and $\hat{\mathbf{f}}$. We introduce an observer to estimate these two items using the measurements of the input \mathbf{v}_{act} and the output \mathbf{v}_C :

$$\begin{bmatrix} \dot{\hat{\mathbf{v}}}_C \\ \dot{\hat{\mathbf{f}}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_C \\ \hat{\mathbf{f}} \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} \mathbf{v}_{act} + \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} (\mathbf{v}_C - \hat{\mathbf{v}}_C) \\ = \begin{bmatrix} -l_1 & 1 \\ -l_2 & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_C \\ \hat{\mathbf{f}} \end{bmatrix} + \begin{bmatrix} l_1 & \mathbf{b} \\ l_2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_C \\ \mathbf{v}_{act} \end{bmatrix}. \quad (4.8)$$

Here the parameters l_1 and l_2 need to be determined to implement the observer. The dynamics of the observer should be faster than the closed-loop response of the cavity. A fast observer is essential to estimate and reject the disturbances before they cause significant errors in the cavity fields. With a proportional controller applied to the cavity, the closed-loop pole is at the location of $\hat{s}_{CL} = -K_p \cdot \omega_{1/2}$, where K_p is the proportional gain. The gain is determined by the desired disturbance suppression ratio for the cavity field. For example, if the phase error caused by microphonics need to be suppressed by 100 times, we should choose $K_p = 100$. A typical design places all the poles of the observer at the same location of

$$\hat{s}_{OBS} = n \cdot \hat{s}_{CL}, \quad (4.9)$$

where n has a typical value between 3 to 10. To place the observer poles to the desired location (4.9), we write down the characteristic polynomial of the state-space equation (4.8) as

$$\begin{vmatrix} \lambda + l_1 & -1 \\ l_2 & \lambda \end{vmatrix} = \lambda^2 + l_1\lambda + l_2.$$

If the two poles of the observer are at the same location, the expected characteristic polynomial can be written as

$$(\lambda - \hat{s}_{OBS})^2 = \lambda - 2\lambda\hat{s}_{OBS} + \hat{s}_{OBS}^2.$$

Comparing the two equations above, we obtain the observer parameters as

$$l_1 = -2\hat{s}_{OBS}, l_2 = \hat{s}_{OBS}^2. \quad (4.10)$$

In the observer (4.8), the error in \mathbf{b} is not avoidable due to the calibration error for \mathbf{g}_{sys} . Fortunately, this error is typically static and can be suppressed by feedbacks. The dynamical disturbances estimated by the observer can be removed using an additional input to the cavity. See Fig. 4.6. Here \mathbf{K} is a PI controller given by (4.1).

Simulations were used to compare the results of a simple proportional controller and an ADRC controller. The following parameters were used in the cavity model: normalized effective shunt impedance $r/Q = 1036 \Omega$, loaded quality factor $Q_L = 3e6$, RF frequency $f_{RF} = 1.3$ GHz, desired cavity voltage $|r| = 25$ MV, beam current $|i_{b0}| = 4$ mA, and the beam is present between 510 μ s to 1110 μ s in the RF pulse. The

Fig. 4.6 An ADRC-based feedback loop. Here \mathbf{r} is the set point phasor of the cavity voltage

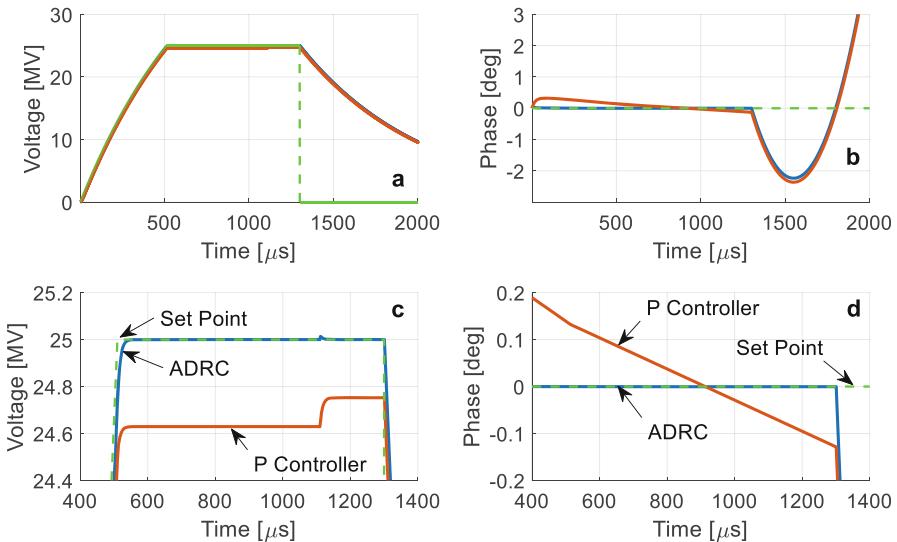
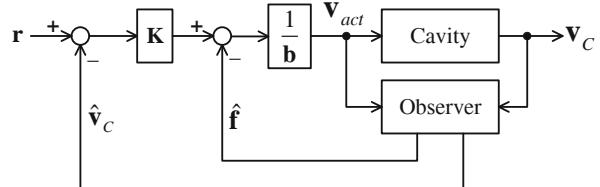


Fig. 4.7 Cavity voltage and phase waveforms resulting from the two controllers. The ADRC controller yields better suppression of the disturbances. (a). cavity voltage; (b). cavity field phase; (c). cavity voltage zoomed in at flattop; (d). cavity field phase zoomed in at flattop

results are depicted in Fig. 4.7. In the simulation, Lorentz force detuning and beam loading were added to the cavity model to check the disturbance rejection performance. For both control loops, the same parameters were used, including the same feedback gain of 100. For the ADRC controller, the pole of the observer was 10 times larger in magnitude than the closed-loop pole of the cavity.

The disturbance observer must respond faster compared to the closed-loop dynamics of the cavity. It results in the pole assignment in (4.9). To explain the results in Fig. 4.7, we shall keep in mind that the disturbances often change fast and affect the cavity fields as additional input signals. A cavity is equivalent to a narrow-band low-pass filter for the input envelope. Therefore, the cavity responds to the disturbances slowly. If we apply a step-change in a disturbance, the resulting cavity voltage error takes time to reach its maximum value. With the nominal PI controller, the feedback only relies on the cavity output measurement. In this case, the controller only applies sufficient corrections in the cavity input after the effects of disturbances become significant in the measurement. Therefore, we cannot compensate for the

disturbances “in time” due to the cavity’s large time constant. However, with the ADRC controller, the disturbances can be estimated much faster by combining the cavity input, output, and dynamics in the observer. Therefore, ADRC can suppress the disturbances in advance before their effects become significant.

4.2.4 Advanced Control Algorithms

PI control and ADRC are often sufficient for most RF control systems. This section collects several advanced control algorithms that are also used in some LLRF systems. We will only introduce these algorithms conceptually and provide reference articles to interested readers. As a general guideline, it is suggested that we only use these advanced algorithms when necessary, such as when the PI or ADRC control algorithms cannot provide the desired performance.

Modern control algorithms are based on state-space equations, which are generally written as

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} + \mathbf{m}, \\ \mathbf{y} &= \mathbf{Cx} + \mathbf{Du} + \mathbf{n},\end{aligned}\tag{4.11}$$

where \mathbf{x} , \mathbf{u} and \mathbf{y} are vectors of states, inputs, and outputs of the system, respectively. The term \mathbf{m} denotes the process disturbances and \mathbf{n} the measurement noise. Note that for a particular system, its state-space model has multiple realizations depending on the definitions of states. For example, if we define a new state vector $\mathbf{x}' = \mathbf{Px}$, where \mathbf{P} is an invertible square matrix, then we obtain a new state-space model with $\mathbf{A}' = \mathbf{PAP}^{-1}$, $\mathbf{B}' = \mathbf{PB}$, $\mathbf{C}' = \mathbf{CP}^{-1}$, $\mathbf{D}' = \mathbf{D}$, $\mathbf{m}' = \mathbf{Pm}$ and $\mathbf{n}' = \mathbf{n}$. From (4.5), one possible realization of the cavity state-space model can be written as

$$\begin{aligned}\mathbf{x} &= [\tilde{v}_{CI} \quad \tilde{v}_{CQ}]^T, \mathbf{u} = [v_{actI} \quad v_{actQ}]^T, \mathbf{y} = [v_{CI} \quad v_{CQ}]^T, \\ \mathbf{m} &= [m_I \quad m_Q]^T, \mathbf{n} = [n_I \quad n_Q]^T, \\ \mathbf{A} &= \begin{bmatrix} -\omega_{1/2} & -\Delta\omega \\ \Delta\omega & -\omega_{1/2} \end{bmatrix}, \mathbf{B} = \omega_{1/2} \begin{bmatrix} g_{sysI} & -g_{sysQ} \\ g_{sysQ} & g_{sysI} \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{D} = \mathbf{0}.\end{aligned}\tag{4.12}$$

Here $\mathbf{v}_C = v_{CI} + jv_{CQ}$, $\mathbf{g}_{sys} = g_{sysI} + jg_{sysQ}$, and $\mathbf{v}_{act} = v_{actI} + jv_{actQ}$. In this particular state-space model, we have defined the state \mathbf{x} as the actual cavity voltage (with \sim above the notations), the input \mathbf{u} as the RF drive voltage, and the output \mathbf{y} the measured cavity voltage. The process disturbances \mathbf{m} are determined by the RF driving chain noise and the Lorentz force detuning, microphonics and beam loading. They are normalized to the cavity input. Since the output \mathbf{y} does not contain any feed-through of the cavity input, we assign the matrix \mathbf{D} to zero. The matrix \mathbf{C} is a unit matrix, so \mathbf{y} is the sum of the actual cavity voltage and the measurement noise \mathbf{n} .

One should keep in mind that the state-space equation is not unique. Specifically, if we identify a black-box model for the RF system, the physical meanings of the state-space model's parameters will not be as obvious as in (4.12).

4.2.4.1 Optimal Control

Optimal control defines a control law by minimizing a cost function. *Linear quadratic Gaussian (LQG) control* is a well-working optimal control algorithm if the system dynamics are linear and known, and the measurement noise and process disturbances are uncorrelated zero-mean Gaussian stochastic processes.

The LQG control problem is to find an optimal input $\mathbf{u}(t)$ to minimize the cost function

$$J = E \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}] dt \right\}, \quad (4.13)$$

where \mathbf{Q} and \mathbf{R} are constant weighting matrices as design parameters and should be semi-positive definitive. The solution of the optimization problem results in a *linear quadratic regulator (LQR)* as a simple state feedback law

$$\mathbf{u}(t) = -\mathbf{K}_r \mathbf{x}(t), \quad (4.14)$$

where \mathbf{K}_r is a constant gain matrix determined by the system model (4.11) and the weighting matrices. Note that the LQR relies on the states \mathbf{x} that are often not well known, but we can find an optimal estimate of the states (denoted as $\hat{\mathbf{x}}$) with a *Kalman filter*. Then the control law becomes $\mathbf{u}(t) = -\mathbf{K}_r \hat{\mathbf{x}}(t)$. The LQR plus the Kalman filter (the combination is called *LQG controller*) can suppress the process disturbances and measurement noise and regulate the output \mathbf{y} around zero. To track a set point \mathbf{r} , we should modify the LQG controller by adding an error term $\mathbf{e} = \mathbf{r} - \mathbf{y}$ to the state feedback law. To do this, we integrate the error \mathbf{e} and obtain an additional state vector $\mathbf{x}_i(t) = \int_0^t \mathbf{e}(\tau) d\tau$. Then an expanded state vector can be defined as

$$\mathbf{x}'(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_i(t) \end{bmatrix}.$$

Using the expanded state, the cost function of the optimization becomes

$$J' = E \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\mathbf{x}'^T \mathbf{Q}' \mathbf{x}' + \mathbf{u}^T \mathbf{R} \mathbf{u}] dt \right\},$$

where \mathbf{Q}' is the weighting matrix for the expanded state \mathbf{x}' . The resulting LQG controller can be then written as

$$\mathbf{u}(t) = [-\mathbf{K}_x \quad -\mathbf{K}_i] \begin{bmatrix} \hat{\mathbf{x}}(t) \\ \mathbf{x}_i(t) \end{bmatrix}.$$

Here $\hat{\mathbf{x}}$ is still the Kalman filter output, and \mathbf{K}_x and \mathbf{K}_i are the optimization results for the new cost function above. This expanded LQG controller can be used for both disturbance rejection and command tracking. The control is optimal in the presence of white noise in the plant and sensors. More details of LQG control can be found in (Skogestad and Postlethwaite 2005).

LQG control was successful in controlling spacecraft but might not perform well in the control of RF systems, because the plant models are not always accurately available, or the assumption of white noise is not always valid.

4.2.4.2 Robust Control and Adaptive Control

RF controllers should be able to deal with the uncertainties in RF system model structures and parameters. In Chap. 3, we have discussed the RF system models thoroughly. The uncertainty of model structures comes mainly from the hidden (high-order) dynamics neglected in our models. For example, the cavity Eq. (4.5) models $\Delta\omega$ as the imaginary part of the cavity transfer function's pole. When designing a feedback controller, we typically neglect $\Delta\omega$ in the cavity dynamics and treat it as a perturbation to be suppressed by the feedback. However, if $\Delta\omega$ is caused by the cavity's mechanical oscillations driven by the Lorentz force, there exist dynamics from \mathbf{v}_{act} to $\Delta\omega$ that are neglected in the cavity model. The uncertainty in model parameters is easier to understand since we can only estimate the parameters like $\omega_{1/2}$ and \mathbf{g}_{sys} to a certain accuracy. Another source of uncertainties is the variations of the system dynamics and parameters due to operating condition changes. For example, the system phase, which should satisfy (4.3) for I/Q feedback stability, may drift due to temperature fluctuations.

An RF controller, especially the one designed based on the model, should be robust against the above model uncertainties. That is, we expect the controller to be still stable and perform acceptably for a real RF system that may behave slightly differently from our model used for the controller design.

Robust control is an approach to design controllers with explicit dealing with uncertainties. Through particular modeling of the uncertainties, such as describing the relative model uncertainty at various frequencies as a weighting function, robust control determines the structure and parameters of the controller. The resulting controller can provide satisfactory performance in the presence of bounded errors in the system model (Pfeiffer et al. 2012). Note that the robust controller might lead to performance degradation for the nominal system (without uncertainties). That is, we have to make a trade-off between robustness and performance.

The ADRC discussed before is a robust controller, which models the model uncertainties as part of the general disturbance term. With the feedback structure in Fig. 4.6, the effects of model uncertainties can be actively rejected.

An alternative method to mitigate the RF system model uncertainties is to employ the *adaptive control*. Instead of using fixed controller structure and parameters, as with robust control, adaptive control varies the controller settings or structure to match the plant variations. It requires a real-time identification of the system parameters, such as the system gain, system phase, and the cavity half-bandwidth and detuning. The gain scheduling (Wibowo et al. 2018), which applies different gains at different parts of the RF pulse, is an example of adaptive control. In the filling period of the cavity voltage pulse, a smaller peak drive power is expected, so we use a smaller gain. Whereas during the RF pulse flattop, lower RF field errors are required for stable beam acceleration, then a higher gain is needed.

4.2.4.3 Model Predictive Control

Model predictive control (MPC) uses the dynamical model of the system, which is typically an empirical model obtained via system identification. MPC optimizes the drive signal \mathbf{u} at a time instant $t = t_0$, taking into account the predicted system response in a limited time range ($t_0 < t < t_0 + \Delta t$, where $\Delta t > 0$) in the future. That is, MPC solves an optimization problem repeatedly to determine the drive signal of each step (i.e., sample) of a discrete system. This makes the MPC suitable for the control of processes with large delays (Blanco Viñuela et al. 2000; Qian et al. 2014). If a feedback loop contains a delay, the effects of control actions are only visible in the system output after the delay time. Before seeing the control action's effects in the measurement, the controller keeps accumulating the system output errors and generates more corrections (assume the controller contains an integrator). Finally, the controller overcorrects for the errors, resulting in a significant overshoot in the system output. The primary benefit of the MPC is that it can predict the overshoot and then lower the control actions correspondingly to reduce it.

In a discrete implementation, the MPC performs an optimization at each measurement step. The cost function for the optimization is similar to that of the LQG control. However, unlike the global optimization in LQG control, MPC only derives the optimal \mathbf{u} for the next several steps. The optimal \mathbf{u} tends to minimize the system output errors predicted by the system model. Note that only the first step of the optimal \mathbf{u} is applied to the system input. Then the system output is sampled again and the optimization process is repeated. Since optimizations are required for each step, MPC needs more computation power than other control algorithms.

One typical application of MPC in accelerator systems is the temperature control of cooling systems in normal-conducting accelerators.

4.3 Self-Excited Loop Control

Self-excited loop (SEL) is another widely used strategy for cavity control (Delayen 1978; Allison et al. 2007; Laverty et al. 2010). It is helpful when starting up a narrow-band cavity after a long-term shutdown. In that case, the resonance frequency of the cavity is often unknown. SEL can also track the cavity resonance frequency, so it is suitable for RF conditioning. A typical SEL, as depicted in Fig. 2.4, connects the cavity output directly to the RF amplifier input. If the system phase and loop gain satisfy the SEL starting conditions, the SEL loop can start up from the noise in the system, and the cavity voltage will increase exponentially. A limiter is needed to limit the RF drive power within a safe level.

4.3.1 Free-Running SEL

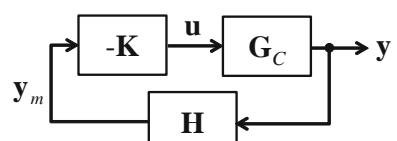
An SEL can be derived from an unstable GDR feedback loop depicted in Fig. 4.2. If the closed-loop is unstable, a small disturbance or noise in the loop (e.g., DAC or amplifier noise) will produce an output y that increases exponentially in magnitude. In this case, the reference input r is not needed any longer and we obtain an SEL as Fig. 4.8. Here we have moved the minus sign in the GDR feedback loop into the controller block for consistency.

In an SEL, we typically use a simple proportional controller $\mathbf{K} = K_P \mathbf{I}$. For a GDR control loop, the system phase should satisfy (4.3) for stable closed-loop operations. Therefore, if the system phase and the feedback gain K_P satisfy

$$\cos \varphi_{\text{sys}} < -1/K_P \text{ and } K_P > 1, \quad (4.15)$$

we will obtain an unstable closed-loop, and the resulting SEL can start up from the wideband noise in the RF system. With the definition of \mathbf{G}_C and \mathbf{H} ($g_{\text{sys}} = 1$) in (4.1), K_P is the maximum loop gain achieved at the cavity resonance frequency. The relation (4.15) is the so-called *SEL starting condition*. Note that φ_{sys} here is still defined at the GDR operating frequency and (4.15) is the condition to switch the GDR loop into an SEL. It implies that to run the loop in the SEL mode, we need a feedback gain higher than 1 and a system phase falls into a particular range. For example, if $K_P = 10$ and φ_{sys} satisfies $95.7^\circ < \varphi_{\text{sys}} < 264.3^\circ$, the SEL can be started. In principle, we obtain a positive feedback loop when it satisfies the SEL starting condition. After started up, the SEL will operate at a frequency where the

Fig. 4.8 An SEL derived from the GDR feedback loop



closed-loop transfer function's gain is maximum. If we only consider the proportional control in \mathbf{K} , the closed-loop transfer function of the GDR loop from \mathbf{r} to \mathbf{y} can be written as

$$\mathbf{T}(j\hat{\omega}) = \frac{\mathbf{L}(j\hat{\omega})}{1 + \mathbf{L}(j\hat{\omega})} = \frac{K_P \omega_{1/2} e^{-j(\varphi_{sys} + \hat{\omega}\Delta t)}}{j(\hat{\omega} - \Delta\omega) + \omega_{1/2} \left(1 + K_P e^{-j(\varphi_{sys} + \hat{\omega}\Delta t)}\right)}, \quad (4.16)$$

where $\hat{\omega}$ is the offset frequency compared to the carrier frequency. The operating frequency of the free-running SEL $\hat{\omega}_{SEL}$ (offset to the carrier frequency) corresponds to the maximum gain of \mathbf{T}_C and satisfies

$$\hat{\omega}_{SEL} = \Delta\omega + K_P \omega_{1/2} \sin(\varphi_{sys} + \hat{\omega}_{SEL} \Delta t).$$

Note that $\varphi_{sys} + \hat{\omega}_{SEL} \Delta t =: \varphi_{sys,SEL}$ is the new system phase defined at the SEL operating frequency, which is different from the GDR operating frequency. Then the formula above can be rewritten as

$$\hat{\omega}_{SEL} = \Delta\omega + K_P \omega_{1/2} \sin \varphi_{sys,SEL}. \quad (4.17)$$

If we want to know the exact $\hat{\omega}_{SEL}$ of a free-running SEL, the equation above should be solved numerically with given $\Delta\omega$, K_P , $\omega_{1/2}$, φ_{sys} and Δt .

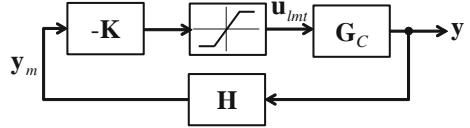
We can use the TESLA cavity as an example, assuming the following parameters: $\omega_{1/2} = 2\pi \times 200$ rad/s, $\Delta\omega = 2\pi \times 300$ rad/s, $K_P = 10$, $\varphi_{sys} = 120^\circ$, and $\Delta t = 0$ or 5 μ s. We can prove that the system phase and loop gain satisfy the SEL starting condition. Solving the equation above numerically and the resulting SEL operating frequency for $\Delta t = 0$ is 2032 Hz and for $\Delta t = 5$ μ s is 1967 Hz. Both frequencies are offset values with respect to the carrier frequency (e.g., 1.3 GHz). It is worth mentioning that the cavity voltage increases faster with a non-zero loop delay because it makes the system “more unstable”.

4.3.2 SEL with Amplitude Limiter

A free-running SEL will drive the cavity voltage to infinity, but in practice, it will not happen because of the limitation of cavity drive power. We typically add an amplitude/power limiter before the cavity input to control the maximum power that goes into the cavity, as illustrated in Fig. 4.9.

The limiter saturates if the input is above a threshold. Since the cavity voltage increases exponentially, the limiter will saturate shortly after the SEL starting up. After the limiter is saturated, the loop gain is no longer equal to K_P , and the SEL operating frequency will be different from the free-running case given by (4.17).

Fig. 4.9 SEL with an amplitude limiter



Assume that the SEL operating frequency becomes $\hat{\omega}'_{SEL}$ after the limiter saturates, and it can be calculated as $\hat{\omega}'_{SEL} = \Delta\omega + K'_P\omega_{1/2} \sin \varphi_{sys,SEL}$, where K_P is the effective loop gain given by $K'_P = |\mathbf{u}_{lmt}|/|\mathbf{y}|$. Here we still denote the system phase at the new SEL operating frequency as $\varphi_{sys,SEL}$. The cavity works at $\hat{\omega}'_{SEL}$ and reaches a steady state shortly after the limiter is saturated. The relationship between the cavity input and output in steady state can be described as $|y| = |\mathbf{u}_{lmt}| \cos \psi$ with ψ the detuning angle. Therefore, we have $K'_P = |\mathbf{u}_{lmt}|/|\mathbf{y}| = 1/\cos \psi$. The detuning angle also satisfies

$$\tan \psi = (\Delta\omega - \hat{\omega}'_{SEL})/\omega_{1/2} = -K'_P \sin \varphi_{sys,SEL}.$$

Furthermore, we get $\sin \varphi_{sys,SEL} = -\sin \psi$, and $K'_P = 1/|\cos \varphi_{sys,SEL}|$. Then the SEL operating frequency with a saturated amplitude limiter can be rewritten as

$$\hat{\omega}'_{SEL} = \Delta\omega + \omega_{1/2} \sin \varphi_{sys,SEL}/|\cos \varphi_{sys,SEL}|. \quad (4.18)$$

For the example of TESLA cavity discussed in the last section, the SEL operating frequency becomes 646 Hz for $\Delta t = 0$ and 631 Hz for $\Delta t = 5 \mu\text{s}$ if an amplitude limiter is used. Therefore, the SEL operating frequency will experience a jump before and after the limiter saturates.

After the limiter is saturated, the SEL loop gain at its operating frequency is 1, which includes K_P and the steady-state cavity gain $\cos \psi$. Furthermore, since $\sin \varphi_{sys,SEL} = -\sin \psi$, the following relationship exists:

$$\varphi_{sys,SEL} + \psi = 180^\circ + k \times 360^\circ, \text{ } k \text{ is an integer.} \quad (4.19)$$

Considering the minus sign of \mathbf{K} in Fig. 4.9, we conclude that the loop phase at the SEL operating frequency after the limiter is saturated is multiple of 360° .

According to (4.19), if we adjust the system phase via the phase shifter in Fig. 2.4, the SEL operating frequency will be changed for proper detuning angles. That is, the SEL will vary its operating point on the cavity resonance curve. On the other hand, if the system phase is fixed, the SEL operating frequency will follow the cavity resonance frequency for a constant detuning angle. Such SEL's capability to track the cavity resonance frequency can also be explained by (4.18) directly, where $\hat{\omega}'_{SEL}$ varies approximately linearly with the cavity detuning $\Delta\omega$.

4.3.3 SEL with Feedback Control

In an RF controller, we may implement two separate loops for the SEL and GDR controls (Allison et al. 2007). See the top diagram of Fig. 4.10. This solution is straightforward for implementation, but the smooth switching between the two parallel loops might be difficult. We need a procedure to switch from SEL to GDR without large transients, e.g., ramping the GDR gain with time to perform the transition. Alternatively, we can apply GDR-based feedback directly to the SEL (Laverty et al. 2010). See the bottom diagram of Fig. 4.10. In this case, the overall SEL is equivalent to a new plant to be controlled.

We will focus on the cascaded loop in Fig. 4.10 (bottom), where a GDR controller \mathbf{K}_C controls the SEL with \mathbf{u}_{FB} applied to the cavity input. We use \mathbf{u}_{SEL} to denote the cavity drive term produced by the SEL. The minus sign in the \mathbf{K} block of the SEL (see Fig. 4.9) has been moved to \mathbf{u}_{SEL} , and then the overall cavity drive is $\mathbf{u} = \mathbf{u}_{FB} - \mathbf{u}_{SEL}$. In the outer loop, the block \mathbf{H}' is written as

$$\mathbf{H}'(\hat{s}) = g_{sys} e^{-j\varphi'_{sys}} e^{-\hat{s}\Delta t},$$

where φ'_{sys} is the system phase of the outer loop that should be adjusted into the stable region given by (4.3). In contrast, the system phase term in the SEL block \mathbf{H} should satisfy (4.15). Here we have assumed that both the inner and outer loops

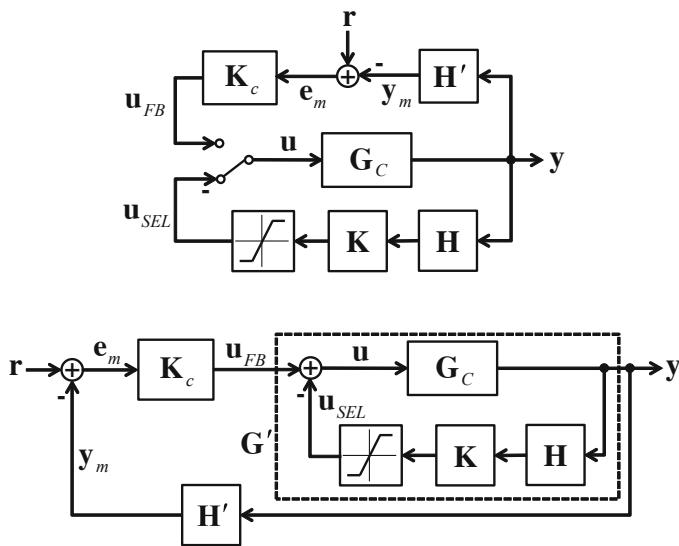


Fig. 4.10 Two structures of RF controller including both GDR and SEL controls. Top: parallel loops; Bottom: cascaded loops with SEL as an inner loop

have the same system gain (typically normalized to 1) and loop delay. Until the amplitude limiter is saturated, the new system \mathbf{G}' is unstable and its transfer function can be written as

$$\mathbf{G}'(\hat{s}) = \frac{\mathbf{G}_C}{1 + \mathbf{G}_C \mathbf{H} \mathbf{K}} = \frac{\omega_{1/2}}{\hat{s} + \omega_{1/2}(1 + K_P e^{-j\varphi_{sys} - \hat{s}\Delta t}) - j\Delta\omega}. \quad (4.20)$$

Here we have used the transfer functions (4.1), and only kept the proportional control in \mathbf{K} , and defined $g_{sys} = 1$. The pole of \mathbf{G}' is located in the right half-plane (RHP) of \hat{s} . Therefore, we must apply feedbacks via \mathbf{K}_C to stabilize this unstable plant. The loop delay in \mathbf{H}' and the RHP pole in \mathbf{G}' will limit the feedback performance. They are summarized as follows:

1. The loop delay Δt limits the maximum closed-loop control bandwidth $\hat{\omega}_B$, which should satisfy $\hat{\omega}_B < 1/\Delta t$ allowing for a phase margin greater than 30° .
2. The closed-loop bandwidth should be larger than twice of the RHP pole's real part to stabilize the plant \mathbf{G}' (amplitude limiter not saturated), which implies $\hat{\omega}_B > -2\omega_{1/2}(1 + K_P \cos \varphi_{sys})$.
3. The RHP pole requires more RF drive power to regulate \mathbf{G}' .

Compared to the GDR control loop in Fig. 4.10 (top), the feedback of a free-running SEL requires faster responses and larger RF drive power for the same performance.

After the amplitude limiter is saturated, the drive term \mathbf{u}_{SEL} has a constant amplitude but a time-varying phase since the SEL operating frequency tracks the cavity resonance frequency. However, because \mathbf{u}_{SEL} is derived from the cavity voltage directly, its phase changes slowly due to the narrow bandwidth of the cavity. Therefore, \mathbf{u}_{SEL} is equivalent to a slow disturbance applied to the cavity input. When closing the outer loop, we often lower the amplitude limiter threshold so that it saturates when the cavity voltage is still small. Then the feedback controller \mathbf{K}_C will dominate the cavity drive. As a result, after the amplitude limiter is saturated, the dynamics of the cascaded feedback loop (Fig. 4.10, bottom) will be very similar to that of the simple GDR control loop (Fig. 4.10, top).

The example below demonstrates the cascaded feedback control of an SEL. We also show the results of a simple GDR control for comparison.

Example 4.1 Cascaded Feedback Control of SEL

A cavity with the following parameters is controlled by an SEL: the half-bandwidth $\omega_{1/2} = 2\pi \times 200$ rad/s, detuning $\Delta\omega = 1.5\omega_{1/2}$, loop gain $K_P = 10$, loop delay $\Delta t = 10 \mu s$, and the system phase is set to $\varphi_{sys} = 120^\circ$. As shown in Fig. 4.11, we use a PI controller to regulate the SEL. The controller's transfer function is.

$$\mathbf{K}_C(\hat{s}) = K_0 \left(1 + \frac{1}{\tau_I \hat{s}} \right).$$

When controlling a free-running SEL, the closed-loop bandwidth of the feedback loop should be between 1.6 kHz (lower limit for stabilizing the RHP pole) and

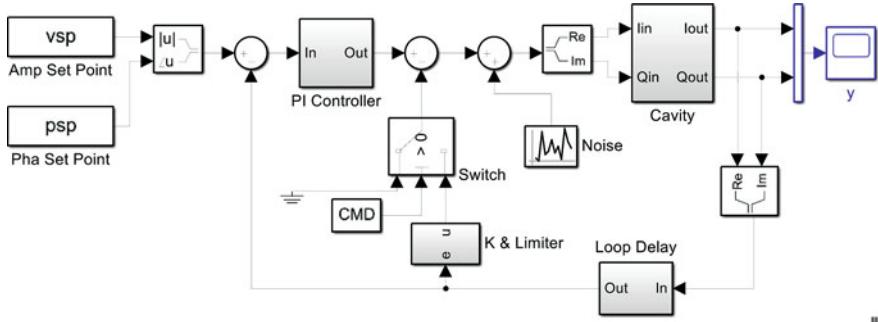


Fig. 4.11 Matlab/Simulink model of the feedback control of an SEL

16 kHz (upper limit imposed by the loop delay). Considering the open-loop transfer function (4.20), the gain K_0 should be between 8 and 80 to achieve the desired closed-loop bandwidth. In the simulation, we select the control parameters as $K_0 = 20$ and $\tau_I = 0.01$ second. The integrator helps to remove the steady-state errors.

With the same parameters of the cavity and controller, the Simulink model was simulated for two configurations: simple GDR, and SEL with cascaded feedback. These two configurations were selected with the “switch” block in the Simulink model. Figure 4.12 illustrates the simulation results of the cavity voltage and cavity drive.

When simulating the cascaded feedback of SEL, we have defined the amplitude limiter threshold to be 1; that is, the limiter saturates when the cavity voltage reaches 0.1 considering the SEL loop gain K_p . The simulation results illustrate that before the limiter is saturated, the controller stabilizes the unstable free-running SEL, resulting in larger errors and higher cavity drive. After the limiter is saturated, the cavity voltage and cavity drive behave similarly to the simple GDR control loop.

4.4 Phase-Locked Loop Control

The phase-locked loop (PLL) is another widely used strategy for RF systems control (Dai et al. 2013; Fu et al. 2020). We often implement PLLs in cavity test stands to test or condition superconducting cavities with very narrow bandwidths. We need to keep the cavity filled with sufficient RF power to perform high power RF testing or conditioning. Since the cavity is often detuned or the tuner is not ready to tune the cavity, the PLL should adjust the RF operating frequency to follow the cavity resonance frequency.

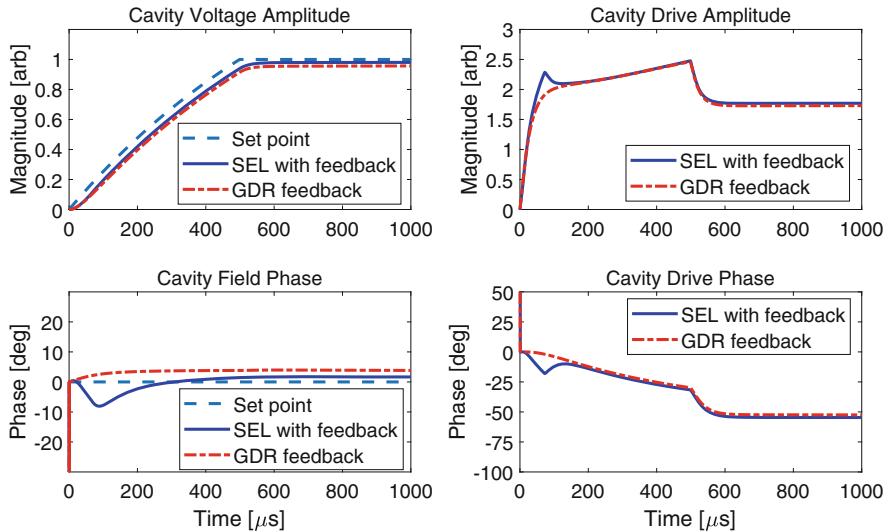


Fig. 4.12 Simulation results of the cavity voltage (left) and cavity drive (right). “SEL with feedback” labels the results of the cascaded loop with the SEL closed, whereas “GDR feedback” stands for the results when the SEL is open and the cavity is controlled only by the outer loop

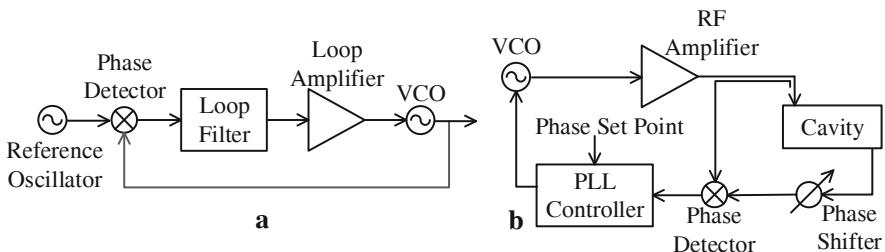


Fig. 4.13 Block diagram of PLLs. (a): a general PLL; (b): a PLL for cavity control

4.4.1 Introduction to PLL

PLL is well known in telecommunication systems for frequency synthesis, RF signal synchronization and clock recovery. Figure 4.13a illustrates a general block diagram of a PLL, in which a frequency-tunable oscillator, such as a voltage-controlled oscillator (VCO), is the source of frequencies. The PLL synchronizes the frequency and phase of the VCO with an external reference input (e.g., a reference oscillator). A phase detector compares the instant phases of the VCO output and the reference. The phase errors are filtered and amplified to adjust the VCO output frequency. After locking, the VCO output will have a desired frequency and phase with respect to the reference.

In a general PLL, as depicted in Fig. 4.13a, the parameter handled by the feedback loop is a phase. The phase is a periodic function; that is, for phase differences $\Delta\varphi$ and $\Delta\varphi + 2\pi$, the phase detector will give the same output. Since the frequencies of the reference signal and the VCO output might be different before the loop is locked, the PLL locking procedure may be non-linear. Furthermore, the reference signal and the VCO output usually have sufficient power so that the phase measurement is deterministic.

The PLL used for cavity control has a slightly different structure. See Fig. 4.13b. We can measure the cavity detuning by comparing the phases between the cavity input and output. Using the phase errors, the controller adjusts the VCO frequency to generate an RF drive filling the cavity in resonance. We use the phase shifter to pre-tune the VCO to produce a frequency close to the cavity resonance frequency before closing the feedback loop. Compared to the general PLL in Fig. 4.13a, the PLL for cavity control has some unique features. First, when the VCO frequency is far from the cavity resonance frequency, the cavity output is close to zero, for which the phase detector output is not defined. Second, the two inputs of the phase detector, i.e., the cavity input and output signals, have the same frequency. From the cavity step response discussed in Sect. 3.3.3, the phase difference between the cavity output and input is between -180° and 180° , taking into account the transient response. This feature avoids the non-linear locking procedure in the general PLL, enabling us to use a simple linear model to study the cavity-control PLL.

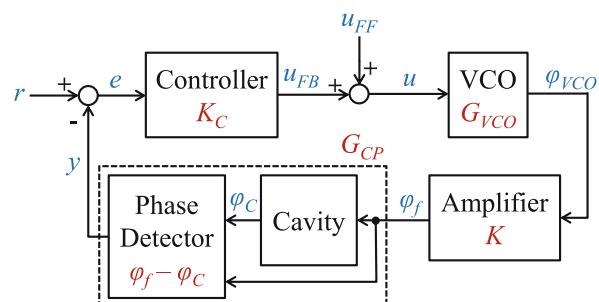
4.4.2 Modeling of PLL for Cavity Control

The PLL controller determines the behavior and performance of the PLL, such as how fast the VCO frequency can track the cavity resonance frequency. In this section, we will derive a model for the cavity control PLL. For this purpose, we redraw the PLL block diagram as Fig. 4.14.

The phase detector output is the phase difference between the cavity input and output, which is given by

$$y = \varphi_f - \varphi_C,$$

Fig. 4.14 Control system block diagram of the PLL for cavity control



where φ_f is the phase of the cavity drive (input) and φ_C of the cavity voltage (output). Such a definition results in a positive gain between the input φ_f and output y of the combined plant G_{CP} . To derive the phase transfer functions, we define the carrier frequency ω_c for the \hat{s} -domain analysis as

$$\omega_c = \omega_{VCO0} := \omega_{VCO}|_{u=0},$$

where ω_{VCO0} is the VCO frequency when its input u is zero. The reference r is the desired value of y . The controller K_C regulates the VCO frequency via the drive term u_{FB} to minimize the phase error e . Generally, we can use the same PI controller as in Example 4.1 for controlling the PLL. The feedforward term u_{FF} is produced by the phase shifter in Fig. 4.13b to pre-tune the VCO frequency. We can also use u_{FF} to scan the VCO frequency in open loop to find out the cavity resonance frequency, which is helpful when the cavity is far detuned. If we neglect the VCO's non-linearity, its output frequency can be calculated as

$$\omega_{VCO} = \omega_{VCO0} + \Omega u = \omega_c + \Omega u,$$

where Ω is the gain of the VCO. Recall the definition of the \hat{s} -domain frequency $\hat{\omega} = \omega - \omega_c$, we define the offset frequency of the VCO as $\hat{\omega}_{VCO} = \Omega u = d\varphi_{VCO}/dt$, where φ_{VCO} is the phase of the VCO output signal concerning the carrier frequency. Therefore, the transfer function of the VCO can be presented as

$$G_{VCO}(\hat{s}) = \frac{\Phi_{VCO}(\hat{s})}{U(\hat{s})} = \frac{\Omega}{\hat{s}}, \quad (4.21)$$

where $\Phi_{VCO}(\hat{s})$ and $U(\hat{s})$ are the Laplace transforms of the baseband signals φ_{VCO} and u , respectively. Since the RF amplifier bandwidth is typically much larger than the cavity bandwidth, we can neglect the amplifier's dynamics and model it as a constant phase shift combined with a group delay. If we further normalize the phase shift to zero, the transfer function of the amplifier can be written as

$$K(\hat{s}) = \frac{\Phi_f(\hat{s})}{\Phi_{VCO}(\hat{s})} = e^{-\hat{s}\Delta t}, \quad (4.22)$$

where Δt is the loop delay of the PLL. The transfer function of the plant G_{CP} relies on the cavity model given by the differential equation

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = \omega_{1/2}\mathbf{v}_f. \quad (4.23)$$

To derive the transfer function of phases, we can write the cavity voltage and cavity drive phasors in the polar format: $\mathbf{v}_C = v_C e^{j\varphi_C}$, $\mathbf{v}_f = v_f e^{j\varphi_f}$. The derivative of the cavity voltage can then be written as $\dot{\mathbf{v}}_C = \dot{v}_C e^{j\varphi_C} + jv_C e^{j\varphi_C} \dot{\varphi}_C$. Apply these

expressions to (4.23) and separate the real and imaginary parts, we get the differential equations of the cavity amplitude and phase

$$\begin{aligned}\dot{v}_C + \omega_{1/2} v_C &= \omega_{1/2} v_f \cos(\varphi_C - \varphi_f), \\ \dot{\varphi}_C - \Delta\omega &= -\frac{\omega_{1/2} v_f}{v_C} \sin(\varphi_C - \varphi_f).\end{aligned}\quad (4.24)$$

The polar differential equations above are not linear. To obtain the phase transfer function G_{CP} , we need to linearize the second differential equation considering the following conditions:

- (a) If the cavity is powered at a frequency close to its resonance frequency and already reaches a steady state, there exist the relations $v_C \approx v_f$ and $\varphi_C \approx \varphi_f$.
- (b) If the cavity is far detuned concerning the VCO frequency, the cavity voltage v_C will be very small and $\varphi_f - \varphi_C$ will be oscillatory in the transient and then approaches $\pm 90^\circ$ after reaching a steady state.

The PLL should avoid working in the condition *b*. Instead, we should disable the PLL feedback if the cavity voltage is too small. In practice, the PLL should be pre-tuned in open loop so that the cavity voltage is large enough for meaningful phase measurement. Then the feedback can be applied to track the cavity resonance frequency automatically. If the PLL operates under the condition *a*, we can simplify the phase differential equation as

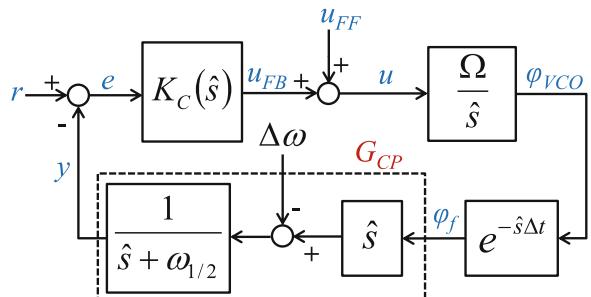
$$\dot{\varphi}_C - \Delta\omega \approx -\omega_{1/2}(\varphi_C - \varphi_f).$$

Considering the definition of y at above, we obtain

$$\dot{y} + \omega_{1/2}y = \dot{\varphi}_f - \Delta\omega. \quad (4.25)$$

Combining the analysis in this section, we can further elaborate the PLL model as Fig. 4.15. Here we replaced the blocks with their transfer functions except for the controller that will be discussed in the next section.

Fig. 4.15 Phase transfer model of the PLL for cavity control



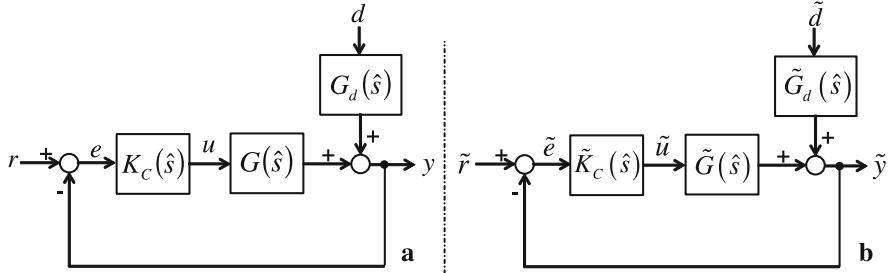


Fig. 4.16 Feedback model of PLL. (a). Nominal loop model; (b). Scaled loop model

4.4.3 Feedback Analysis of PLL

The design of a controller should take into account the desired feedback performance and the disturbances. Here we only consider the detuning $\Delta\omega$ as a disturbance. In this section, we will follow the controller design procedure (Skogestad and Postlethwaite 2005) for analyzing and designing K_c .

The PLL transfer model can be converted to a format suitable for feedback design, see Fig. 4.16a. We denote the disturbance as d to represent the cavity detuning. The feedforward term u_{FF} has been removed because it does not affect the feedback performance. The transfer functions of G and G_d are presented as

$$G(\hat{s}) = \frac{\Omega e^{-\hat{s}\Delta t}}{\hat{s} + \omega_{1/2}}, \quad G_d(\hat{s}) = \frac{1}{\hat{s} + \omega_{1/2}}, \quad d = -\Delta\omega. \quad (4.26)$$

We will focus on defining the control goals and analyzing the *input-output controllability* in terms of limitations of loop parameters. They are the basis for the controller design. The concrete design can be done with software tools like Matlab.

4.4.3.1 Control Goals Analysis

Control goals are defined in terms of the maximum allowed deviations of the values of r , e , u , y and d , such as the maximum allowed phase error (e) or VCO input (u).

The largest expected change of the reference r (denoted as r_{\max}) describes how far we may run the cavity off resonance. From the cavity model, it should satisfy

$$r_{\max} = \pi/2, \quad \hat{\omega}_r = h_r \omega_{1/2}, \quad (4.27)$$

where $\hat{\omega}_r$ is the maximum frequency of r . Note that r is usually a constant to keep the cavity operating at a particular location on the resonance curve. Still, it may be time varying to modulate the RF frequency. We define the maximum expected frequency

of r in terms of h_r times the cavity bandwidth. Similarly, the largest expected magnitude of the disturbance d is defined as

$$d_{\max} = k_d \omega_{1/2}. \quad (4.28)$$

Here we define that the maximum cavity detuning (i.e., d) caused by the disturbances is k_d times the cavity bandwidth. For superconducting cavities, the disturbance comes mainly from the Lorentz force detuning and microphonics. The frequency of microphonics is determined by the external vibration frequencies and the mechanical oscillation modes of the cavity, which may be up to several kilohertz or even higher. Here we assume that the disturbance d is a wideband noise for simplicity because it is anyway filtered by the narrow-band block G_d before it is added to the cavity output.

As a control goal, we need to define the maximum allowed phase error e_{\max} . The maximum error typically happens when the reference and disturbance change with the largest magnitudes and highest frequencies. Particularly, if we allow the largest tracking error between the VCO frequency and the cavity resonance frequency to be $\eta \omega_{1/2}$, then e_{\max} can be calculated as

$$e_{\max} = \arctan \eta. \quad (4.29)$$

In a practical PLL, the VCO drive u is typically generated by a DAC or a voltage amplifier. The maximum available drive u_{\max} is a critical constraint when designing the controller.

With the maximum values above, we can scale the loop variables to normalize their maximum magnitudes to 1:

$$\begin{aligned} \tilde{e} &= e/e_{\max}, \tilde{u} = u/u_{\max}, \tilde{d} = d/d_{\max} \\ \tilde{r} &= r/e_{\max}, \tilde{y} = y/e_{\max}. \end{aligned} \quad (4.30)$$

Note that the variables r and y are scaled with the maximum error, so their normalized magnitudes may be larger than 1. The largest value of \tilde{r} is $R = r_{\max}/e_{\max}$. With these scaled variables, we can normalize the transfer functions as

$$\tilde{G}(\hat{s}) = \frac{u_{\max}}{e_{\max}} G(\hat{s}), \tilde{G}_d(\hat{s}) = \frac{d_{\max}}{e_{\max}} G_d(\hat{s}), \tilde{K}_C(\hat{s}) = \frac{e_{\max}}{u_{\max}} K_C(\hat{s}). \quad (4.31)$$

Then the PLL feedback loop changes to Fig. 4.16b.

In terms of scaled variables we have $|\tilde{r}| \leq R$ and $|\tilde{d}| \leq 1$, and our PLL control goal is to manipulate \tilde{u} with $|\tilde{u}| < 1$ such that $|\tilde{e}| < 1$. The control goal should be satisfied at all concerned frequencies of the reference and disturbance.

4.4.3.2 Controllability Analysis

The input-output controllability is represented by the limitations of loop parameters. The major concern includes the required closed-loop bandwidth, the maximum allowed loop delay and the minimum VCO gain.

The closed-loop bandwidth $\hat{\omega}_B$ is a measure of the PLL's capability for command tracking and disturbance rejection. First, the controller should be able to suppress the disturbances up to the crossover frequency $\hat{\omega}_d$ of $\tilde{G}_d(\hat{s})$, which is defined by $|\tilde{G}_d(j\hat{\omega}_d)| = 1$. If $d_{\max}/e_{\max} \gg \omega_{1/2}$, then we have $\hat{\omega}_d \approx d_{\max}/e_{\max}$ approximately. Therefore, for disturbance rejection, $\hat{\omega}_B$ must satisfy

$$\hat{\omega}_B > \hat{\omega}_d \approx d_{\max}/e_{\max}. \quad (4.32)$$

For example, assume the cavity detuning caused by the microphonics is a wideband disturbance with its magnitude up to $d_{\max} = 10\omega_{1/2}$. If we expect the control error to be below $e_{\max} = 0.1$ rad for the disturbances at all frequencies, we need a closed-loop bandwidth larger than $100\omega_{1/2}$ approximately.

The achievable closed-loop bandwidth is limited by the loop delay. For an expected $\hat{\omega}_B$, the loop delay should satisfy

$$\Delta t < 1/\hat{\omega}_B, \quad (4.33)$$

which implies a phase-lag of 57.3° at the frequency $\hat{\omega}_B$ caused by the delay. Typically, the phase margin of the PLL should be larger than 30° to avoid overshoots and oscillations in the step response. This requires that the phase-lag caused by the loop delay should be smaller than 60° in addition to the phase-lag (90°) caused by the cavity dynamics. To track the \tilde{r} changes at all concerned frequencies with $|\tilde{e}| < 1$, we need the loop gain $|\tilde{K}_C \tilde{G}| > R - 1$ up to the frequency $\hat{\omega}_r$. That is, we need sufficient loop gains at all possible frequencies of \tilde{r} . Here R is the maximum value of \tilde{r} . It imposes a constraint to the gains of the VCO and controller:

$$\Omega |\tilde{K}_C(j\hat{\omega}_r)| > (R - 1) \frac{e_{\max}}{u_{\max}} \sqrt{\hat{\omega}_r^2 + \omega_{1/2}^2}. \quad (4.34)$$

Here we have assumed that the gain of \tilde{K}_C becomes smaller at larger frequencies.

The gain of the plant $\tilde{G}(\hat{s})$ should allow suppressing the disturbances and tracking the reference changes with an input \tilde{u} bounded by $|\tilde{u}| < 1$. Therefore, the following relations are required

$$|\tilde{G}(j\hat{\omega})| > |\tilde{G}_d(j\hat{\omega})|, \quad \forall \hat{\omega} < \hat{\omega}_d; \quad |\tilde{G}(j\hat{\omega})| > R - 1, \quad \forall \hat{\omega} < \hat{\omega}_r,$$

which further require the VCO gain to be

$$\Omega > d_{\max}/u_{\max} \text{ and } \Omega > (R - 1) \frac{e_{\max}}{u_{\max}} \sqrt{\hat{\omega}_r^2 + \omega_{1/2}^2}. \quad (4.35)$$

The above controllability analysis provides a guideline for the controller design. We will use an example to demonstrate the design procedure of the PLL controller.

Example 4.2 PLL Feedback Controller Design

In this example, a PLL is used to condition a TESLA cavity with a half-bandwidth $\omega_{1/2} = 2\pi \times 200 \text{ rad/s}$. Our goal is to control the relative phase between the cavity input and output within 0.1 rad ($e_{\max} = 0.1$). The maximum VCO input should be smaller than 1 Volt ($u_{\max} = 1$). The magnitude of the setpoint is up to $\pi/2$ ($r_{\max} = \pi/2$) with its frequency up to the cavity half-bandwidth ($\hat{\omega}_r = \omega_{1/2}$). The cavity detuning may be up to 10 times the cavity half-bandwidth ($d_{\max} = 10\omega_{1/2}$) due to strong microphonics. Here $R = r_{\max}/e_{\max} = 5\pi$.

The scaled transfer functions are written as.

$$\tilde{G}(\hat{s}) = \frac{10\Omega e^{-\hat{s}\Delta t}}{\hat{s} + \omega_{1/2}}, \tilde{G}_d(\hat{s}) = \frac{100\omega_{1/2}}{\hat{s} + \omega_{1/2}}.$$

The controllability analysis imposes the following limitations to the feedback loop

- From (4.32), the closed-loop bandwidth should satisfy $\hat{\omega}_B > \hat{\omega}_d \approx 100\omega_{1/2}$.
- From (4.33), the loop delay Δt should be smaller than $7.9 \mu\text{s}$.
- The VCO gain should satisfy $\Omega > 10\omega_{1/2}$ according to (4.35). It means that a 1 V input will generate a frequency change more than 2 kHz.

Here we select a VCO with a gain $\Omega = 20\omega_{1/2} \text{ rad/s/V}$. Considering the typical delay in a digital implementation, we assume the delay $\Delta t = 3 \mu\text{s}$. This assumption imposes an upper limit to the closed-loop bandwidth as $\hat{\omega}_B < 265\omega_{1/2}$. Therefore, the closed-loop bandwidth should satisfy $100\omega_{1/2} < \hat{\omega}_B < 265\omega_{1/2}$.

We choose a PI controller $\tilde{K}_C(\hat{s}) = K_0[1 + 1/(\tau_I \hat{s})]$. The crossover frequency of the open-loop transfer function $L(\hat{s}) = \tilde{K}_C(\hat{s})\tilde{G}(\hat{s})$ approximates the closed-loop bandwidth, which can be estimated by solving $|L(j\hat{\omega}_B)| = 1$. If the effect of the integral control is negligible at $\hat{\omega}_B$, then we get $\hat{\omega}_B \approx 10\Omega K_0 = 200K_0\omega_{1/2}$. According to the range of $\hat{\omega}_B$, the gain K_0 should satisfy $0.5 < K_0 < 1.32$. Here we choose $K_0 = 0.6$. It can be proven that (4.34) is satisfied regardless the integral time constant τ_I . Here we choose $\tau_I = 500 \mu\text{s}$ that provides more suppression of the disturbances slower than $1.6\omega_{1/2}$.

The Bode plots of the transfer functions $\tilde{G}(\hat{s})$, $\tilde{G}_d(\hat{s})$ and $L(\hat{s})$ are illustrated in Fig. 4.17. The closed-loop bandwidth is about $120\omega_{1/2}$. The gain margin is over 10 dB and the phase margin is larger than 60° .

We used a Simulink model to verify the control performance, see Fig. 4.18. It is a direct realization of the loop in Fig. 4.15. The variables here are not scaled. The PLL controller has been converted to the unscaled version by $K_C(\hat{s}) = \tilde{K}_C(\hat{s})u_{\max}/e_{\max}$. The largest error of the PLL occurs when both the reference and

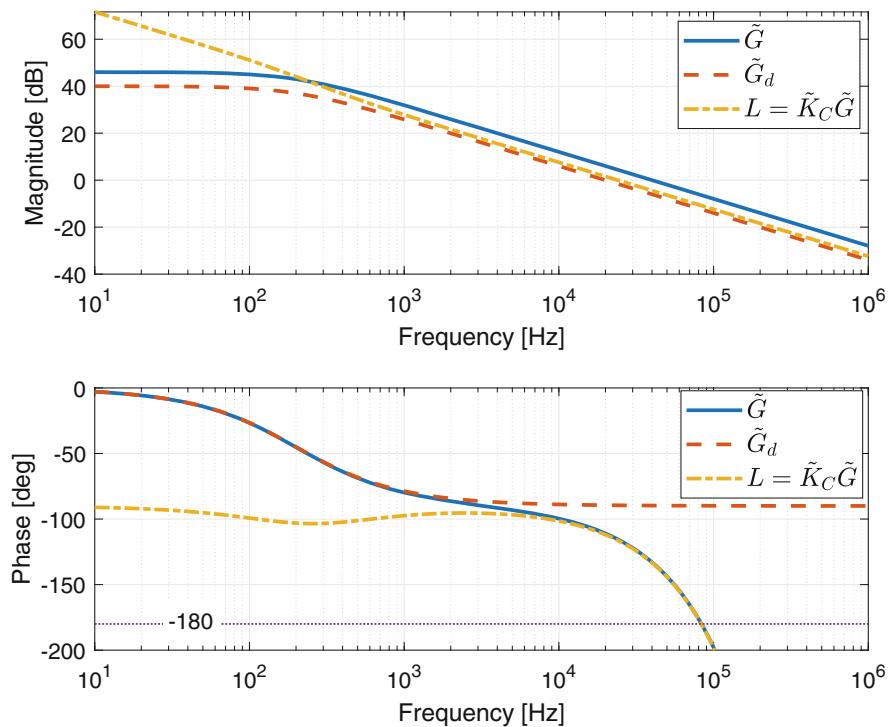


Fig. 4.17 Bode plots of the open-loop transfer functions of the PLL

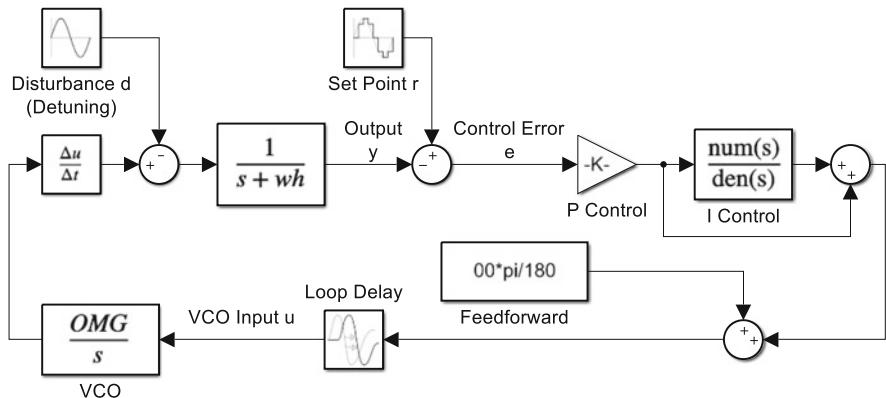


Fig. 4.18 Simulink model of the PLL

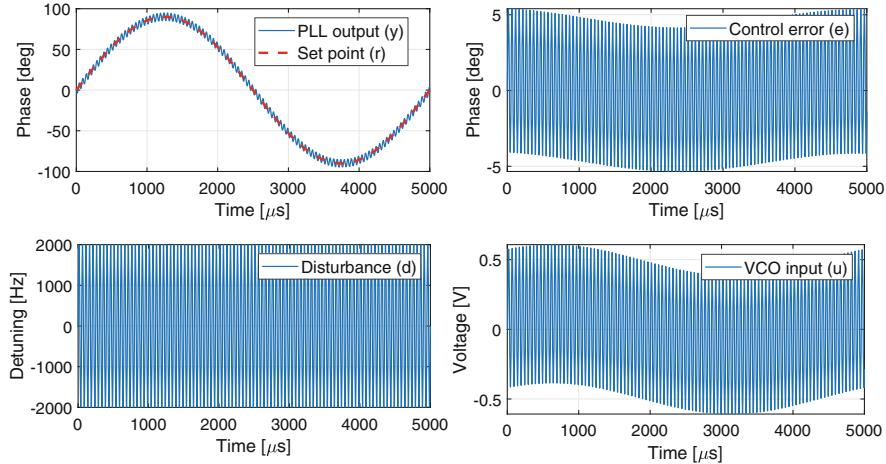


Fig. 4.19 Simulation of PLL with reference and disturbance at largest magnitude and frequency

disturbance change by the maximum magnitudes at the highest frequencies. Figure 4.19 shows the PLL output, control error and VCO input under such extreme conditions. The reference changes at $\hat{\omega}_r$ with an amplitude of r_{max} and the disturbance changes at $\hat{\omega}_d$ with an amplitude of d_{max} . The results show that the control error is smaller than e_{max} (5.7°) while keeping the VCO input below 1 V.

4.5 Adaptive Feedforward

Adaptive feedforward (AFF) is often used in LLRF systems operating in the pulsed mode. If the pulses and the disturbances are repetitive, then the feedforward control can be continuously improved by learning from previous pulses. The Lorentz force detuning and beam loading are well-known repetitive disturbances in a cavity.

AFF tries to determine the system inputs based on the output errors, aiming at achieving a desired RF pulse shape. In most cases, constant amplitude and phase in the pulse are expected for an equal acceleration of a bunch train. AFF often involves an “inversion” of the RF system model (Pfeiffer et al. 2020). Typically, a direct inversion of the system model is not practical. It requires an exact system model and often involves calculating the derivative that is sensitive to the noise. This section introduces several AFF algorithms that approximately “inverse” the system model to adapt the feedforward signals.

Generally, AFF is presented as an iterative algorithm

$$\begin{aligned}\mathbf{u}(n+1, k) &= f(\mathbf{u}(n, k), \mathbf{e}(n, k)) \\ \mathbf{e}(n, k) &= \mathbf{r}(k) - \mathbf{y}(n, k)\end{aligned} \quad k = 0, 1, 2 \dots K-1; \quad n = 0, 1, 2 \dots \quad (4.36)$$

where \mathbf{u} is the input and \mathbf{y} is the output of the RF system. The RF field error \mathbf{e} is with respect to a setpoint \mathbf{r} . The term n is the RF pulse index and k the sample index in a pulse. Here we assume that each pulse contains K samples. AFF takes the input and the output error of the n^{th} pulse and determines the input for the $(n+1)^{\text{th}}$ pulse using an algorithm f .

4.5.1 Adapt Feedforward with Feedback Actuation

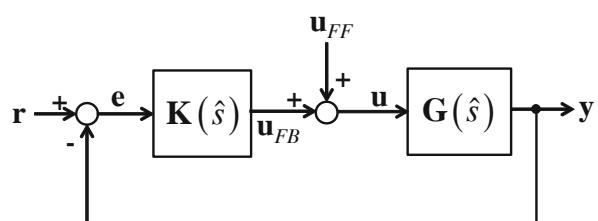
If the intra-pulse feedback is applied, the feedforward signal can be corrected iteratively with the feedback controller's actuation signal (Brandt 2007). A feedback controller regulates the system output to follow the setpoint. Due to limited feedback gain, we can only expect the feedback to suppress random errors. Without AFF, the feedback also partly compensates for the repetitive errors. If we add the feedback output of each pulse to the feedforward signal, then the repetitive errors to be suppressed are smaller in the next pulse. After several iterations, the feedforward signal can compensate for most of the repetitive errors. This concept is depicted in Fig. 4.20.

Figure 4.20 describes the control of an RF system \mathbf{G} with a feedback controller \mathbf{K} and a feedforward input \mathbf{u}_{FF} . The variables are time-domain phasors and we will use their capital version to represent their Laplace transforms for each RF pulse. The feedforward signal is updated after each RF pulse by adding the output of the feedback controller, \mathbf{u}_{FB} . Assume that the original feedforward is zero ($\mathbf{U}_{FF}(0) = \mathbf{0}$), then the AFF algorithm can be written in transfer functions as

$$\begin{aligned}\mathbf{U}_{FB}(n) &= \frac{\mathbf{K}}{1 + \mathbf{GK}} \mathbf{R} - \frac{\mathbf{GK}}{1 + \mathbf{GK}} \mathbf{U}_{FF}(n), \\ \mathbf{U}_{FF}(n+1) &= \mathbf{U}_{FF}(n) + \mathbf{U}_{FB}(n).\end{aligned} \quad n = 0, 1, 2 \dots \quad (4.37)$$

Here we dropped the variable \hat{s} but only kept the pulse index n for simplicity. At the beginning of each pulse, we assume that the states and outputs of the RF system are

Fig. 4.20 Adaptive feedforward with feedback controller output



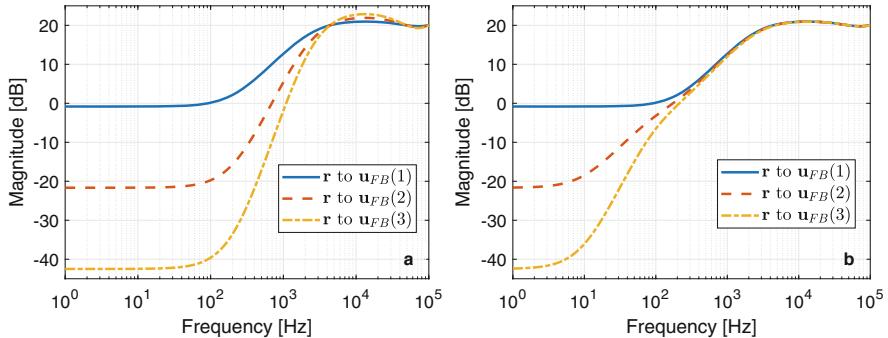


Fig. 4.21 Bode plots of the transfer function from \mathbf{r} to $\mathbf{u}_{FB}(n)$. In this example, the cavity transfer function follows \mathbf{G}_C in (4.1) with a 200 Hz half-bandwidth and a zero detuning. \mathbf{K} contains only proportional control with $K_P = 10$ and the loop delay is set to 10 μs

completely reinitialized: no RF fields in the cavity and no RF drive. This allows us to calculate the feedback controller output for each pulse with the above closed-loop transfer functions without considering the previous pulses, see the first equation of (4.37). However, if we adapt the feedforward signal following the second equation of (4.37), the system behavior in a pulse will rely on all previous pulses.

The iterative algorithm (4.37) has been proved not stable. The transfer function from \mathbf{r} to $\mathbf{u}_{FB}(n)$ can be derived recursively. It contains a high-order term $(1 + \mathbf{GK})^{-n}$ that leads to a poor phase margin resulting in instability. This is demonstrated with the example in Fig. 4.21a, which shows the transfer function from \mathbf{r} to $\mathbf{u}_{FB}(n)$ for $n = 1, 2$ and 3 . It depicts that the gain from \mathbf{r} to \mathbf{u}_{FB} generally decreases when n increases, which means less correction is required from the feedback to achieve the setpoint. This also implies that the feedforward signal itself can generate the desired output. However, the gain at frequencies around the closed-loop bandwidth of feedback increases with n , causing a blow-up of the \mathbf{u}_{FB} after several iterations.

Note that the transfer function from \mathbf{r} to $\mathbf{u}_{FB}(n)$ obtains one term $(1 + \mathbf{GK})^{-1}$ after each iteration, which increases the phase lag. If we can find a filter that provides a phase lead, we may apply the filter to \mathbf{u}_{FB} before adding it to the feedforward signal. We expect such a phase lead to compensate for the phase lag and achieve a stable AFF algorithm. Since we apply the filter to the stored waveform of \mathbf{u}_{FB} , the filter needs not to be causal. Such a filter can be realized with a *time-reversed low-pass filter* as follows:

1. Read the waveform of the feedback signal \mathbf{u}_{FB} of the previous pulse.
2. Reverse \mathbf{u}_{FB} in time, the time-reversed signal is denoted as \mathbf{u}'_{FB} . After time reversing, the last sample of \mathbf{u}_{FB} becomes the first sample of \mathbf{u}'_{FB} .
3. Apply a low-pass filter to \mathbf{u}'_{FB} .
4. Reverse again the filtered \mathbf{u}'_{FB} and then add it to the feedforward signal.

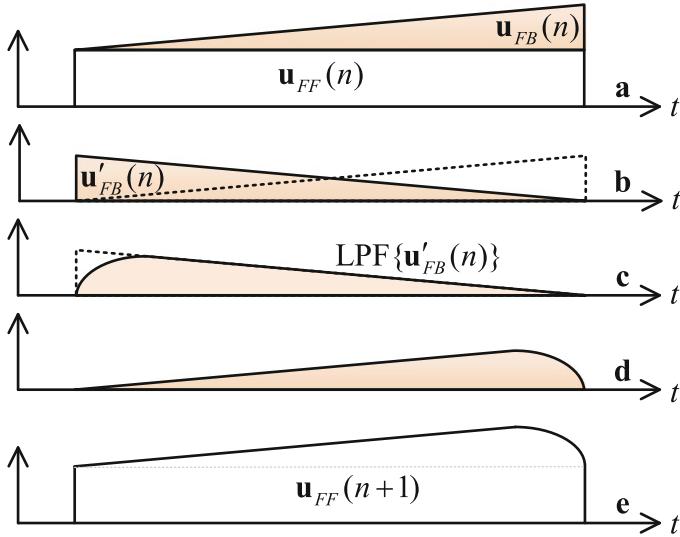


Fig. 4.22 Time-reversed low-pass filter for adaptive feedforward. (a). Identify the feedback control output; (b). time reverse the feedback output; (c). apply low-pass filter to the time-reversed feedback output; (d). reverse the filter output; (e). add the result to the current feedforward signal

The procedure above is illustrated in Fig. 4.22.

A simple implementation of the low-pass filter in step 3 is a first-order filter

$$\mathbf{H}_f(\hat{s}) = \frac{\omega_{LP}}{\hat{s} + \omega_{LP}}, \quad (4.38)$$

where ω_{LP} is the low-pass cutoff frequency. The time-reversed filtering procedure above is equivalent to applying $\mathbf{H}_f(-\hat{s}) = \omega_{LP}/(-\hat{s} + \omega_{LP})$ directly to the feedback signal \mathbf{u}_{FB} . The minus sign before \hat{s} comes from the time reversion process by substituting t with $-t$. Then the feedforward adaption algorithm in (4.37) becomes

$$\mathbf{U}_{FF}(n+1) = \mathbf{U}_{FF}(n) + \mathbf{H}_f(-\hat{s})\mathbf{U}_{FB}(n). \quad n = 0, 1, 2, \dots \quad (4.39)$$

Here $\mathbf{H}_f(-\hat{s})$ low-pass filters the magnitude and provides a phase lead up to 90° . Note that $\mathbf{H}_f(-\hat{s})$ contains an RHP pole, so we cannot directly apply it to \mathbf{u}_{FB} . The time-reversed filtering procedure can achieve the same goal (low-pass filtering in magnitude and providing a phase lead) while keeping the calculation stable. This is because we only reverse the time of the input signal and apply the nominal stable filter (4.38). With the time-reversed filter, we can plot again the transfer function from \mathbf{r} to \mathbf{u}_{FB} in Fig. 4.21b. Compared to Fig. 4.21a, the peaks at the frequencies around the closed-loop bandwidth do not increase with more iterations.

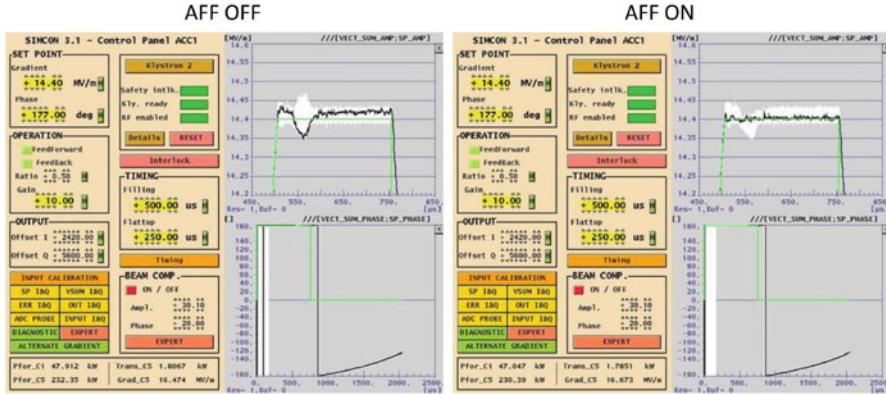


Fig. 4.23 Compensate of beam loadings with AFF based on time-reversed low-pass filter (Courtesy of V. Ayvazyan)

The algorithm with a time-reversed low-pass filter is stable and very efficient to suppress repetitive errors. Figure 4.21b shows that each iteration introduces 20 dB more suppression of low-frequency repetitive errors. The algorithm has been implemented in FLASH. Figure 4.23 shows an example to compensate beam loadings.

4.5.2 Iterative Learning Control

If a linear model of the RF system exists, then a model-based iterative learning control (ILC) algorithm (Kichhoff et al. 2008; Rezaeizadeh 2016) can be used to realize (4.36), given by

$$\begin{aligned} \mathbf{u}(n+1) &= \mathbf{u}(n) + \mathbf{L}\mathbf{e}(n) & n = 0, 1, 2 \dots \\ \mathbf{e}(n) &= \mathbf{r} - \mathbf{y}(n) \end{aligned} \quad (4.40)$$

Here \mathbf{L} is a $K \times K$ constant matrix derived from the system model. It generates a correction to the feedforward signal from the errors of the previous pulse. The other variables are all waveforms with K samples, which are defined as

$$\begin{aligned} \mathbf{u}(n) &:= [\mathbf{u}(n, 0) \quad \mathbf{u}(n, 1) \quad \cdots \quad \mathbf{u}(n, K-1)]^T, \\ \mathbf{y}(n) &:= [\mathbf{y}(n, 0) \quad \mathbf{y}(n, 1) \quad \cdots \quad \mathbf{y}(n, K-1)]^T, \\ \mathbf{e}(n) &:= [\mathbf{e}(n, 0) \quad \mathbf{e}(n, 1) \quad \cdots \quad \mathbf{e}(n, K-1)]^T, \\ \mathbf{r} &:= [\mathbf{r}(0) \quad \mathbf{r}(1) \quad \cdots \quad \mathbf{r}(K-1)]^T. \end{aligned}$$

Various ILC algorithms have been developed with different implementations of \mathbf{L} . The simplest realization employs a diagonal matrix, equivalent to applying the pulse-to-pulse integral feedback to every point in the RF pulse independently. This algorithm works well when the time constant of the RF system is small compared to the RF measurements sampling period. Such an algorithm has been successfully used to flatten the output pulses of klystrons (Geng 2017a, b). If the RF system response is slow, such as with a narrow-band cavity, the realization of \mathbf{L} should consider the RF system's dynamics. Here we will introduce an algorithm based on the impulse response of the RF system.

4.5.2.1 FIR Model of RF System

For a discrete linear system with a finite pulse width, its output and input of the n th pulse are connected by the system's impulse response:

$$\mathbf{y}(n, k) = \sum_{m=0}^{M-1} \mathbf{h}(m) \mathbf{u}(n, k - m) \quad k = 0, 1, 2 \dots K - 1 \quad (4.41)$$

Here \mathbf{u} and \mathbf{y} are the input and output, and \mathbf{h} is a finite-length series of the impulse response of the RF system. Note that \mathbf{h} is formed by complex numbers because \mathbf{u} and \mathbf{y} are both phasors. The series \mathbf{h} contains M points with $M \leq K$. Equation (4.41) shows that the k th point of the output pulse is determined by M previous input points, including the k th input point. The number M should be big enough so that the series \mathbf{h} contains most of the significant points of the impulse response. For a fast system, the impulse response decays rapidly and a smaller M may be selected.

By stacking the points in the input and output waveforms, we can rewrite (4.41) with a transfer matrix

$$\mathbf{y}(n) = \mathbf{Gu}(n), \text{ where } \mathbf{G} = \begin{bmatrix} \mathbf{h}(0) & 0 & \cdots & 0 \\ \mathbf{h}(1) & \mathbf{h}(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \mathbf{h}(M-1) & \cdots & \mathbf{h}(0) \end{bmatrix} \quad (4.42)$$

Here we have assumed that \mathbf{u} is zero for $k < 0$ for the n th RF pulse. To construct the transfer matrix \mathbf{G} , we assign each column starting from the diagonal element with the impulse response series \mathbf{h} (first M points).

The impulse response can be identified with empirical methods such as the least-square fitting of the input and output measurements. Here we will introduce a technique to derive the impulse response from the cavity model. First, we rewrite the cavity equation as

$$\dot{\mathbf{y}}(t) = -(\omega_{1/2} - j\Delta\omega)\mathbf{y}(t) + \omega_{1/2}\mathbf{u}(t), \quad (4.43)$$

where the half-bandwidth $\omega_{1/2}$ and detuning $\Delta\omega$ can be estimated using the method discussed in Chap. 9. If the sampling time T_s is much smaller than the time constant of the cavity, then (4.43) can be discretized as

$$\frac{\mathbf{y}(k) - \mathbf{y}(k-1)}{T_s} = -(\omega_{1/2} - j\Delta\omega)\mathbf{y}(k-1) + \omega_{1/2}\mathbf{u}(k). \quad (4.44)$$

Its discrete transfer function (z -transform) is given by

$$\mathbf{G}_C(z) = \frac{T_s\omega_{1/2}}{1 - z^{-1}[1 - T_s(\omega_{1/2} - j\Delta\omega)]}. \quad (4.45)$$

With the inverse z -transform of $\mathbf{G}_C(z)$, the impulse response of the cavity can be written as

$$\mathbf{h}(k) = T_s\omega_{1/2}[1 - T_s(\omega_{1/2} - j\Delta\omega)]^k \quad k = 0, 1, 2, \dots \quad (4.46)$$

A cavity is, in principle, an Infinite-Impulse Response (IIR) system, and its impulse response decays with time. However, only the first K samples (i.e., the RF pulse width) of the impulse response are of interest when applying it to (4.42). Equation (4.46) can be viewed as the impulse response of the entire RF system if the cavity bandwidth is much smaller than that of other components in the RF system.

4.5.2.2 ILC Algorithm

Assume that the vectors (i.e., waveforms) $\mathbf{u}(n)$ and $\mathbf{y}(n)$ are known from the measurement of the n th pulse. We need to determine a new drive vector $\mathbf{u}(n+1)$ that tends to minimize the output error of the $(n+1)$ th pulse. We also want to avoid too large changes in the drive signal. Then the problem can be formalized as follows:

Find a drive signal $\mathbf{u}(n+1)$ for the $(n+1)$ th pulse to minimize the cost function

$$J = \mathbf{e}(n+1)^H \mathbf{P} \mathbf{e}(n+1) + \Delta\mathbf{u}(n+1)^H \mathbf{Q} \Delta\mathbf{u}(n+1). \quad (4.47)$$

where " H " denotes the conjugate transpose of a vector or a matrix. The output error vector of the $(n+1)$ th pulse is given by

$$\mathbf{e}(n+1) = \mathbf{r} - \mathbf{y}(n+1) \approx \mathbf{r} - \mathbf{y}(n) - \mathbf{G}\Delta\mathbf{u}(n+1), \quad (4.48)$$

Here we have estimated the output of the $(n+1)$ th pulse using (4.42) as $\mathbf{y}(n+1) \approx \mathbf{y}(n) + \mathbf{G}\Delta\mathbf{u}(n+1)$. The changes in the drive signal correspond to the feedforward correction signal

$$\Delta \mathbf{u}(n+1) = \mathbf{u}(n+1) - \mathbf{u}(n). \quad (4.49)$$

The matrix \mathbf{P} and \mathbf{Q} are positive-definite weight matrices.

When J is minimized, its gradient concerning $\mathbf{u}(n+1)$ should be zero, so we need to satisfy $\nabla_{\mathbf{u}(n+1)} J = 0$, resulting in a solution of $\mathbf{u}(n+1)$ as

$$\mathbf{u}(n+1) = \mathbf{u}(n) + (\mathbf{Q} + \mathbf{G}^H \mathbf{P} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{P} \mathbf{e}(n). \quad (4.50)$$

This is the ILC update law. Therefore, the update matrix in (4.40) is

$$\mathbf{L} = (\mathbf{Q} + \mathbf{G}^H \mathbf{P} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{P}. \quad (4.51)$$

Note that we can calculate \mathbf{L} offline with a given transfer matrix \mathbf{G} . Adjustment of the weight matrices will result in different penalties to the output errors and input changes. A larger \mathbf{P} requires a smaller error in the next pulse and then increases the convergence speed, but it may allow a bigger change in the drive signal. Usually, we can assign \mathbf{P} and \mathbf{Q} as simple constant diagonal matrices, which are sufficient in most cases.

We summarize the ILC procedure as follows:

1. Identify the impulse response \mathbf{h} of the RF system.
2. Determine the setpoint waveform \mathbf{r} of the RF system output.
3. Determine the system transfer matrix \mathbf{G} as in (4.42) according to the impulse response \mathbf{h} and the number of samples in the pulse.
4. Determine the weight matrices \mathbf{P} and \mathbf{Q} and calculate \mathbf{L} from (4.51). The weight matrices are often determined empirically, which is demonstrated in Example 4.3 below.
5. Apply the ILC law (4.50) to adapt the feedforward signal of the RF system. If the performance is not satisfactory, tune \mathbf{P} , \mathbf{Q} and repeat from step 4.

In practice, several other points should be taken care of when implementing the ILC algorithm. First is the group delay between the RF system output and input. The feedforward correction signal, which is computed from the output error, should be shifted to an earlier time before applying to the feedforward signal. The time shift is necessary to compensate for the group delay to keep the iterative algorithm stable. Second is the gain and phase shift between the cavity input \mathbf{u} and the RF controller output. The feedforward correction signal refers to \mathbf{u} and should be scaled and rotated so that it refers to the same plane as the feedforward signal in the RF controller. Finally is that we may only apply a fraction of the feedforward correction to avoid instability. Therefore, a gain smaller than 1 is multiplied on the feedforward correction signal in the ILC law (4.50).

The following example illustrates the application of ILC to a TESLA cavity.

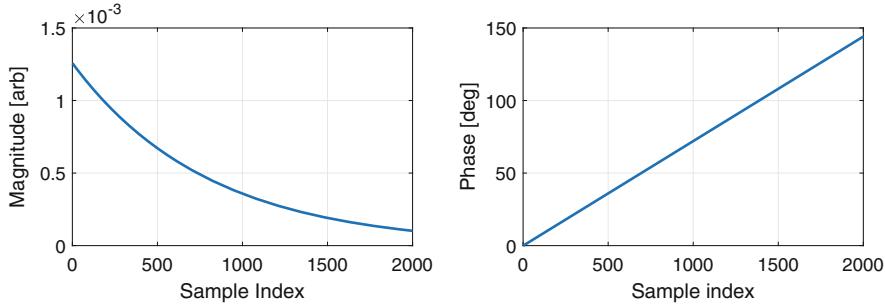


Fig. 4.24 Impulse response of the cavity in magnitude and phase

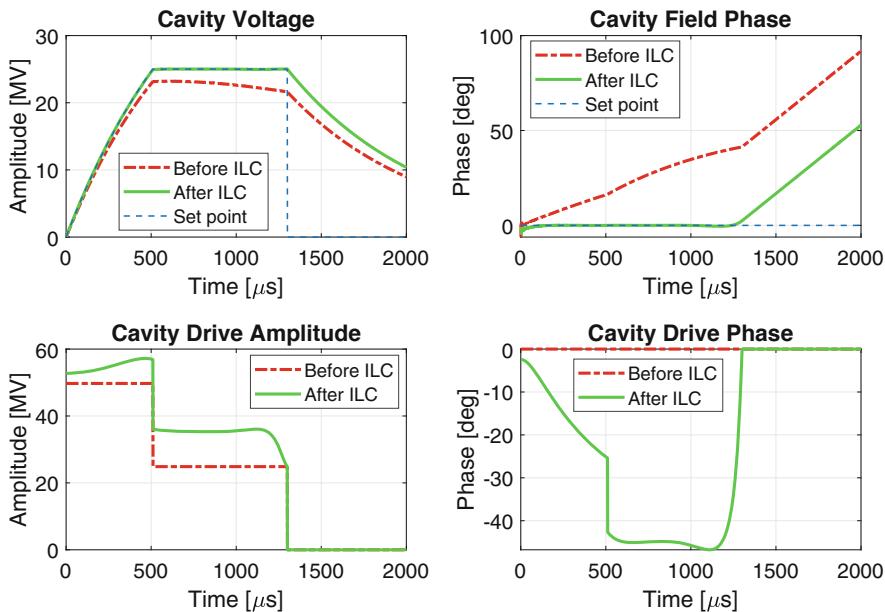


Fig. 4.25 Amplitude and phase waveforms of the cavity input and output. The waveforms before the ILC and after 20 iterations of ILC are both displayed

Example 4.3 ILC for Cavity Control

In this example, we use the ILC algorithm to flatten the RF pulses of a TESLA cavity. The half-bandwidth of the cavity is $\omega_{1/2} = 2\pi \times 200 \text{ rad/s}$ and the cavity is detuned by $\Delta\omega = \omega_{1/2}$. The sampling rate is assumed 1 MHz.

As shown in Fig. 4.25, with a basic feedforward signal, the cavity output errors are large compared to the setpoint. See the waveforms labeled as “Before ILC”. The errors come from the cavity parameter perturbations (e.g., detuning for this example). The cavity impulse response is calculated following (4.46), see Fig. 4.24. Then the ILC algorithm described above was applied. After 20 iterations, the cavity output is very close to the setpoints of both amplitude and phase (Fig. 4.25).

In this example, the weight matrices \mathbf{P} and \mathbf{Q} are diagonal matrices, and each has the same diagonal elements with values 10 and 1, respectively. This causes a fast convergence and relaxes the limits on the cavity drive signal. The matrix \mathbf{P} and \mathbf{Q} may also be more complex (e.g., with patterns in diagonal elements), but the tuning of them is unfortunately still empirical.

The convergence of the ILC algorithm (4.50) has been demonstrated. The detailed proof can be found in the article (Rezaeizadeh 2016). Like the LQG control (see Sect. 4.2.4.1), the ILC algorithm is optimal if the system model is accurate and the measurement noise and process disturbances are Gaussian processes.

In practice, an adaptive feedforward algorithm may be executed only for a limited number of iterations. A long-term continuous operation may lead to instability due to the missing dynamics in the model or the time-alignment error between the feedforward and feedforward correction signals. Typically, the AFF helps to define a desired operating point of the RF system with flattened pulses. Then we close the feedback to suppress the random errors.

4.6 Cavity Resonance Control

From Sects. 4.2, 4.3, 4.4 and 4.5, we have focused on regulating the RF field by manipulating the RF drive signal. If the cavity is not well tuned, the RF field regulation may become difficult. First, more RF drive power is required to maintain the desired cavity voltage, which may saturate the RF amplifiers. Second, we need a larger loop gain to suppress the perturbations introduced by the unwanted detuning. The detuning of a cavity comes mostly from deformations of the cavity shape. In a normal-conducting cavity, the temperature change of the cooling system is the major source of detuning, while for superconducting cavities, the detuning is mainly caused by the Lorentz force and microphonics. See the discussion in Chap. 1.

The effects of detuning depend on its magnitude compared to the cavity bandwidth and on its changing frequency. Time-varying detuning is more harmful to a narrow-band cavity. The resonance control is to regulate the cavity resonance frequency for stabilizing the RF field and reducing the required RF drive power. In synchrotrons, we may also adjust the cavity resonance frequency to track the changes in the particle revolution frequency.

This section is organized as follows. First, we introduce some methods to measure the cavity detuning in various working conditions. Then the cavity tuners and their transfer functions will be discussed. Finally, we describe several widely used resonance control strategies, including the initial cavity tuning and the tuning control with feedback or (adaptive) feedforward algorithms.

4.6.1 Detuning Measurement

Measuring the cavity resonance frequency is essential to realize the resonance control. When tuning a cavity at startup, its resonance frequency can be measured by scanning the input frequency. However, during beam operation, we need non-invasive measurement methods, which typically use the measured cavity signals to estimate the detuning by solving the cavity equations.

4.6.1.1 RF Frequency Scanning

After a long-term shutdown, the cavity, especially a superconducting cavity, might be far detuned from the RF operating frequency. In this case, the cavity reflects all the input RF power and cannot establish measurable RF fields. Therefore, initial tuning of the cavity is necessary. To do it, we can scan the RF drive frequency and measure the amplitude response of the cavity (Ayyazyan et al. 2011). By searching resonance peaks in the cavity response, we can estimate its resonance frequencies. The frequency scanning can be performed with the standard LLRF hardware and software. Here we assume that the RF detector bandwidth is much larger than the possible cavity detuning.

To change the RF drive frequency, we introduce a time-varying phase in the RF controller output. In a digital LLRF system, we can program the RF controller to produce a phase step $\Delta\varphi$ for each DAC sample, which yields a frequency shift

$$\Delta\omega_{RF} = \Delta\varphi/T_s \quad (4.52)$$

in the RF drive, where T_s is the DAC sampling time. Since we only need the cavity's amplitude response, the phase slope in the cavity output is not relevant. Figure 4.26 illustrates the amplitude responses of eight cavities measured at FLASH. For multi-

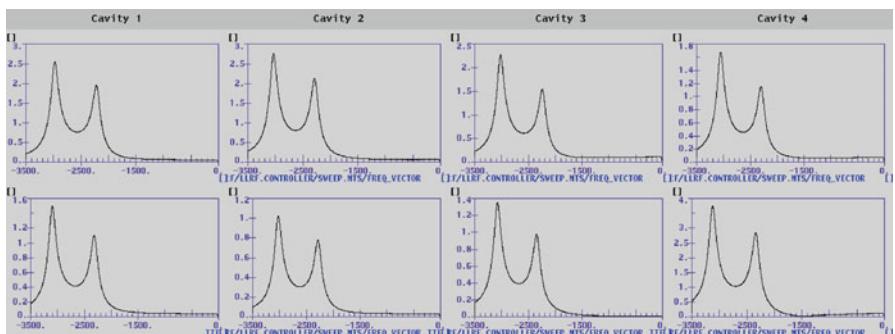


Fig. 4.26 Frequency-amplitude responses of eight cavities at FLASH. Both the π -mode and $8\pi/9$ -mode are visible

cell cavities, multiple passband modes may appear in the frequency response if the scan range is broad enough. To tune the cavity, we need to identify the particular passband mode used for beam acceleration.

Special care needs to be taken when using this method for the scanning of a superconducting cavity. The RF drive power should be small, or the Lorentz force detuning will affect the results. As we have known, the Lorentz force deforms the cavity and decreases its resonance frequency. With a large gradient, the cavity resonance frequency may be changed significantly due to the Lorentz force. In this case, the frequency scanning results cannot tell the cavity's exact resonance frequency. We will discuss more the Lorentz force detuning in Sect. 4.6.4.

4.6.1.2 Phase Slope at RF Pulse Decay

When a standing-wave cavity operates in the pulsed mode, the cavity voltage decays exponentially after the RF drive is off. During the decay stage, the RF field oscillates at the cavity resonance frequency. Therefore, the cavity detuning can be estimated through the slope of the time-varying phase. At the decay stage, the cavity equation is written as

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = 0,$$

and its solution is given by

$$\mathbf{v}_C(t) = \mathbf{v}_{C0}e^{-(\omega_{1/2}-j\Delta\omega)t} = |\mathbf{v}_{C0}|e^{-\omega_{1/2}t}e^{j(\Delta\omega t + \angle\mathbf{v}_{C0})} \quad (4.53)$$

where \mathbf{v}_{C0} is the cavity voltage at the moment (defined as $t = 0$) when turning off the RF drive. The phase of the cavity voltage at the decay stage is a linear function of time given by $\angle\mathbf{v}_C(t) = \Delta\omega t + \angle\mathbf{v}_{C0}$. Then we can estimate the detuning $\Delta\omega$ as the rate of phase change. It is also seen that the cavity half-bandwidth $\omega_{1/2}$ can be determined by fitting an exponential function to the amplitude decay. For normal-conducting cavities, both $\omega_{1/2}$ and $\Delta\omega$ are typically constants during an RF pulse. It implies that we can obtain a complete cavity model by calculating its $\omega_{1/2}$ and $\Delta\omega$ from the RF pulse decay. However, in superconducting cavities, the $\omega_{1/2}$ may vary within an RF pulse, probably due to quenches. As discussed before, the Lorentz force and microphonics will introduce time-varying detuning within the RF pulse. The knowledge of the intra-pulse detuning is useful to compensate for the Lorentz force detuning with feedforward controls (see Sect. 4.6.3).

4.6.1.3 Solving Cavity Equation

To measure the time-varying cavity detuning during RF pulses non-invasively, we need to measure the waveforms of the cavity input \mathbf{v}_f and output \mathbf{v}_C . Then we can solve the cavity equation

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = \omega_{1/2}\mathbf{v}_f$$

and get

$$\Delta\omega = \text{Im}\{(\dot{\mathbf{v}}_C - \omega_{1/2}\mathbf{v}_f)/\mathbf{v}_C\}. \quad (4.54)$$

Here the operator “Im” takes the imaginary part of a complex number. We calculate $\omega_{1/2}$ from the decay stage of the RF pulse and assume it a constant during the pulse. Note that we neglected the beam loading in the cavity equation. The calculation of $\Delta\omega$ in the presence of beam loading is more complicated and will be discussed in Chap. 9.

In a practical LLRF system, the cavity voltage and cavity drive are measured through separate RF detector channels. The relative gain and phase shift between them are unknown due to arbitrary offsets from the cables. Calibration is required to convert the two signals to a common reference plane so that they can be used in (4.54). Usually, we use the cavity voltage as a reference and scale and rotate the measurement of the cavity drive signal.

We may follow the procedure below to calibrate the cavity drive signal:

1. Calculate $\omega_{1/2,decay}$ and $\Delta\omega_{decay}$ from the RF pulse decay according to (4.53).
2. Estimate the cavity drive signal, which is expected to refer to the same reference plane as \mathbf{v}_C , using $\omega_{1/2,decay}$ and $\Delta\omega_{decay}$ by solving the cavity equation:

$$\tilde{\mathbf{v}}_f = [\dot{\mathbf{v}}_C + (\omega_{1/2,decay} - j\Delta\omega_{decay})\mathbf{v}_C]/\omega_{1/2,decay}. \quad (4.55)$$

3. Compare the measured cavity drive signal $\mathbf{v}_{f,m}$ to the estimate above, we can obtain a calibration coefficient

$$\mathbf{c}_f = \tilde{\mathbf{v}}_f(k_0)/\mathbf{v}_{f,m}(k_0), \quad (4.56)$$

then calibrate the cavity drive signal as

$$\mathbf{v}_f = \mathbf{c}_f \mathbf{v}_{f,m}. \quad (4.57)$$

Here k_0 should be the index of the waveform sample just before the RF power is turned off. This is because (4.55) is accurate only at the moment when turning off the RF drive power since the actual $\omega_{1/2}$ and $\Delta\omega$ at other times in the pulse are different from $\omega_{1/2,decay}$ and $\Delta\omega_{decay}$.

As an example, Fig. 4.27a and b illustrate the waveforms of a TESLA cavity. The cavity drive signal has been calibrated from the klystron output signal using the algorithm above. The calculated intra-pulse detuning of the cavity is depicted in Fig. 4.27c. It is a typical Lorentz-force detuning curve in a superconducting cavity.

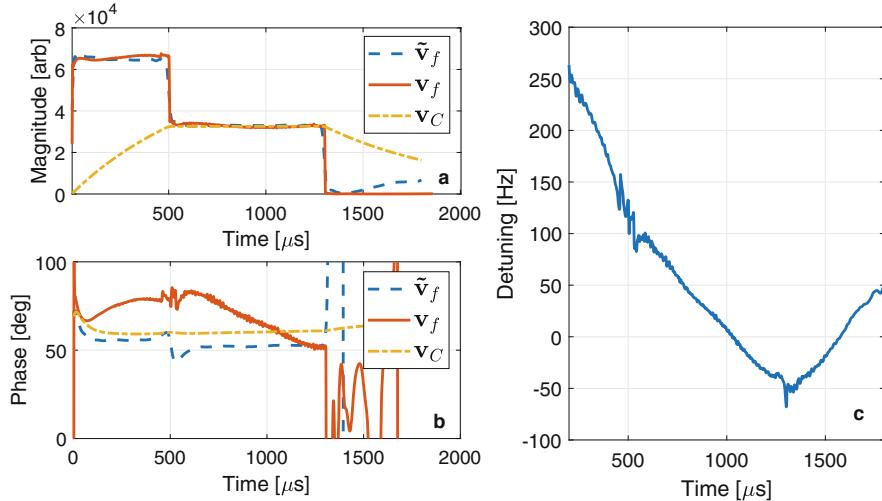


Fig. 4.27 The cavity voltage and the calibrated drive signal of a TESLA cavity (plot **a** and **b**), and the calculated intra-pulse detuning (plot **c**)

4.6.1.4 Cavity Input-Output Phase Shift

A cavity operating in the CW mode is most of the time in the steady state. In Chap. 3, we have defined the cavity detuning angle as $\psi = \tan^{-1}(\Delta\omega/\omega_{1/2})$, which is the steady-state cavity phase shift. It implies that the cavity phase shift, i.e., the phase difference between the cavity output and input, can represent the detuning. In general, the cavity phase shift is calculated as

$$\theta = \varphi_C - \varphi_f,$$

where φ_f and φ_C have the same meanings as in the PLL model (Fig. 4.14). Note that ψ is the value of θ after the cavity reaches the steady state. Also, note that the definition of θ here has an opposite sign than the definition of y in the PLL model. Then we can rewrite the cavity phase Eq. (4.25) as

$$\dot{\theta} + \omega_{1/2}\theta = \Delta\omega - \dot{\varphi}_f. \quad (4.58)$$

This equation approximates the cavity phase equation in (4.24). It is valid when the detuning is small and the cavity voltage is close to the steady state. Note that (4.58) works only when there is no beam in the cavity. In the presence of beam loading, the phase φ_f should be the phase of the vector sum of the RF and beam drives. In a tuning control system, frequency tuners are adjusted to compensate for the undesired cavity detuning. Therefore, the term $\Delta\omega$ in (4.58) contains both the tuning disturbances and the detuning changes caused by the tuners.

In a practical tuning control loop, we often use θ as a measure of the cavity detuning $\Delta\omega$. However, it is accurate (i.e., θ is proportional to $\Delta\omega$) only in the steady state. Equation (4.58) indicates that θ is filtered by the cavity bandwidth and is affected by the derivative of φ_f . It implies that the relationship between θ and $\Delta\omega$ is dynamic in the transient state. Therefore, if the cavity input and output phases are varied by a phase feedback loop, the tuning control loop will be affected if the control goal is to stabilize θ . In this case, the cavity tuner may be adjusted undesirably even if the actual cavity resonance frequency does not change. To understand this situation, we write down the differential equation describing the phase transfer function as the basis of phase feedback design:

$$\dot{\varphi}_C + \omega_{1/2}\varphi_C = \Delta\omega + \omega_{1/2}\varphi_f. \quad (4.59)$$

The phase feedback loop regulates the cavity output phase φ_C by adjusting the input phase φ_f and treating $\Delta\omega$ as a disturbance. The relationship between θ and $\Delta\omega$ can be studied in the following scenarios:

- (i) *The phase feedback loop stabilizes the cavity phase perfectly against the phase fluctuations caused by detuning.*

We assume that the cavity phase φ_C is maintained to be constant by the phase feedback loop. By calculating the derivative of (4.59) at both sides, we get

$$\Delta\dot{\omega} + \omega_{1/2}\dot{\varphi}_f = 0. \quad (4.60)$$

Substitute it into (4.58), we obtain

$$\dot{\theta} + \omega_{1/2}\theta = (\Delta\dot{\omega} + \omega_{1/2}\Delta\omega)/\omega_{1/2}. \quad (4.61)$$

It implies that the phase shift θ can represent exactly the detuning $\Delta\omega$ with a constant factor $\omega_{1/2}$:

$$\theta(t) = \Delta\omega(t)/\omega_{1/2}. \quad (4.62)$$

- (ii) *The phase feedback loop is open.*

In this case, the cavity drive phase does not change and its derivative vanishes in (4.58), then the equation changes to

$$\dot{\theta} + \omega_{1/2}\theta = \Delta\omega. \quad (4.63)$$

It is a simple low-pass system function, which means, θ is related to $\Delta\omega$ through the cavity dynamics.

- (iii) *The detuning remains constant, but the cavity drive phase is changed by the operators to adjust the cavity phase.*

After the cavity voltage approaches the steady state again, the value of θ recovers. However, during the transient of the cavity voltage, θ is time varying and its change is not caused by the detuning change. According to (4.58), a step change in the cavity drive phase, $\Delta\varphi_f$, at $t = 0$, will generate a θ described by $\theta(t) = \theta_0 - \Delta\varphi_f e^{-\omega_1/2t}$, where θ_0 is the original value of θ . In this situation, the tuning loop will make actions during the transient state of the cavity even there are no tuning disturbances. The false tuning corrections will cause cavity field errors and disturb the RF field control loops. To mitigate this situation, we may switch off the tuning loop when adjusting the cavity phase in a broad range. At least, we should avoid changing the cavity phase too fast if the tuning loop is on. For example, a slow ramping results in less perturbation than a sharp step change in the cavity drive phase.

In reality, we usually encounter the first situation, in which the phase shift θ is a good approximation of $\Delta\omega$. Then the tuning control regulates θ by actuating on the cavity tuners. Most tuning loops aim at keeping the cavity on-resonance, which is equivalent to maintaining the value of θ close to zero. However, if in the presence of large beam loading and the beam phase is off-crest, we need to detune the cavity to reduce the reflected power. Then a question arises: assume a tuning loop has been tuned to maintain θ to zero when there is no beam in the cavity, will it automatically generate the optimal detuning after injecting an off-crest beam? Let us analyze this situation. We will only consider the steady-state case.

In Chap. 3, we know that the required RF drive signal for a desired cavity voltage and beam loading is given by (3.36), which is:

$$i_{for} = \frac{v_{C0}}{2R_L \cos \psi \cdot e^{j\psi}} + i_{b0} e^{-j\varphi_b}.$$

Note that the phase of the cavity voltage has been normalized to zero. If the beam phase is off crest ($\varphi_b \neq 0$), we shall introduce a detuning angle that satisfies $\tan \psi = -2R_L i_{b0} \sin \varphi_b / v_{C0}$ to minimize the required RF drive power. Then we get

$$i_{for,opt} = \frac{v_{C0}}{2R_L} + i_{b0} \cos \varphi_b. \quad (4.64)$$

It can be seen that the phase of the drive signal (4.64) is also zero. This means, if we always maintain θ at zero even after the beam is injected, an optimal detuning will be achieved automatically by the tuning loop.

The discussion above helps to define a procedure to set up a tuning control loop. First, we should establish a desired cavity voltage when there is no beam. At the same time, the cavity should be tuned to minimize the reflected power. With the reflected power minimized, we use the phase difference between the measurements of the cavity output and input signals as a reference. Then we close the tuning control loop and maintain the phase difference close to the reference value. Finally, we switch on the beam. As mentioned before, the cavity will be automatically detuned by the tuning loop to compensate for the off-crest beam for minimizing the required RF drive power.

4.6.2 Cavity Tuners

Tuners are actuators for the cavity resonance control. They vary the cavity resonance frequency by deforming the cavity shape or changing its effective capacitance or inductance. From the viewpoint of LLRF, the dynamics between the inputs of the tuner to the cavity detuning is of interest. Many cavities are equipped with motor tuners for slow tuning. In superconducting cavities, piezo tuners are also used to compensate for the Lorentz force detuning and microphonics. Some normal-conducting cavities have no tuner equipped, and their resonance frequencies may be tuned by changing the cooling water temperature.

4.6.2.1 Motor Tuner

A motor-driven tuner is typically slow. It is often used for the coarse tuning of the cavity or compensating for the slow drifts in the cavity resonance frequency. The motor may be either a stepping motor or a DC motor. Some tuners may move a plunger in the RF field space in the cavity to change the resonance frequency, while other tuners may move a lever that pushes a tuning plate to deform the cavity. The second method is mainly used for cavities with thin walls that can be easily deformed.

If we only consider the slow resonance control with a motor tuner, the dynamics of the motor and the cavity mechanical oscillations can be neglected. Then we obtain a simplified model of a motor tuner as in Fig. 4.28.

Here the input and output of the motor-tuner model are the desired motor speed and the resulting cavity detuning. The motor position, which determines the amount of tuning, is the integral of motor speed over time. We assume that the motor speed regulation has a much higher bandwidth than the tuning control loop; that is, we can neglect the dynamics of the motor driver. The relation between the motor position and the cavity detuning can be approximated by a linear function with a constant gain k_p . Then the transfer function of the motor tuner (MT) can be written as

$$G_{MT}(\hat{s}) = \frac{k_p}{\hat{s}} e^{-\hat{s}\Delta t}, \quad (4.65)$$

where Δt is the delay between the tuner movement and the resulting changes of the cavity resonance frequency.

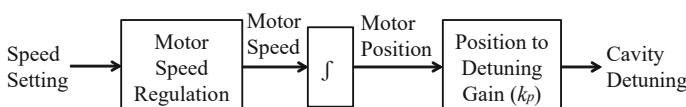


Fig. 4.28 Simplified block diagram of a motor tuner

Fig. 4.29 Piezo tuner installed in the cryostat for a superconducting cavity

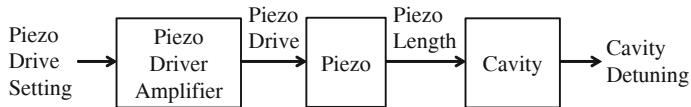
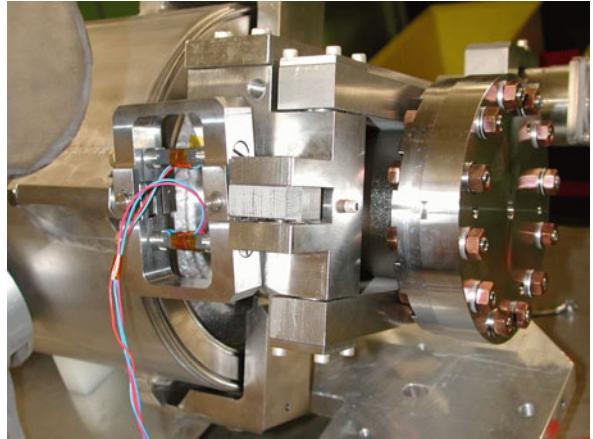


Fig. 4.30 Simplified block diagram of a piezo tuner

4.6.2.2 Piezo Tuner

Piezo tuners (Grecki et al. 2018) perform fast small-range tuning of a cavity. Figure 4.29 (Simrock 2003) shows a typical installation of a piezo tuner in a superconducting cavity. When applied a voltage, the piezo changes its length (on a micrometer level) to deform the cavity. The tuning range of a piezo tuner is typically up to 1 or 2 kHz.

For feedback control, the piezo drive signal is controlled by a DAC. Then it is amplified to the required voltage and current to drive the piezo. Figure 4.30 is a simplified block diagram of a piezo tuner. The input is the piezo drive voltage and the output is the cavity detuning.

To simplify the discussion, we assume the piezo drive amplifier is wideband and neglect its dynamics. Both the piezo and the cavity contain various mechanical resonance modes. If the piezo tuner is used in a feedback loop, the resonance modes may cause instability and limit the achievable closed-loop bandwidth. The transfer function of a piezo tuner (PT) can be written in the form of

$$G_{PT}(\hat{s}) = \left(\frac{k_0}{\tau\hat{s} + 1} + \sum_{i=1}^N \frac{\omega_i^2 k_i}{\hat{s}^2 + 2\xi_i \omega_i \hat{s} + \omega_i^2} \right) e^{-\hat{s}\Delta t}, \quad (4.66)$$

The input and output of the transfer function are the piezo amplifier input and the cavity detuning, respectively. The transfer function consists of a low-pass term,

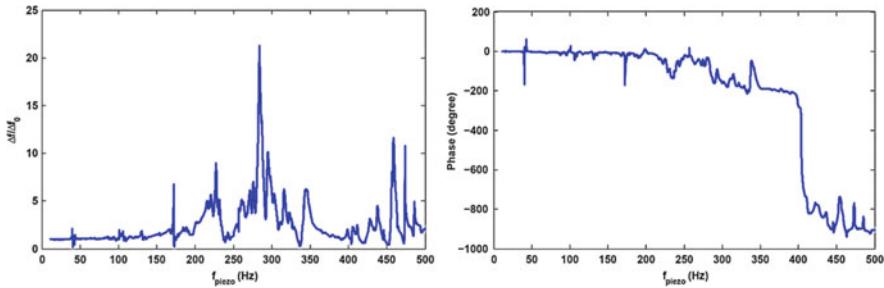


Fig. 4.31 Measurement of frequency responses of a piezo tuner installed in a TESLA cavity (Courtesy of A. Neumann et al)

several resonance terms, and a delay term. For the low-pass term, k_0 is the DC gain, and τ is the time constant. We consider N mechanical resonance modes. For each mode, we use ω_i , ξ_i , and k_i to denote the mechanical resonance frequency, the damping constant, and the DC gain, respectively. The group delay Δt represents the propagation time of the piezo actuation to the cavity deformation. When controlling the RF field in a cavity, we often apply both the RF feedback and the tuning control simultaneously. A mechanical disturbance with a frequency over the closed-loop bandwidth of the RF feedback is generally not relevant to the RF stability. Therefore, in (4.66), we only consider the mechanical modes below the RF control's closed-loop bandwidth.

Figure 4.31 illustrates the frequency responses of a piezo tuner of a TESLA cavity (Neumann et al. 2010). It is identified by sweeping the frequency of a sinusoidal drive signal applied to the piezo and measuring the resulting cavity detuning. The cavity detuning will oscillate at the same frequency as the sinusoidal input after reaching a steady state. For each frequency, the gain is calculated as the ratio between the cavity detuning amplitude and the piezo drive amplitude, whereas the phase response is the phase difference between the cavity detuning and the piezo drive. Note that the gain and phase responses are calculated after the piezo tuner reaches a steady state under each drive frequency.

To obtain the model of a piezo tuner, we may fit the empirical frequency responses with some software tools like Matlab (e.g., using the function `tfeest`). Particularly, by fitting a linear function in the low-frequency region of the phase response, the group delay Δt can be estimated.

4.6.2.3 Cooling Water Temperature

In some normal-conducting cavities, there are no mechanical tuners installed. The cavity tuning is performed by changing the cooling water temperature. The cooling system typically has a feedback loop for temperature regulation. When using the cooling water temperature as a tuner in a tuning feedback loop, the closed-loop bandwidth should be smaller than the temperature regulation loop to keep the cascaded loops stable.

The response of the tuning control loop using the cooling water temperature as a tuner is typically slow. The settling time for a new temperature setpoint may be up to several minutes. Some methods have been developed to tune the cavity faster. For example, in the article (Pfeiffer et al. 2018), the RF pulse width is modulated to change the RF power dissipation in the cavity to regulate its temperature.

4.6.3 Tuning Control Scheme

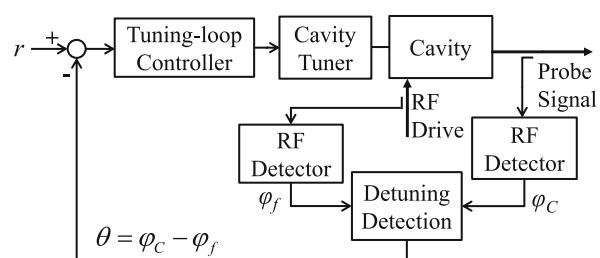
To mitigate the cavity tuning errors, we should reduce the environmental disturbances like the temperature fluctuations and mechanical vibrations. Furthermore, the cavity mechanical rigidity should be improved for lower sensitivities to external disturbances. For a normal-conducting cavity, a careful design of the cooling water system is essential for the tuning stability. The cooling regulation loop should be robust against heat-load changes (e.g., during an interlock trip) to avoid significant temperature fluctuations. In a superconducting cavity, microphonics is the major concern for tuning control. Microphonics has many sources, such as the helium bath pressure, water/helium/vacuum pumps and ground motions. The mechanical design of the cryogenic system should reduce the transmission of vibrations to the cavities.

Besides the mitigation strategies at the design stage, active controls of cavity detuning may be implemented after the cavities are installed. Figure 4.32 illustrates a block diagram of the cavity tuning control. Measurements of the detuning usually involve the cavity drive and probe signals. Here we only show the detuning measurement using the cavity phase shift, but other methods discussed in Sect. 4.6.1 may be implemented when applicable. The tuning-loop controller implements the control algorithms that are the major topic of this section.

4.6.3.1 Tuning Control with Feedback

Feedback controllers are widely used in the resonance control of both normal-conducting and superconducting cavities (Simrock 2003). They compensate for the slow fluctuations in the cavity resonance frequency. In storage rings, feedback is also used to tune the cavity resonance frequency to track the changes of beam revolution frequency.

Fig. 4.32 A general block diagram of cavity tuning control



In a normal-conducting cavity, the resonance frequency fluctuations come mainly from the cooling water temperature changes. For superconducting cavities, static Lorentz force (for CW operation) and helium pressure fluctuations are the major sources of low-frequency tuning errors.

Coupling with Phase Feedback Loop

Most cavities are equipped with motor tuners. With the motor tuner shown in Fig. 4.28, a proportional-derivative (PD) controller may be used. Of course, if the motor controller regulates the motor position instead of the speed, i.e., accepting the destination position as the set point input, the PI controller is more suitable. The transfer function model of the tuning feedback loop (also denoted as the “tuning loop”) with a motor tuner and a PD controller is depicted in Fig. 4.33a.

In Fig. 4.33a, we placed a speed limiter for the motor, which is either a voltage limiter (for DC motors) or a pulse-rate limiter (for stepping motors). One major purpose of this limiter is to avoid deforming the cavity too fast. To model the cavity, we have used the cavity phase Eq. (4.58). The detuning caused by the external environment acts as a disturbance $\Delta\omega_{dist}$ to the tuning loop. Note that the derivative of the cavity drive phase φ_f also appears as a disturbance term. Changes of φ_f come mainly from the phase feedback loop (also denoted as the “phase loop”, see Fig. 4.33b), which follows the model (4.59). The variables φ_f and $\Delta\omega_{all}$ (the overall detuning caused by the disturbance and the tuner) couple the tuning and phase loops. If we are interested in the disturbance rejection capability of the tuning loop, we can evaluate the closed-loop transfer function from $\Delta\omega_{dist}$ to θ . The coupling between the tuning and phase loops should be taken into account.

First, we calculate the output θ caused only by the disturbance $\Delta\omega_{dist}$ in the tuning loop. We denote this closed-loop transfer function as $T_{TL}(\Delta\omega_{dist} \rightarrow \theta)$, where the subscript “TL” stands for the tuning loop. Then we should calculate the value of φ_f induced by $\Delta\omega_{dist}$ via the coupling between the tuning and phase loops. According to Fig. 4.33, we have the following relationships

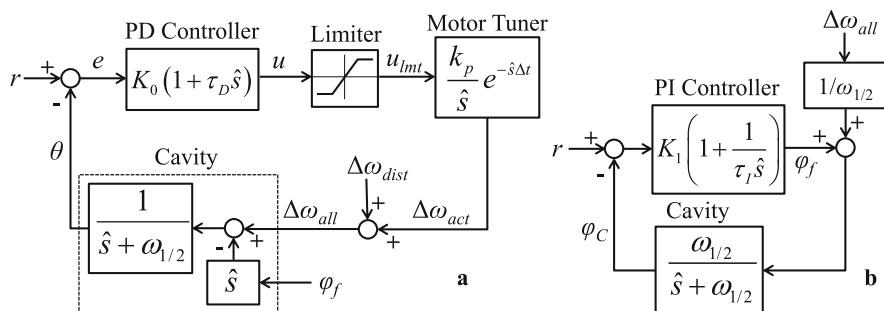


Fig. 4.33 (a). Block diagram of a tuning feedback loop with a motor tuner and a PD controller; (b). Block diagram of the phase feedback loop with a PI controller

$$\begin{aligned}\Delta\omega_{all} &= T_{TL}(\Delta\omega_{dist} \rightarrow \Delta\omega_{all}) \cdot \Delta\omega_{dist} + T_{TL}(\varphi_f \rightarrow \Delta\omega_{all}) \cdot \varphi_f, \\ \varphi_f &= T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f) \cdot \Delta\omega_{all},\end{aligned}$$

where $T_{TL}(\Delta\omega_{dist} \rightarrow \Delta\omega_{all})$ and $T_{TL}(\varphi_f \rightarrow \Delta\omega_{all})$ are the closed-loop transfer functions describing how the overall detuning depends on the tuning disturbance and the cavity drive phase. Here we have assumed the setpoints of both loops are zero. The subscript “PL” stands for the phase loop, and $T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f)$ is the closed-loop transfer function representing the resulting cavity drive phase from the overall detuning. Combine the two equations above and eliminate $\Delta\omega_{all}$, we get

$$\varphi_f = \frac{T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f)T_{TL}(\Delta\omega_{dist} \rightarrow \Delta\omega_{all})}{1 - T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f)T_{TL}(\varphi_f \rightarrow \Delta\omega_{all})} \cdot \Delta\omega_{dist}.$$

Since the output θ in the tuning loop is a superposition of the contributions from φ_f and $\Delta\omega_{dist}$, it can be written as

$$\theta = T_{TL}(\varphi_f \rightarrow \theta) \cdot \varphi_f + T_{TL}(\Delta\omega_{dist} \rightarrow \theta) \cdot \Delta\omega_{dist}.$$

Here $T_{TL}(\varphi_f \rightarrow \theta)$ is the closed-loop transfer function from the cavity drive phase to the tuning output. With the analysis above, we can write the overall transfer function from the tuning disturbance $\Delta\omega_{dist}$ to the tuning output θ as

$$\begin{aligned}T(\Delta\omega_{dist} \rightarrow \theta) &= T_{TL}(\Delta\omega_{dist} \rightarrow \theta) + \\ &\frac{T_{TL}(\varphi_f \rightarrow \theta)T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f)T_{TL}(\Delta\omega_{dist} \rightarrow \Delta\omega_{all})}{1 - T_{PL}(\Delta\omega_{all} \rightarrow \varphi_f)T_{TL}(\varphi_f \rightarrow \Delta\omega_{all})}.\end{aligned}\quad (4.67)$$

The concrete transfer functions can be derived from the block diagrams in Fig. 4.33. Here we only give an example of the typical frequency responses of the overall $T(\Delta\omega_{dist} \rightarrow \theta)$ and the tuning-loop transfer function $T_{TL}(\Delta\omega_{dist} \rightarrow \theta)$. See Fig. 4.34, where the tuning disturbance $\Delta\omega_{dist}$ has been scaled by dividing the cavity half-bandwidth $\omega_{1/2}$. It can be seen from both frequency responses that the tuning loop can suppress the low-frequency tuning disturbances. If the phase feedback is disabled, i.e., φ_f does not change, the effects of $\Delta\omega_{dist}$ on the cavity detuning (i.e., θ) follows the frequency response $T_{TL}(\Delta\omega_{dist} \rightarrow \theta)$. In contrast, the frequency response $T(\Delta\omega_{dist} \rightarrow \theta)$ will be followed when the phase feedback is applied. It implies that more high-frequency tuning disturbances will contribute to the cavity resonance frequency fluctuations. In other words, with the phase feedback off, $\Delta\omega_{dist}$ will be low-pass filtered by the open-loop bandwidth of the cavity. However, if the phase feedback is on, $\Delta\omega_{dist}$ will be filtered by the closed-loop bandwidth of the phase loop, which is typically larger than the cavity’s open-loop bandwidth.

If the phase loop is closed, the tuning loop’s open-loop gain increases at higher frequencies. This is due to the equivalent increase of the cavity bandwidth under the control of the closed phase loop. Therefore, for a particular tuning-loop delay, the

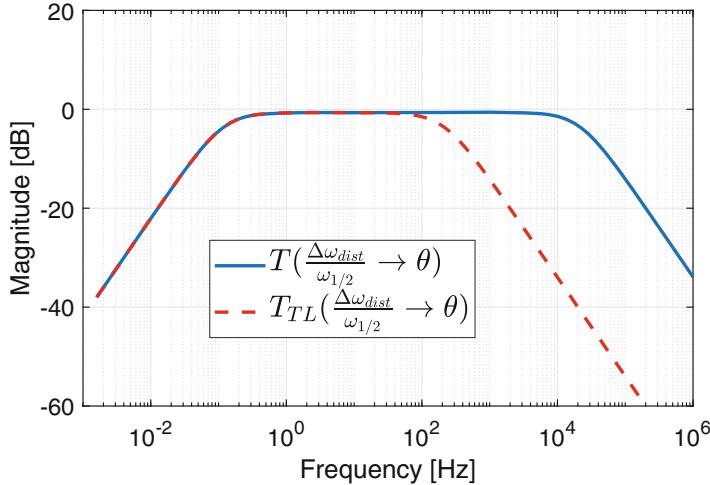


Fig. 4.34 Contribution of tuning disturbances of different frequencies to cavity tuning errors. A negative value in dB represents the disturbance suppression. The following parameters are used: cavity $\omega_{1/2} = 2\pi \times 200$ rad/s; motor tuner $k_p = 10$ and $\Delta t = 0$; tuning loop controller $K_0 = 10$ and $\tau_D = 0.1$ s; phase loop controller $K_1 = 100$ and $\tau_I = 0.001$ s

gain margin of the tuning loop will be reduced. That is, at the frequency where the phase-lag of the open-loop transfer function of the tuning loop crosses -180° , the loop gain is more likely to be over 1 due to the equivalently increased cavity bandwidth. We should keep this in mind when optimizing the feedback gains of the tuning and phase loops. Nevertheless, the closed-loop bandwidth of the tuning loop is typically small (e.g., due to slow motor tuners or the resonance modes of piezo tuners) compared to the upper limits imposed by the gain margin. In practice, the RF (i.e., I/Q or amplitude/phase) and tuning feedback loops are often designed separately without considering the coupling between them.

Issues of Feedback with Piezo Tuners

Superconducting cavities are typically equipped with both motor tuners and piezo tuners. If a tuning feedback loop is formed with a motor tuner for a superconducting cavity, the model in Fig. 4.33a also applies. A tuning feedback loop may also make use of the piezo tuner controlled by a PI controller. In this case, we still use the loop model in Fig. 4.33a for the analysis but replace the PD controller with a PI controller and replace the motor tuner with the piezo tuner transfer function (4.66).

The closed-loop bandwidth of a tuning feedback loop using a piezo tuner is limited by the resonance modes of the tuner. At each resonance frequency of the piezo tuner, a phase-lag close to -180° is introduced that may cause instability. As seen in Fig. 4.31, the resonance mode at around 30–40 Hz is the first mode that

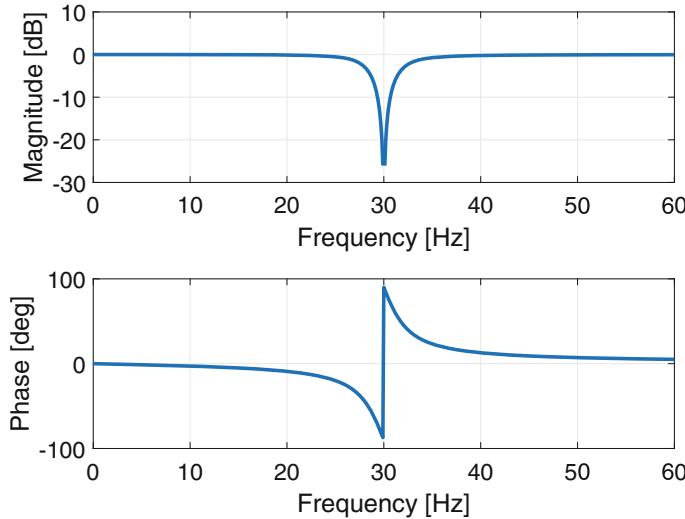


Fig. 4.35 Frequency response of a notch filter with $\omega_0 = 2\pi \times 30$ rad/s and $\omega_h = 2\pi \times 2$ rad/s

causes instability, which imposes a limit on the closed-loop bandwidth. However, if we place a notch filter at the frequency of the unstable resonance mode in the feedback loop, we can reduce the gain at this frequency and avoid instability. This method has been widely used to increase the closed-loop bandwidth of a tuning feedback loop. The transfer function of a second-order notch filter is given by

$$H_{notch}(\hat{s}) = \frac{\hat{s}^2 + \omega_0^2}{\hat{s}^2 + 2\omega_h\hat{s} + \omega_0^2}, \quad (4.68)$$

where ω_0 and ω_h are the central frequency and half-bandwidth of the notch filter. The typical frequency response of a notch filter is shown in Fig. 4.35. Note that the notch filter also introduces a phase-lag at lower frequencies that may aggravate the original phase-lag at the resonance mode. Of course, if the notch filter can reduce the loop gain to below 1 at the frequency where the overall phase-lag crosses -180° , the loop will be stable. Typically, we should tune the notch filter bandwidth to achieve enough phase margin (e.g., larger than 30°). Figure 4.36 shows the bode plots of the open-loop transfer function of a tuning feedback loop equipped with a piezo tuner. The piezo tuner contains five mechanical resonance modes at 280, 341, 460, 487 and 618 Hz. We use a PI controller in the feedback loop and assume a loop delay of 200 μ s. As seen from the bode plots, the loop without notch filters is unstable due to the mechanical modes, which result in a loop gain larger than one when the loop phase crosses -180° . With the notch filters, the closed-loop becomes stable, and we achieve a gain margin larger than 5 dB and a phase margin over 37° . A closed-loop simulation illustrates successful suppression of low-frequency

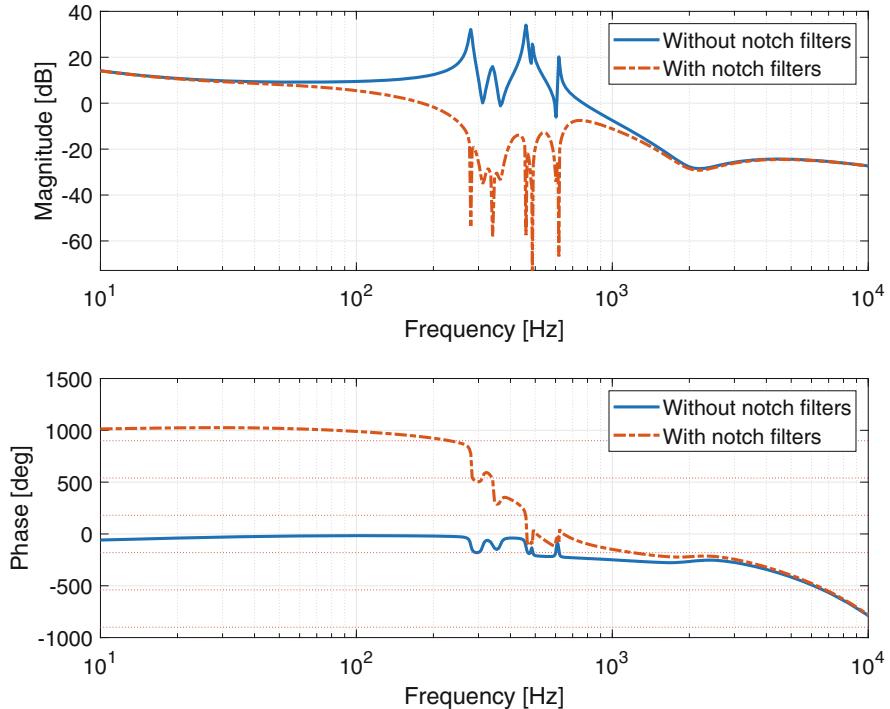


Fig. 4.36 Bode plots of the open-loop transfer function of a tuning feedback loop without/with notch filters. With notch filters, the gain margin is 5.17 dB and the phase margin is 37.5°

disturbances, as in Fig. 4.37. In the simulation, the disturbances include white noise and several sinusoidal perturbations at the frequencies of the piezo tuner's mechanical modes. Note that the piezo tuner model in Fig. 4.30 describes the dynamics between the piezo drive voltage and the resulting cavity detuning. Therefore, when we refer to the piezo tuner's mechanical modes, we include the mechanical modes of the piezo, the cavity mechanical frame, and the cavity body.

Note that the feedback control cannot suppress the sinusoidal disturbances because the loop gains at these frequencies have been reduced to below 1 by the notch filters. We will handle these sinusoidal disturbances in the next section.

4.6.3.2 Tuning Control with Adaptive Noise Cancellation

A sinusoidal disturbance with predictable amplitude and phase in the time scale of the control actions can be compensated for by a sinusoidal input at the same frequency. The input should have a proper magnitude and phase determined by the system response at the corresponding frequency. This concept has led to the well-known *adaptive noise cancellation* (ANC) algorithm shown in Fig. 4.38 (Miljkovic 2016).

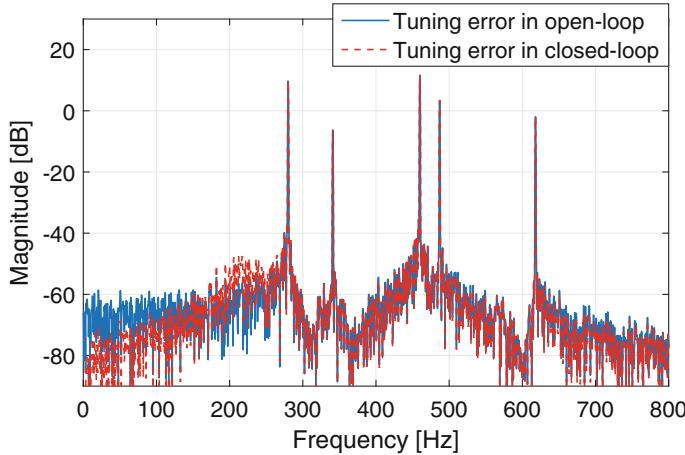
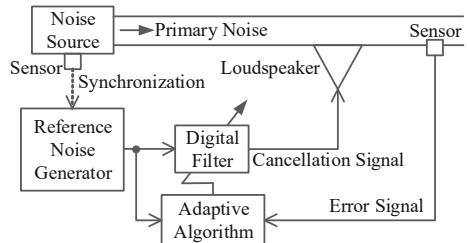


Fig. 4.37 Closed-loop simulation of the tuning feedback loop with notch filters

Fig. 4.38 Block diagram of the narrow-band adaptive noise cancellation algorithm



In Fig. 4.38, we need to cancel the noise produced by a noise source to achieve a quiet environment downstream. We assume that the noise is narrowband and consists of several dominant frequencies. For canceling the noise of a specific frequency, a signal generator is used to generate a reference noise at the same frequency. The reference noise is processed by a digital filter to obtain a cancellation signal. Compared to the primary noise, the cancellation signal has the same amplitude but the opposite phase. The cancellation signal cancels the primary noise downstream via a loudspeaker. The error signal is monitored to determine the digital filter coefficients with adaptive algorithms like the Least Mean Square (LMS) algorithm.

Here we will use a similar algorithm to compensate for the sinusoidal tuning disturbances in a cavity. The control architecture is depicted in Fig. 4.39 (Carcagno et al. 2003; Kandil et al. 2004; Neumann 2008; Rybaniec et al. 2017).

Before applying the ANC, we measure the cavity detuning and identify the disturbance frequencies by spectral analysis, e.g., using the FFT technology. Then we set up multiple frequency generators in the digital controller according to the identified frequencies. The output of each frequency generator is processed by an FIR filter with its coefficients adapted by the LMS algorithm. Each LMS block

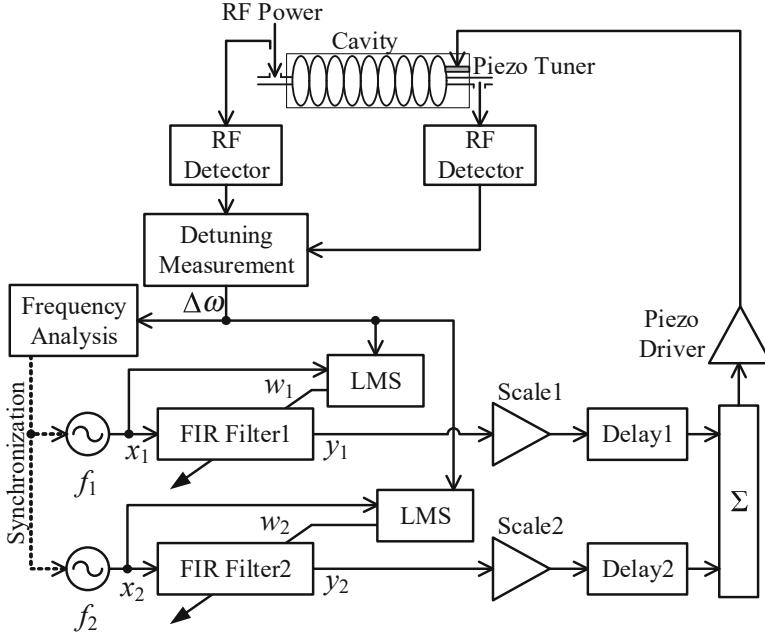


Fig. 4.39 Compensation of narrow-band tuning disturbances with an ANC algorithm based on adaptive FIR filters. Here two noise frequencies are displayed as example

accepts the detuning measurement and the corresponding frequency generator output as inputs. The FIR filter outputs are scaled, phase-shifted (by delay stages), and summed up to control the piezo driver. Each ANC channel consists of a frequency generator, an FIR filter, an LMS block, and a scale and delay stage, suppressing the tuning disturbance of a particular frequency.

Let us have a close look into the LMS algorithm. Assume that a digital frequency generator (e.g., numerically controlled oscillator) produces a discrete sine wave

$$x(k) = \sin(\omega k T_s) \quad k = 0, 1, 2, \dots, \quad (4.69)$$

where k is the index of samples and T_s the sampling time. We use an M -order FIR filter, and its coefficients can be written in a vector

$$\mathbf{w} = [w_0 \quad w_1 \quad \cdots \quad w_{M-1}]^T. \quad (4.70)$$

At the k th sample where $k > M$, the LMS algorithm gives the following update law to adapt the FIR coefficients:

$$\mathbf{w}(k) = \mathbf{w}(k-1) + \mu \Delta\omega(k) \mathbf{x}(k), \quad (4.71)$$

where $\Delta\omega(k)$ is the detuning at sample k , and \mathbf{x} is a vector formed by the buffered samples of the reference sine wave

$$\mathbf{x}(k) = [x(k) \quad x(k-1) \quad \cdots \quad x(k-M+1)]^T. \quad (4.72)$$

Then the FIR filter output at sample k can be calculated as

$$y(k) = \mathbf{w}(k)^T \mathbf{x}(k). \quad (4.73)$$

In (4.71), we typically define $\mathbf{w}(0) = \mathbf{0}$ to avoid generating large initial output by the FIR filter. The update speed parameter μ should be between 0 and 1. With a larger μ , the LMS algorithm converges faster but with the risk of oscillations or instability. Determination of μ is empirical, and we often start from a smaller value for more stable convergence.

Note that for each ANC channel, the FIR filter output y only cancels a sinusoidal tuning disturbance in the measurement of the cavity detuning. Therefore, we still need to determine the required piezo drive to produce a detuning that equals y to cancel the corresponding tuning disturbance. To do this, we must inverse the frequency response of the piezo tuner. The inverse of the frequency response at the frequency of this ANC channel determines the required scaling factor and phase shift on y to obtain the corresponding piezo drive signal. This is needed to have an effective compensation of the sinusoidal tuning disturbance at each frequency. In Fig. 4.39, we used a delay stage to shift the phase of y .

We typically apply the LMS adaption process to multiple ANC channels simultaneously. The FIR order and the adaption speed should be chosen carefully to avoid oscillations. Specifically, the FIR filter order should cover at least a 1/4 period of the frequency to be canceled.

We applied the ANC algorithm to the example of Fig. 4.37. The simulation results are shown in Fig. 4.40. It can be seen that the ANC algorithm suppressed the five sinusoidal disturbances simultaneously.

In reality, the sinusoidal disturbances, mostly caused by microphonics, may have different frequencies depending on the cavity operating conditions. For example, if we adjust the helium flow rate, the microphonics caused by the helium pump may have different frequencies. Based on this consideration, the frequencies of the numerical oscillators in Fig. 4.39 should be updated from time to time with the disturbance frequency changes.

Note that the ANC algorithm introduced in this section works only for the cancellation of sinusoidal disturbances. Similar algorithms with slightly different implementations have been developed for compensating for wide-band disturbances. Readers can refer to the reference articles given in this section.

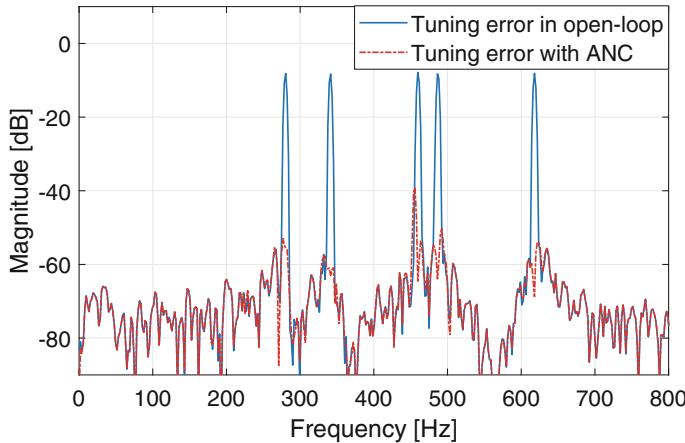


Fig. 4.40 Simulation results of the compensation of sinusoidal tuning disturbances using the ANC algorithm

4.6.3.3 Feedforward Control for Pulsed Cavities

The Lorentz force detuning (LFD) follows the pattern in Fig. 4.27c in a superconducting cavity operating in the pulsed mode. As we have learned in Chap. 1, the Lorentz force reduces the cavity resonance frequency, resulting in a negative change in $\Delta\omega$.

If the cavity drive power and phase repeat for each pulse, the Lorentz force detuning is also repetitive. An RF pulse produces an impulse force on the cavity and causes mechanical oscillations. Therefore, the RF pulse rate should be small (e.g., < 10 Hz) to allow the oscillations to damp before the next pulse. The cavity detuning induced by the microphonics is typically less significant than the Lorentz force detuning, especially when the cavity operates at a high gradient. The repetitive Lorentz force detuning for each RF pulse enables us to compensate for it with a piezo tuner driven by a feedforward voltage (Paparella 2007; Pischalnikov et al. 2012).

In an RF pulse, the particle beam is accelerated at the flattop, where the cavity resonance frequency should be kept as constant as possible. As seen in Fig. 4.27c, to compensate for the time-varying detuning at the flattop (between 500 μ s to 1300 μ s), the piezo tuner should generate a resonance-frequency variation with a positive slope. Note that the underlying source of detuning is the cavity mechanical modes stimulated by the pulsed Lorentz force. The periods of the mechanical oscillations of a cavity are typically much longer than the RF pulse width. Therefore, the Lorentz force detuning during the flattop often changes approximately linearly, as shown in Fig. 4.27c. To suppress the Lorentz force detuning, we may apply a pulsed drive to the piezo tuner. See Fig. 4.41. Here we have chosen a half-sine wave for the shape of the piezo drive pulse. Various pulse shapes, such as triangle, rectangular or full-sine wave, may also be used. The tunable parameters of the half-sine wave are its

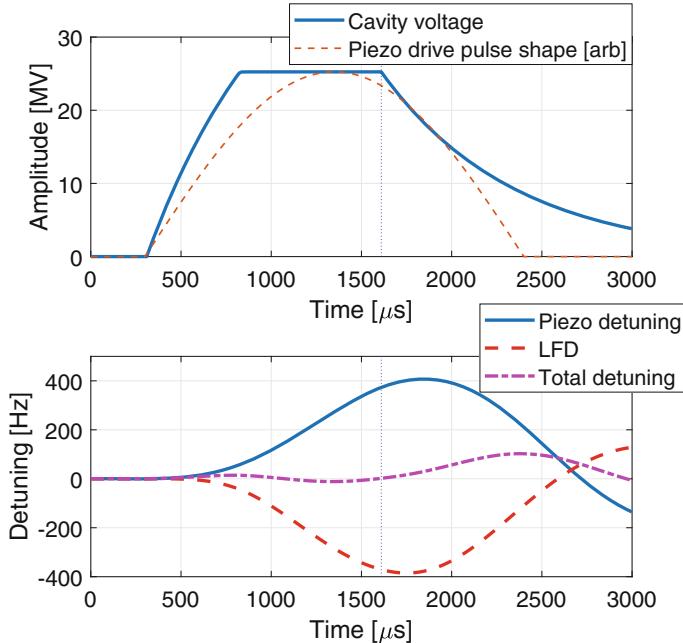


Fig. 4.41 LFD compensation with a piezo tuner in a cavity operating in the pulsed mode. The pulse shape of the piezo drive signal (with arbitrary unit) is also displayed in the top plot

magnitude, delay and pulse width. The pulsed drive of the piezo produces an additional mechanical oscillation in the cavity. We can adjust the piezo drive parameters to create a resonance frequency change opposite to the Lorentz force detuning during the RF pulse. Finally, we obtain a smaller overall detuning at the flattop. This algorithm is depicted in Fig. 4.41, where a TESLA cavity model is used.

The selection of the shape and parameters of the piezo drive signal is often empirical. Of course, if we have identified the piezo tuner's transfer function, simulation is a good start to determine the piezo drive signal. We may optimize a parameter of the piezo drive signal by scanning it in a range. For example, we scan the timing of the piezo drive pulse to find an optimal delay. When scanning a piezo drive parameter, the overall cavity detuning may be worsened, therefore, we should reduce the magnitude of the piezo drive signal during parameter scanning.

4.6.4 Ponderomotive Effects

The mechanical and electromagnetic oscillations in a superconducting cavity are coupled with each other by the radiation pressures. The resulting phenomena are denoted as the *ponderomotive effects*. The RF field in a cavity produces radiation pressures (i.e., Lorentz force) on the cavity wall, resulting in variations in the cavity's

RF resonance frequency. On the other hand, the resonance frequency fluctuation will affect the RF field strength in the cavity. The coupling between the RF and mechanical oscillations is nonlinear and may lead to *ponderomotive instability*, which has two possibilities: static and oscillatory instability (Schulze 1972; Delayen 2005). The static instability may occur when the RF operating frequency is below the cavity resonance frequency, causing jumps in the cavity voltage. In contrast, the oscillatory instability, which appears as oscillations in the cavity voltage and resonance frequency, may happen at RF operating frequencies higher than the cavity resonance frequency. Ponderomotive instability is typically observed in the cavities operating in the CW mode with a high gradient. This topic will be discussed in this section.

4.6.4.1 Introduction to Ponderomotive Instability

The static and oscillatory instability is illustrated conceptually in Fig. 4.42.

For the operating scenario of Fig. 4.42a, we have $\omega_{rf} < \omega_{cav}$, where ω_{rf} is the RF operating frequency and ω_{cav} the cavity resonance frequency. In this scenario, the Lorentz force detuning (LFD) and the cavity voltage form a positive feedback loop. We should keep in mind that a higher cavity voltage causes a larger reduction of ω_{cav} , i.e., shifting the cavity resonance curve further to the left (lower frequency side). If ω_{cav} changes (e.g., due to the microphonics), such as to a higher value, the cavity voltage will drop since the cavity resonance curve moves to the right (higher frequency side). A lower cavity voltage results in a smaller LFD that increases furtherly the ω_{cav} . The positive feedback causes instability if the coupling gain between the LFD and cavity voltage is larger than 1 (i.e., strong coupling). A strong coupling means that a cavity voltage variation causes a change in LFD, which then causes a larger cavity voltage change in the same direction.

As a comparison, Fig. 4.42b illustrates the oscillatory instability between the LFD and cavity voltage at $\omega_{rf} > \omega_{cav}$. The LFD and cavity voltage form a negative feedback loop, which can avoid the static instability but may cause oscillations between the cavity voltage and the resonance frequency. Understanding the

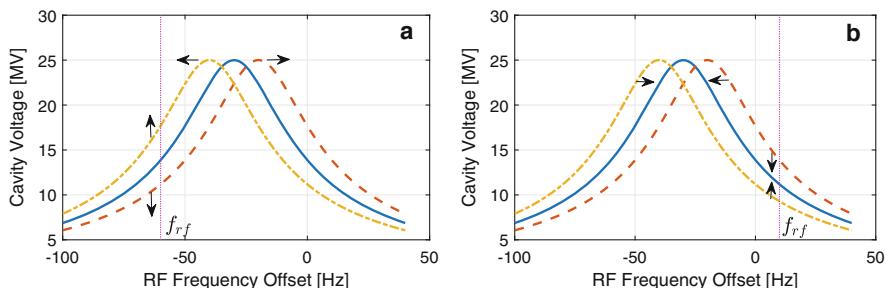


Fig. 4.42 Ponderomotive instability. (a). static instability for $\omega_{rf} < \omega_{cav}$; (b). oscillatory instability for $\omega_{rf} > \omega_{cav}$

oscillatory instability involves non-linear system analysis based on the differential equations of the cavity voltage (3.15) and the cavity mechanical modes (3.44). We will not go deep into this topic. Interested readers can refer to (Leewe and Fong 2019). Typically, the oscillatory instability happens when the maximum LFD is larger than the bandwidths of the cavity's mechanical resonance modes.

4.6.4.2 Static Instability Analysis

The static instability can be approximately explained with the linear equations. Let us only consider the steady-state relations between the cavity voltage and LFD:

$$\mathbf{v}_C = \frac{\omega_{1/2} v_{C0}}{\omega_{1/2} - j\Delta\omega}, \omega_{cav} = -K|\mathbf{v}_C|^2, \Delta\omega = \omega_{cav} - \omega_{rf}, \quad (4.74)$$

where v_{C0} is the cavity voltage achieved on resonance with a particular drive power. Here ω_{rf} and ω_{cav} are defined as offset frequencies relative to the cavity resonance frequency with zero voltage (i.e., without LFD). From (4.74), we obtain

$$|\mathbf{v}_C|^2 = \frac{\omega_{1/2}^2 v_{C0}^2}{\omega_{1/2}^2 + (K|\mathbf{v}_C|^2 + \omega_{rf})^2}. \quad (4.75)$$

Then we get

$$\omega_{rf} = -K|\mathbf{v}_C|^2 \pm \frac{\omega_{1/2}}{|\mathbf{v}_C|} \sqrt{v_{C0}^2 - |\mathbf{v}_C|^2}, \quad (4.76)$$

which gives the required ω_{rf} to achieve a particular cavity voltage $|\mathbf{v}_C|$. Specifically, if we want to achieve the maximum cavity voltage v_{C0} , then ω_{rf} should be $-Kv_{C0}^2$. We calculate the required ω_{rf} for various $|\mathbf{v}_C|$ and plot the results in Fig. 4.43a. The results form a tilted cavity resonance curve compared to the one without LFD. The resonance peak moves toward lower frequencies due to stronger Lorentz force with

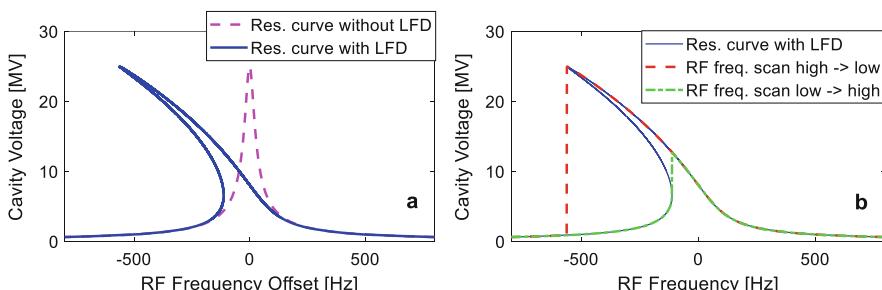


Fig. 4.43 (a). Cavity resonance curves with LFD; (b). static instability when scanning ω_{rf}

increasing cavity voltage. We may measure the cavity resonance curve by scanning the frequency of a constant RF drive power. The results differ from each other, depending on the direction of frequency scanning.

When we scan ω_{rf} from higher to lower values, the measured resonance curve follows the dashed line in Fig. 4.43b. Let us start the scanning from $\omega_{rf} > 0$. After we build up a nonzero cavity voltage (i.e., $|v_C| > 0$), the cavity resonance frequency ω_{cav} decreases due to the LFD. When reducing ω_{rf} while keeping the RF drive power constant, ω_{rf} moves closer to ω_{cav} and we obtain a larger $|v_C|$. It then generates stronger LFD and pushes ω_{cav} further lower. This procedure repeats when further reducing ω_{rf} until we reach the maximum cavity voltage v_{C0} at $\omega_{rf} = -K|v_{C0}|^2$. At this critical point, a further reduction of ω_{rf} will reduce the cavity voltage, reducing the LFD, and increasing the ω_{cav} . Therefore, the positive feedback condition is established at this point, causing a sudden drop in the cavity voltage. Similarly, if we scan ω_{rf} from lower to higher values, the measured resonance curve follows the dash-dotted line. It is seen that the cavity voltage jumps at a different frequency.

Note that the jumps (i.e., static instability) occur only for a large v_{C0} with the maximum LFD larger than the cavity bandwidth. For a small v_{C0} , the coupling gain between the LFD and cavity voltage mentioned in the last section is smaller than one, and the static instability does not happen.

The analysis above is based on the scanning of ω_{rf} . In practice, ω_{rf} is fixed, and our goal is to achieve the desired cavity voltage with a minimum RF drive power. When starting up a cavity, we first tune it on resonance when the cavity voltage is low with the LFD negligible. Then we ramp the cavity voltage slowly and use a motor tuner to compensate for the LFD simultaneously. This procedure is equivalent to pushing the entire tilted resonance curve (Fig. 4.43a) to higher frequencies. When the cavity voltage reaches its maximum value v_{C0} , the tuner should have compensated for the LFD completely. Then the cavity operates at the peak point of the tilted resonance curve. As we have analyzed, the cavity voltage may drop suddenly if the cavity resonance frequency increases due to microphonics.

4.6.4.3 Mitigation of Ponderomotive Instability

The following methods can be used to mitigate the static instability:

- (a) *Choose a cavity operating point away from the resonance peak.* We may operate the cavity with ω_{cav} slightly smaller than ω_{rf} . The operating point should guarantee that the microphonics do not push ω_{cav} to above ω_{rf} .
- (b) *Apply fast piezo-based tuning feedback control.* The tuning feedback loop helps to reduce the variation of ω_{cav} due to microphonics. It can also reduce the coupling gain between the LFD and cavity voltage, thus avoids instability (Bellandi et al. 2020). For example, a cavity voltage variation causes a change in the LFD. If the tuning control can compensate for the LFD variation, the resulting further changes in the cavity voltage can be reduced. That is, the coupling gain between the LFD and cavity voltage is reduced.

- (c) *Apply a feedforward piezo drive proportional to $|v_C|^2$.* This method has been successfully demonstrated at the Cornell University (Schappert et al. 2015).

The methods b) and c) above both reduce the effective LFD (actual LFD plus the piezo-induced resonance frequency change). If the effective LFD is smaller than the cavity bandwidth, the static instability can be avoided.

The mitigation of the oscillatory instability relies on the amplitude and tuning feedback loops. Feedbacks with more vigorous disturbance rejection can help damp the oscillations or prohibit the oscillations from starting. For example, a stable cavity voltage produces less LFD variation, and a stable cavity resonance frequency results in little cavity voltage fluctuations. Both help to prohibit the oscillations in the cavity voltage and cavity resonance frequency.

4.7 Summary

In this chapter, the control algorithms are discussed for cavity field control and resonance control. The basic controller design (e.g., loop shaping, internal model control, etc.) is not discussed, and the readers should refer to control theory textbooks. The GDR stability issues are emphasized. We introduce the ADRC controller because it illustrates excellent performance with reasonable complexity. We also discuss some details of the SEL and PLL principles. The analysis of PLL demonstrates how to model and analyze the controllability of a practical control problem. For adaptive feedforward, the iterative learning control is popular in recent years; however, the time-reversed filter-based method is simple and practical for RF systems controlled by intra-pulse feedbacks. The resonance control part covers basic concepts and algorithms, which can be the basis for developing more advanced algorithms. We also introduce the ponderomotive instability, which has been well studied and can be mitigated by mature methods, but knowing the concepts will be helpful to understand the cavity behavior observed during operation.

References

- Allison T, Delayen J, Hovater C et al (2007) A digital self excited loop for accelerating cavity field control. In: Proceedings of the 22nd IEEE particle accelerator conference, Albuquerque
- Ayvazyan V, Koprek W, Kostin D et al (2011) Design and implementation of automatic cavity resonance frequency measurement and tuning procedure for FLASH and European XFEL cryogenic modules. In: Proceedings of the 2ed international particle accelerator conference, San Sebastian
- Bellandi A, Branlard J, Eichler A et al (2020) Integral resonance control in continuous wave superconducting particle accelerators. IFAC-PapersOnLine 53(2):361–367. <https://doi.org/10.1016/j.ifacol.2020.12.186>

- Blanco Viñuela E, Casas Cubillos J, de Prada MC (2000) Linear model-based predictive control of the LHC 1.8 K cryogenic loop. In: Shu QS (ed) Advances in cryogenic engineering. Advances in cryogenic engineering. Springer, Boston. https://doi.org/10.1007/978-1-4615-4215-5_119
- Brandt A (2007) Development of a finite state machine for the automated operation of the LLRF control at FLASH. Ph.D. thesis, Hamburg University
- Carcagno R, Bellantoni L, Berenc T et al (2003) Microphonics detuning compensation in 3.9 GHz superconducting RF cavities. In: Proceedings of the 11th workshop on RF superconductivity, Luebeck/Travemuender
- Dai J, Zhang J, Huang H et al (2013) LLRF and data acquisition systems for spoke012 cavity vertical test at IHEP. In: Proceedings of the 16th international conference on RF superconductivity, Paris
- Delayen JR (1978) Phase and amplitude stabilization of superconducting resonators. Ph.D. thesis, California Institute of Technology
- Delayen JR (2005) Ponderomotive instabilities and microphonics – a tutorial. In: Proceedings of the 12th international workshop on RF superconductivity, Ithaca
- Fafara P, Jalmuzna W, Koprek W et al (2007) FPGA-based implementation of a cavity field controller for FLASH and X-FEL. Meas Sci Technol 18(8):2365–2371. <https://doi.org/10.1088/0957-0233/18/8/010>
- Fu X, Fong K, Yin Z et al (2020) A digital phase-locked loop based LLRF system. Nucl Instrum Methods Phys Res A962:163688. <https://doi.org/10.1016/j.nima.2020.163688>
- Geng Z (2017a) Superconducting cavity control and model identification based on active disturbance rejection control. IEEE Trans Nucl Sci 64(3):951–958. <https://doi.org/10.1109/TNS.2017.2663660>
- Geng Z (2017b) RF control optimization and automation for normal conducting linear accelerators. IEEE Trans Nucl Sci 64(8):2361–2368. <https://doi.org/10.1109/TNS.2017.2723508>
- Grecki M, Przygoda K, Weddig H et al (2018) Piezo control at DESY, from FLASH to XFEL. Paper presented at the 2ed topical workshop on cryomodule microphonics and resonance control, Brookhaven National Laboratory and Jefferson Lab, 25–26 October 2018
- Han J (2009) From PID to active disturbance rejection control. IEEE Trans Ind Electron 56(3): 900–906. <https://doi.org/10.1109/TIE.2008.2011621>
- Herbst G (2013) A simulative study on active disturbance rejection control (ADRC) as a control tool for practitioners. Electronics 2(3):246–279. <https://doi.org/10.3390/electronics2030246>
- Kandil T, Grimm TL, Hartung W et al (2004) Adaptive feedforward cancellation of sinusoidal disturbances in superconducting RF cavities. In: Proceedings of the 22ed international linear accelerator conference, Luebeck
- Kichhoff S, Schmidt C, Lichtenberg G et al (2008) An iterative learning algorithm for control of an accelerator based free electron laser. In: Proceedings of the 47th IEEE conference on decision and control, Cancun
- Laverty M, Fong K, Zheng Q et al (2010) Design and testing of the TRIUMF ISACII high-beta RF control system. In: Proceedings of the 25th linear accelerator conference, Tsukuba
- Leewe R, Fong K (2019) Simulation analysis of Lorentz force induced oscillations in RF cavities in vector sum and CW operation. In: Proceedings of the 19th international conference on RF superconductivity, Dresden
- Michizono S, Katagiri H, Matsumoto T et al (2009) Vector-sum control of superconducting RF cavities at STF. In: Proceedings of the 23rd IEEE particle accelerator conference, Vancouver
- Miljkovic D (2016) Active noise control: from analog to digital – last 80 years. Paper presented in the 36th international convention on information and communication technology, electronics and microelectronics, Opatija, Croatia, 30 May-3 June 2016
- Neumann A (2008) Compensating microphonics in SRF cavities to ensure beam stability for future free-electron-lasers. Ph.D. thesis, Humboldt University of Berlin
- Neumann A, Anders W, Kugeler O et al (2010) Analysis and active compensation of microphonics in continuous wave narrow-bandwidth superconducting cavities. Phys Rev ST Accel Beams 13(8):082001. <https://doi.org/10.1103/PhysRevSTAB.13.082001>

- Paparella R (2007) Fast frequency tuner for high gradient SC cavities for ILC and XFEL. Ph.D. thesis, University of Milan
- Pfeiffer S, Lichtenbert G, Schmidt C et al (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
- Pfeiffer S, Butkowski L, Hensler O et al (2018) Precision feedback control of a normal conducting standing wave resonator cavity. *Phys Rev Accel Beams* 21(9):092802. <https://doi.org/10.1103/PhysRevAccelBeams.21.092802>
- Pfeiffer S, Eichler A, Schlarb H (2020) Model-based feed-forward control for time-varying systems with an example for SRF cavities. *IFAC PapersOnLine* 53(2):1331–1336
- Pischalnikov Y, Cancelo G, Chase B et al (2012) Lorentz force compensation for long pulses in SRF cavities, Proceedings of the 3rd international particle accelerator conference, New Orleans
- Qian XP, Yao ZE, Wang Q (2014) Model-predictive control of power supply for particle accelerators. *Nucl Sci Tech* 25(5):050203. <https://doi.org/10.13538/j.1001-8042/nst.25.050203>
- Rezaeizadeh A (2016) Automatic control strategies for the Swiss free electron laser. Ph.D. thesis, Eidgenössische Technische Hochschule Zürich
- Rybaniec R, Przygoda K, Cichalewski W et al (2017) FPGA-based RF and piezocontrollers for SRF cavities in CW mode. *IEEE Trans Nucl Sci* 64(6):1382–1388. <https://doi.org/10.1109/TNS.2017.2687981>
- Schappert W, Holzbauer J, Pischalnikov Y et al (2015) Resonance control for narrow-bandwidth, superconducting RF applications. In: Proceedings of the 17th international conference on RF superconductivity, Whistler
- Schilcher T (1998) Vector sum control of pulsed accelerating fields in Lorentz force detuned superconducting cavities. Ph.D. thesis, Hamburg University
- Schulze D (1972) Ponderomotive stability of RF resonators and resonator control systems. Ph.D. thesis, Universitt Karlsruhe
- Simrock S (2003) Control of microphonics and Lorentz force detuning with a fast mechanical tuner. In: Proceedings of the 11th workshop on RF superconductivity, Luebeck/Travemuender
- Skogstad S, Postlethwaite I (2005) Multivariable feedback control: analysis and design, 2nd edn. Wiley, New York
- Vincent J, Morris D, Usher N et al (2011) On active disturbance rejection based control design for superconducting RF cavities. *Nucl Instrum Methods Phys Res A* 643(1):11–16. <https://doi.org/10.1016/j.nima.2011.04.033>
- Vogel E (2007) High gain proportional RF control stability at TESLA cavities. *Phys Rev ST Accel Beams* 10(5):052001. <https://doi.org/10.1103/PhysRevSTAB.10.052001>
- Wibowo SB, Matsumoto T, Michizono et al (2018) Digital low level RF control system for the international linear collider. *Phys Rev Accel Beams* 21(8):082004. <https://doi.org/10.1103/PhysRevAccelBeams.21.082004>

Chapter 5

RF Detection and Actuation



RF detectors transform the RF frequency signals to baseband signals in the format of complex envelopes (i.e., phasors) or amplitude and phase. RF actuators perform the inverse transformations. This chapter starts with an introduction to several widely used analog or digital RF detection schemes. Then we focus on the digital solutions and provide a few popular RF detection algorithms. The non-I/Q demodulation algorithm will be discussed in detail. Furthermore, some advanced topics of RF detection involving reference tracking will be discussed. Finally, we present some basic RF actuation schemes, such as the direct up-conversion, the single sideband (SSB) up-conversion, and the intermediate frequency (IF) up-conversion. These up-conversion schemes are also frequently found in commercial RF transmitters.

5.1 RF Detection Schemes

The RF detector input is an RF signal to be measured, and the outputs are its amplitude and phase or I/Q components. The RF detector output represents the complex envelope (also known as a *baseband signal*) of the input RF signal. Classical analog RF detectors typically operate at the RF frequency and yield the baseband outputs directly. A digital RF detector uses an ADC to sample the RF signal or an *intermediate frequency (IF) signal* down converted from the original RF signal and calculates the baseband outputs with digital signal processing algorithms.

The performance requirements of an RF detector are typically specified by its *bandwidth*, *accuracy*, *linearity*, *signal-to-noise ratio (SNR)*, and *latency*.

The bandwidth describes how fast changes in the amplitude or phase of the RF signal can be detected. For example, to detect an RF pulse envelope with a rise time $t_r = 20 \text{ ns}$ (t_r is the time for the envelope to change from 10 to 90% of its final value), we require the RF detection bandwidth $f_B > 0.35 / t_r \approx 17.5 \text{ MHz}$ (Buzuayene 2008). Here f_B is the bandwidth of the baseband signal; that is, it is the half-bandwidth of the RF signal if we consider both sidebands around the RF carrier frequency.

The accuracy indicates the steady-state error of the RF detector, e.g., for a phase change $\Delta\varphi$ in the RF signal, how accurate the detected phase change $\Delta\varphi_m$ is. The accuracy of analog RF detectors is often an issue (Geng et al. 2005), whereas, for digital RF detectors, the accuracy is generally remarkably high.

The linearity of an RF detector determines the RF detection error when the RF input level reaches the saturation of the RF detector, which will be discussed in Chap. 7. The noise level, often quantified by the SNR, is a primary concern when designing an RF detector. We must reduce the RF detection noise within the closed-loop bandwidth of the RF feedback loop to achieve the desired field stability. The noise in RF detectors and its effects on RF field stability will be discussed in Chap. 6. Finally, the RF detector latency contributes to the loop delay when used in a feedback loop. As discussed in Chap. 4, we prefer a small loop delay compared to the cavity time constant to improve the RF feedback performance.

Specifying the requirements of RF detectors is a crucial task when designing an LLRF system. We will cover the noise and linearity specifications in the later chapters (Chaps. 6 and 7). This chapter will mainly focus on the principles and algorithms of RF detection.

In this chapter, the *regular frequency* f will be used instead of the *angular frequency* ω (they satisfy $\omega = 2\pi f$). We have used ω when describing the RF system models and RF controllers, but f is more used to describe RF circuits and RF signal processing.

5.1.1 Amplitude and Phase Detectors

The simplest implementation of an RF detector is a *diode detector* for amplitude detection and a *mixer* for phase detection. See Fig. 5.1a.

Since the amplitude and phase detectors are separated, we use a power splitter to supply inputs to them. For this purpose, 3-dB Hybrids (low loss) or Wilkinson divider (3-dB power loss, but better decoupling) can be used. A good decoupling of the two outputs of the power splitter ensures that the impedance changes of the amplitude or phase detector (e.g., with temperature) do not affect the other measurement channel.

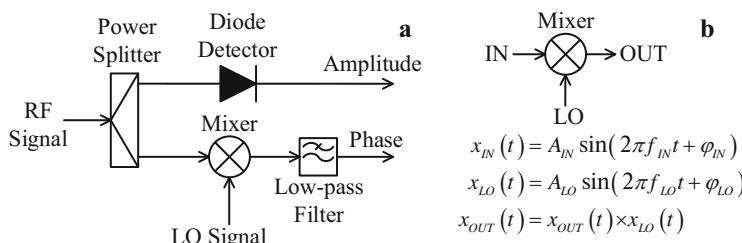


Fig. 5.1 (a) An RF detector consisting of a diode detector and a mixer; (b) Ideal RF mixer model

The diode detector can detect the amplitude envelope of an RF signal with a very high resolution. It has excellent SNR in amplitude detection compared to other RF detection schemes. Diode detectors are still widely used in radio communication systems. Many integrated circuits are available on the market ranging from temperature compensated diode detectors to complex logarithmic detectors.

RF Mixers are 3-port active or passive devices. They are designed to yield both the sum and the difference frequency at a single output port when two distinct input frequencies are inserted into the other two ports. An ideal mixer can be viewed as a multiplier, as in Fig. 5.1b.

With the input signals shown above, the output of the mixer is given by

$$\begin{aligned} x_{OUT}(t) &= x_{IN}(t) \cdot x_{LO}(t) \\ &= \frac{A_{IN}A_{LO}}{2} \underbrace{(\cos [2\pi(f_{IN} - f_{LO})t + (\varphi_{IN} - \varphi_{LO})])}_{\text{lower sideband}} \\ &\quad - \underbrace{\cos [2\pi(f_{IN} + f_{LO})t + (\varphi_{IN} + \varphi_{LO})])}_{\text{upper sideband}} \end{aligned} \quad (5.1)$$

Both a *lower sideband* and an *upper sideband* appear in the mixer output with the difference and sum of the two input frequencies. In a mixer, the local oscillator (LO) input is a reference signal, and the IN port accepts either an RF signal to be detected (for a receiver) or an IF signal to be modulated (for a transmitter). If we obtain an output frequency lower than the input frequency in the mixing process, we call it *down-conversion*, otherwise, we call it *up-conversion*. For down-conversion, we often label the IN/OUT ports in Fig. 5.1b as RF/IF and insert a low-pass filter in the output port; whereas, for up-conversion, we label them as IF/RF and insert a high-pass or band-pass filter in the output.

A mixer used for phase detection works in down-conversion mode with the frequencies of LO and RF inputs identical. After low-pass filtering the mixer output and tuning the LO phase to $\pi/2$ (with a phase shifter), we get the IF output

$$x_{IF}(t) = \frac{A_{RF}A_{LO}}{2} \sin \varphi_{RF} \approx \frac{A_{RF}A_{LO}}{2} \varphi_{RF} \quad (\text{for small } \varphi_{RF}). \quad (5.2)$$

Here the mixer output is proportional to the phase of the RF input signal. Note that the linear approximation in (5.2) is only valid for $\varphi_{RF} \ll 1$ rad. The accuracy of phase detection will get worse for a large phase variation due to stronger nonlinearity. Furthermore, fluctuations in LO and RF amplitudes will contribute to the phase measurement error. In practice, the LO port operates in saturation, and the LO signal amplitude A_{LO} can be viewed as a constant. The noise introduced by a state-of-the-art mixer is typically very low. Due to their simplicity, mixer-based phase detectors are widely used in phase-locked loops and analog LLRF systems.

5.1.2 Analog I/Q Demodulator

An analog I/Q demodulator consists of two mixers operating in quadrature (i.e., 90° out of phase), which has been introduced in Sect. 3.2.2 and illustrated in Fig. 3.2.

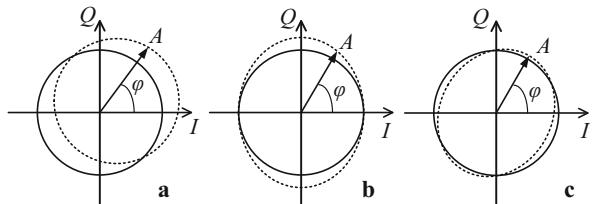
With identical LO and RF frequency, the I/Q demodulator forms an RF detector for *direct down-conversion*. That is, it down-converts an RF signal directly to baseband without intermediate stages. The outputs of the low-pass filters in the I and Q ports are the real and imaginary parts of the RF signal's complex envelope. For convenience, we normalize the LO phase to zero and assume the LO port operating in saturation, such that the LO amplitude jitter is isolated for the outputs. As discussed in Sect. 3.2.2, the I/Q demodulator outputs form a phasor of the input RF signal represented as

$$\mathbf{x}_{RF}(t) = I(t) + jQ(t) = \frac{A_{RF}A_{LO}}{4} e^{j\varphi_{RF}}. \quad (5.3)$$

Here A_{RF} and A_{LO} are the amplitudes of the RF and LO signals, respectively. The LO phase $\varphi_{LO} = 0$ and φ_{RF} is the RF signal phase. The outputs of the I/Q demodulator are often sampled by two ADCs, and the amplitude and phase can be calculated in software (or firmware). The I/Q demodulator can detect the phasor of the input RF signal directly. Compared to a single mixer, the phase detection range covers a full cycle, i.e., 360°. Since the outputs are two slowly changing baseband signals, ADCs with low sampling rates and higher resolution can sample them. This also imposes less critical requirements on the ADC clock jitter. When an ADC samples an IF signal of a higher frequency, the clock jitter becomes more critical because the time jitter converts directly to the phase of the IF frequency. We will discuss the jitter issues in Chap. 6. Furthermore, since an I/Q demodulator is constructed with only passive components and mixers, its added noise is very low.

In practice, several weak points of direct down-conversion may limit its usage (Hoffmann 2008). First, the baseband signal might be in the same frequency range as the environmental noise (e.g., ripples or spurious frequencies in the power supplies of the crate or ADC boards, or the electrical noise picked up from the environment, such as the klystron modulator noise coupled via grounding or space radiation). It is difficult to remove such noise with frequency filters. Another source of RF detection errors is the LO leakage, amplitude and phase imbalances in the I/Q demodulator. They are due to technological limitations in manufacturing practical I/Q demodulators, such as the PCB trail length mismatch and the conversion loss deviation in the two mixers. At the I/Q demodulator mixers, the LO signal may leak to the RF port and mix with the LO signal itself, resulting in DC offsets in the I and Q outputs. Furthermore, the two branches of the I/Q demodulator may not be balanced in amplitude or phase, both causing static amplitude and phase measurement errors. We can illustrate the consequences of LO leakage and imbalances graphically in Fig. 5.2.

Fig. 5.2 RF detection errors caused by the DC offsets (due to LO leakage, plot a) and the amplitude (plot b) and phase (plot c) imbalances of an I/Q demodulator



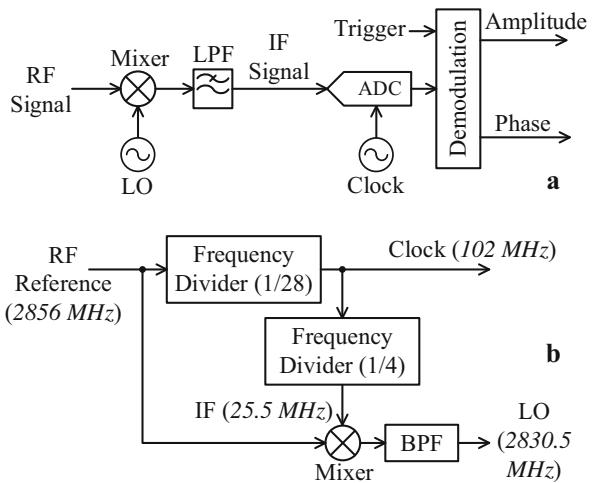
The I/Q demodulator outputs form a circle in the I/Q plane if we fix the LO signal and rotate the RF signal phase while keeping its amplitude constant. The LO leakage and the amplitude and phase imbalances deform the response circle, causing errors dependent on the RF input phase. These errors are systematic errors resulting in poor RF detection accuracy. In an RF feedback loop, such errors are at DC frequency and are always transferred to the loop output, the RF field in our case.

Systematic errors can often be calibrated. In a digital LLRF system, we may calibrate the distorted I and Q outputs of the I/Q demodulator in software or firmware. For example, we can add offsets to the measured phasor to compensate for the LO leakage, and scale and rotate the results to eliminate the effects of imbalances. However, this requires a procedure to determine all the necessary parameters for the calibration. The temperature sensitivity of some of these parameters, such as the offsets, may require additional temperature compensation algorithms. In this book, the calibration of I/Q demodulators will not be discussed because the direct down-conversion scheme is rarely used in LLRF systems nowadays. Readers who are interested in the calibration can refer to the article (Geng et al. 2005).

5.1.3 Down-Converter and IF Sampling

A general architecture of the RF detectors in digital LLRF systems is depicted in Fig. 5.3a. The RF signal to be measured is down-converted by mixing with an LO signal with a frequency offset ($f_{LO} = f_{RF} - f_{IF}$, where $-f_{IF}$ is the frequency offset). A low-pass filter selects the lower sideband of the mixing product, resulting in an IF signal with a frequency f_{IF} . We have chosen the LO frequency lower than the RF frequency such that the phases of the RF and IF signals rotate in the same direction. It means, an increase in the RF phase should result in the same increase in the IF phase. Then we use an ADC to sample the IF signal and demodulate the samples in software or firmware to derive the complex envelope or the amplitude and phase of the IF signal. As seen from (5.1), the amplitude of the IF signal is proportional to the amplitude of the RF signal ($A_{IF} = A_{RF}A_{LO}/2$), and the phase of the IF signal has a constant offset ($-\varphi_{LO}$) to the phase of the RF signal. Here we have assumed the amplitude (A_{LO}) and phase (φ_{LO}) of the LO signal constant.

Fig. 5.3 General architecture of the RF detectors in digital LLRF systems (plot a) and a typical LO and clock generator (plot b, the frequencies are examples)



Various digital demodulation algorithms have been developed, which will be discussed in Sect. 5.2. For pulsed RF signals, a trigger is often used to start the demodulation process. The LO and clock signals are typically synchronized to the RF signal, that is, they are derived from the same RF reference frequency. Figure 5.3b is an example of a synchronous LO and clock generator. The scheme with synchronous LO and clock can simplify the demodulation algorithms and is selected in many LLRF systems. In some cases, asynchronous LO or clock signals (e.g., from separate oscillators) may be used. They require specific algorithms that are covered in Sect. 5.2.5.

5.1.4 Direct RF Sampling

With state-of-the-art fast ADCs, it is possible to sample the RF signal directly without down-conversion. For example, the ADS5474 (14 bits, 400 MSPS, and 1.4 GHz analog input bandwidth) can be used to sample an RF signal of 1.3 GHz (Geng and Simrock 2008; Habib 2013).

For direct sampling, the clock frequency is typically lower than the RF frequency. This forms the so-called *subsampling* or *undersampling* scheme, which aliases the RF frequency to the clock's Nyquist band. The clock frequency should be selected so that the relationship between the clock and aliased RF frequencies follows the ratio required by specific demodulation algorithms (see Sect. 5.2.2). Then the same digital demodulation algorithm can be applied to the samples.

Since the direct sampling scheme does not need down converters, it can significantly reduce the complexity of the RF detector hardware. However, it is more sensitive to the clock jitter due to the high-frequency inputs to ADCs. The phase measurement error $\Delta\varphi_{RF}$ caused by the clock jitter Δt_{CLK} is given by $\Delta\varphi_{RF} = 2\pi f_{RF} \Delta t_{CLK}$, where f_{RF} is the ADC input frequency. We will provide more details of the noise analysis in Chap. 6.

5.2 RF Detection Algorithms

Digital RF detectors demodulate the samples of the IF (RF) signals in digital signal processors (e.g., FPGA). Several algorithms, such as the I/Q demodulation, non-I/Q demodulation, and digital down-conversion, have been developed for LLRF systems. They correspond to various relationships between the clock frequency and the frequency of ADC inputs.

5.2.1 I/Q Demodulation

5.2.1.1 I/Q Sampling and Demodulation Algorithm

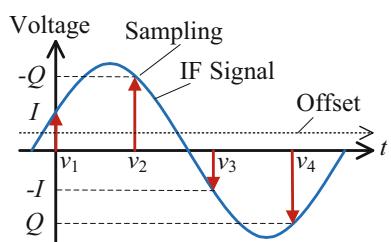
I/Q demodulation requires the clock frequency to be four times the IF frequency (or the aliased RF frequency for subsampling cases; here, we only consider the IF sampling for simplicity). Such a sampling scheme is known as the *I/Q sampling* (Ziomek and Corredoura 1995). See Fig. 5.4. We define the first sample as the real part (i.e., the *I* component) of the complex envelope (phasor) of the RF signal. Then the followed samples stand for $-Q$, $-I$, Q , I ..., and so on. Here Q is the imaginary part of the phasor. We can demodulate the samples of the IF signal using every two neighboring points as

$$\mathbf{x}_{IF}(1) = v_1 - jv_2, \mathbf{x}_{IF}(2) = -v_3 - jv_2, \mathbf{x}_{IF}(3) = -v_3 + jv_4, \dots \quad (5.4)$$

Compared to an analog I/Q demodulator, the I/Q sampling scheme can eliminate the errors in Fig. 5.2 and achieve accurate RF signal measurement. The implementation of the demodulation algorithm is simple. It introduces a group delay of one clock cycle, which is negligible for a clock frequency of tens of MHz.

Note that the measured phase of the IF signal depends on the starting point defined as *I*. Shifting the starting point by one sample earlier will change the phase value by -90° , whereas shifting by one sample later causes 90° phase change. Another issue we should handle is the DC offset in the IF samples. The DC offset causes errors if we directly apply (5.4) to the ADC samples. To compensate for the offset, we use the average of four samples:

Fig. 5.4 I/Q sampling of the IF signal



$$\begin{aligned}\mathbf{x}_{IF}(1) &= \frac{v_1 - v_3}{2} - j\frac{v_2 - v_4}{2}, \mathbf{x}_{IF}(2) = -\frac{v_3 - v_5}{2} - j\frac{v_2 - v_4}{2}, \\ \mathbf{x}_{IF}(3) &= -\frac{v_3 - v_5}{2} + j\frac{v_4 - v_6}{2}, \dots\end{aligned}\quad (5.5)$$

Here we have assumed the DC offset a constant. Note that the updated algorithm (5.5) introduces a group delay of two clock cycles.

5.2.1.2 Harmonics Aliasing of I/Q Sampling

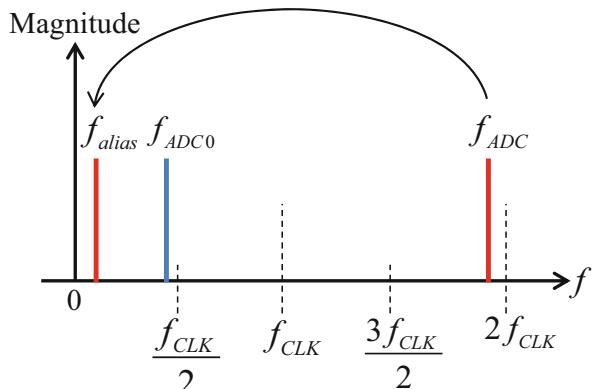
The I/Q sampling scheme is widely used in early generations of digital LLRF systems. It provides accurate amplitude and phase measurements for a pure sinusoidal signal. However, due to the nonlinearities of mixers and ADCs, the ADC samples contain both the IF signal and its harmonics. For IF harmonics with frequencies higher than half of the clock frequency (i.e., Nyquist frequency), they are aliased to below the Nyquist frequency by the sampling process. It will be shown that some harmonics overlap with the IF frequency after being aliased, resulting in errors in the measurements of the RF signal (Schilcher 2008).

Let us look at the *frequency aliasing* in a more general sense. Assume the clock frequency is f_{CLK} , and then the Nyquist band is from DC (0 Hz) to $f_{CLK}/2$. An ADC input with a frequency $f_{ADC0} < f_{CLK}/2$ can be sampled without frequency aliasing. See Fig. 5.5. If the ADC input frequency satisfies $f_{ADC} \geq f_{CLK}/2$, it will be aliased to a frequency given by

$$f_{alias} = |f_{ADC} - pf_{CLK}|, \quad (5.6)$$

where p is an integer and should be selected to minimize the value of f_{alias} . For example, if $f_{CLK} = 100$ MHz and $f_{ADC} = 180$ MHz, then $f_{alias} = 20$ MHz with $p = 2$. In an RF detector, as shown in Fig. 5.3a, the harmonics of the IF signal up to the analog bandwidth of ADC will show up in the ADC samples. These harmonics will

Fig. 5.5 Frequency aliasing when sampling a signal with a frequency smaller (f_{ADC0}) or larger (f_{ADC}) than the Nyquist frequency $f_{CLK}/2$



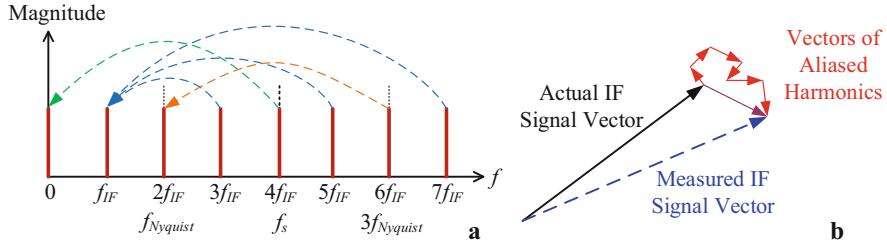


Fig. 5.6 (a) Frequency aliasing of the IF harmonics with the I/Q sampling scheme; (b) RF detection errors caused by the aliased harmonics

be all aliased to the Nyquist band of the ADC clock frequency. In particular, for I/Q sampling, the odd-ordered harmonics of the IF frequency will overlap with the IF frequency after being aliased, as depicted in Fig. 5.6a.

Let us check the demodulation results of the ADC samples. Each aliased harmonic can be viewed as a small phasor added to the phasor of the IF fundamental frequency. See Fig. 5.6b. When the phase of the IF signal varies (e.g., due to microphonics in the cavity), the phases of the higher-order harmonics will change faster (i.e., n times faster for the n th harmonic) than that of the IF signal. That is, the aliased harmonics introduce time-varying errors in the RF measurements. When the RF detector is used in a feedback loop as a sensor, such time-varying RF detection errors will cause deviations in the cavity field, which must be avoided. This leads to the non-I/Q sampling scheme discussed in the next section.

5.2.2 Non-I/Q Demodulation

5.2.2.1 Non-I/Q Sampling

To overcome the aliasing of harmonics in I/Q sampling, we shift the clock frequency slightly away from four times the IF frequency. Then we get a new scheme called *non-I/Q sampling* (Doolittle et al. 2006), for which the frequencies have the following relationship

$$f_{CLK} = \frac{n}{m} f_{IF}, \quad (5.7)$$

where n and m are integers. Typically, we choose $n > 2m$ for IF sampling such that $f_{IF} < f_{CLK}/2$, resulting in no frequency aliasing. The relationship (5.7) corresponds to sampling m IF periods using n samples. See Fig. 5.7 as an example.

In direct RF sampling, the ADC input frequency (f_{RF}) is usually larger than the clock frequency, for which we may use the following frequency relationship

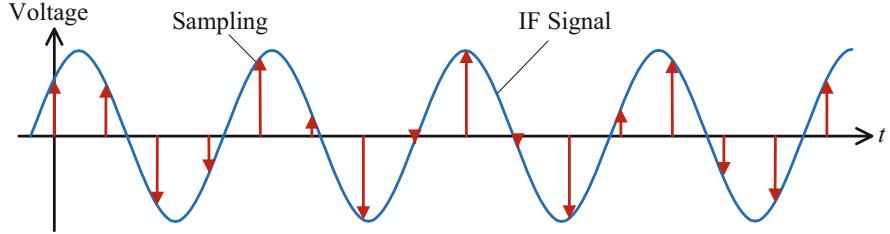


Fig. 5.7 Example of non-I/Q sampling with $m = 4$ and $n = 15$

$$f_{CLK} = \frac{f_{RF}}{p + \frac{m}{n}}, \quad p = 0, 1, 2, \dots \quad (5.8)$$

Then the RF frequency is aliased to

$$f_{RF,alias} = f_{RF} - pf_{CLK} = \frac{m}{n}f_{CLK}. \quad (5.9)$$

Here the frequency aliasing formula (5.6) has been used, and we still require $n > 2m$. Note that the value of p should minimize $f_{RF} - pf_{CLK}$ under the constraint $f_{RF} > pf_{CLK}$, which guarantees that the phases of the RF signal and its aliased signal rotate in the same direction (i.e., an increase of the RF phase causes a corresponding phase increase of the aliased signal). For example, assume we directly sample an RF signal $f_{RF} = 1.3$ GHz with $m = 4$ and $n = 15$, the f_{CLK} may be selected to be 178.899 or 157.258 MHz if we limit it to $150 < f_{CLK} < 200$ MHz. In particular, if $f_{CLK} = 178.899$ MHz, then the RF frequency is aliased to 47.706 MHz.

With the sampling Scheme (5.7) and (5.8), we can use the same algorithm to demodulate the ADC samples.

5.2.2.2 Harmonics Aliasing of Non-I/Q Sampling

According to (5.6), the h th ($h > 1$) harmonic of the IF frequency (our discussion is based on (5.7) and the results also apply to (5.8)) is aliased to

$$f_{hf_{IF},alias} = |hf_{IF} - pf_{CLK}| = \left| \frac{hm}{n} - p \right| f_{CLK} = \left| h - \frac{pn}{m} \right| f_{IF}. \quad (5.10)$$

It overlaps with f_{IF} when $|h - pn/m| = 1$, which helps to determine the lowest-order harmonic that overlaps with the IF fundamental frequency. If n and m are coprime, the lowest order of the harmonic aliased to the IF frequency is $(n - 1)$ for $p = m$. For example, if we select $m = 4$ and $n = 15$, the first overlap happens at $h = 14$. That is, the IF harmonics with an order less than 14 are all separated with the IF frequency in the spectrum. The frequency aliasing of this example is depicted in Fig. 5.8.

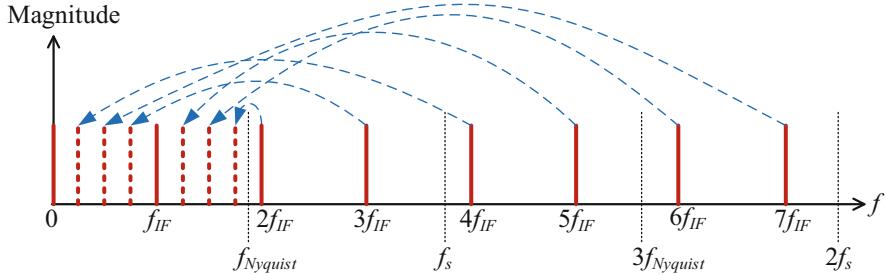


Fig. 5.8 Aliasing of IF harmonics with non-I/Q sampling ($m = 4$, $n = 15$). Only the first seven harmonics are shown

The number of IF harmonics in the ADC samples is determined by the analog bandwidth and nonlinearity of the RF detector. Typically, we only need to consider the first few harmonics with larger magnitudes. A proper combination of n and m should be selected to separate the concerned harmonics from the IF fundamental frequency. Of course, the group delay of the non-I/Q demodulation algorithm is another factor that determines the n and m values. We will discuss this topic in the next part.

5.2.2.3 Non-I/Q Demodulation Algorithm

With the non-I/Q sampling, the IF phase change in one clock period is

$$\Delta\varphi = 2\pi m/n = 2\pi f_{IF}/f_{CLK}. \quad (5.11)$$

We present the IF signal as a sinusoidal function $x(t) = A\sin(2\pi f_{IF}t + \varphi)$, and the k th sample is given by

$$\begin{aligned} x_k &= A \sin(2\pi f_{IF}kT_s + \varphi) \\ &= A \sin(k\Delta\varphi + \varphi) \\ &= I \sin(k\Delta\varphi) + Q \cos(k\Delta\varphi), \quad k = 0, 1, 2, \dots, \end{aligned} \quad (5.12)$$

where $T_s = 1/f_{CLK}$ is the sampling period and the I and Q are defined by

$$I := A \cos \varphi, \quad Q := A \sin \varphi. \quad (5.13)$$

Here we assume that the amplitude A and phase φ of the IF signal are constant within m IF periods to simplify the demodulation algorithm.

To derive the I and Q values from the ADC samples (5.12), we may consider two possibilities: demodulate using two samples or using n samples. Let us look at these algorithms. Note that we assume the clock and the IF signal are synchronized (i.e., derived from the same frequency source).

Demodulation with Two Samples

If we use two adjacent samples, x_k and x_{k-1} , the I and Q values can be derived by solving the linear equations below

$$\begin{aligned} x_k &= I \sin(k\Delta\varphi) + Q \cos(k\Delta\varphi) \\ x_{k-1} &= I \sin[(k-1)\Delta\varphi] + Q \cos[(k-1)\Delta\varphi]. \end{aligned} \quad (5.14)$$

Then we obtain

$$\begin{bmatrix} I \\ Q \end{bmatrix} = \frac{1}{\sin \Delta\varphi} \begin{bmatrix} \cos[(k-1)\Delta\varphi] & -\cos(k\Delta\varphi) \\ -\sin[(k-1)\Delta\varphi] & \sin(k\Delta\varphi) \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix}, \quad (5.15)$$

which requires $\sin \Delta\varphi \neq 0$; that is, the IF phase step $\Delta\varphi$ between two ADC samples should be kept away from multiple π . For example, if $\Delta\varphi = \pi$, it corresponds to sampling the IF signal with a clock frequency $f_{CLK} = 2f_{IF}$, which is not enough to derive the amplitude and phase information of the IF signal.

Note that the algorithm (5.15) is applicable if the relationship between the IF and clock frequencies are known. They do not need to follow a special relationship like (5.7) or (5.8). We usually calculate a pair of I/Q values using each sample and the one before it. This calculation induces a group delay of one clock cycle. Note that the measured phase of the IF signal depends on the starting point of demodulation. Shifting the starting point by one sample earlier will change the phase value by $-\Delta\varphi$, whereas shifting it by one sample later will cause a phase change of $\Delta\varphi$.

Demodulation with n Samples

Since the IF sample (5.12) is periodic every n points, we can write it as a sum of n discrete sinusoidal series (discrete Fourier series)

$$x_k = \frac{1}{n} \sum_{i=0}^{n-1} X_i e^{j2\pi k i / n}, \quad (5.16)$$

where k is the index of the sample, and i is the index of the sinusoidal series with a frequency $f_i = if_{CLK}/n$. The coefficient X_i is given by

$$X_i = \sum_{l=0}^{n-1} x_l e^{-j2\pi l i / n}, \quad (5.17)$$

which is a complex number representing the amplitude and phase of the i th sinusoidal series. Here we have used l as the index for summing to avoid confusion. From (5.11) and (5.12), x_k contains only one frequency $f_{IF} = mf_{CLK}/n$. Therefore, X_i has only two non-zero components with $i = m$ and $i = n - m$. Note that when $i = n - m$, the frequency of the sinusoidal series is $f_{n-m} = (n - m)f_{CLK}/n = f_{CLK} - mf_{CLK}/n = f_{CLK} - f_{IF}$. It will also be aliased to f_{IF} if we convert all the sinusoidal series

frequencies on the right-hand side of (5.16) into the Nyquist band. The non-zero coefficients of the Fourier series are given by

$$\begin{aligned} X_m &= \sum_{l=0}^{n-1} x_l e^{-j2\pi lm/n}, \\ X_{n-m} &= \sum_{l=0}^{n-1} x_l e^{-j2\pi l(n-m)/n} = \sum_{l=0}^{n-1} x_l e^{j2\pi lm/n} \end{aligned} \quad (5.18)$$

Substitute (5.18) into (5.16) and notice X_{n-m} is conjugate to X_m , we get

$$x_k = \frac{1}{n} \left(X_m e^{j2\pi km/n} + X_{n-m} e^{-j2\pi km/n} \right) = \frac{2}{n} \operatorname{Re} \left\{ X_m e^{j2\pi km/n} \right\}, \quad (5.19)$$

where $\operatorname{Re} \{ \cdot \}$ represents the real part of a complex number. To find out the formula to calculate the I and Q components in (5.12), we rewrite it to

$$x_k = I \sin(k\Delta\varphi) + Q \cos(k\Delta\varphi) = \operatorname{Re} \left\{ (Q - jI) e^{j2\pi km/n} \right\}. \quad (5.20)$$

Here we have used the relationship of (5.11). It can be seen that the I and Q components are directly related to the coefficient X_m , then we obtain

$$I = \frac{2}{n} \sum_{l=0}^{n-1} x_l \sin(l\Delta\varphi), Q = \frac{2}{n} \sum_{l=0}^{n-1} x_l \cos(l\Delta\varphi). \quad (5.21)$$

Equation (5.21) describes how to demodulate the IF signal using the first n samples. Practically, we calculate a pair of I/Q values for each clock cycle using the last n samples, including the newest one. Then the algorithm can be rewritten as

$$I_k = \frac{2}{n} \sum_{l=k-n+1}^k x_l \sin(l\Delta\varphi), Q_k = \frac{2}{n} \sum_{l=k-n+1}^k x_l \cos(l\Delta\varphi). \quad (5.22)$$

Note that the demodulation coefficients $\sin(l\Delta\varphi)$ and $\cos(l\Delta\varphi)$ are periodic for every n points. The algorithm (5.22) is valid only when $k \geq n - 1$, that is, we need at least n samples to perform a meaningful demodulation. We also notice that the I/Q demodulation algorithm (5.5) is a particular case of (5.22) with $\Delta\varphi = 90^\circ$.

The non-I/Q demodulation algorithm (5.22) induces a group delay of $n/2$ clock cycles. Note that the measured phase of the IF signal also depends on the starting point of demodulation. Shifting the starting point by one sample earlier will change the phase value by $-\Delta\varphi$, whereas moving it by one sample later will cause a phase change of $\Delta\varphi$.

5.2.2.4 Frequency Response of Non-I/Q Demodulation

As discussed in Sect. 5.2.2.2, non-I/Q sampling can separate the IF fundamental frequency and its harmonics up to the order of $n-2$. This feature helps reduce the RF detection errors caused by the overlapping of harmonics, which has been a problem for the I/Q sampling. However, we still need to remove these aliased IF harmonics from the ADC samples; that is, we need a filter.

The demodulation algorithm (5.22) is equivalent to multiplying the ADC samples with a sine series and a cosine series and averaging the products for every n points. The frequencies of the sine and cosine series equal the IF frequency f_{IF} . The ADC samples consist of the IF frequency and its aliased harmonics, whose frequencies are given by (5.10). The multiplications in (5.22) yield the following frequency components derived from the ADC samples:

- The IF signal in the ADC samples is converted to DC and $2f_{IF}$. The DC component is our expected I or Q value, whereas the component at $2f_{IF}$ (we denote it as x_{mul+}) is aliased to the Nyquist band according to (5.10).
- The original DC component and aliased IF harmonics in the ADC samples are shifted by $\pm f_{IF}$ and aliased again into the Nyquist band following (5.10). These signals are unwanted and are viewed as spurious signals, and their frequencies locate at qf_{CLK}/n , where q is an integer and $1 \leq q \leq n/2$.

In (5.22), the moving average of n points of the multiplication is equivalent to applying to the products an FIR filter with identical coefficients $c_l = 2/n$ ($l = 0, 1, \dots, n-1$). The frequency response of the moving-average filter is given by

$$\mathbf{H}_{mvavg}(f) = \frac{2}{n} \sum_{l=0}^{n-1} e^{-j2\pi lf/f_{CLK}}, 0 \leq f \leq f_{CLK}/2. \quad (5.23)$$

The magnitude of \mathbf{H}_{mvavg} is zero (notch) at frequencies $f = qf_{CLK}/n$, where q is an integer and $1 \leq q \leq n/2$. For these frequencies, n samples will cover full cycles of them such that the sum of (5.23) is zero. Therefore, the moving average can perfectly filter out the original DC component and aliased IF harmonics in the ADC samples and the x_{mul+} term induced by the multiplications in (5.22). Note that the DC gain of \mathbf{H}_{mvavg} is 2, which is necessary to calculate the correct magnitude of the original IF signal. The multiplications in (5.22) convert half of the IF magnitude to DC and another half to the double frequency term x_{mul+} . For this statement, one may refer to the mixer Eq. (5.1). The term x_{mul+} is eventually filtered out by the moving-average filter, and the DC term is used as the demodulation output. Therefore, we need to double the magnitude of the DC term (i.e., $|\mathbf{H}_{mvavg}(0)| = 2$) to obtain the correct magnitude of the original signal.

The bandwidth of the non-I/Q demodulation is defined as the lowest frequency f_B , at which the gain of \mathbf{H}_{mvavg} drops by -3 dB. The magnitude of \mathbf{H}_{mvavg} is given by

$$|\mathbf{H}_{mvavg}(f)| = \frac{2}{n} \left| \frac{\sin(\pi n f / f_{CLK})}{\sin(\pi f / f_{CLK})} \right|, 0 \leq f \leq f_{CLK}/2, \quad (5.24)$$

from which the bandwidth f_B can be calculated numerically. Alternatively, we may use an empirical formula to estimate the bandwidth

$$f_B \approx 0.443 f_{CLK} / \sqrt{n^2 - 1}, \quad (5.25)$$

where n should be larger than 1. The phase response of \mathbf{H}_{mvavg} is equivalent to that of a pure delay term of $nT_s/2$, which is given by

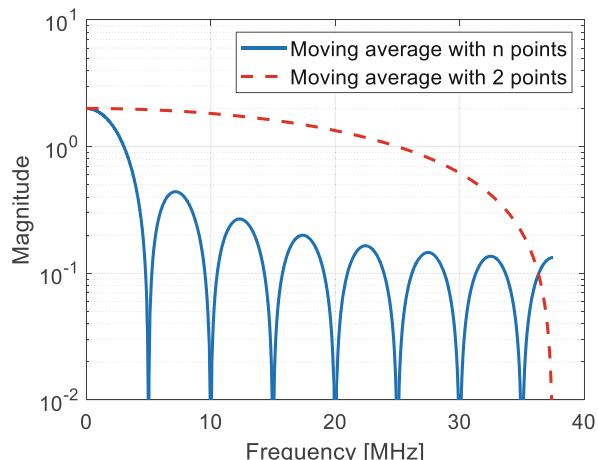
$$\angle \mathbf{H}_{mvavg}(f) = -360^\circ \times fnT_s/2, \quad (5.26)$$

where $T_s = 1/f_{CLK}$.

For example, Fig. 5.9 shows the frequency response of a non-I/Q demodulation realization with the following parameters: $n = 15$, $m = 4$, $f_{IF} = 20$ MHz, and $f_{CLK} = 75$ MHz. The frequency response notches at multiple of 5 MHz. Since the IF harmonics are aliased to multiple of $f_{CLK}/n = 5$ MHz, they can be perfectly filtered out by the moving-average filter. Furthermore, f_B can be estimated from the frequency response plot to be about 2.2 MHz, close to the value calculated by (5.25).

As seen in Fig. 5.9, the bandwidth of the non-I/Q demodulation algorithm (5.22) is small compared to the clock frequency. In some cases, larger RF detection bandwidth is expected, such as when detecting fast RF amplitude or phase changes. Then instead of using the algorithm (5.22), we may adopt the two-sample demodulation algorithm (5.15), achieving the maximum f_B for a given f_{CLK} . The frequency response of the two-sample demodulation algorithm is also shown in Fig. 5.9. We can see that its bandwidth is $n/2$ times larger than that of the algorithm (5.22). However, the demodulation algorithm (5.15) cannot filter out the aliased harmonics, which should be handled separately (e.g., using notch filters) if necessary.

Fig. 5.9 Frequency responses of the non-I/Q demodulation moving-average filters. Both the demodulation algorithms with n points and with 2 points are displayed



5.2.2.5 Non-I/Q Demodulation for Transient RF Measurement

The non-I/Q demodulation algorithm (5.22) assumes that the amplitude and phase of the IF signal within n samples are constant. Therefore, it only applies to the measurement of slow-varying signals. When using it to measure an RF transient, such as the rising or falling edge of an RF pulse, large errors (i.e., amplitude and phase jumps) are observed (Geng and Kalt 2019), as shown by the solid lines in Fig. 5.10.

The jumps shown in Fig. 5.10 are artifacts caused by the algorithm (5.22) when the IF amplitude or phase changes are fast. To understand the situation, we assume an IF signal with time-varying amplitude and phase as $x(t) = A(t)\sin[2\pi f_{IF}t + \varphi(t)]$. Then the non-I/Q demodulation of the IF samples can be written as

$$\begin{aligned} I_k &= \frac{2}{n} \sum_{l=k-n+1}^k A_l \sin(2\pi f_{IF}lT_s + \varphi_l) \cdot \sin(2\pi f_{IF}lT_s), \\ Q_k &= \frac{2}{n} \sum_{l=k-n+1}^k A_l \sin(2\pi f_{IF}lT_s + \varphi_l) \cdot \cos(2\pi f_{IF}lT_s), \end{aligned} \quad (5.27)$$

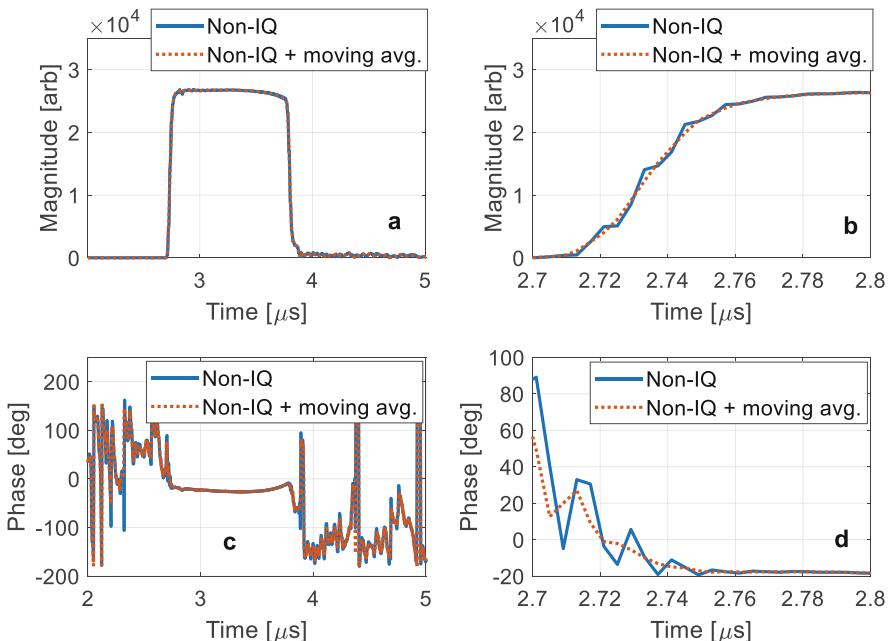


Fig. 5.10 Transient detection with non-IQ demodulation algorithm with an IF frequency $f_{IF} = 41.65$ MHz and clock frequency $f_{CLK} = 6f_{IF} = 249.9$ MHz. (a) Amplitude waveform; (b) Amplitude zoom-in at rising edge; (c) Phase waveform; (d) Phase zoom-in at rising edge

where $A_l = A(lT_s)$ and $\varphi_l = \varphi(lT_s)$. Here we have used the relation $l\Delta\varphi = 2\pi f_{IF}lT_s$. We can further work out (5.27) as

$$\begin{aligned} I_k &= \frac{1}{n} \sum_{l=k-n+1}^k A_l \cos \varphi_l - \frac{1}{n} \sum_{l=k-n+1}^k A_l \cos (4\pi f_{IF}lT_s + \varphi_l), \\ Q_k &= \frac{1}{n} \sum_{l=k-n+1}^k A_l \sin \varphi_l + \frac{1}{n} \sum_{l=k-n+1}^k A_l \sin (4\pi f_{IF}lT_s + \varphi_l). \end{aligned} \quad (5.28)$$

The first terms on the right-hand side of both equations in (5.28) are the I and Q values averaged within n samples, which are the expected outputs of the demodulation algorithm. The second terms are at the frequency of $2f_{IF}$. If the amplitude and phase of the IF signal are constant, the second terms will be zeroed out by averaging n samples. The non-I/Q demodulation algorithm is developed for this situation. However, if the IF signal has a time-varying amplitude or phase, the second terms will remain in the demodulation result. This causes RF detection errors near $2f_{IF}$. In the time domain, these errors appear as jumps in the transients, see Fig. 5.10.

The spectrum of the demodulated RF pulse of Fig. 5.10 is shown in Fig. 5.11a. At frequencies near $2f_{IF}$, the noise level is high, which agrees to (5.28). The analysis above implies that a notch filter at $2f_{IF}$ can filter out such errors. Since the clock is synchronized to the IF signal, a moving-average filter can provide an accurate notch at $2f_{IF}$. As shown in Fig. 5.11b, a moving-average filter with $n/2$ taps (n is even) or n taps (n is odd) can provide a notch at $2f_{IF}$. A 3-tap moving-average filter has been applied to the RF pulse amplitude and phase waveforms in Fig. 5.10. The resulting baseband signal and its spectrum are shown as the dotted-line plots in Figs. 5.10 and 5.11a, respectively.

The moving-average filter is easy to implement in digital processors like DSP or FPGA. Of course, for RF detection, the moving-average filter also generates additional attenuation of the magnitude near DC (see Fig. 5.11b) and causes a group delay. If these issues are significant concerns, a higher-order narrow-band notch filter can be used instead of the moving-average filter.

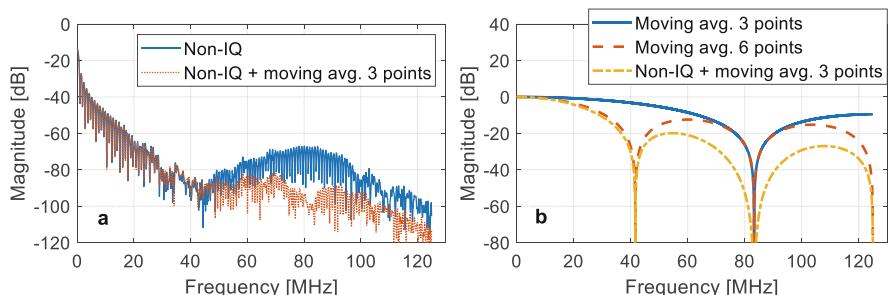
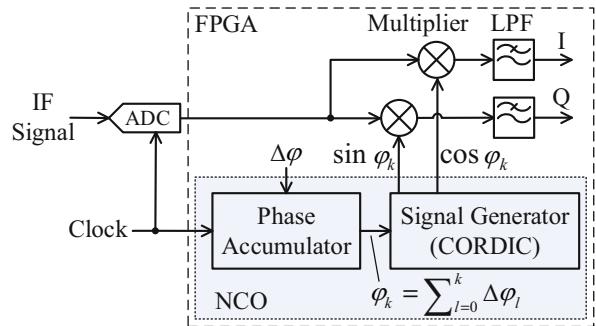


Fig. 5.11 (a) Spectra of the baseband RF pulse; (b) Frequency responses of the 3-tap moving average filter, the 6-tap moving average filter (i.e., the non-I/Q demodulation filter), and the cascaded filter with the non-I/Q demodulation + 3-tap moving average filter

Fig. 5.12 Digital down-conversion for RF detection



5.2.3 Digital Down-Conversion

A more general architecture for demodulating the IF samples is the so-called *digital down-conversion (DDC)* scheme (Schilcher 2008). See Fig. 5.12. The DDC implements the function of an I/Q demodulator (see Fig. 3.2) numerically, performing a second stage down-conversion of the IF samples. Compared to the I/Q and non-I/Q demodulation algorithms, it does not need particular ratios of the clock and IF frequencies. Therefore, we may even not require the LO and clock to be synchronous to the RF signal to be measured. This feature offers great flexibility. However, we may need to measure the RF reference signal and obtain the relative phase between the RF signal to be measured and the RF reference signal. This is called reference tracking, which will be discussed in Sect. 5.2.5.

The DDC is typically implemented in digital processors like DSP or FPGA. It consists of a numerically controlled oscillator (NCO), two multipliers and two low-pass filters. The NCO generates a sine series and a cosine series at the IF frequency, corresponding to the LO signal in Fig. 3.2 split by a quadrature splitter. The multipliers play the role of mixers. The digital low-pass filters pick up the lower sideband of the multiplication products. Note that the non-I/Q demodulation algorithm (5.22) is a particular case of DDC, where the NCO is implemented with two pre-defined sine and cosine series (with a fixed frequency, i.e., fixed $\Delta\varphi$, and we only store n points of the series in circular buffers). Furthermore, the moving-average filters in (5.22) implement the low-pass filters required by the DDC.

A general block diagram of the NCO can be also found in Fig. 5.12. The input is the phase increment step $\Delta\varphi$, which is accumulated at each clock cycle, and the accumulated phase is used to generate the sine and cosine series. Practically, the sine and cosine series generators are implemented as look-up tables or using the coordinate rotation digital computer (CORDIC) algorithm. Note that the NCO output phases are arbitrary after the NCO is restarted or the clock is interrupted for an unknown period of time. The clock interruption happens quite often in practical LLRF systems during maintenance (e.g., power cycling of clock generator). The phase uncertainty is multiple of $\Delta\varphi$. We have discussed this issue when describing the I/Q and non-I/Q demodulation algorithms, where, as we have mentioned, the measured phase depends on the starting point of demodulation.

In addition to the moving-average filter, the low-pass filter in Fig. 5.12 may also be implemented as other types of digital filters, including the finite impulse response (FIR) filter, the infinite impulse response (IIR) filter, and the cascaded integrator-comb (CIC) filter. These filters have been discussed in detail in general textbooks of digital signal processing. Here we will not discuss them in-depth.

5.2.4 Handling of Time-varying Frequency

The RF operating frequency of some RF stations may vary with time. For example, in a proton circular accelerator, the RF operating frequency should follow the increasing revolution frequency of the particles with increasing beam energy. In this case, the RF detectors should follow the frequency changes.

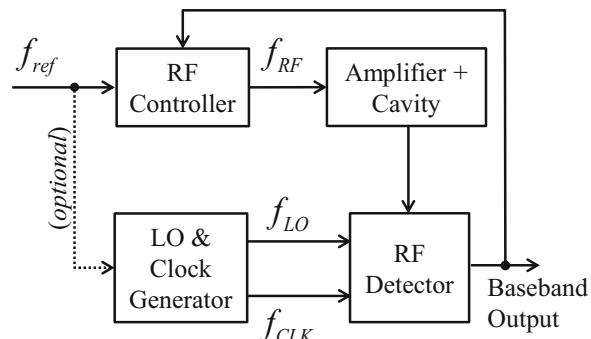
An RF station operating at a time-varying frequency may be controlled as shown in Fig. 5.13. The following combinations of the RF reference frequency f_{ref} and the RF operating frequency f_{RF} may be used:

- Constant f_{ref} and constant f_{RF} , and $f_{RF} = f_{ref}$. We have assumed this scenario as a default.
- Constant f_{ref} but time-varying f_{RF} . In this case, the RF controller generates a time-varying f_{RF} by modulating the phase of the RF reference signal. This situation is often found for RF stations fed with a constant master oscillator frequency but requiring different operating frequencies.
- Time-varying f_{ref} and time-varying f_{RF} , but $f_{RF} = f_{ref}$. It corresponds to changing the master oscillator frequency. This strategy is widely used in the RF stations of storage rings.

The RF detectors are often designed for the case *a*. However, we need to handle cases *b* and *c* properly to extract meaningful information from the RF signals with a time-varying frequency.

Let us first study the situation when the LO and clock frequencies are derived from f_{ref} with constant frequency ratios:

Fig. 5.13 An RF station consisting of LLRF and HPRF with possibly different f_{ref} and f_{RF}



- In the case *b* above, the LO and clock frequencies are constant, and we obtain a time-varying IF frequency $f_{IF} = f_{RF} - f_{LO}$. Suppose we still use the algorithm (5.22), providing a constant $\Delta\varphi$ derived from a predefined IF frequency f_{IF0} . Because f_{IF} varies with time and differs from f_{IF0} , eq. (5.22) yields a time-varying phase in the output. The output phase contains a slope described as a function of time: $\varphi(t) = 2\pi(f_{IF} - f_{IF0})t + \varphi_0$, where φ_0 is the initial phase related to the starting point of demodulation.
- For the case *c*, both f_{LO} and f_{CLK} change with $f_{RF} = f_{ref}$. Because f_{LO} and f_{CLK} are both derived from f_{ref} and reserve proper ratios to f_{ref} , the frequency changes are transparent to the digital demodulation algorithm (5.22). That is, we can still obtain the baseband signal of the IF signal with a time-varying frequency without changing the digital demodulation part of the RF detector. The main problem here is the time-varying f_{CLK} . ADCs and FPGAs do not like run-time changes in the clock frequency because they may cause metastability issues. For example, the latency of the FPGA routing between two registers is typically optimized for a particular clock period. If we increase the clock frequency, the value on the routing path between two registers may not reach a stable 0 or 1 state within a clock cycle, resulting in data transfer errors. Of course, if we optimize the FPGA firmware using the upper bounds of the clock frequency, this problem can be partially mitigated.

In practice, flexible LO and clock frequencies are used to deal with the time-varying RF operating frequency. The LO frequency can track the RF signal frequency changes, resulting in a constant IF frequency (Ayyazyan et al. 2011). A time-varying IF frequency can be sampled by a constant clock frequency and then demodulated by DDC with a time-varying NCO frequency (Li et al. 2015; Tamura et al. 2019). With the reference tracking algorithms discussed in Sect. 5.2.5, we can even use asynchronous LO and clock signals that are not locked to the RF reference signal.

5.2.5 RF Detection with Reference Tracking

When measuring the RF pick-up signals from a HPRF system, we may face the following difficulties:

- a. A time-varying f_{IF} will cause difficulties for RF detectors with a constant clock frequency. It requires varying the NCO frequency in the DDC (or the $\Delta\varphi$ in the non-I/Q demodulation coefficients) to track the f_{IF} changes.
- b. The LO and clock signals may be produced by independent oscillators, and therefore, are not synchronized with the RF signal. This situation is often found in laboratories where a dedicated LO and clock generator is not available.

- c. The phase measurement result is arbitrary in any of the following conditions:
 - i. The starting point of the I/Q or non-I/Q demodulation is changed. A time jitter of the trigger edge could cause this problem if it determines the demodulation starting point.
 - ii. Rebooting the LO and clock generator (see Fig. 5.3b) or interrupting its reference input. The phases of the LO and clock are arbitrary due to the restarting of the frequency dividers.
 - iii. Rebooting the NCO (see Fig. 5.12) or interrupting its clock input.

The RF signals picked from the HPRF system are derived from an RF reference signal. We may measure this reference signal using one RF detector channel. We assume that all the RF detector channels have the same characteristics, including identical LO and clock signals. Then we can use the measurement results of the RF reference signal to deal with the above difficulties. This method is called *reference tracking*. When applying reference tracking to the cavity probe signal, we can detect the phase of the RF field relative to the RF reference and cancel the RF detector channels' common errors. Therefore, in addition to mitigating the phase detection ambiguity in the above situations, the reference tracking can also reduce the drifts common in all RF detector channels (e.g., the drifts in electronics, LO and clock phases). We will introduce several reference tracking strategies in this section.

5.2.5.1 Reference Tracking with PLL

When a DDC-based RF detector is used to measure an IF signal with a time-varying frequency, the NCO output frequency should follow exactly the IF frequency. To track the IF frequency and lock to its phase, we may adopt a digital PLL like Fig. 5.14. Here we used a PI controller. To accelerate the initial locking process, we also implemented a feedforward path. The feedforward controller analyzes the frequency of the IF signal (e.g., via FFT) and determines a feedforward input $\Delta\varphi_0$ close to what is required for the NCO input.

The IF signal of the RF reference signal can be sampled (reference IF samples) and used as the PLL reference input. The phase difference between the IF signal and

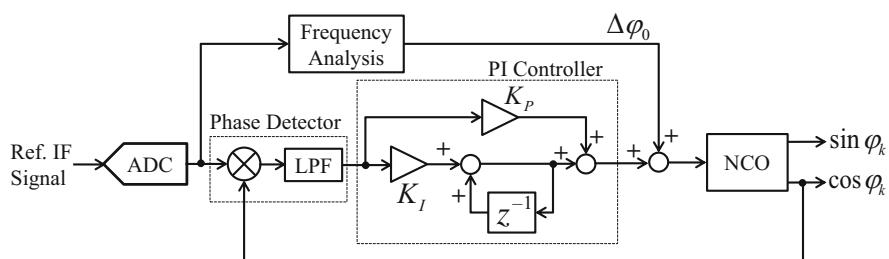
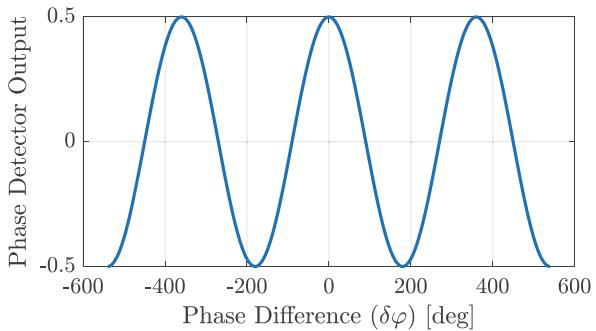


Fig. 5.14 PLL to lock the NCO frequency and phase with the reference IF signal

Fig. 5.15 Response of the phase detector based on a multiplier and a low-pass filter



the NCO output is detected by a multiplier and a low-pass filter, similar to the mixer-based phase detector discussed in Sect. 5.1.1. According to (5.1), the phase detector output is a cosine function $\cos[\delta\varphi(t)]$, where $\delta\varphi(t) = \varphi_{IF}(t) - \varphi_{NCO}(t)$. Here we have $\varphi_{IF}(t) = 2\pi f_{IF}t + \varphi_{IFO}$ and $\varphi_{NCO}(t) = 2\pi f_{NCOr}t + \varphi_{NCO0}$, and they are the instant phases of the IF signal and NCO output signal, respectively. Figure 5.15 is an example of the phase detector response. The gain of the phase detector (from $\delta\varphi(t)$ to the output) is positive in the range of $-180^\circ < \delta\varphi(t) < 0$ (or shifted by multiple of 360°). A positive gain is necessary for a stable feedback loop, therefore, when the PLL is in a steady state, $\delta\varphi(t)$ should be in the above range. For example, after the loop is locked and operates stably, the phase detector output is typically close to zero, corresponding to $\delta\varphi(t) \approx -90^\circ$. However, when initially locking the PLL, the value of $\delta\varphi(t)$ might be outside the positive-gain range, resulting in an oscillatory transient in the NCO output.

Figure 5.16 is an example depicting the transient response of a PLL. In this example, $f_{CLK} = 10$ MHz, $f_{IF} = 1$ MHz, a PI controller is used with $K_P = 0.2$ and $K_I = 0.01$, and we assigned a feedforward input $\Delta\varphi_0 = 30^\circ$. During the transient state, the phase difference between the NCO output and the reference IF input, $\delta\varphi(t)$, varies within the full range of $\pm 180^\circ$. After entering the steady state, $\delta\varphi(t)$ is stabilized at -90° , corresponding to a zero output from the phase detector, just as we have expected.

In most practical LLRF systems, the IF frequency is constant or only changes slowly. Therefore, we typically choose a small PLL closed-loop bandwidth to reduce the noise transfer from the reference to the output. The reference IF samples contain the noise contributions from the RF reference signal and the RF detector. These noise sources are transferred to the PLL output within its closed-loop bandwidth. The phase detector in Fig. 5.14 is also sensitive to the amplitude noise in the reference IF samples; that is, both the amplitude and phase noise in the reference IF samples will cause phase noise in the NCO output. Therefore, to reduce the noise, the PLL is typically configured to only track the slow IF frequency changes.

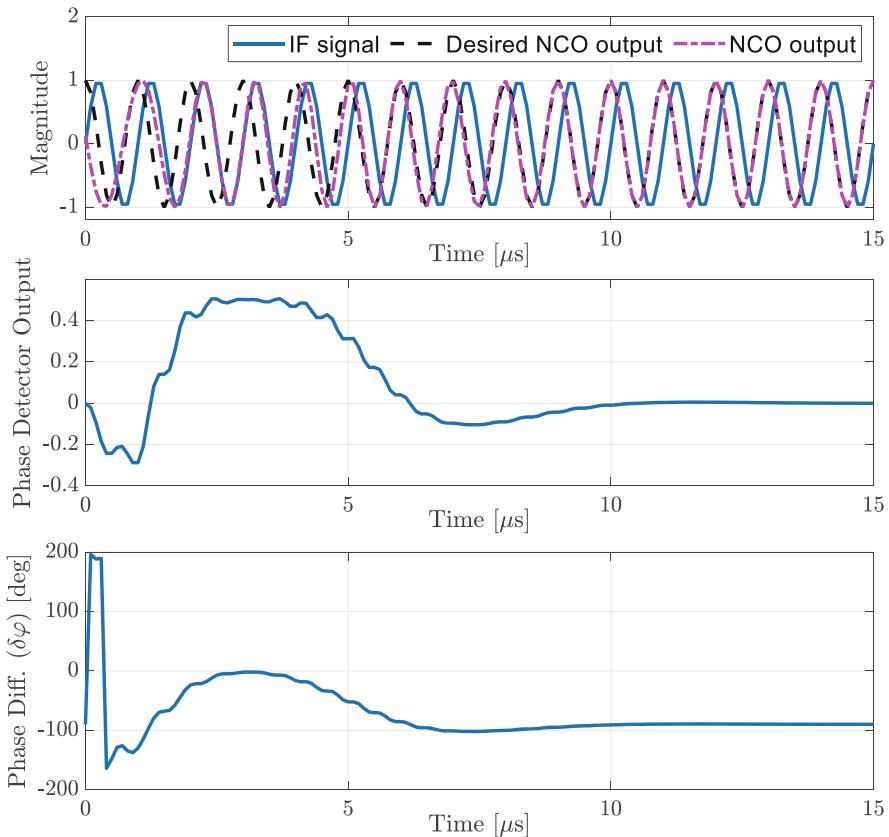


Fig. 5.16 Transient response of a PLL tracking the NCO output to a reference IF signal

5.2.5.2 Reference Tracking with Hilbert Transform

Although the PLL can lock the NCO output to a reference IF signal, we still face the risk of losing the lock due to possible feedback instability. Since our goal is to generate a sine and a cosine series that are frequency- and phase-locked to the reference IF signal, we may consider deriving the desired series from the reference IF samples directly. Let us consider the configuration in Fig. 5.17 (Geng and Kalt 2019).

In the digital signal processing part, the reference IF samples are first filtered by a band-pass filter (BPF) to remove the DC offset and higher-order harmonics. The samples are then processed in two branches with in-phase (0° phase-shifted with a proper delay) and quadrature (90° phase-shifted) phases. The outputs from such a digital $0^\circ/90^\circ$ splitter are what we expect from an NCO. Note that the IF phase interval between two adjacent ADC samples is arbitrary if we use asynchronous LO or clock signals. The 90° phase shifter in the quadrature branch should also work for such arbitrary relations of the IF and clock frequencies; therefore, we adopt the Hilbert transform to generate the expected 90° phase shift. See Fig. 5.18.

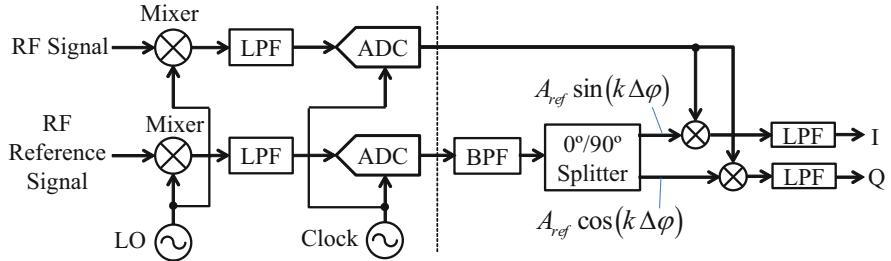


Fig. 5.17 Generate the sine and cosine series from the samples of a reference IF signal

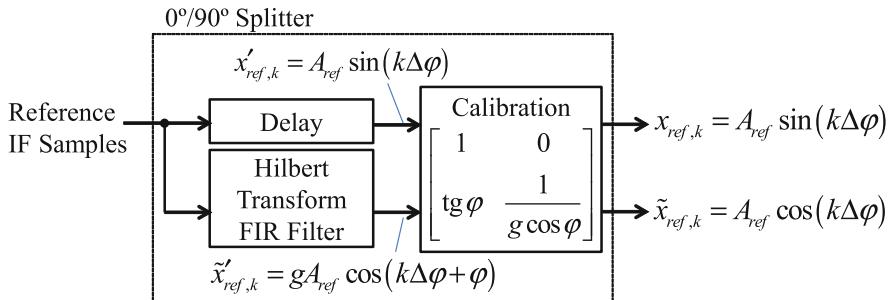


Fig. 5.18 Implementation of the 0°/90° splitter with Hilbert transform

The Hilbert transform of a real time-domain signal $x(t)$ is defined as

$$H[x(t)] := x^H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau. \quad (5.29)$$

The Hilbert transform $x^H(t)$ is also a real time-domain signal that is the convolution of $x(t)$ with a function $1/(it)$. In the frequency domain, the Hilbert transform introduces a phase shift of 90° to every Fourier component of $x(t)$. Full implementation of the Hilbert transform for a time series adopts fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT), which are not suitable for real-time processing. It is more practical to use an FIR filter to implement the Hilbert transform approximately. Figure 5.19 shows the frequency response of a 16-tap Hilbert transform FIR filter designed with the ‘Filter Designer’ of Matlab.

The FIR-based implementation of the Hilbert transform has a perfect phase response, but there is a large ripple in the amplitude response that should be 0 dB at all frequencies. The error in the Hilbert transform implementation at the IF frequency can be modeled as an amplitude imbalance g and phase imbalance φ , as in Fig. 5.18. These imbalances are systematic errors for a specific IF frequency and can be compensated for using the matrix shown in Fig. 5.18, which is written as

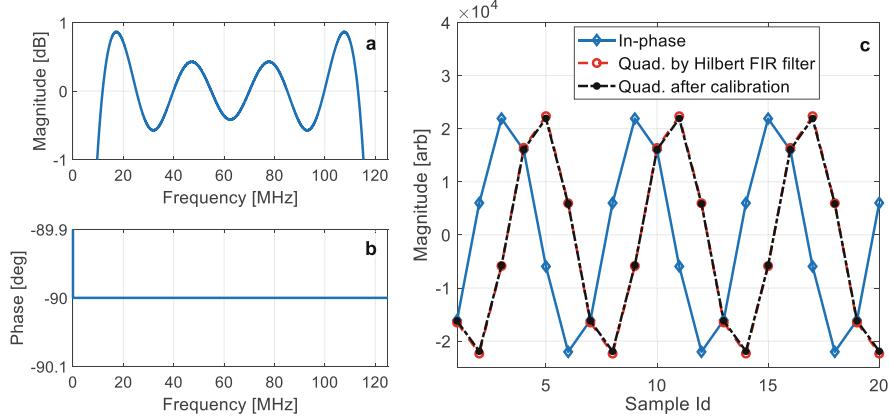


Fig. 5.19 (a, b) Frequency response of a Hilbert transform FIR filter ($f_{CLK} = 249.9$ MHz). The phase lag caused by the group delay has been removed from the phase response; (c) Generation of the in-phase and quadrature signals of the reference IF signal with the FIR-based Hilbert transform and the calibration of imbalances

$$\begin{bmatrix} x_{ref,k} \\ \tilde{x}_{ref,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \operatorname{tg}\varphi & \frac{1}{g \cos\varphi} \end{bmatrix} \begin{bmatrix} x'_{ref,k} \\ \tilde{x}'_{ref,k} \end{bmatrix}, \quad (5.30)$$

where $x_{ref,k}$ and $\tilde{x}_{ref,k}$ are the k th sample of the expected in-phase and quadrature copies of the reference IF signal. The term $x'_{ref,k}$ is the delay stage output and $\tilde{x}'_{ref,k}$ represents the output of the Hilbert transform FIR filter with imbalance errors (see Fig. 5.18).

To determine the values of g and φ for a given Hilbert transform FIR filter, we can sample a stable reference IF signal, from which the series $x'_{ref,k}$ and $\tilde{x}'_{ref,k}$ can be calculated. They form a complex signal $\mathbf{x}'_{ref,k} = x'_{ref,k} + j\tilde{x}'_{ref,k}$. Then we can estimate the phase increment step $\Delta\varphi$ for each sample by calculating the phase slope of $\mathbf{x}'_{ref,k}$. We rewrite the two series in a general form

$$\begin{aligned} x'_{ref,k} &= A_{ref} \sin(k\Delta\varphi + \varphi_{ref}) + A_{ref0}, \\ \tilde{x}'_{ref,k} &= \tilde{A}_{ref} \cos(k\Delta\varphi + \tilde{\varphi}_{ref}) + \tilde{A}_{ref0}, \end{aligned} \quad (5.31)$$

where A_{ref} , φ_{ref} , A_{ref0} , \tilde{A}_{ref} , $\tilde{\varphi}_{ref}$, and \tilde{A}_{ref0} are unknown constant parameters that can be estimated with multi-variable linear fitting. Let us use the in-phase signal $x'_{ref,k}$ as an example to explain the fitting algorithm. First, we rewrite $x'_{ref,k}$ as

$$x'_{ref,k} = a \sin(k\Delta\varphi) + b \cos(k\Delta\varphi) + c, \quad k = 0, 1, \dots, N-1, \quad (5.32)$$

where

$$a = A_{ref} \cos \varphi_{ref}, b = A_{ref} \sin \varphi_{ref}, c = A_{ref} 0. \quad (5.33)$$

Here we have used N samples ($N > 3$) resulting in N equations about the unknown parameters a , b , and c . This is a typical least-square problem and can be solved with Matlab. Then the amplitude and phase of the in-phase signal can be calculated as

$$A_{ref} = \sqrt{a^2 + b^2}, \varphi_{ref} = \operatorname{tg}^{-1}(b/a). \quad (5.34)$$

The same algorithm can be applied to the quadrature signal $\tilde{x}'_{ref,k}$ to estimate its amplitude and phase. Then the imbalances in (5.30) can be calculated as

$$g = \tilde{A}_{ref}/A_{ref}, \varphi = \tilde{\varphi}_{ref} - \varphi_{ref}. \quad (5.35)$$

The algorithm above was applied to a reference IF signal and the results are depicted in Fig. 5.19c. The estimated amplitude and phase imbalances are $g = 1.02$ (or 0.17 dB) and $\varphi = -1.52\text{e-}4^\circ$. They agree well with the frequency responses at the IF frequency (41.65 MHz).

Similar to the discussion in Sect. 5.2.5.1, the BPF in Fig. 5.17 is often chosen to be a narrow-band filter. This can avoid transferring too much noise in the measured reference IF signal to the digital 0°/90° splitter output. With the configuration of Fig. 5.17, the reference IF samples are directly involved in the measurements of other RF signals. Both the amplitude and phase noise in the reference IF samples will contribute to the RF measurement errors.

The noisy IF samples of an RF reference signal and an RF signal to be measured are given by

$$\begin{aligned} x_{ref,k} &= A_{ref} (1 + \alpha_{ref,k}) \sin(2\pi f_{IF} k T_s + \Delta\varphi_{ref,k}), \\ x_{rf,k} &= A_{rf} (1 + \alpha_{rf,k}) \sin(2\pi f_{IF} k T_s + \varphi_0 + \Delta\varphi_{rf,k}), \end{aligned} \quad (5.36)$$

where A_{ref} , α_{ref} and $\Delta\varphi_{ref}$ are the amplitude, relative amplitude noise and phase noise of the RF reference signal; and A_{rf} , α_{rf} , φ_0 and $\Delta\varphi_{rf}$ are the amplitude, relative amplitude noise, initial phase and phase noise of the RF signal to be measured. Assume that the digital 0°/90° splitter can generate a perfect quadrature signal of x_{ref} . Then after demodulation, we obtain a baseband signal of the RF signal to be measured in terms of I and Q :

$$\begin{aligned} I_k &= \frac{A_{ref} A_{rf}}{2} (1 + \alpha_{ref,k} + \alpha_{rf,k}) \cos(\varphi_0 + \Delta\varphi_{rf,k} - \Delta\varphi_{ref,k}), \\ Q_k &= \frac{A_{ref} A_{rf}}{2} (1 + \alpha_{ref,k} + \alpha_{rf,k}) \sin(\varphi_0 + \Delta\varphi_{rf,k} - \Delta\varphi_{ref,k}). \end{aligned} \quad (5.37)$$

Here we have assumed α_{ref} and $\Delta\varphi_{ref}$ are constant for a few clock cycles, that is, we only consider the noise in a bandwidth much narrower than the clock frequency.

The baseband output contains amplitude and phase noise from both the RF reference signal and the RF signal to be measured. The noise may come from the common LO and clock signals, resulting in correlated noise; or from the RF detector electronics like mixers and ADCs, resulting in uncorrelated noise. The noise in (5.37) are summarized as follows:

- (1) *The amplitude noise adds up.* Low-frequency variations of α_{ref} and α_{rf} are often correlated and change in the same direction. Therefore, their magnitudes sum up, and the total amplitude noise increases. At high frequencies, they are typically uncorrelated, and their powers sum up.
- (2) *The phase noise subtracts.* Typically, the low-frequency variations of $\Delta\varphi_{ref}$ and $\Delta\varphi_{rf}$ are caused by the common LO and clock drifts, and they cancel each other via the subtraction. At high frequencies, $\Delta\varphi_{ref}$ and $\Delta\varphi_{rf}$ are uncorrelated, and the subtraction results in a sum in the noise power.

Equation (5.37) implies that the amplitude noise always gets worse in the structure of Fig. 5.17. The phase noise cancels each other at lower frequencies but also gets worse at higher frequencies. A narrow-band BPF filter helps to reduce the high-frequency noise. To reduce the low-frequency amplitude noise, we may calculate the amplitude of the digital $0^\circ/90^\circ$ splitter output, and then correct the amplitude measurement of the RF signal.

As discussed, we tend to limit the reference IF signal bandwidth in both the PLL-based (Fig. 5.14) and the filter-based (Fig. 5.17) reference tracking schemes. This limits the tracking performance in case of a fast phase jump in the reference IF signal. We will address this issue using a direct reference phase tracking approach.

5.2.5.3 Direct Reference Phase Tracking

If we stay with the I/Q or non-I/Q demodulation scheme, we can still implement the *direct reference phase tracking* as in Fig. 5.20. We subtract the phase of the RF reference signal directly from the phase of the RF signal to be measured. The amplitude measurement is not affected.

The phase noise analysis in (5.37) is still valid for the direct reference phase tracking. We benefit from the suppression of the common-mode phase drifts in RF detectors but suffer from the sum of the uncorrelated noise. Typically, we low-pass filter the reference phase before subtracting it from the phases of other RF signals.

Moreover, direct reference phase tracking offers a possibility to avoid unwanted phase measurement errors (jumps). The errors may happen when the demodulation is started by a trigger in an RF system operating in the pulsed mode. If the rising edges of the trigger and ADC clock occur almost simultaneously, a small trigger timing jitter may select a different sample as the demodulation starting point. This is denoted as a *race condition*. As we have mentioned, the changes in I/Q or non-I/Q demodulation starting point will result in phase measurement errors of multiple $\Delta\varphi$ (the IF phase change within a clock cycle). They will cause RF field phase errors if the measurement result is used in a feedback loop. When the race condition occurs,

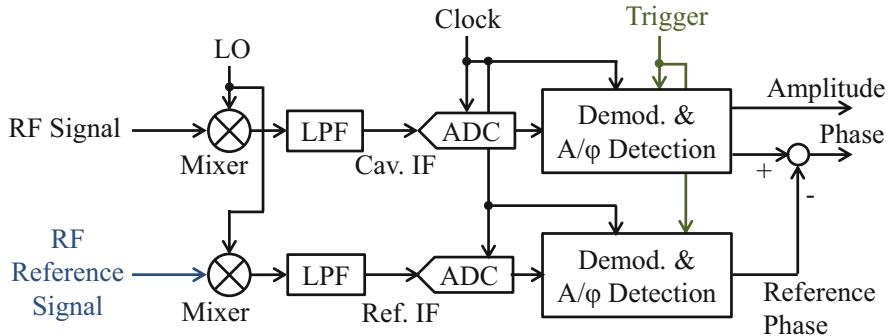


Fig. 5.20 Direct reference phase tracking for I/Q or non-I/Q demodulation

the same phase measurement error appears in all RF signals, including the RF reference signal. Since we know that the RF reference signal has a stable phase, the phase detection error can be identified and immediately removed from the measurements of other RF signals. This feature needs to be implemented in the LLRF real-time firmware or software to remove fast phase jump errors induced by the race condition.

5.3 RF Actuation Schemes

RF actuation is an inverted process of RF detection. The input to an RF actuator is a baseband signal, namely, the desired amplitude and phase (or I/Q components). The output is an RF frequency signal. From the viewpoint of an RF controller, the RF actuator controls the amplitude, phase, or frequency of an RF reference signal.

Similar to RF detectors, the performance of an RF actuator is also characterized by its *bandwidth*, *accuracy*, *linearity*, *signal-to-noise ratio (SNR)*, and *latency*.

A larger bandwidth is required if fast amplitude or phase changes are expected in the RF actuator output. The formula between the required baseband bandwidth (f_B) and the desired RF envelope rise time (t_r) is still valid, i.e., $f_B > 0.35/t_r$.

When an RF actuator is used in a feedback loop, its steady-state error can be corrected by the feedback; that is, its accuracy is often not a critical concern. However, since the RF actuator is in the RF driving chain, we expect it to contribute less noise (especially high-frequency noise since slow drifts are controllable by feedback) and less latency. We will discuss the RF actuator noise and its effects on RF field stability in Chap. 6. The nonlinearity of the RF driving chain, including the RF actuator, amplifiers, and klystron, will be covered in Chap. 7.

Here we will mainly focus on the principles and algorithms of RF actuation.

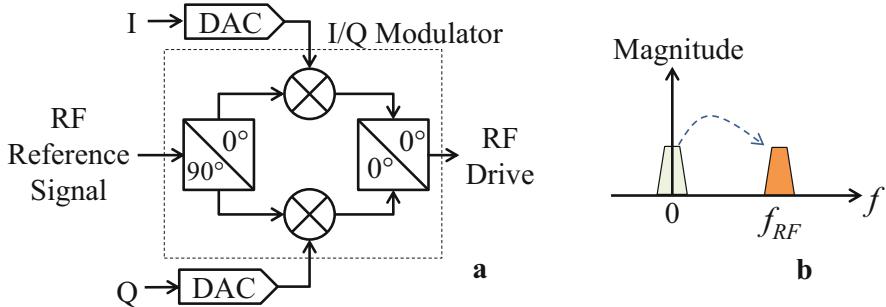


Fig. 5.21 Direct up-conversion. (a) architecture; (b) frequency-domain conversion

5.3.1 Direct Up-Conversion

Direct up-conversion is a conjugate of the direct down-conversion described in Sect. 5.1.2. It is formed by an I/Q modulator and two DACs, see Fig. 5.21a.

For direct up-conversion, the RF reference signal input to the I/Q modulator has the same frequency as the output RF signal. We denote the reference signal as

$$x_{ref}(t) = A_{ref} \sin(2\pi f_{RF} t), \quad (5.38)$$

where A_{ref} and f_{RF} are constant. Here we have normalized the phase of x_{ref} to zero. Then the I/Q modulator output is

$$\begin{aligned} x_{RF}(t) &= I \frac{A_{ref}}{\sqrt{2}} \sin(2\pi f_{RF} t) + Q \frac{A_{ref}}{\sqrt{2}} \cos(2\pi f_{RF} t) \\ &= A_{RF} \sin(2\pi f_{RF} t + \varphi_{RF}), \end{aligned} \quad (5.39)$$

where the amplitude and phase of the output signal are given by

$$A_{RF} = A_{ref} \sqrt{I^2 + Q^2} / \sqrt{2}, \quad \varphi_{RF} = \tan^{-1}(Q/I). \quad (5.40)$$

This implies that the baseband I/Q inputs modulate the RF reference signal, changing its amplitude and phase. In the frequency-domain, (5.39) corresponds to shifting the baseband signal directly to the RF reference frequency, see Fig. 5.21b.

With direct up-conversion, the RF output phase remains constant when the clock is interrupted for an arbitrary period or when the race condition occurs between the trigger and the clock. Since the DACs generate only the baseband signals, the phase noise of the RF output signal is not sensitive to the DAC clock jitter. Compared to other RF actuation schemes, the direct up-conversion is simple to implement. In the digital part, we directly produce the base-band signals via DACs and do not need to implement digital modulations in FPGAs or DSPs that often require an NCO. A simpler implementation typically implies better reliability in operation.

However, the DAC output baseband signals may be sensitive to environmental noise falling in the same frequency range (DC to several MHz). For example, the electromagnetic interference (EMI) caused by the high power RF components (e.g., klystron modulator) or the FPGA board's power supply may couple to some elements in the signal path and end up finally at the DAC outputs.

Another issue for the direct up-conversion is the leakage of RF reference signal to the output and the amplitude and phase imbalances in the I/Q modulator. The imbalances come from the I/Q modulator imperfections, such as the amplitude or phase errors in the power splitters, the different conversion losses and phase shifts in the two mixers, and the different attenuations and phase shifts in the signal paths. The leakage and imbalances will generate static RF actuation errors, reducing the accuracy of amplitude and phase adjustments. As we have mentioned, such static errors are not relevant if the RF actuator is used in a feedback loop.

5.3.2 Single Sideband Up-Conversion

To mitigate the influence of environmental noise in the direct up-conversion, we may consider generating IF signals via the DACs, leading to the *single sideband (SSB) up-conversion* in Fig. 5.22 and the *IF up-conversion* in Fig. 5.24.

SSB up-conversion performs two stages of up-conversion, one in the digital part and the other via an analog I/Q modulator. The baseband I/Q signals, often the outputs of a digital RF controller, are modulated to the IF frequency digitally. An NCO is typically used to generate the reference IF series $\sin(2\pi f_{IF}kT_s)$ and $\cos(2\pi f_{IF}kT_s)$. The modulated IF signal and its quadrature signal are generated with two DACs and then filtered by BPFs. The BPFs reduce the environmental noise and remove the higher-order images of the DAC outputs in higher Nyquist zones. The analog I/Q modulator has an LO input at the frequency $f_{LO} = f_{RF} - f_{IF}$. The DAC outputs (after BPFs) can be written as

$$\begin{aligned} x_I(t) &= -Q \sin(2\pi f_{IF}t) + I \cos(2\pi f_{IF}t) = A_{IF} \cos(2\pi f_{IF}t + \varphi_{IF}) \\ x_Q(t) &= I \sin(2\pi f_{IF}t) + Q \cos(2\pi f_{IF}t) = A_{IF} \sin(2\pi f_{IF}t + \varphi_{IF}) \end{aligned} \quad (5.41)$$

where $A_{IF} = \sqrt{I^2 + Q^2}$, $\varphi_{IF} = \tan^{-1}(Q/I)$. With the LO signal described by $x_{LO}(t) = A_{LO} \sin(2\pi f_{LO}t)$, we obtain the output of the I/Q modulator as

$$\begin{aligned} x_{RF}(t) &= \frac{A_{LO}}{\sqrt{2}} [x_I(t) \sin(2\pi f_{LO}t) + x_Q(t) \cos(2\pi f_{LO}t)] \\ &= \frac{A_{LO}A_{IF}}{\sqrt{2}} \sin(2\pi f_{RF}t + \varphi_{IF}) \end{aligned} \quad (5.42)$$

The result (5.42) implies that we obtain the desired upper-sideband at $f_{LO} + f_{IF}$ of the mixing product and the lower-sideband at $f_{LO} - f_{IF}$ is suppressed. The frequency-

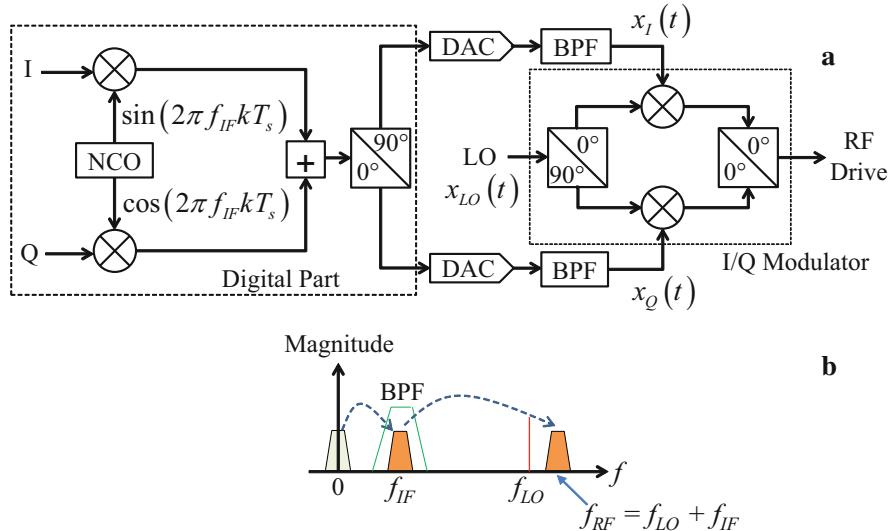


Fig. 5.22 Single sideband up-conversion. (a) architecture; (b) frequency-domain conversion

domain interpretation of the SSB up-conversion is depicted in Fig. 5.22b. Note that we can also select the lower-sideband if we exchange the I and Q inputs of the I/Q modulator.

In the I/Q modulator, the LO signal leakage to the output and the amplitude and phase imbalances may degrade the SSB up-conversion performance. In the frequency-domain, the LO leakage introduces a LO-frequency component in the output, and the imbalances result in a poor suppression of the unwanted sideband. These errors can be calibrated by pre-distorting the baseband I/Q signals in the digital part. One who is interested in the calibration can refer to the article (Geng and Hong 2016).

Here we use DACs to generate IF signals instead of baseband signals, which imposes more strict requirements on the DAC clock jitter. The relationship between the IF phase jitter $\Delta\varphi_{IF}$ and the DAC clock jitter Δt_{CLK} is given by $\Delta\varphi_{IF} = 2\pi f_{IF} \Delta t_{CLK}$. For example, if $f_{IF} = 40$ MHz, then we need $\Delta t_{CLK} < 0.7$ ps RMS if we want $\Delta\varphi_{IF} < 0.01^\circ$ RMS. The higher the IF frequency, the less DAC clock jitter is required.

Since an NCO is used in the digital part, its phase ambiguity caused by clock interruption may also be a problem for the phase continuity of the SSB up-conversion; that is, after the clock is interrupted for an arbitrary period, the IF phases are uncertain. This has been a severe problem for some practical LLRF systems, which require many efforts to restore the correct phase after rebooting the LLRF electronics. To mitigate this problem, we can also apply reference tracking as in the RF detection, for which we need to measure an RF reference signal. We can apply the PLL introduced in Fig. 5.14 directly to the NCO in the digital part of

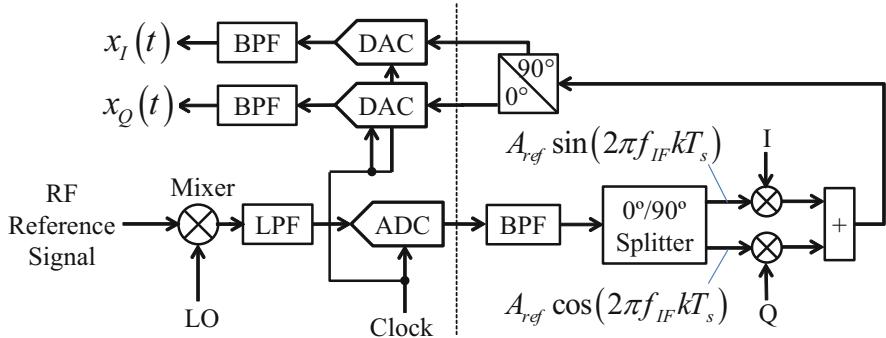


Fig. 5.23 IF generation with reference tracking

Fig. 5.22. Alternatively, a similar structure of Fig. 5.17 can be used, as shown in Fig. 5.23.

With the strategy in Fig. 5.23, we introduce more noise than the nominal structure of Fig. 5.22. The additional amplitude noise comes from the amplitude jitter of the RF reference signal and the added amplitude noise by the RF detector (mixer and ADC). As for additional phase noise, the phase jitter of the LO and clock signal will be canceled out because they appear in both the down-conversion process (i.e., RF reference signal measurement) and the up-conversion process (i.e., actuation via DACs and I/Q modulator). Here we have assumed the LO and clock are synchronous to the RF reference signal. Therefore, the additional phase noise comes mainly from the uncorrelated phase jitter added by the RF detector.

Interesting is that the structure of Fig. 5.23 can relax the requirements of the added phase noise by the LO and clock generator. Assume the RF reference signal is stable but the LO or clock signal has large phase fluctuations, the reference tracking in Fig. 5.23 may yield lower phase noise than the configuration of Fig. 5.22. We will discuss in more detail the noise in Chap. 6. To avoid much influence from the additional noise, we also require the PLL in Fig. 5.14 and the BPF in the digital part of Fig. 5.23 to be narrowband.

5.3.3 IF Up-Conversion

IF up-conversion is an alternative to SSB up-conversion by replacing the I/Q modulator with a mixer, see Fig. 5.24a. The digital part remains the same, and we use a single DAC to generate the IF signal. The IF signal is mixed with the LO signal and a second BPF is used to pick up the desired sideband from the mixing product. The frequency conversion relation is depicted in Fig. 5.24b.

Compared to the SSB up-conversion, the IF up-conversion is simpler in architecture. The BPF2 removes the LO leakage and the image sideband frequency. The major problem is that the design of BPF2 becomes difficult if $f_{LO} > > f_{IF}$, which

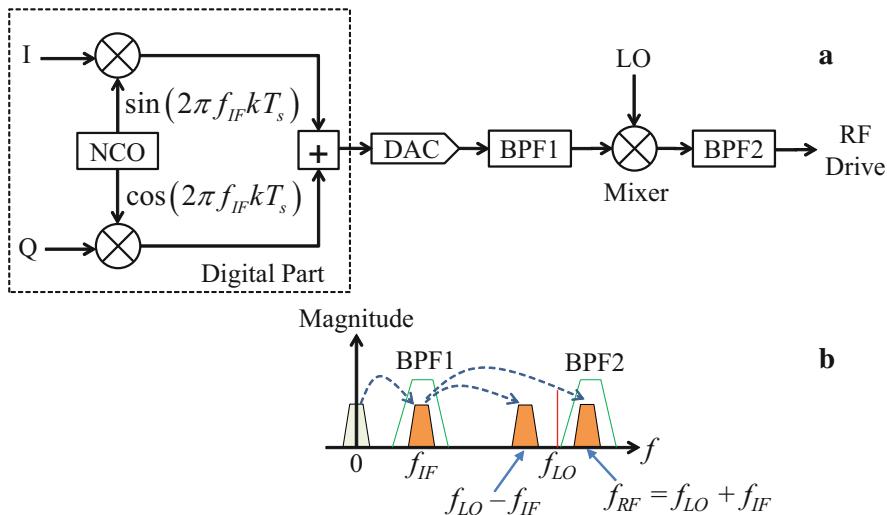


Fig. 5.24 IF up-conversion. (a) architecture; (b) frequency-domain conversion

requires a high-quality-factor BPF (e.g., cavity filter). Such filters might be costly and sensitive to temperature changes.

5.4 Summary

Some widely used strategies and algorithms for RF signal detection and actuation are discussed in this chapter. We also studied several advanced topics on RF signal detection, such as RF transient measurement, time-varying RF frequency, and reference tracking. These algorithms guarantee an accurate RF field measurement that is critical for stable RF field controls. The RF actuation is relatively simpler, and we only introduced the basic principles and algorithms. The noise in the RF signal detection and actuation will be covered in the next chapter.

References

- Ayvazyan V, Koprek W, Kostin D et al (2011) Design and implementation of automatic cavity resonance frequency measurement and tuning procedure for FLASH and European XFEL cryogenic modules. In: Proceedings of the 2nd international particle accelerator conference, San Sebastian, Spain, 4–9 September 2011
- Buzuayene M (2008) Rise time vs. bandwidth and applications. <https://interferencetechnology.com/rise-time-vs-bandwidth-and-applications/>. Accessed 26 May 2021
- Doolittle L et al (2006) Digital low-level RF control using non-IQ sampling. In: Proceedings of the 23rd linear accelerator conference, Knoxville, TN, USA, 20–25 August 2006

- Geng Z, Hong B (2016) Design and calibration of an RF actuator for low-level RF systems. *IEEE Trans Nucl Sci* 63(1):281–287. <https://doi.org/10.1109/TNS.2015.2507204>
- Geng Z, Kalt R (2019) Advanced topics on RF amplitude and phase detection for low-level RF systems. *Nucl Sci Tech* 30:146. <https://doi.org/10.1007/s41365-019-0670-7>
- Geng Z, Simrock S (2008) Evaluation of fast ADCs for direct sampling RF field detection for the European XFEL and ILC. In: Proceedings of the 24th linear accelerator conference, Victoria, BC, Canada, 29 September–3 October 2008
- Geng Z, Gu P, Hou M et al (2005) Design and calibration of a phase and amplitude detector. In: Proceedings of the 2005 particle accelerator conference, Knoxville, TN, USA, 16–20 May 2005
- Habib SB (2013) Performance evaluation of 8-channel card for direct sampling of 1.3 GHz signals. *IEEE Trans Nucl Sci* 60(5):3597–3602. <https://doi.org/10.1109/TNS.2013.2273718>
- Hoffmann M (2008) Development of a multichannel RF field detector for the low-level RF control of the free-electron laser at Hamburg. Ph.D. thesis, Technische Universität Hamburg
- Li X, Sun H, Long W et al (2015) Design and performance of the LLRF system for CSNS/RCS. *Chin Phys C* 39(2):027002. <https://doi.org/10.1088/1674-1137/39/2/027002>
- Schilcher T (2008) RF applications in digital signal processing. CERN Accelerator School: Digital Signal Processing. <https://doi.org/10.5170/CERN-2008-003.249>
- Tamura F, Sugiyama Y, Yoshii M et al (2019) Development of next-generation LLRF control system for J-PARC rapid cycling synchrotron. *IEEE Trans Nucl Sci* 66(7):1242–1248. <https://doi.org/10.1109/TNS.2019.2899358>
- Ziomek C, Corredoura P (1995) Digital I/Q demodulator. In: Proceedings of the 16th particle accelerator conference, Dallas, TX, USA, 1–5 May 1995

Chapter 6

Noise in RF Systems



The noise in an RF system limits the achievable RF field stability. The HPRF noise affects the cavity fields and must be suppressed by the RF controller. The RF detector noise determines the RF field measurement precision and resolution and affects the RF field stability via feedbacks. A systematic analysis of the noise, including where it comes from and how it affects the RF field, is required to specify the RF system performance requirements. In this chapter, we first introduce some basic concepts and methodologies for noise analysis. The noise models of some basic RF components (e.g., amplifier, mixer, ADC and DAC) and the noise transfer relationships in RF control loops will be discussed. We will establish a top-to-down procedure to specify the noise budgets for the RF stations in an accelerator and for each component of an RF station. Then we discuss in detail the noise model of a single RF station. Finally, we will use the measurements at an RF station of SwissFEL as an example to validate the RF station noise model.

6.1 General Description of Noise

6.1.1 Basic Concepts of Noise

Noise is any unwanted electrical disturbance existing in the RF signal processing and control devices. The noise is either generated by internal electronics or coupled in from the external environment. Such unwanted disturbances in an RF system are harmful to the system's performance. The noise in an RF driving component will generate adverse amplitude and phase jitter in the RF field, resulting in a beam jitter. The noise in RF detectors will limit the RF signal measurement resolution, which also affects the RF field stability via an RF feedback loop.

Noise is random and should be analyzed using statistical methods (Howard 2002). In this book, we assume that the noise always has the same statistical properties and is a *stationary stochastic process*. Therefore, we can calculate its

statistical properties using the time-domain measurement of the noise. In this chapter, we will describe the noise as a real-valued time-domain quantity defined at $t \geq 0$. The noise is a *power signal* with a finite power but infinite energy (an *energy signal*, e.g., a signal consisting of one pulse, has finite energy).

Here we introduce several statistical properties describing the noise. Mathematically, noise is represented as a function of time, such as $x(t)$ for the noise x . In practice, we deal with the sampled series of the noise $x_d(n) = x(nT_s)$, where T_s is the sampling time. The statistical properties will be presented by both x and x_d . In this chapter, we use x , y , and z to denote arbitrary noise. In the literature about phase noise, $x(t)$ often denotes the *phase time fluctuation*, and $y(t)$ the *fractional frequency fluctuation* (Rubiola 2021). These two concepts are not used in this book. Readers should distinguish the meanings of x and y when reading other materials.

Mean Value The mean value of the noise x is defined as

$$\bar{x} := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x_d(n). \quad (6.1)$$

In this book, we assume that the mean value of the noise is zero. The definition above requires calculating the limit when N approaches infinity, but in practice, we use a finite number of x_d to calculate the mean value. If N is large enough, the calculated mean value will be close to the real value.

Root-Mean-Square (RMS) Value The RMS value of the noise x is defined as

$$x_{rms} := \lim_{T \rightarrow \infty} \sqrt{\frac{1}{T} \int_0^T [x(t)]^2 dt} = \lim_{N \rightarrow \infty} \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} [x_d(n)]^2}. \quad (6.2)$$

Like the calculation of the mean value, we use a finite number of x_d to calculate x_{rms} , which approaches the real value when increasing N . Electrical noise can be presented as a voltage v or current i , then the average noise power applied to a resistance R can be calculated as

$$P_{avg} = \frac{v_{rms}^2}{R} = i_{rms}^2 R \quad (6.3)$$

For example, a voltage noise has an RMS value of 10 mV, and it is applied to a 50Ω load. We assume the load matches the noise source; that is, all the noise power, which is estimated to be -27 dBm, is consumed by the load.

Autocorrelation Function Autocorrelation is the correlation of a noise with the delayed copy of itself as a function of delay. It describes the similarity between the noise values at different time instants, which helps to find the repeating patterns in the noise. The autocorrelation of a noise x can be defined as

$$\begin{aligned} R_x(\tau) &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t + \tau)dt, \text{ or} \\ R_x(k) &:= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+k). \end{aligned} \quad (6.4)$$

The first definition is based on the continuous representation $x(t)$, and the second is based on the discrete series $x_d(n)$. Here τ is the delay in time, and k is the delay in the number of samples. Note that both $R_x(\tau)$ and $R_x(k)$ are real-valued even functions, i.e., $R_x(\tau) = R_x(-\tau)$ and $R_x(k) = R_x(-k)$. At $\tau = 0$ (i.e., $k = 0$), R_x reaches its maximum value $R_x(0) = x_{rms}^2$, which is proportional to the power of the noise.

Power Spectral Density (PSD) A noise is a random process in the time domain. In the frequency domain, its frequency, magnitude, and phase are also random variables. We use the concept of PSD to describe the average power of the noise within a bandwidth of 1 Hz at various frequencies. The PSD of the noise $x(t)$ can be defined as the average of the square of the Fourier transform magnitude over a large time interval, which is given by

$$S'_x(f) := \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_0^T x(t)e^{-j2\pi ft} dt \right|^2 \quad (6.5)$$

Here we used the continuous form of the noise and the continuous Fourier transform to define the PSD. The calculation of PSD using the discrete Fourier transform (DFT) will be discussed in Sect. 6.1.2. The Wiener-Khinchin theorem relates the autocorrelation and the PSD via Fourier transforms:

$$S'_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau, \quad R_x(\tau) = \int_{-\infty}^{\infty} S'_x(f) e^{j2\pi f\tau} df \quad (6.6)$$

Note that the frequency in (6.5) and (6.6) is in the range of $-\infty \leq f \leq \infty$, so $S'_x(f)$ is a *two-sided PSD*. Since $S'_x(f)$ is a real-valued even function of f , we typically use the *one-sided PSD* $S_x(f)$ in practice:

$$S_x(f) := \begin{cases} 2S'_x(f) & f \geq 0 \\ 0 & f < 0 \end{cases} \quad (6.7)$$

Later, if not particularly mentioned, we will assume the PSDs are one-sided and only consider the non-negative frequencies. The overall noise power within a frequency range f_1 to f_2 (e.g., passband of an RF amplifier) can be calculated as

$$P_{x,f_1 \rightarrow f_2} = \int_{f_1}^{f_2} S_x(f) df \quad (6.8)$$

For example, the noise x has a constant PSD $S_x = -174$ dBm/Hz at all frequencies. We use a band-pass filter with a passband from 100 to 200 MHz to filter the noise. Then at the output, we obtain a noise power of $P_x = -174 + 10\log_{10}(200e6 - 100e6) = -94$ dBm. In LLRF systems, such calculations are frequently performed for RF detectors. Later discussion shows that -174 dBm/Hz is the thermal noise PSD at the room temperature (290 K). The calculation above implies that for an RF detector with a bandwidth of 100 MHz, the thermal noise contribution in the RF measurement is -94 dBm. The noise power increases with a larger bandwidth. This gives a hint when designing RF detectors, whose bandwidth should be limited (of course, they should also satisfy other RF measurement requirements) to avoid accepting too much noise power.

Signal-to-Noise Ratio (SNR) SNR is defined as the ratio between the signal power and noise power:

$$SNR := \frac{P_s}{P_n} = \left(\frac{A_{s,rms}}{A_{n,rms}} \right)^2 \quad (6.9)$$

where P_s is the power of the useful signal, such as the cavity probe RF signal, and P_n is the noise power. Here $A_{s,rms}$ and $A_{n,rms}$ are the RMS values of the signal and noise voltages or currents measured across the same impedance. As described above, the signal and noise powers or RMS values are defined for a specified bandwidth. When defining SNR, the bandwidth of the system should be specified. Another concept that is useful for noise analysis is the *noise propagation through linear systems*. Suppose x is a noise input to a linear system with a transfer function $G(s)$, and y is the output, then the PSD of y can be derived following

$$S_y(f) = S_x(f)|G(j2\pi f)|^2. \quad (6.10)$$

6.1.2 Estimation of PSD and SNR

6.1.2.1 Discrete Fourier Transform

In digital systems, a noisy signal to be measured is sampled by an analog-to-digital converter (ADC). The *discrete Fourier transform (DFT)* can be applied to estimate the PSD and SNR of the measurement.

Suppose that we sample a noisy signal $y(t)$ at a sampling rate f_s and use $y(n)$ ($n = 0, \dots, N-1$) to denote the series of the samples at $t = nT_s$, where $T_s = 1/f_s$ is the sampling time. The DFT of $y(n)$ calculated from N samples is given by

$$Y(f_k) = \sum_{n=0}^{N-1} y(n)e^{-j2\pi f_k n/f_s}, \text{ where } f_k = kf_s/N, \quad (6.11)$$

$$k = 0, \dots, N-1.$$

The DFT yields a spectrum of y at N discrete frequencies f_k with $0 \leq f_k < f_s$. We call each value of $Y(f_k)$ a *DFT bin*. The difference between f_{k+1} and f_k is the so-called *DFT bin spacing* denoted as $\Delta f_{bin} := f_s/N$. For a real-valued signal y , we have $Y(f_k) = \text{conj}\{Y(f_{N-k})\}$ for $k = 1, \dots, N-1$, where $\text{conj}\{\}$ is an operator to calculate the complex conjugate. That is, the magnitude $|Y(f_k)|$ is symmetric to the Nyquist frequency $f_{Nyquist} := f_s/2$, excluding the one at $k = 0$. We define the frequency range $0 \leq f \leq f_{Nyquist}$ as the *Nyquist zone*. Note that if N is an odd number, the bin at $f_{Nyquist}$ does not exist. To simplify the discussion, we select N to be an even number when calculating DFT.

In digital signal processing, the *fast Fourier transform (FFT)* algorithm is widely used to calculate the DFT. Sometimes we also use FFT as an alias name of DFT if the meaning is clear.

DFT uses a finite set of data that should cover full periods of a periodic signal. Otherwise, the endpoints of the data set are discontinuous, resulting in false high-frequency components in the DFT results. These high-frequency components are not present in the original signal, and therefore, cause errors in the spectrum measurement. It appears as if the energy at one frequency leaks into other frequencies. This phenomenon is known as the *spectral leakage*, which causes the signal's spectral lines to spread into neighbor DFT bins. We can minimize the spectral leakage effects using a technique called windowing. Windowing reduces the discontinuities at the boundaries of the data set. It multiplies the data set with a finite-length window with a magnitude varying smoothly and gradually toward zero at the edges. Fig. 6.1a shows several typical window functions used in DFT calculations. The DFT results of an example non-integer-period data set without and with windowing are depicted in Fig. 6.1b.

Windowing will reduce the spectral leakage of DFT but will also reduce the magnitude of the calculated spectral density. Proper scaling should be performed in DFT to compensate for the windowing effects, especially when measuring the PSD or SNR of the data. If possible (e.g., the ADC clock frequency is synchronized to the RF frequency), we should use a data set covering integer times period of the input signal to calculate DFT. This is the so-called *coherent sampling*. Windowing is not needed in this case, and a significant benefit is that the spectrum measurement resolution is as good as a single DFT bin. It enables the noise measurement close to the carrier frequency. Fig. 6.1c is the close-to-carrier spectrum comparing the cases with windowing and coherent sampling.

6.1.2.2 PSD and SNR Calculation

Suppose a noisy signal $y(t)$ is the sum of a sinusoidal signal $s(t)$ and a noise $x(t)$; that is, $y(t) = s(t) + x(t)$. We often need to estimate the PSD and SNR of y .

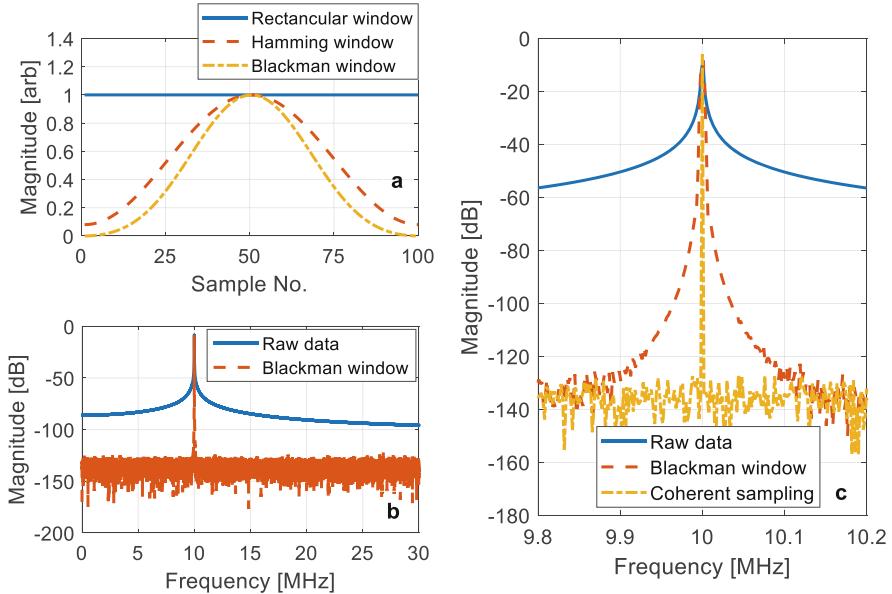


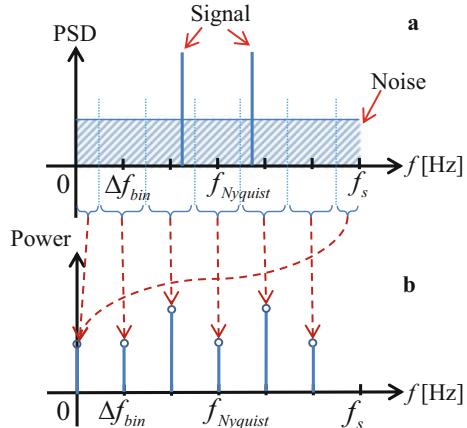
Fig. 6.1 (a). Typical window functions for 100 samples. The rectangular window corresponds to the case with no window; (b). DFT spectrum of a non-integer-period data set without and with windowing. The Blackman window is used here; (c). Spectrum comparison between windowing and coherent sampling.

The two-sided PSD of $y(t)$, which is denoted as $S'_y(f)$, is defined over a continuous frequency in the range of $-\infty \leq f \leq \infty$, as in (6.5). If we discretize $y(t)$ at a sampling rate f_s , the power of y will be aliased to the frequency range of $0 \leq f < f_s$ as a continuous-frequency PSD denoted as $S'_{y,\text{alias}}(f)$ (see Fig. 6.2a). The PSD of the sinusoidal signal is $P_s/2[\delta(f-f_{\text{signal}}) + \delta(f-f_s + f_{\text{signal}})]$, where P_s is the signal power and f_{signal} is the signal frequency. The delta-functions indicate that the signal produces two spectral lines at f_{signal} (assuming $f_{\text{signal}} < f_{\text{Nyquist}}$) and $f_s - f_{\text{signal}}$ (equivalent to $-f_{\text{signal}}$).

In practice, we calculate the PSD using the DFT of finite number (N) of samples. Hence, we can only obtain the PSD values at N discrete frequencies f_k ($k = 0, \dots, N-1$), each is the average PSD within a bandwidth of the bin spacing Δf_{bin} . The algorithm to calculate the PSD using DFT is described as follows. First, we calculate the power in each DFT bin by integrating $S'_{y,\text{alias}}(f)$ within Δf_{bin} around f_k :

$$\begin{aligned} P'_0 &= \int_0^{\Delta f_{\text{bin}}/2} S'_{y,\text{alias}}(f) df + \int_{f_s - \Delta f_{\text{bin}}/2}^{f_s} S'_{y,\text{alias}}(f) df, \\ P'_k &= \int_{f_k - \Delta f_{\text{bin}}/2}^{f_k + \Delta f_{\text{bin}}/2} S'_{y,\text{alias}}(f) df, k = 1, \dots, N-1. \end{aligned} \quad (6.12)$$

Fig. 6.2 (a). Continuous-frequency PSD of the noisy signal aliased to the frequency range of $0 \leq f < f_s$. Here we used $N = 6$ points to calculate the DFT. (b). Power spectrum at DFT bin frequencies (each point is the total power in a DFT bin spacing).



Note that $S'_{y,\text{alias}}(f)$ is symmetric to f_{Nyquist} . If we map the values of $S'_{y,\text{alias}}(f)$ in the frequency range $f_{\text{Nyquist}} < f < f_s$ to negative frequencies, we can define

$$S'_{y,\text{alias}}(-f) = S'_{y,\text{alias}}(f_s - f), 0 < f < f_{\text{Nyquist}},$$

so $S'_{y,\text{alias}}(f)$ is equivalent to a two-sided PSD defined in the frequency range $-f_{\text{Nyquist}} < f \leq f_{\text{Nyquist}}$. Equation (6.12) indicates that the power in each DFT bin is the integrated power in a frequency band from $f_k - \Delta f_{\text{bin}}/2$ to $f_k + \Delta f_{\text{bin}}/2$, see Fig. 6.2b. The DFT bin power at 0 Hz is an exception, for which we need to integrate within two separate frequency bands from 0 to $\Delta f_{\text{bin}}/2$ and from $f_s - \Delta f_{\text{bin}}/2$ to f_s . As discussed above, the power spectrum depicted in Fig. 6.2b is equivalent to a two-sided power spectrum. For a real-valued signal, we often use the one-side power spectrum defined in the Nyquist zone:

$$P_0 = P'_0, P_{N/2} = P'_{N/2}, P_k = 2P'_k, k = 1, \dots, N/2 - 1. \quad (6.13)$$

Note that the powers at 0 Hz and f_{Nyquist} do not double. Since each P_k ($k = 0, \dots, N/2$) is the total power within a bandwidth Δf_{bin} around f_k , the central frequency of the k th DFT bin, the one-sided PSD at f_k can be estimated as

$$S_y(f_k) = P_k / \Delta f_{\text{bin}}, k = 0, \dots, N/2. \quad (6.14)$$

We notice that if the spectral line of the signal s is located near a DFT bin frequency f_k , the signal power will be included in P_k . Then the signal power will be averaged within Δf_{bin} and contributes to $S_y(f_k)$ according to (6.14). This also implies that the frequency of the signal on the power spectrum has an uncertainty up to $\pm \Delta f_{\text{bin}}/2$. To increase the frequency resolution for detecting the signal frequency, more data points (i.e., larger N) should be used for the DFT calculation.

The approach above to estimate the PSD is called the *periodogram method*, which can be presented in a more practical form as

$$S_y(f_k) = \begin{cases} \frac{2|Y(f_k)|^2}{Nf_s W}, & 0 < k < N/2 \\ \frac{|Y(f_k)|^2}{Nf_s W}, & k = 0, N/2 \end{cases} \quad (6.15)$$

where $Y(f_k)$ is the DFT of $y(n)$ given by (6.11). Here W is a scale factor to compensate for the effects of the window function $w(n)$ used in the DFT calculation, which is given by $W = 1/N \cdot \sum_{n=0}^{N-1} [w(n)]^2$. Using the PSD calculated above, we can estimate the signal power P_s and the noise power P_n , and then the SNR. Fig. 6.2b shows that P_s is contained in the DFT bin nearest to the signal in frequency. We denote this bin as the *signal bin* and its frequency as f_{ks} . We will neglect the noise power in the signal bin to simplify the calculation. Besides, if there exists spectral leakage, the signal power will spread into the neighbor bins. Therefore, we may obtain multiple bins containing signal power (all denoted as *signal bins*) if the coherent sampling is not available. In this case, P_s should be calculated including all the bins containing signal power. Similarly, P_n should be calculated by summing up the powers in all other bins containing only noise power (denoted as *noise bins* and their frequencies as f_{kn}). Therefore, the SNR can be calculated as

$$\begin{aligned} SNR &= P_s/P_n, \text{ where } P_s = \sum [S_y(f_{ks}) \Delta f_{bin}] \text{ and } P_n \\ &= \sum [S_y(f_{kn}) \Delta f_{bin}]. \end{aligned} \quad (6.16)$$

As an example, Fig. 6.3 is the PSD of a sinusoidal signal with white noise simulated with Matlab. The amplitude and frequency of the signal are 1 and 10 MHz, respectively, and the RMS value of the noise is 4.46e-5. The sampling rate is $f_s = 100$ MHz. We used the algorithm above to calculate the PSD and SNR, where a Kaiser window is used in the DFT calculation. The plot illustrates that we have spectral leakage since the signal power spreads into multiple bins around its frequency, 10 MHz. The SNR calculated from the PSD is 83.96 dB, which is close to the value (84 dB) given by the simulation settings.

6.1.3 Correlation of Noise

The correlation of multiple noises determines their total impact when they are summed up. Noises from different sources are often uncorrelated, whereas correlated noises typically have a common (or partly common) origination.

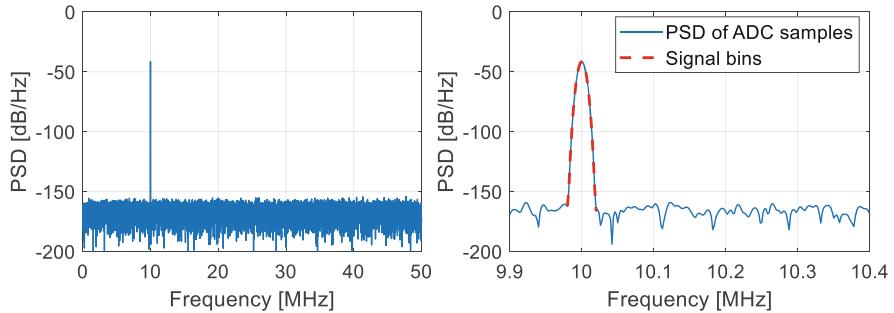


Fig. 6.3 PSD and SNR estimation using the periodogram method. The plot on the right-hand side is the zoom in around the signal frequency

6.1.3.1 Description of Correlation

The correlation of the noise $x(n)$ and $y(n)$ (already discretized) can be characterized by the so-called *correlation coefficient*, which is defined as

$$\rho := \frac{\sum_{n=0}^{N-1} x(n)y(n)}{\sqrt{\sum_{n=0}^{N-1} [x(n)]^2 \cdot \sum_{n=0}^{N-1} [y(n)]^2}}, \quad (6.17)$$

where N is the number of samples. Note that we have assumed that the mean values of x and y are zero. The value of ρ is in the range of $-1 \leq \rho \leq 1$. If ρ approaches 0, x and y are uncorrelated. If ρ is close to 1 (or -1), it indicates a strong positive (or negative) correlation between x and y . See the examples in Fig. 6.4. The correlation coefficient shows how strong the two noises change together in the same or opposite direction. They should be sampled at the same time instant. Note that the correlation coefficient is defined with the assumption of a linear correlation between the noises. If the correlation follows other patterns (e.g., a sine function), the definition (6.17) is no longer applicable.

Another concept describing the correlation of two noises is the *cross-correlation function*, which is defined similarly as autocorrelation as

$$\begin{aligned} R_{xy}(\tau) &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t + \tau)dt, \text{ or} \\ R_{xy}(k) &:= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n+k). \end{aligned} \quad (6.18)$$

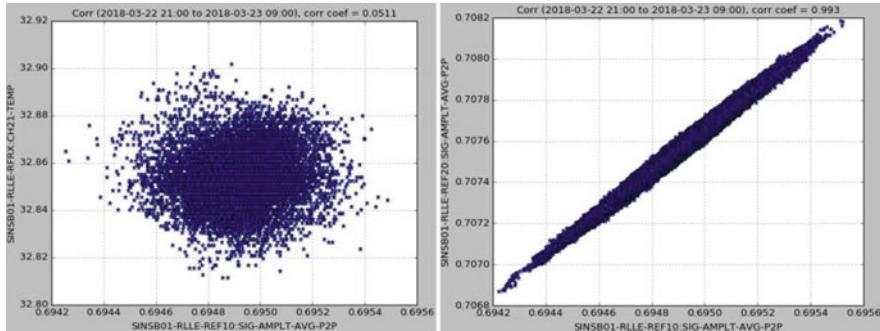


Fig. 6.4 Example of correlation coefficients with weak (left) and strong (right) correlations

It reflects the correlation between the noise x and the noise y delayed by τ (or k samples). The advantage of R_{xy} is that it can capture the correlation between two noises even if they have different delays in the measurements. The cross-correlation functions satisfy the relation below

$$R_{xy}(\tau) = R_{yx}(-\tau), R_{xy}(k) = R_{yx}(-k). \quad (6.19)$$

When x and y are uncorrelated, $R_{xy} = 0$ for all τ (or k).

Similar to the PSD definition (6.6), we can also define a *cross power spectral density (CPSD)* of two noises using the cross-correlation function:

$$S'_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau} d\tau, R_{xy}(\tau) = \int_{-\infty}^{\infty} S'_{xy}(f) e^{j2\pi f\tau} df. \quad (6.20)$$

Note that $S'_{xy}(f)$ is a complex-valued function of the frequency f . Its real part is an even function, while its imaginary part is an odd function. The CPSD is defined for both positive and negative frequencies.

In practice, we may also use the periodogram method to calculate the CPSD of two noise series $x(n)$ and $y(n)$:

$$S'_{xy}(f_k) = \frac{X(f_k)Y^*(f_k)}{Nf_s W}, \quad k = 0, N-1 \quad (6.21)$$

where $X(f_k)$ and $Y(f_k)$ are the DFTs of the windowed series $x(n)$ and $y(n)$. The superscript “*” represents the conjugate of the complex-valued spectrum, and other parameters have the same meanings as (6.15). Because f_k ($k = N/2 + 1, \dots, N-1$) corresponds to negative frequencies, (6.21) is equivalent to a two-sided CPSD defined in the frequency range $-f_{Nyquist} < f \leq f_{Nyquist}$.

6.1.3.2 PSD of the Sum of Two Noises

Summing up multiple noises is a practical topic for LLRF noise analysis. Some algorithms may require summing or subtracting two noisy signals (e.g., calculating the phase difference between two RF signals). Given two noises, $x(t)$ and $y(t)$, we need to find out how to estimate the PSD of their sum $z(t) = x(t) + y(t)$.

Using the definition of autocorrelation (6.4) and the definition of cross-correlation (6.18), we obtain

$$\begin{aligned} R_z(\tau) &= R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{yx}(\tau) \\ &= R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{xy}(-\tau) \end{aligned} \quad (6.22)$$

where $R_z(\tau)$, $R_x(\tau)$ and $R_y(\tau)$ are the autocorrelations of z , x and y , respectively, and $R_{xy}(\tau)$ and $R_{yx}(\tau)$ are the cross-correlations between x and y . We obtain an equation in terms of two-sided PSDs and CPSDs by applying Fourier transform to both sides of (6.22):

$$S'_z(f) = S'_x(f) + S'_y(f) + S'_{xy}(f) + S'_{xy}(-f). \quad (6.23)$$

The property of the cross-correlation tells that $S'_{xy}(f) + S'_{xy}(-f) = 2 \operatorname{Re}\{S'_{xy}(f)\}$, where $\operatorname{Re}\{\cdot\}$ fetches the real part of a complex number. We can define an *equivalent one-sided CPSD* as

$$S_{xy}(f) := \begin{cases} 2 \operatorname{Re}\{S'_{xy}(f)\} & f \geq 0 \\ 0 & f < 0 \end{cases}. \quad (6.24)$$

Note that S_{xy} is real valued and its sign may be positive or negative. We can rewrite (6.23) in terms of one-sided PSDs and CPSDs as

$$S_z(f) = S_x(f) + S_y(f) + 2S_{xy}(f). \quad (6.25)$$

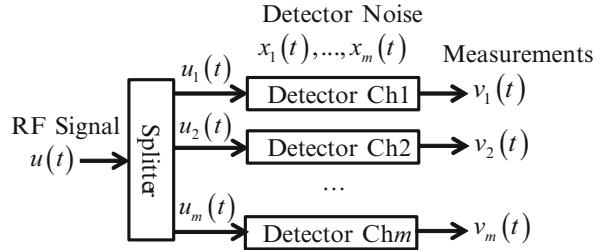
If x and y are uncorrelated, then $S_{xy}(f) = 0$, and the PSD of $x + y$ is the sum of the PSDs of x and y ; otherwise, the S_{xy} term will be non-zero to include the effects of correlation. Consider an extreme case, in which x and y are identical, then $S_x = S_y = S_{xy}$ and the PSD of $x + y$ is four times the PSD of x or y .

Let us use an example to demonstrate the summing up of uncorrelated noises.

Example 6.1 RF Detection Noise Reduction with Multiple Channels

If the RF detectors have uncorrelated noise in different channels, we can measure the same RF signal using multiple channels and average the results to reduce the detector noise effects. This measurement strategy is shown in Fig. 6.5.

Fig. 6.5 Measurement of the same RF signal using multiple detector channels



We assume that the split signals $u_1(t), \dots, u_m(t)$ are identical and use $u_c(t)$ to denote them. We also assume that the noise of different detector channels is uncorrelated to each other but has the same RMS value x_{rms} . The output of the i th channel can be written as

$$v_i(t) = u_i(t) + x_i(t) = u_c(t) + x_i(t), \quad i = 1, \dots, m.$$

The presence of the noise $x_i(t)$ will reduce the measurement resolution. Here we average the outputs of the m channels and yield

$$v_{avg}(t) = \frac{1}{m} \sum_{i=1}^m v_i(t) = u_c(t) + \frac{1}{m} \sum_{i=1}^m x_i(t),$$

where the second term is the average of m uncorrelated noises. As discussed in this section, the total noise PSD is the sum of the PSDs of all the m noises. Therefore

$$PSD \left\{ \frac{1}{m} \sum_{i=1}^m x_i(t) \right\} = \frac{1}{m^2} \sum_{i=1}^m S_{x_i}(f).$$

Then we derive the relationship in terms of RMS values as

$$RMS \left\{ \frac{1}{m} \sum_{i=1}^m x_i(t) \right\} = \sqrt{\frac{1}{m^2} \sum_{i=1}^m x_{i,rms}^2} = x_{rms}/\sqrt{m}.$$

It implies that the RMS value of the noise term of v_{avg} is $1/\sqrt{m}$ of that in the single-channel measurement.

In practice, the noise added by RF detectors consists of both correlated noise (e.g., LO and clock noise) and uncorrelated noise (e.g., thermal noise). The method described above will reduce the uncorrelated noise but not the correlated noise.

6.1.4 Additive Noise and Parametric Noise

Noise is coupled into a signal via two typical ways: summing or modulating. If the noise is added to a signal by summing, it is called *additive noise*. The power of additive noise is independent of the signal power. Therefore, we may improve the SNR by increasing the signal level. For example, if we use an ADC to sample an IF signal, we require the IF signal level to be close to the full scale of ADC input to mitigate the ADC's additive noise.

On the other hand, if the noise is modulated onto the amplitude or phase of the RF signal, it is called *parametric noise*. A sinusoidal signal with amplitude and phase modulations can be written as

$$\begin{aligned} y(t) &= A[1 + \alpha(t)] \sin(2\pi f_c t + \Delta\varphi(t)) \\ &\approx \underbrace{A \sin(2\pi f_c t)}_{s(t)} + \underbrace{A\alpha(t) \sin(2\pi f_c t)}_{x_\alpha(t)} + \underbrace{A\Delta\varphi(t) \cos(2\pi f_c t)}_{x_\varphi(t)}, \end{aligned} \quad (6.26)$$

where A is the amplitude, f_c is the carrier frequency of the signal, and both are constant. The terms $\alpha(t)$ and $\Delta\varphi(t)$ are the random amplitude and phase modulations assumed to be much smaller than 1. They are the so-called amplitude and phase noise that will be discussed in Sect. 6.1.7. Note that we use $\Delta\varphi(t)$ to denote the phase noise to distinguish with the phase $\varphi(t)$, which has been used to denote the absolute phase of the RF signal related to the beam phase. The absolute phase can also be time-varying if the beam phase varies with time. The magnitudes of the induced noise terms $x_\alpha(t)$ and $x_\varphi(t)$ are proportional to A . Therefore, even if we increase the signal level, the noise level rises proportionally, resulting in a constant SNR.

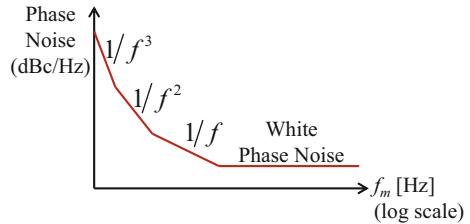
6.1.5 White Noise and 1/f Noise

White noise has a constant PSD at all frequencies and its time-domain value follows Gaussian distribution. The *thermal noise* is approximately a white noise up to thousands of GHz. It is caused by the thermal movement of electrons or other charge carriers in conductive materials when the absolute temperature is higher than 0 Kelvin (K). The thermal noise PSD delivered to a matched load from a resistor is

$$S_{th}(f) = kT, \quad (6.27)$$

where k is the Boltzmann's constant ($1.374\text{e-}23 \text{ J/K}$) and T is the absolute temperature of the resistor (in K). Note that S_{th} is independent of the resistance of the resistor. An important value worth remembering is the *thermal noise PSD at room temperature* (290 K), -174 dBm/Hz . The thermal noise is not avoidable for any practical devices, limiting the best achievable SNR. In RF detectors, we should

Fig. 6.6 Phase noise PSD of an RF oscillator (f_m is the offset frequency relative to the carrier)



consider the thermal noise if the RF signal level is small or the system bandwidth is large. The $1/f$ noise is frequency dependent. It has a larger PSD at lower frequencies and has a typical PSD like $S_{1/f}(f) \propto 1/f^\eta$. For example, the $1/f$ phase noise in an RF oscillator comes from various sources, corresponding to different η values: $\eta = 3$, flicker frequency noise; $\eta = 2$, white frequency noise; $\eta = 1$, flicker phase noise, and $\eta = 0$, white phase noise. Figure 6.6 shows a typical PSD distribution of the phase noise in an RF oscillator.

6.1.6 Noise Factor

The *noise factor* is a concept frequently used in the specifications of RF components like amplifiers and mixers. It specifies the amount of white noise added by the RF component.

The noise factor (F) of an RF component is defined as

$$F = \frac{S_{out}}{G_p S_{th}}, \quad (6.28)$$

where S_{th} is the thermal noise PSD at 290 K, and S_{out} is the PSD of the total white noise at the RF component output when the input is only the thermal noise. The power gain of the RF component is G_p , and $G_p S_{th}$ is the contribution of the input thermal noise at the output. For a noiseless RF component, we have $F = 1$; that is, the output noise comes only from the input thermal noise. The noise factor is often written in the logarithm scale as the *noise figure* (NF):

$$NF = 10 \log_{10} F. \quad (6.29)$$

For example, we consider the noise figure of a 10 dB passive attenuator at room temperature. The power gain of the attenuator is $G_p = 0.1$, and since both the input and output noise PSDs are all S_{th} , we conclude that $NF = 10$ dB. Generally, the noise figure of any passive attenuator equals the attenuation (in dB) of itself.

In some literature, the noise figure is also defined as the SNR degradation at the output compared to the input SNR of the RF component ($SNR_{out} - SNR_{in}$, in dB).

When using this definition, one should keep in mind that the input noise is only the thermal noise.

When multiple (n) RF components are cascaded serially, the total noise factor is calculated with the Friis formula

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_n - 1}{G_1 G_2 \dots G_{n-1}}, \quad (6.30)$$

where F_i and G_i are the noise factor and power gain of the i th component. Note that G_i is a linear gain. We can see that a larger gain and a lower noise factor in the first stage will reduce the total noise factor. Note that the noise factor only specifies the wideband white noise (noise floor) added by the RF component. In LLRF systems, the 1/f noise close to carrier frequency is more interested because the RF cavities are typically narrow-band. When selecting an RF component (e.g., RF amplifier), the noise figure is usually specified in the datasheet, but the close-to-carrier noise is not. Measurement is essential to evaluate the noise performance of the RF component. As an example (Teytelman 2015), Fig. 6.7 shows the phase noise measurement of two different amplifiers, SKY65162 ($NF = 4.45$ dB) and MGA-31589 ($NF = 1.15$ dB). The amplifier with a lower noise figure shows bigger low-frequency phase noise than the one with a higher noise figure.

6.1.7 Phase Noise and Amplitude Noise

In LLRF systems, the signals to be processed are typically narrow-band RF signals around a carrier frequency. The phase noise (Rubiola 2008) and amplitude noise are critical factors that should be considered in the design of LLRF electronics and control algorithms.

6.1.7.1 Phase Noise

For a sinusoidal signal with constant amplitude A and frequency f_c , a random phase $\Delta\varphi(t)$ causes *phase noise*. The sine signal with phase noise is presented as

$$y(t) = A \sin(2\pi f_c t + \Delta\varphi(t)) \approx \underbrace{A \sin(2\pi f_c t)}_{s(t)} + \underbrace{A \Delta\varphi(t) \cos(2\pi f_c t)}_{x_\varphi(t)}, \quad (6.31)$$

where $s(t)$ is the desired signal and $x_\varphi(t)$ describes the phase noise. Here we have assumed that $|\Delta\varphi(t)| \ll 1$ rad so that $\cos\Delta\varphi(t) \approx 1$ and $\sin\Delta\varphi(t) \approx \Delta\varphi(t)$. Fig. 6.8a depicts the PSD of the noisy signal $y(t)$. Suppose the noisy signal $y(t)$ is applied to a resistance R , then the power of the signal $s(t)$ is $P_s = A^2/(2R)$. Assume the RMS value of the phase is $\Delta\varphi_{rms}$, and we can calculate the total noise power (within the

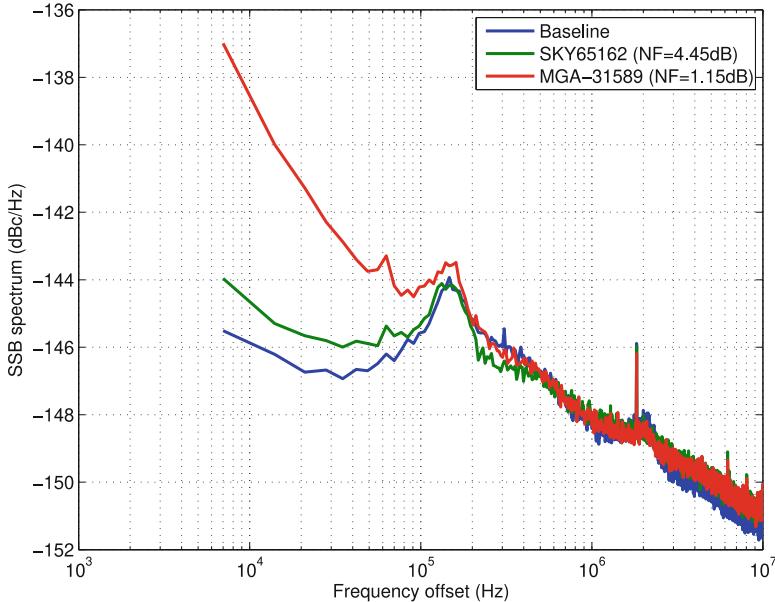


Fig. 6.7 Phase noise of two amplifiers with different noise figures (Courtesy of D. Teytelman)

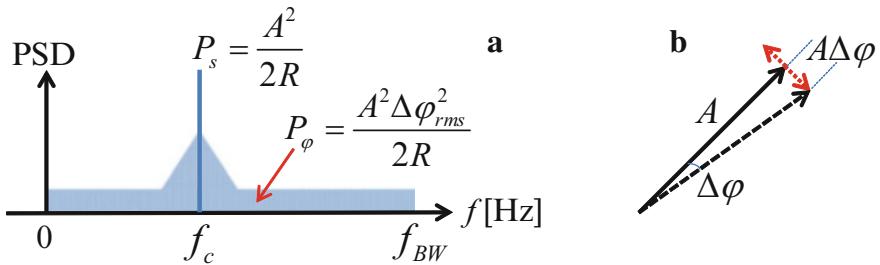


Fig. 6.8 (a). PSD of $y(t)$ consisting of $s(t)$ and $x_\varphi(t)$. The PSD of $s(t)$ is $S_s(f) = P_s \delta(f - f_c)$, and the PSD of $x_\varphi(t)$ is symmetric around f_c ; (b). vector diagram of the phase noise.

bandwidth of the noisy signal, f_{BW}) as $P_\varphi = A^2 \Delta\varphi_{rms}^2 / (2R)$ according to the expression of $x_\varphi(t)$ in (6.31). Therefore, we can use $\Delta\varphi_{rms}^2 = P_\varphi / P_s$ to estimate the phase jitter. That is, the relative value between the noise power and the signal power can represent the phase noise.

The phase noise is defined within the carrier frequency's higher sideband ($f > f_c$) as a relative PSD to the carrier signal power. It is typically described in the logarithmic scale with a unit dBc/Hz. The phase noise is a function of the offset frequency f_m , where $f_m = f - f_c > 0$, and we use $\mathcal{L}(f_m)$ to denote the phase noise. Such a definition is the *single-sideband (SSB) phase noise*. The RMS phase jitter within an offset-frequency range between f_1 and f_2 (both larger than 0) can be calculated as

$$\Delta\varphi_{rms} = \sqrt{2 \int_{f_1}^{f_2} \mathcal{L}(f_m) df_m}. \quad (6.32)$$

Note that when calculating the integral above, $\mathcal{L}(f_m)$ should use linear units instead of dBc/Hz. In (6.32), we double the SSB phase noise to include the noise power in the lower sideband ($f < f_c$), assuming that the phase noise is symmetric in the lower and higher sidebands.

If we view $\Delta\varphi(t)$ as a general noise and denote its PSD as $S_\varphi(f_m)$, we obtain

$$S_\varphi(f_m) = 2\mathcal{L}(f_m). \quad (6.33)$$

Here S_φ is the *double-sideband (DSB) phase noise*. It is equivalent to folding the spectrum of Fig. 6.8a around f_c and adding the lower sideband noise power to the higher sideband. In other words, (6.33) illustrates that the noise power of the random phase is distributed equally into both sidebands of the carrier frequency. The unit of S_φ is “rad²/Hz” (linear scale) or “dBrad²/Hz” (logarithmic scale). A commercial phase noise analyzer (e.g., Agilent E5052B) typically provides results as SSB phase noise.

We may also present the carrier frequency signal and the phase noise as vectors. See Fig. 6.8b. The phase noise vector is orthogonal to the carrier signal vector. The RMS values of the vector magnitudes are $A/\sqrt{2}$ (for signal) and $A\Delta\varphi_{rms}/\sqrt{2}$ (for phase noise), respectively. Therefore, if only phase noise is present in the signal, we can calculate the RMS phase jitter from the SNR directly as

$$\Delta\varphi_{rms} = \frac{A\Delta\varphi_{rms}/\sqrt{2}}{A/\sqrt{2}} = \frac{1}{\sqrt{SNR_\varphi}}, \quad (6.34)$$

where SNR_φ is the SNR when only considering the phase noise power. In some cases, the timing jitter is of more interest, which can be calculated as

$$\Delta t_{rms} = \frac{\Delta\varphi_{rms}}{2\pi f_c}. \quad (6.35)$$

6.1.7.2 Amplitude Noise

If the amplitude of the sinusoidal signal has a random component $\alpha(t)$, we get a signal with *amplitude noise*

$$y(t) = A[1 + \alpha(t)] \sin(2\pi f_c t) = \underbrace{A \sin(2\pi f_c t)}_{s(t)} + \underbrace{A\alpha(t) \sin(2\pi f_c t)}_{x_\alpha(t)}, \quad (6.36)$$

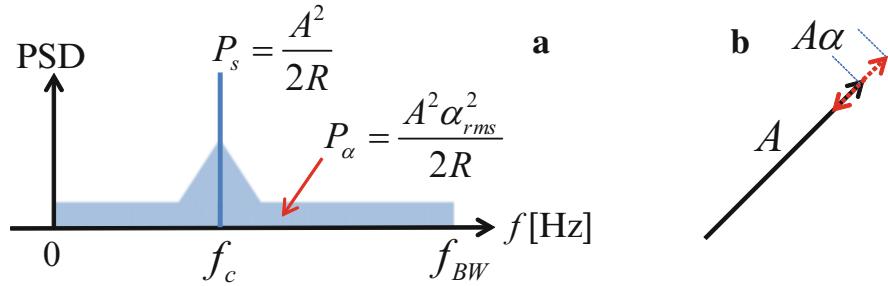


Fig. 6.9 (a). PSD of $y(t)$ consisting of $s(t)$ and $x_\alpha(t)$. The PSD of $s(t)$ is $S_s(f) = P_s\delta(f-f_c)$, and the PSD of $x_\alpha(t)$ is symmetric around f_c ; (b). vector diagram of the amplitude noise

where $s(t)$ is the desired signal and $x_\alpha(t)$ represents the amplitude noise. Similar to Fig. 6.8, we plot the PSD of (6.36) and its vector diagram in Fig. 6.9. The ratio between the amplitude noise power P_α and the signal power P_s describes the relative jitter of the amplitude: $\alpha_{rms}^2 = P_\alpha/P_s$. Note that the noise power is calculated within the bandwidth of the noisy signal (i.e., bandwidth of the filter applied to the signal). The amplitude noise is typically represented as a DSB PSD of $\alpha(t)$, $S_\alpha(f_m)$. Similar to the DSB phase noise, S_α is equivalent to folding the spectrum of Fig. 6.9a around f_c and adding the lower sideband noise power to the higher sideband. The unit of S_α is dB/Hz in logarithmic scale. The RMS relative amplitude jitter within an offset-frequency range between f_1 and f_2 (both larger than 0) can be calculated as

$$\alpha_{rms} = \left(\frac{\Delta A}{A} \right)_{rms} = \sqrt{\int_{f_1}^{f_2} S_\alpha(f_m) df_m}. \quad (6.37)$$

Here $\Delta A(t) = A\alpha(t)$ is the absolute amplitude error. When calculating the integral above, $S_\alpha(f_m)$ should use linear units instead of dB/Hz. The vector diagram of the amplitude noise is depicted in Fig. 6.9b. The relative amplitude jitter can also be calculated using the SNR if the amplitude noise is the only noise in the signal:

$$\alpha_{rms} = \frac{A\alpha_{rms}/\sqrt{2}}{A/\sqrt{2}} = \frac{1}{\sqrt{SNR_\alpha}}, \quad (6.38)$$

where SNR_α is the SNR when only considering the amplitude noise power.

6.1.7.3 Signal with Amplitude and Phase Noise

A sinusoidal signal containing both amplitude and phase noise is described by (6.26). The total noise PSD at an offset frequency f_m can be calculated as

$$S_n(f_m) = S_\varphi(f_m) + S_\alpha(f_m). \quad (6.39)$$

Here α and $\Delta\varphi$ do not need to be uncorrelated because they are orthogonal to each other according to Figs. 6.8b and 6.9b. Note that S_n , S_φ and S_α are all DSB relative PSDs to the carrier signal power. The amplitude and phase jitter has the following relationship with the overall noise PSD and SNR:

$$\alpha_{rms}^2 + \Delta\varphi_{rms}^2 = \int_{f_1}^{f_2} S_n(f_m) df_m = \frac{1}{SNR}, \quad (6.40)$$

where f_1 and f_2 (both larger than 0 and $f_2 > f_1$) define the offset-frequency range where the noise is interested. In practice, (6.34), (6.38) and (6.40) are very useful to estimate the amplitude or phase jitter using the SNR. For example, if an ADC can achieve an SNR of 80 dB, the measurement resolution can be estimated to be $\Delta\varphi_{rms} = 0.0057^\circ$ (if only phase noise is present) or $\alpha_{rms} = 1e-4$ (if only amplitude noise is present). These are the upper bounds of the phase and amplitude jitter for this example.

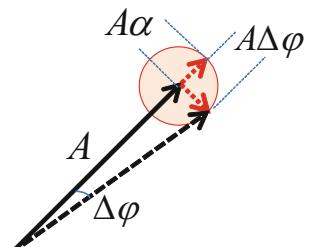
6.1.8 Additive Noise and RF Jitter

In some cases, white noise (with constant PSD) is added by the RF components in the RF signal chain. We are interested in how the amplitude and phase noise will be affected by such additive noise. Suppose we have

$$y(t) = A \sin(2\pi f_c t) + n(t), \quad (6.41)$$

where A and f_c are constant, and $n(t)$ is white noise introducing a random vector located in a disk on the signal vector. See Fig. 6.10. The noise vector is projected in two directions, one is the same as the signal vector, and the other is perpendicular. They correspond to the resulting amplitude and phase noise, respectively. The amplitude noise $\alpha(t)$ and phase noise $\Delta\varphi(t)$ caused by $n(t)$ are correlated to each other, and each obtains half of the power of $n(t)$. That is, the following relationships exist:

Fig. 6.10 Vector diagram of the additive noise



$$S_\varphi(f_m) = S_\alpha(f_m) = S_n(f_m)/2, \alpha_{rms} = \Delta\varphi_{rms} = \frac{1}{\sqrt{2SNR}}. \quad (6.42)$$

6.1.9 Frequency-Domain Meaning of RMS Value

We often calculate the noise's RMS value, such as the amplitude or phase RMS jitter, using a number of samples. To interpret the RMS value properly, we need to clarify its physical meaning, i.e., the frequency range in which the noise power is included.

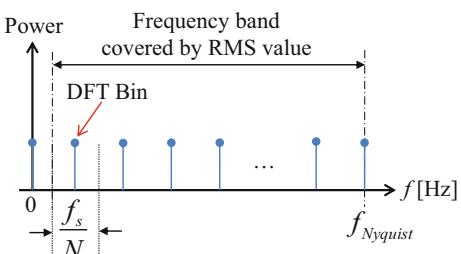
Assume that we sample a noise $x(t)$ at a rate f_s and use N samples to calculate the RMS value. The mean value of $x(t)$ is excluded from the RMS value calculation. That is, the DFT bin at 0 Hz (DC) does not contribute to the RMS value, but other DFT bins, which are derived from the N samples (see Fig. 6.11), will be included. Therefore, the upper boundary of the frequency band in which the noise power contributes to the RMS value is $f_{Nyquist}$. The low-frequency limit for the noise power covered by the RMS value is half the bin spacing, $f_s/(2N)$. This is because the noise power between 0 Hz and $f_s/(2N)$ is included in the DC bin.

Sampling the noise at rate f_s will first alias the noise power into the first Nyquist zone (DC to $f_s/2$). Therefore, *calculating the RMS value using N samples corresponds to summing up the aliased noise power in the frequency range from $f_s/(2N)$ to $f_s/2$* . For example, assume we sample a noise at $f_s = 119$ MHz and calculate the RMS value using $N = 32,768$ samples. The noise power is aliased to the Nyquist zone (DC to 59.5 MHz) by the sampling process, and the RMS value of the 32,768 samples includes the aliased noise power from 1.82 kHz to 59.5 MHz. If we want to measure the noise at lower-frequencies, N should be increased.

6.1.10 Drift and Jitter

The concepts *drift* and *jitter* are frequently used when describing the variations of a parameter like the amplitude or phase of the cavity voltage. The definition of drift

Fig. 6.11 Power spectrum of N samples aliased into the first Nyquist zone. The plot is for an even number of N . If N is an odd number, the bin at $f_{Nyquist}$ does not exist



and jitter are rough, expressing the long-term and short-term variations, respectively. In this section, we will clarify these concepts in the frequency domain.

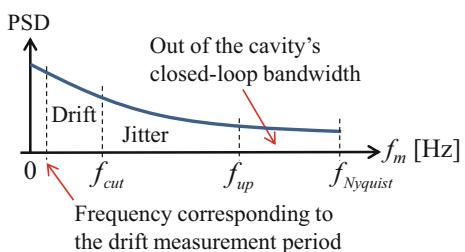
The drift is typically caused by slow disturbances like the environmental temperature or humidity changes. The drift of a parameter is described as the peak-to-peak (relative) variation slower than a cutoff frequency f_{cut} (e.g., 1 Hz) and measured in a specific period (e.g., 24 hours). The measurement period should typically cover a full cycle of the temperature or humidity changes. The jitter sources are usually fast disturbances, such as the klystron high voltage jitter, ADC clock jitter and microphonics. The jitter of a parameter can be quantified by the (relative) RMS variation faster than f_{cut} and below an upper frequency, f_{up} . In an RF system, the upper frequency for jitter calculation is often the cavity's closed-loop bandwidth. In this case, the jitter will represent the fast cavity voltage variation experienced by the beam.

The cutoff frequency splitting the definitions of drift and jitter is subjective. In practice, we may choose such a cutoff frequency, below which the cavity voltage variations can be compensated for by the beam-based feedbacks. Then the LLRF system can focus on dealing with the RF field jitter. When defining the tolerances of drift and jitter, the cutoff frequency f_{cut} , the measurement period for drift, and the upper jitter frequency f_{up} , should always be specified. This helps to define drift and jitter requirements that can be validated via practical measurements.

Figure 6.12 shows the frequency bands in which the noise power should be integrated to calculate the drift and jitter. Typically, we select $f_{up} < f_{Nyquist}$ (i.e., we choose a Nyquist frequency larger than the closed-loop bandwidth of the cavity) to avoid aliasing. The noise power above f_{up} is not relevant for the beam acceleration.

For an RF station operating in the pulsed mode, the amplitude and phase are calculated from the RF pulse envelope. Usually, we average over a time segment within the RF pulse to derive the amplitude and phase for every pulse, corresponding to limiting the measurement within the cavity bandwidth. See Sect. 6.3.2 for details. The pulse-to-pulse amplitude and phase sequences are equivalent to sampling the cavity voltage (band-limited by the cavity bandwidth via averaging) at the pulse repetition rate f_{pul} . In this case, the noise power up to the cavity bandwidth will be aliased to the Nyquist zone of f_{pul} (DC to $f_{pul}/2$). Therefore, for such a pulsed RF station, the pulse-to-pulse jitter calculation should choose $f_{up} = f_{pul}/2$.

Fig. 6.12 Drift and jitter representing the integrated noise power in the frequency domain



6.2 Noise Model of Basic RF Components

In this section, we will introduce the (amplitude and phase) noise model of some basic RF components (Hoffmann 2008; Gallo 2011). The model describes how the input noise transfers to the output and the additional noise introduced by the RF component. To simplify the discussion, we assume the RF component operates at the room temperature (290 K).

6.2.1 Two-Port Passive RF Components

Two-port passive RF components, such as attenuators, transmission lines, and passive filters, do not add noise to the input RF signal. However, if all the ports of the RF component are matched by resistance impedances (e.g., $50\ \Omega$), thermal noise will be added to the output RF signal.

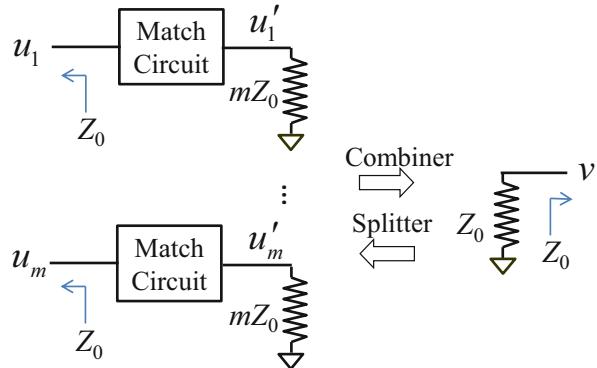
Suppose the two-port passive RF component has a frequency-dependent power gain $G_p(f_m)$, where f_m is the offset frequency relative to the carrier frequency. We have $G_p(f_m) \leq 1$ for passive components. When a noisy RF signal passes through the component, the carrier signal will experience a gain of $G_p(0)$, whereas the amplitude and phase noise will experience a gain of $G_p(f_m)$. Assume the input RF signal power is $P_{s,in}$ and its amplitude and phase noise PSDs are $S_{\alpha,in}(f_m)$ and $S_{\varphi,in}(f_m)$, respectively. The thermal noise PSD is denoted as S_{th} that is a constant ($-174\ \text{dBm/Hz}$) within the frequency range in which the LLRF system operates. Then the output amplitude and phase noise PSDs ($S_{\alpha,out}$ and $S_{\varphi,out}$) can be expressed as

$$S_{n,out}(f_m) = \frac{G_p(f_m)}{G_p(0)} S_{n,in}(f_m) + \frac{S_{th}}{G_p(0)P_{s,in}}, \text{ with } n : \alpha, \varphi. \quad (6.43)$$

Here $S_{\alpha,out}$ and $S_{\varphi,out}$ follow the same formula and we use $S_{n,out}$ to denote either of them. The first term on the right side of (6.43) describes the transfer of the input noise. When $G_p(f_m)$ is frequency-independent, the input noise $S_{n,in}$ (i.e., $S_{\alpha,in}$ or $S_{\varphi,in}$) will be fully transferred to the output; that is, both the signal and the noise are attenuated by the same factor. The second term is from the thermal noise of the resistive output impedance. The thermal noise is additive white noise, so half of its power goes into the amplitude noise and another half into the phase noise. That is, both the amplitude noise and phase noise will receive an additional power $S_{th}/2$ at both sidebands. Therefore, when calculating the DSB PSDs, we obtain S_{th} (divided by the output signal power) in both $S_{\alpha,out}$ and $S_{\varphi,out}$.

Equation (6.43) implies that a two-port passive RF component (with a constant gain at various frequencies) will preserve the amplitude and phase noise of the input RF signal. However, the thermal noise will be added that requires the output power, $G_p(0)P_{s,in}$, to be large to avoid degrading the SNR significantly.

Fig. 6.13 Equivalent circuit of an m -way power splitter (combiner)



6.2.2 Power Splitter and Combiner

An RF power splitter divides an RF signal into several branches, and a power combiner performs the opposite process. In this section, we will build a noise model for the RF power splitters/combiners.

The equivalent circuit (Szczepaniak 2010) of an m -way lossless power splitter (combiner) is shown in Fig. 6.13. It is a power splitting (summing) device, and all the ports are matched to the external transmission lines. If the ports u_i ($i = 1, \dots, m$) are used as inputs, it works as a power combiner, whereas if v is the input, it is a power splitter.

When used as a power combiner, the input voltages u_i ($i = 1, \dots, m$) come from matched sources with an impedance Z_0 , which equals the characteristic impedance of the transmission line. At each input port, the impedance is converted to mZ_0 so that the m shunt inputs match the output impedance. Then we have $u'_i = \sqrt{m} \cdot u_i$. The current flows from the inputs to the output should be consistent. It implies that $\sum_{i=1}^m u'_i / (mZ_0) = v/Z_0$. Therefore, the input and output voltages should satisfy

$$v = \frac{1}{\sqrt{m}} \sum_{i=1}^m u_i. \quad (6.44)$$

Assume that the inputs u_i are RF signals with the same carrier frequency f_c and phase (normalized to 0) and contain uncorrelated amplitude and phase noise. Referring to (6.26), we can write the i th input as

$$u_i(t) = A_i \sin(2\pi f_c t) + A_i \alpha_i(t) \sin(2\pi f_c t) + A_i \Delta\varphi_i(t) \cos(2\pi f_c t), \quad (6.45)$$

where A_i , α_i and $\Delta\varphi_i$ are the amplitude, amplitude noise and phase noise of the i th input, respectively. Following (6.44), we obtain the output as

$$v(t) = \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \right) \sin(2\pi f_c t) + \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \alpha_i(t) \right) \sin(2\pi f_c t) + \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \Delta\varphi_i(t) \right) \cos(2\pi f_c t). \quad (6.46)$$

Let us consider a particular case. Assume that the m inputs have the same amplitude A_u (i.e., $A_i = A_u$, $i = 1, \dots, m$) and uncorrelated noise with the same PSDs. We denote the common RMS amplitude jitter as $\alpha_{u,rms}$ and the common RMS phase jitter as $\Delta\varphi_{u,rms}$. From the assumption, we have $\alpha_{i,rms} = \alpha_{u,rms}$ and $\Delta\varphi_{i,rms} = \Delta\varphi_{u,rms}$ for $i = 1, \dots, m$. The amplitude of the output $v(t)$ becomes $\sqrt{m}A_u$, which implies that $P_v = mP_u$, where P_v is the output power and P_u is the power of each input. The relative amplitude variation and the phase variation of the output are given by

$$\alpha_v(t) = \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \alpha_i(t) \right) / \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \right) = \sum_{i=1}^m \alpha_i(t) / m,$$

$$\Delta\varphi_v(t) = \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \Delta\varphi_i(t) \right) / \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m A_i \right) = \sum_{i=1}^m \Delta\varphi_i(t) / m.$$

With the assumption above, the output amplitude and phase RMS jitter can be calculated as

$$\alpha_{v,rms} = \alpha_{u,rms} / \sqrt{m}, \Delta\varphi_{v,rms} = \Delta\varphi_{u,rms} / \sqrt{m}.$$

Therefore, the noise PSD at the *power combiner* output can be written as

$$S_{n,v}(f_m) = S_{n,u}(f_m) / m + S_{th}/P_v, \text{ with } n : \alpha, \varphi, \quad (6.47)$$

where $S_{n,v}$ and $S_{n,u}$ denote the amplitude or phase noise in the output and each input, respectively. Here we have neglected the frequency dependency in the gain of the power combiner. The thermal noise effects in the output port are also included. As a conclusion, if the amplitude (or phase) noise of the m inputs is uncorrelated, the amplitude (or phase) noise in the output is about m times lower than the input, assuming similar input noise PSDs. If the circuit in Fig. 6.13 is used as a *power splitter*, the voltage eq. (6.44) is still valid. Both the signal power and noise power in the input signal v will be split into m branches. Then we obtain the following noise PSD in the i th output

$$S_{n,ui}(f_m) = S_{n,v}(f_m) + S_{th}/P_{ui}, \text{ with } n : \alpha, \varphi, i = 1, \dots, m, \quad (6.48)$$

where P_{ui} is the power at the i th output port of the power splitter. The result implies that the input noise is preserved (except for the additional thermal noise) in the power splitter output.

6.2.3 RF Amplifier

An RF amplifier amplifies both the carrier RF signal and the amplitude and phase noise. It also introduces additional noise, which may come from internal noise sources or the external environment and could be additive or parametric. The internal noise, including both white noise and flicker noise, is essential to determine the performance of the amplifier.

The PSD of the added amplitude and phase noise by the amplifier (Boudot and Rubiola 2012) in terms of white noise and flicker noise can be expressed as

$$S_{n,amp}(f_m) = b_0 + b_{-1}/f_m, \text{ with } n : \alpha, \varphi. \quad (6.49)$$

The white amplitude or phase noise b_0 is a wide-band additive noise. It is typically calculated with the noise factor F . From (6.28), the total white noise PSD in the amplifier output is $S_{white} = FG_p S_{th}$. As we know, half of S_{white} contributes to the phase noise and another half to the amplitude noise. They add to both sidebands of the carrier frequency. Therefore, the full S_{white} normalized by the amplifier output power $P_{s,out} = G_p P_{s,in}$, should be added in the DSB noise PSDs. Then we obtain

$$b_0 = S_{white}/P_{s,out} = FS_{th}/P_{s,in}, \quad (6.50)$$

where $P_{s,in}$ is the amplifier input power. The term b_0 contributes to the wideband noise floor of the amplitude and phase noise. To reduce b_0 , we should choose an amplifier with a small noise factor and increase the input RF power. The flicker amplitude or phase noise can be modeled as a $1/f$ noise existing near the carrier frequency. Its PSD follows a distribution like b_{-1}/f_m . The $1/f$ noise comes from the low frequency (near DC) flicker noise in the electronics modulated onto the carrier frequency. Since the $1/f$ noise power is proportional to the carrier signal power, b_{-1} is independent of $P_{s,in}$. We typically determine b_{-1} experimentally. Figure 6.14 depicts a typical spectrum of the added phase noise by an amplifier. If we change the input RF power, b_{-1} keeps unchanged but b_0 changes accordingly. The corner frequency f_c , at which $b_{-1}/f_c = b_0$, is given by $f_c = b_{-1}P_{s,in}/(FS_{th})$.

The overall noise PSDs in the amplifier output can be written as

$$S_{n,out}(f_m) = \frac{G_p(f_m)}{G_p(0)} S_{n,in}(f_m) + S_{n,amp}(f_m), \text{ with } n : \alpha, \varphi, \quad (6.51)$$

where $G_p(f_m)$ is the frequency-dependent power gain of the amplifier, and $S_{n,out}$ and $S_{n,in}$ denote the amplitude or phase noise in the output and input, respectively. Here we have assumed that the noise in the input RF signal and the noise added by the amplifier are uncorrelated. Note that if the amplifier operates closed-to-saturation, the input amplitude noise will be compressed at the output. Multiple amplifiers may be connected in serial or in parallel, see Fig. 6.15.

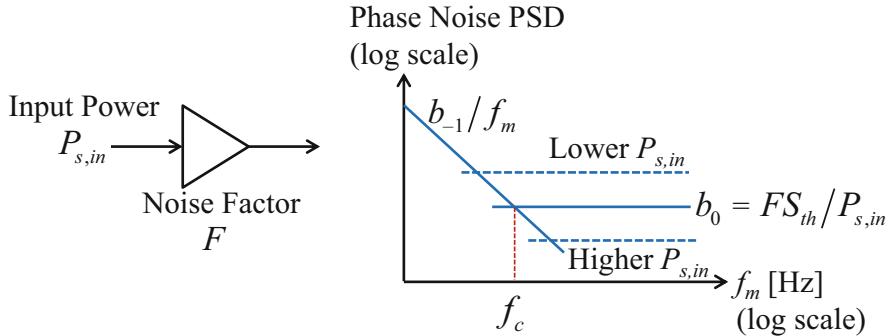


Fig. 6.14 Added noise PSD of the amplifier

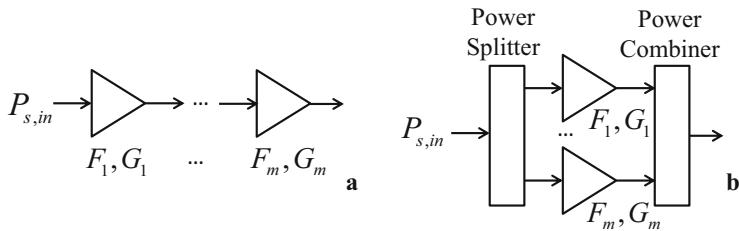


Fig. 6.15 (a). Cascaded m -stage amplifiers; (b). parallel network of m amplifiers.

The noise factor of a series of cascaded amplifiers is given by (6.30). If the gain of the first stage is large, the overall noise factor will be dominated by the first amplifier. Then the total added white noise at the output of the serial chain is

$$b_{0,serial} = \left(F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_m - 1}{G_1 G_2 \dots G_{m-1}} \right) \frac{S_{th}}{P_{s,in}}, \quad (6.52)$$

where F_i and G_i are the noise factor and power gain of the i th amplifier, respectively. The overall added $1/f$ noise can be directly calculated as

$$b_{-1,serial} = \sum_{i=1}^m b_{-1,i}, \quad (6.53)$$

where $b_{-1,i}$ is the added $1/f$ noise by the i th amplifier. Since $b_{-1,i}$ is parametric noise and independent of the signal power, $b_{-1,serial}$ does not depend on the gains of the amplifiers. Equation (6.53) implies that if we cascade two equal amplifiers, the flicker noise will be 3 dB higher, whereas (6.52) indicates that the white noise will be only slightly increased if G_1 is large. Connecting multiple amplifiers in parallel is often used when a single amplifier cannot provide the required RF power. For example, the RF system of a storage ring requires hundreds of kW power that can

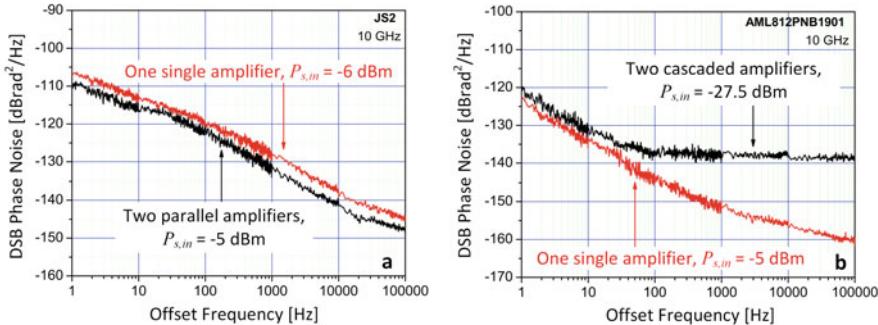


Fig. 6.16 The total phase noise of two amplifiers connected in parallel (a) and serial (b) (Courtesy of E. Rubiola and R. Boudot with CC BY license). For the serial case, the input power is reduced to avoid saturating the amplifier, which increases the noise floor significantly, as implied by (6.50)

be produced by a parallel network of many solid-state amplifiers. To simplify the discussion, we assume that the m amplifiers in Fig. 6.15b have the same gains (i.e., $G_i = G$), noise factors (i.e., $F_i = F$) and flicker noise PSDs (i.e., $b_{-1,i} = b_{-1}$); and that the power splitter and combiner are lossless and have the same power ratios and phases in the m branches.

We first analyze a single amplifier in the parallel network. Its input power is $P_{s,in}/m$ and the added white noise PSD is $b'_0 = mF_{Sth}/P_{s,in}$ according to (6.50). From the power combiner's noise model (6.47), we can calculate the output noise of the parallel amplifier network as

$$\begin{aligned} b_{0,parallel} &= b'_0/m = FS_{th}/P_{s,in}, \\ b_{-1,parallel} &= b_{-1}/m. \end{aligned} \quad (6.54)$$

Note that b'_0 includes the thermal noise from the matched resistive impedance of the amplifier input and output ports (consider the noise factor definition). The expression of $b_{-1,parallel}$ shows that the parallel amplifier network features a flicker-noise reduction of a factor m , assuming perfect symmetry and no dissipative losses in the splitter/combiner network. As an example, Fig. 6.16 depicts the measured phase noise of two amplifiers connected in parallel and serial, respectively.

6.2.4 Mixer

We have introduced the mixer in Sect. 5.1.1. With noisy inputs x_{IN} and x_{LO} given by

$$\begin{aligned} x_{IN}(t) &= A_{IN}[1 + \alpha_{IN}(t)] \sin(2\pi f_{IN}t + \Delta\varphi_{IN}(t)), \\ x_{LO}(t) &= A_{LO}[1 + \alpha_{LO}(t)] \sin(2\pi f_{LO}t + \Delta\varphi_{LO}(t)), \end{aligned} \quad (6.55)$$

the mixer output can be calculated as

$$x_{OUT}(t) \approx \frac{A_{IN}A_{LO}}{2}[1 + \alpha_{IN}(t) + \alpha_{LO}(t)] \\ (\underbrace{\cos[2\pi(f_{IN} - f_{LO})t + \Delta\varphi_{IN}(t) - \Delta\varphi_{LO}(t)]}_{\text{lower sideband}}) \\ - \underbrace{\cos[2\pi(f_{IN} + f_{LO})t + \Delta\varphi_{IN}(t) + \Delta\varphi_{LO}(t)]}_{\text{upper sideband}}. \quad (6.56)$$

For down-conversion, the input x_{IN} is the RF signal x_{RF} and the lower sideband (i.e., the IF signal x_{IF}) is selected, whereas for up-conversion, $x_{IN} = x_{IF}$, and we chose the upper sideband (i.e., x_{RF}) as the output. Equation (6.56) implies that the two input amplitude variations sum up ($\alpha_{IN} + \alpha_{LO}$) and are transferred to the output. The phase noise transfer relation depends on whether the mixer is used as a frequency down-converter ($\Delta\varphi_{IN} - \Delta\varphi_{LO}$) or up-converter ($\Delta\varphi_{IN} + \Delta\varphi_{LO}$). Since the amplitude or phase variations of x_{IN} and x_{LO} might be (partly) correlated, we derive the noise PSDs at the mixer output using (6.25), including the noise added by the mixer:

$$S_{\alpha,OUT}(f_m) = S_{\alpha,IN}(f_m) + S_{\alpha,LO}(f_m) + 2S_{\alpha,IN \times LO}(f_m) + S_{\alpha,mix}(f_m), \\ S_{\varphi,OUT}(f_m) = S_{\varphi,IN}(f_m) + S_{\varphi,LO}(f_m) \pm 2S_{\varphi,IN \times LO}(f_m) + S_{\varphi,mix}(f_m), \quad (6.57)$$

where $S_{\alpha,OUT}$, $S_{\alpha,IN}$ and $S_{\alpha,LO}$ are the amplitude noise PSDs of the output, input and LO signals, respectively; $S_{\alpha,mix}$ is the amplitude noise added by the mixer, and $S_{\alpha,IN \times LO}$ is the CPSD of the α_{IN} and α_{LO} . The corresponding notations with the subscript φ represent the phase noise PSD or CPSD. In the second equation, the positive $2S_{\varphi,IN \times LO}$ term is for up-conversion and the negative case is for down-conversion. The noise added by the mixer is dominated by the additive white noise, which can be calculated with the single-sideband IEEE noise factor F as:

$$S_{\alpha,mix}(f_m) = S_{\varphi,mix}(f_m) = FS_{th}/P_{s,in}, \quad (6.58)$$

where $P_{s,in}$ is the power of x_{IN} . For a noiseless mixer (e.g., a passive mixer), F equals its conversion loss (in linear units, e.g., 3 dB conversion loss corresponds to $F = 2$). Nowadays, the noise added by (active) mixers is typically small and can often be neglected in the noise analysis. For the results (6.57), we have the following comments:

- (a) If the LO port of the mixer operates in saturation, the amplitude jitter in the LO signal will be suppressed, so the terms $S_{\alpha,LO}$ and $S_{\alpha,IN \times LO}$ can be neglected. We focus more on the phase noise in a mixer.
- (b) In LLRF systems, $\Delta\varphi_{IN}$ and $\Delta\varphi_{LO}$ are often correlated (e.g., x_{IN} and x_{LO} are from the same reference signal x_{REF}). In practice, we deal with the correlation differently instead of using CPSD because it is difficult to measure the CPSD of the phase noise in two signals of different frequencies. Let us use the down-conversion (i.e., $x_{IN} = x_{RF}$) as an example. If $\Delta\varphi_{RF}$ and $\Delta\varphi_{LO}$ are partly correlated, we can split them into correlated and uncorrelated parts, such as

$\Delta\varphi_{RF} = \Delta\varphi_{REF} + \Delta\varphi_{RFadd}$ and $\Delta\varphi_{LO} = k\Delta\varphi_{REF} + \Delta\varphi_{LOadd}$. Here $\Delta\varphi_{REF}$ is the phase noise in the common reference signal x_{REF} , $k = f_{LO}/f_{REF}$ is the ratio between the LO frequency and the reference frequency, and $\Delta\varphi_{RFadd}$ and $\Delta\varphi_{LOadd}$ are the uncorrelated phase variations in x_{RF} and x_{LO} . Then the phase noise in the mixer output IF signal can be calculated as $\Delta\varphi_{IF} = \Delta\varphi_{RF} - \Delta\varphi_{LO} = (1 - k)\Delta\varphi_{REF} + \Delta\varphi_{RFadd} - \Delta\varphi_{LOadd}$. The three terms in the result are uncorrelated, so we obtain a more compact equation of PSDs as

$$S_{\varphi,IF} = (1 - k)^2 S_{\varphi,REF} + S_{\varphi,RFadd} + S_{\varphi,LOadd} + S_{\varphi,mix}. \quad (6.59)$$

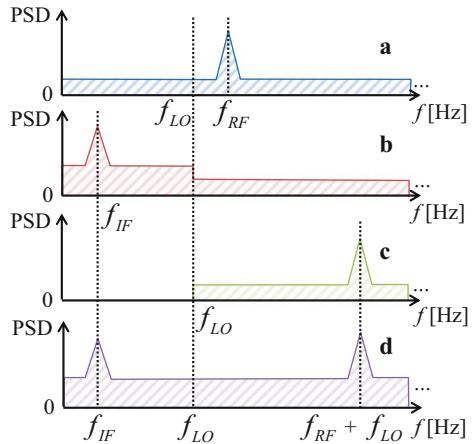
Note that $\Delta\varphi_{REF}$ may transfer to the RF and LO signals in different bandwidths. Equation (6.59) is valid only in the frequency range where $\Delta\varphi_{REF}$ transfers to both $\Delta\varphi_{RF}$ and $\Delta\varphi_{LO}$. At the frequencies where $\Delta\varphi_{REF}$ only appears in $\Delta\varphi_{RF}$ (or $\Delta\varphi_{LO}$), the RF reference phase noise term in (6.59) becomes $S_{\varphi,REF}$ (or $k^2 S_{\varphi,REF}$). At frequencies above both transfer bandwidths, the $S_{\varphi,REF}$ term can be neglected. In practical LLRF systems, we will directly use (6.59) because the noise transfer bandwidths from $\Delta\varphi_{REF}$ to $\Delta\varphi_{RF}$ and $\Delta\varphi_{LO}$ are typically much larger than the closed-loop bandwidth of the cavity.

- (c) When deriving (6.57), we have assumed that both x_{IN} and x_{LO} have been filtered by a perfect band-pass filter with a half-bandwidth smaller than f_{IF} . Without such band-limits, the noise floors of both sidebands in the mixer output (6.56) will overlap (Rubiola 2006), resulting in 3 dB higher in the noise floor of the signals at $f_{IN} - f_{LO}$ and $f_{IN} + f_{LO}$. This situation is explained in Fig. 6.17 using the phase noise in the down-conversion case as an example. To simplify the discussion, we assume the LO signal is noiseless. The RF signal phase noise (plot a) is down-converted to $f_{IF} = f_{RF} - f_{LO}$ (plot b) and up-converted to $f_{RF} + f_{LO}$ (plot c). In the lower sideband of the mixer output (plot b), the RF signal noise power between DC and f_{LO} is imaged to the positive frequency and doubles the noise floor below f_{LO} . In the higher sideband (plot c), the RF signal noise does not contribute to the frequency range between DC and f_{LO} . When there are no filters, the mixer output contains both sidebands and the overall noise follows plot d. In this case, the noise floor in the plot d is 3 dB higher than the RF signal noise floor in the plot a.

6.2.5 Frequency Divider and Multiplier

Frequency dividers and multipliers are widely used in frequency synthesizers. An ideal frequency divider (or multiplier) divides (or multiplies) the instantaneous frequency of the input RF signal. The amplitude noise of a frequency divider or multiplier is typically not relevant in most applications.

Fig. 6.17 Overlap of the noise floors of the lower and upper sidebands. (a). RF signal noise (LO signal is assumed noiseless); (b). lower-sideband noise; (c). higher-sideband noise; (d). overall noise in mixer output



A sinusoidal signal x with phase noise input to a frequency divider/multiplier produces an output y . The input and output can be written as

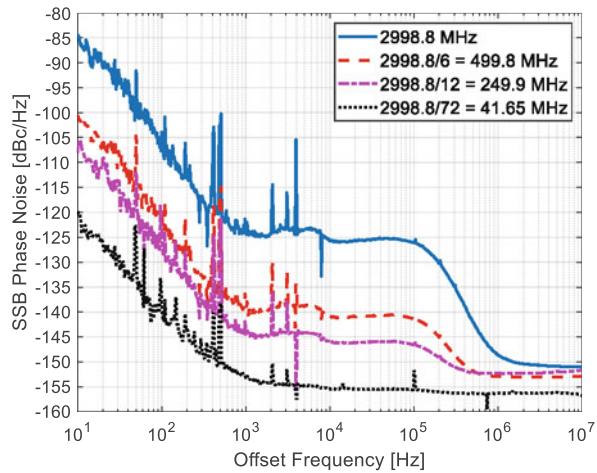
$$\begin{aligned} x(t) &= A \sin(2\pi f_c t + \Delta\varphi(t)), \\ y(t) &= A \sin(\gamma(2\pi f_c t + \Delta\varphi(t))), \end{aligned} \quad (6.60)$$

where $\gamma = 1/N$ or $\gamma = N$ ($N > 1$ is an integer) for a frequency divider or multiplier, respectively. Then the phase noise PSD of the output can be written as

$$S_{\varphi,out}(f_m) = \gamma^2 S_{\varphi,in}(f_m) + S_{\varphi,freq}(f_m), \quad (6.61)$$

where $S_{\varphi,in}$ and $S_{\varphi,out}$ are the phase noise PSDs of the input and output, and $S_{\varphi,freq}$ is the added phase noise (additive white noise) by the frequency divider/multiplier. The frequency divider/multiplier does not change the timing jitter of the input RF signal, so the input phase jitter will be scaled by the ratio between the output and input frequencies. Therefore, a frequency multiplier scales the input phase noise PSD by a factor of N^2 at all frequencies. In practice, we should avoid using frequency multipliers when N is greater than a few. Instead, a phase-locked loop can be used to multiply the input frequency (see Sect. 8.2.2). In contrast, a frequency divider reduces the phase noise PSD to $1/N^2$. When N is large, the noise floor of the frequency divider output will be dominated by $S_{\varphi,freq}$. For example, Fig. 6.18 shows the phase noise of a 2998.8 MHz input signal and that of the outputs of three frequency dividers ($N = 6, 2$, and 6) connected in series. The close-to-carrier phase noise is much larger than $S_{\varphi,freq}$ and follows the $1/N^2$ policy (i.e., 15.56 dB lower by 1/6, and 6 dB lower by 1/2 frequency dividers, respectively). However, the noise floors of the frequency dividers' outputs do not follow the $1/N^2$ policy due to the phase noise of the frequency dividers.

Fig. 6.18 Phase noise of the RF signals at the input and outputs of frequency dividers. (Courtesy of A. Dietrich)



6.2.6 Analog-to-Digital Converter

6.2.6.1 ADC Noise Model

The noise of an ADC is critical for precise RF detection. Nowadays, low-noise ADCs used in RF detectors have resolutions up to 16 bits and sampling rates up to several hundred MSPS (Mega-Samples Per Second). The effective number of bits (ENOB) can be up to 12–14 bits, corresponding to an SNR of 70–80 dB, and the analog input bandwidths can be up to more than 1 GHz.

An ADC accepts an analog input and a clock and generates a sequence of samples discretized in both time and values. We need to work out an ADC noise model describing how the noise of the analog input and clock transfers to the samples (Neu 2017). The noise added by the ADC itself is another aspect of the noise model.

In an RF detector, the IF signal x_{IF} is sampled by an ADC at each zero-crossing time of the clock x_{CLK} . To simplify the discussion, we assume that x_{CLK} is a sinusoidal signal and the IF signal is sampled at the clock's rising edges. Both signals can be noisy described as

$$\begin{aligned} x_{IF}(t) &= A_{IF}[1 + \alpha_{IF}(t)] \sin(2\pi f_{IF}t + \Delta\varphi_{IF}(t)), \\ x_{CLK}(t) &= A_{CLK} \sin(2\pi f_{CLK}t + \Delta\varphi_{CLK}(t)), \end{aligned} \quad (6.62)$$

where α_{IF} and $\Delta\varphi_{IF}$ are the amplitude and phase variations of the IF signal. For the clock, we only consider its phase variation $\Delta\varphi_{CLK}$. The rising-edge zero-crossing time of the clock satisfies $2\pi f_{CLK}t_k + \Delta\varphi_{CLK}(t_k) = 2k\pi$, $k = 0, 1, 2, \dots$, which yields $t_k = [2k\pi - \Delta\varphi_{CLK}(t_k)]/(2\pi f_{CLK})$. Then the IF signal values at around $t = kT_s$ (where $T_s = 1/f_{CLK}$) can be written as

$$\begin{aligned} x_{IF}(k) := & x_{IF}(T_k) = A_{IF}[1 + \alpha_{IF}(t_k)] \\ & \sin\left(2\pi f_{IF} k T_s - \frac{f_{IF}}{f_{CLK}} \Delta\varphi_{CLK}(t_k) + \Delta\varphi_{IF}(t_k)\right). \end{aligned} \quad (6.63)$$

Therefore, the amplitude and phase variations in the ADC output (i.e., the sequence of ADC samples) caused by the noise in the IF signal and the clock can be written as

$$\begin{aligned} \alpha_{out0}(k) &= \alpha_{IF}(t_k), \\ \Delta\varphi_{out0}(k) &= \Delta\varphi_{IF}(t_k) - \frac{f_{IF}}{f_{CLK}} \Delta\varphi_{CLK}(t_k), \quad k = 0, 1, 2, \dots \end{aligned} \quad (6.64)$$

To derive the relationship between the noise PSDs of the ADC inputs (IF and clock) and output (ADC samples), we assume (6.64) is also valid in continuous time (replace t_k with t). Then for an equivalent “continuous” ADC output, we have

$$\begin{aligned} S_{\alpha,out}(f_m) &= S_{\alpha,IF}(f_m) + S_{\alpha,adc}(f_m), \\ S_{\varphi,out}(f_m) &= S_{\varphi,IF}(f_m) + \left(\frac{f_{IF}}{f_{CLK}}\right)^2 S_{\varphi,CLK}(f_m) \\ &\quad - 2 \frac{f_{IF}}{f_{CLK}} S_{\varphi,IF \times CLK}(f_m) + S_{\varphi,adc}(f_m), \end{aligned} \quad (6.65)$$

where $S_{\alpha,IF}$ and $S_{\varphi,IF}$ are the DSB amplitude and phase noise PSDs of the IF signal, $S_{\varphi,CLK}$ is the phase noise PSD of the clock, and $S_{\alpha,adc}$ and $S_{\varphi,adc}$ are added by the ADC itself (see next section). The term $S_{\varphi,IF \times CLK}$ is the CPSD between $\Delta\varphi_{IF}$ and $\Delta\varphi_{CLK}$ modeling the correlations between them. In practice, we will still follow the method introduced in the comment (b) of Sect. 6.2.4 to deal with the correlations. Note that $S_{\alpha,out}$ and $S_{\varphi,out}$ in (6.65) are defined within the analog bandwidths of the IF and clock signals. The noise of the discrete ADC output is determined by $S_{\alpha,out}$ and $S_{\varphi,out}$ after aliased to the Nyquist zone (i.e., from 0 Hz to $f_{Nyquist} = f_{CLK}/2$). To determine the aliased frequency of the noise power at the offset frequency f_m , we convert $S_{\alpha,out}(f_m)$ and $S_{\varphi,out}(f_m)$ to both sidebands of f_{IF} . Each sideband obtains $S_{\alpha,out}(f_m)/2$ and $S_{\varphi,out}(f_m)/2$ at absolute frequencies of $f_{IF} \pm f_m$. Then we use the absolute frequency to determine the aliased frequency in the Nyquist zone following the rule given in Sect. 5.2.1.2.

Figure 6.19 illustrates how the phase noise of the IF and clock signals transfers to the ADC samples (Brannon 2004). We define the analog bandwidths of the IF and clock input circuits to be $f_{BW,IF}$ and $f_{BW,CLK}$, respectively. They are often much larger than $f_{Nyquist}$. If the IF or clock frequency is not at the center of its input bandwidth, we use the method in Fig. 6.20 to derive its equivalent DSB PSD within half of the input bandwidth (Lapointe 2016). To illustrate the frequency aliasing, we included a spurious signal (denoted as “spur” later) in the clock phase noise. The phase noise of the equivalent “continuous” ADC output is shown in Fig. 6.19c, which is the sum of the $S_{\varphi,IF}$ in plot a and the scaled $S_{\varphi,CLK}$ in plot b, assuming they are uncorrelated.

Fig. 6.19 Aliasing of the phase noise in the analog IF input and clock to the ADC samples

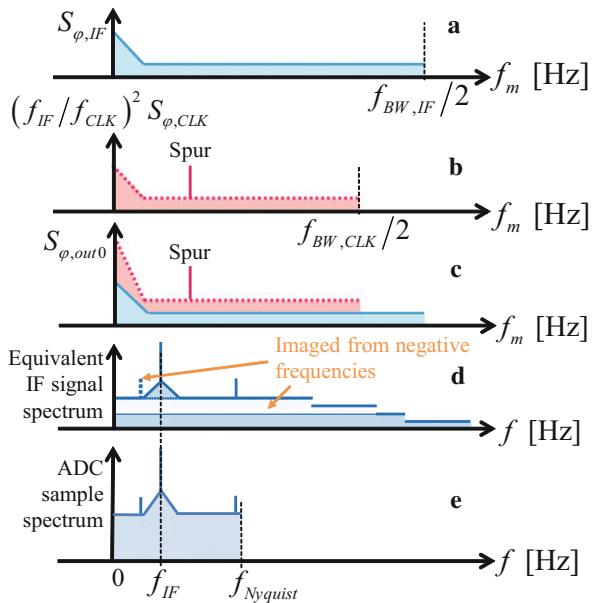
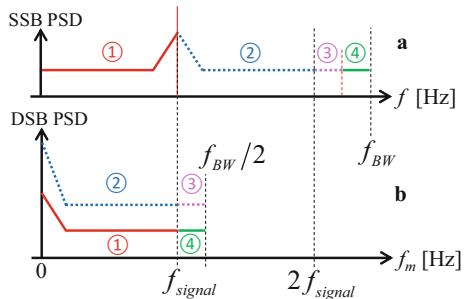


Fig. 6.20 Equivalent DSB PSD derived from the asymmetric SSB PSD in the input bandwidth (f_{BW}) of the IF or clock signal. The DSB PSD is then defined between 0 and $f_{BW}/2$



We denote it as $S_{\phi,out0}$, which does not include the ADC added phase noise compared to the $S_{\phi,out}$ in (6.65). Note that when adding the noise PSDs, they should use linear units instead of dB rad^2/Hz . Figure 6.19d has converted $S_{\phi,out0}$ to the two sidebands of f_{IF} at absolute frequencies, where the dotted part of the spectrum represents the image of the negative frequencies. Here the spur at the lower sideband falling into a negative frequency has been imaged back to a positive frequency. Note that when converting the DSB noise to the two sidebands of f_{IF} , each sideband contains half of the noise power. The discretizing process aliases the plot d spectrum into the Nyquist zone as in plot e. In this example, if we demodulate the IF signal samples and calculate the phase, we will obtain ripples at two different frequencies that are caused by the same clock spur via frequency aliasing.

The analysis above defines a procedure for analyzing the ADC noise:

1. Consider all ADC noise sources, including the noise in the external inputs and the noise added by the ADC internally, then calculate the “continuous” output noise with (6.65). The results are DSB amplitude and phase noise PSDs defined at the offset frequencies with respect to a carrier frequency. Note that the magnitudes of the PSDs are relative to the carrier signal power. See Figs. 6.19a to 6.19c.
2. Modulate the DSB noise PSDs obtained from the last step to both sidebands of the IF frequency, which then results in SSB PSDs defined over absolute frequencies. In this step, the negative-frequency noise power should be imaged and added to the positive frequency. See Fig. 6.19d.
3. Alias the SSB noise PSDs calculated at the last step into the Nyquist zone. The results are the SSB noise PSDs of the ADC samples (the ADC output) defined over absolute frequencies between 0 Hz and $f_{Nyquist}$. See Fig. 6.19e. When the IF signal samples are demodulated, the aliased noise will be processed by the filter defined by the demodulation algorithm (e.g., non-I/Q demodulation) and contributes to the amplitude and phase measurement errors.

In practice, we often filter the IF and clock inputs with band-pass filters (BPFs) to avoid aliasing wideband noise into the Nyquist zone. The half-passband (i.e., 1/2 of the full passband) of the BPF is typically smaller than the minimum value of f_{IF} and $f_{Nyquist} - f_{IF}$, which can avoid both the frequency imaging in Fig. 6.19d and the frequency aliasing in Fig. 6.19e. Note that the BPFs can only be applied to the ADC external inputs. The noise added by the ADC ($S_{\alpha,adc}$ and $S_{\varphi,adc}$) will be coupled into the ADC samples within the bandwidths of the analog and clock input circuits.

6.2.6.2 Noise Added by ADC

An ADC introduces additional noise to both the clock and the analog inputs. The typical noise added by an ADC is summarized as follows:

- *Parametric noise*: such as the clock aperture jitter, the clock spurious frequencies and the analog input spurious frequencies.
- *Additive noise*: such as the analog input thermal noise and the quantization noise.

The procedure introduced in Sect. 6.2.6.1 can also be used to analyze the noise added by the ADC. However, since it is hard to measure the wideband noise added by the ADC directly, we will make the analysis in a different way. The final goal is to estimate the DSB amplitude and phase noise PSDs introduced by the ADC, i.e., the terms $S_{\alpha,adc}(f_m)$ and $S_{\varphi,adc}(f_m)$ in (6.65). First, we obtain the SSB amplitude and phase noise PSDs added by the ADC within the Nyquist zone ($0 \leq f \leq f_{Nyquist}$), which are denoted as $D_{\alpha}(f)$ and $D_{\varphi}(f)$, respectively. The noise added by the ADC can either be estimated with the information given by the ADC datasheet or be measured indirectly using the method introduced in Sect. 6.2.6.3. Then we can convert the obtained SSB noise PSDs to DSB with respect to the IF frequency.

The noise added by the ADC to the clock includes the aperture jitter and the spurious frequencies (spur). As implied by (6.63), they introduce additional phase modulation to the IF signal and affect the phase noise of the ADC samples.

The aperture jitter is caused by the additive white noise in the clock input circuit. It can be reduced by increasing the clock slew rate (Neu 2010). This requires either the sinusoidal clock to have a large magnitude or a clock with a sharp rising edge (e.g., a square wave). To enable a square wave, the clock input circuit typically has a large bandwidth to allow for several higher-order harmonics of f_{CLK} . Then more clock noise will be passed to the ADC samples (see Fig. 6.19b). We should find a compromise for the clock input circuit design as a tradeoff between the slew rate and bandwidth requirements. The ADC datasheets often specify the aperture jitter as a constant value $\Delta t_{apt,rms}$, corresponding to a white phase noise added to the clock. The clock aperture jitter will add to the IF signal samples a white phase noise, which has a constant PSD in the Nyquist zone, calculated as

$$D_{\varphi,apt} = \left(\frac{f_{IF}}{f_{CLK}} \right)^2 \frac{(2\pi f_{CLK} \Delta t_{apt,rms})^2}{f_{Nyquist}} = \frac{(2\pi f_{IF} \Delta t_{apt,rms})^2}{f_{Nyquist}}. \quad (6.66)$$

The spurs added to the clock come from the sources generating periodic noise, such as the DC-DC converters in the power supplies. A clock spur at an offset frequency $f_{m,spur}$ will be modulated onto the IF signal to the frequencies $f_{IF} \pm f_{m,spur}$. They will be aliased to the Nyquist zone in the spectrum of the ADC samples. We use f_{spa} and f_{spb} to denote their aliased frequencies. Therefore, the SSB phase noise introduced by a clock spur (Brannon 2004) to the IF samples can be written as

$$D_{\varphi,clkspur}(f) = \left(\frac{f_{IF}}{f_{CLK}} \right)^2 \frac{P_{spur}}{2P_{CLK}} [\delta(f - f_{spa}) + \delta(f - f_{spb})], \quad (6.67)$$

where P_{spur}/P_{CLK} is the relative power of the spur in the clock's DSB phase noise PSD. Note that P_{spur} should be the total power of the spurious-frequency signal. If the spur is additive, it appears only in one sideband of the clock frequency. In contrast, if the spur is modulated onto the clock signal, it appears in both sidebands of the clock frequency, and in this case, P_{spur} should include the powers from both sidebands. The delta-functions in eq. (6.67) indicate that the spurs only exist at $f = f_{spa}$ and $f = f_{spb}$ in the ADC samples. The noise added by the ADC to the analog input includes the thermal noise, the quantization noise and also spurs.

The thermal noise is additive white noise existing in the ADC analog input circuits (e.g., the termination resistance, buffer amplifiers and sampling switch). The quantization noise comes from the discretization of an analog signal with a limited vertical resolution. For an N -bit ADC, the SNR caused by the quantization noise is approximately $SNR_{qu} = 6.02 N + 1.76$ dB, which is an upper limit of the achievable SNR. For example, for a 16-bit ADC, $SNR_{qu} = 98.1$ dB. The quantization noise is typically much smaller than other noise sources and we often neglect it. In the ADC samples, the thermal noise yields a white noise term D_{th} distributed in

the Nyquist zone, contributing equally to both the amplitude noise ($D_{\alpha,th}$) and phase noise ($D_{\varphi,th}$):

$$D_{\alpha,th} = D_{\varphi,th} = \frac{D_{th}}{2P_{IF,out}}, \quad (6.68)$$

where D_{th} should be calculated with the samples of the thermal noise, and $P_{IF,out}$ is the IF signal power in the ADC samples. The spurs in the analog input circuit can be either additive (e.g., caused by external radiations) or parametric (e.g., caused by power supplies). An additive spur contributes half of its power to the amplitude noise and another half to the phase noise of the ADC samples. In contrast, a parametric spur may affect the amplitude or phase noise or both. In some cases, we also view the higher-order harmonics of the IF frequency as spurs. However, if we adopt the non-I/Q demodulation algorithm (see Chap. 5), the IF harmonics up to a particular order can be filtered. A parametric spur generates spectral lines symmetric to the IF frequency, whereas an additive spur can exist at any single frequency. In the ADC samples, the spurs will be aliased to the Nyquist zone.

Considering the noise sources discussed above, we can estimate the PSDs of the amplitude and phase noise added by the ADC in (6.65) as

$$\begin{aligned} S_{\alpha,adc}(f_m) &= \zeta(f_m)D_{\alpha,th} + \sum_i D_{\alpha,ifspur,i}, \\ S_{\varphi,adc}(f_m) &= \zeta(f_m)(D_{\varphi,th} + D_{\varphi,apt}) + \sum_j D_{\varphi,clkspur,j} + \sum_k D_{\varphi,ifspur,k}. \end{aligned} \quad (6.69)$$

Here f_m should satisfy $0 < f_m \leq \max(f_{IF}, f_{Nyquist} - f_{IF})$. The values of $S_{\alpha,adc}$ and $S_{\varphi,adc}$ are zero beyond this frequency range. The factor ζ is a function of f_m defined as

$$\zeta(f_m) = \begin{cases} 2, & 0 < f_m \leq \min(f_{IF}, f_{Nyquist} - f_{IF}) \\ 1, & \min(f_{IF}, f_{Nyquist} - f_{IF}) < f_m \leq \max(f_{IF}, f_{Nyquist} - f_{IF}) \end{cases}.$$

The functions $\min(x_1, x_2)$ and $\max(x_1, x_2)$ return the minimum and maximum of the two inputs x_1 and x_2 , respectively. In (6.69), $S_{\alpha,adc}$ and $S_{\varphi,adc}$ are DSB PSDs, therefore, we double the SSB white noise PSDs, $D_{\alpha,th}$, $D_{\varphi,th}$ and $D_{\varphi,apt}$, at the offset frequencies contributed from both sidebands of the IF frequency. The ADC sample spurs (6.67), which is induced by a clock spur, are mapped to the offset frequencies $f_{m,spa} = |f_{IF} - f_{spa}|$ and $f_{m,spb} = |f_{IF} - f_{spb}|$. In (6.69), we also included the spurs added by the ADC analog input circuit as $D_{\alpha,ifspur}$ and $D_{\varphi,ifspur}$. Their frequencies have been converted to offset frequencies relative to the IF frequency.

The wideband noise added by the ADC (thermal noise and clock aperture jitter) will be aliased to the Nyquist zone, adding to the IF signal samples. Therefore, if we increase the $f_{Nyquist}$, the spectral density of the noise added by the ADC in the Nyquist zone will be smaller (consider pouring the same amount of water in two bathtubs with different sizes). This is the so-called *oversampling* (Candy and Temes

1991), corresponding to sampling more points in the same period and doing average. Furthermore, a higher f_{CLK} implies that the phase noise of the ADC samples is less sensitive to the clock jitter, as depicted in (6.65).

6.2.6.3 Measurement of Noise Added by ADC

The noise added by the ADC can also be evaluated by means of measurements with the setups in Fig. 6.21 (Cárdenas-Olaya et al. 2017).

If we terminate the ADC analog input with a matched load (e.g., $50\ \Omega$ termination, see Fig. 6.21a), the ADC samples will be a direct measurement of the thermal noise. When evaluating the ADC thermal noise, we often calculate its relative power with respect to a virtual sinusoidal signal, whose amplitude equals the ADC output's full scale (FS). For example, for a 16-bit ADC, the full-scale amplitude is $A_{FS} = 2^{15}$. To calculate the relative noise power, we divide the ADC samples of the thermal noise by $A_{FS}/\sqrt{2}$ before doing DFT. The unit of the relative PSD to a full-scale signal is dBFS/Hz. Figure 6.22 shows the measurement of the thermal noise of a 16-bit ADC with $f_{CLK} = 249.9$ MHz. The noise floor corresponds to the item $D_{th}/P_{IF,out}$ in (6.68), from which we can estimate the values of $D_{\alpha,th}$ and $D_{\varphi,th}$ (they are equal and 3 dB lower than the noise floor in Fig. 6.22). We can see several spurs that are either additive or parametric. If an IF signal is sampled by the ADC, the parametric spurs will modulate the amplitude or phase of the IF signal. The value of $D_{\alpha,th}$ gives an estimate of the ADC added amplitude noise $S_{\alpha,adc}$ according to (6.69). Note that we typically do not need to know the exact magnitudes of the resulting spurs in the ADC samples, but only be sure that they are small enough not to cause significant

Fig. 6.21 Setups to measure the ADC added thermal noise (a) and phase noise (b)

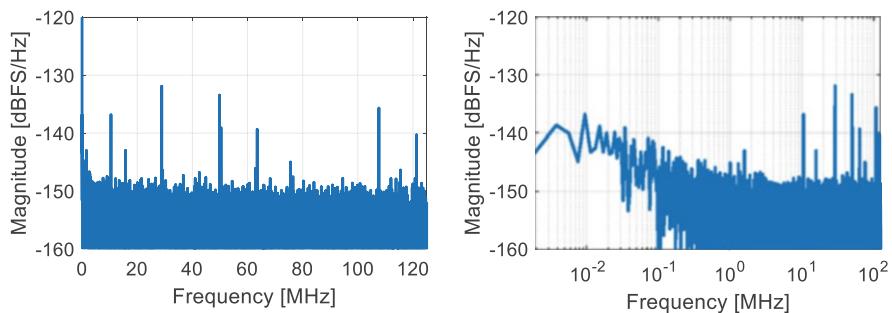
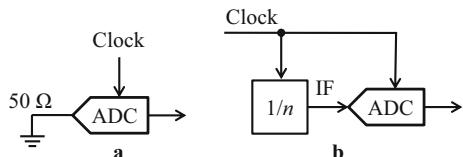


Fig. 6.22 ADC thermal noise measurement with linear (left) and logarithmic (right) frequency scales

amplitude or phase measurement errors. For example, the maximum PSD of the spurs in Fig. 6.22 is about -132 dBFS/Hz, from which we can calculate the power of the spur by integrating the PSD in the DFT bin spacing (1.907 kHz, see the explanation at the end of this section) as $-132 + 10\log_{10}(1.907e3) = -99.2$ dBFS. For an IF input signal with its amplitude close to the ADC full scale, the spur will cause an amplitude jitter about 1.1e-5 RMS or a phase jitter about 6.3e-4 degree RMS. For many LLRF systems, the measurement errors caused by this spur can often be neglected. There also exists flicker noise near 0 Hz, which is more clearly visible in the plot with logarithmic frequencies (see Fig. 6.22b). The flicker noise comes from the power supply of the ADC board and will modulate the IF input signal.

With the setup in Fig. 6.21b, we can estimate the total phase noise added by the ADC. If the IF signal is derived from the clock using a low-noise frequency divider, their phase noise will be canceled in the ADC samples. Assume the clock has a phase noise $\Delta\varphi_{CLK}$. If we neglect the added phase noise by the frequency divider, we will obtain the IF signal phase noise $\Delta\varphi_{IF} = \Delta\varphi_{CLK} / n$. Then (6.64) implies that in the ADC “continuous” output, the total phase noise caused by the clock and IF signal is $\Delta\varphi_{out0} = \Delta\varphi_{IF} - (f_{IF}/f_{CLK})\Delta\varphi_{CLK} = 0$. Therefore, the phase jitter in the ADC samples will be an estimate of the added phase noise by the ADC itself. Under this configuration, the amplitude and phase noise PSDs of the IF signal samples, as given by (6.65), are simplified to $S_{\alpha,out} = S_{\alpha,IF} + S_{\alpha,adc}$, $S_{\varphi,out} = S_{\varphi,adc}$, where $S_{\alpha,adc}$ and $S_{\varphi,adc}$ are described in (6.69). Figure 6.23 shows the measurement of the added phase noise of the same ADC. Here the IF frequency is $f_{IF} = f_{CLK}/6$. The results provide the following information:

- Compared to Fig. 6.22, the noise floor of Fig. 6.23a (with IF input) increases due to the clock aperture jitter ($D_{\varphi,apt}$) and the IF signal amplitude noise ($S_{\alpha,IF}$). The absolute noise power caused by these two terms is proportional to the input IF signal power. Therefore, when the IF signal is small, the effects of $D_{\varphi,apt}$ and $S_{\alpha,IF}$ are negligible and the thermal noise (with a constant absolute PSD) dominates the total noise power. Under this condition, increasing the IF signal power will improve the SNR. However, when the IF signal power is large and $D_{\varphi,apt}$ and $S_{\alpha,IF}$ are much larger than the thermal noise, the SNR will approach a constant and will not benefit from increasing the IF signal power.
- The harmonics of the IF signal is large because the IF signal is a square wave in this test.
- The spectrum of the IF signal samples (Fig. 6.23a) contains some spurs at the same frequencies as that in the thermal noise spectrum (Fig. 6.22). These should be additive spurs.
- We used the non-I/Q demodulation algorithm (5.15) to calculate the amplitude and phase of the IF signal samples. The noise PSDs in Fig. 6.23b, c are directly derived from the measured amplitude and phase, so they are DSB PSDs. The amplitude noise (Fig. 6.23b) contains contributions from both the IF signal amplitude noise and the ADC added amplitude noise. The phase noise (Fig. 6.23c) contains only the ADC added phase noise and is an estimate of

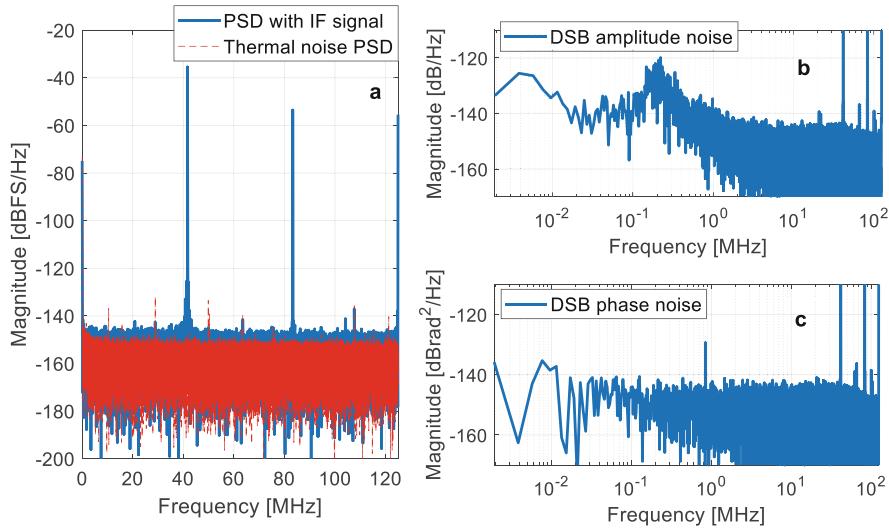


Fig. 6.23 ADC added phase noise measurement. (a). Overall noise in ADC samples (dBFS/Hz); (b). Relative amplitude noise in ADC samples (dB/Hz); (c). Phase noise in ADC samples ($dBrad^2/Hz$)

$S_{\varphi,adc}$, which is a superposition of the contributions of the thermal noise, aperture jitter, and various spurs coming from the clock or analog input circuits.

In Fig. 6.23, the powers of the IF signal, its harmonics and the spurs are all distributed in a DFT bin (here $f_{CLK} = 249.9$ MHz and the DFT was calculated with 2^{17} samples, so the bin spacing is $\Delta f_{bin} = 249.9e6 / 2^{17} = 1.907$ kHz). Therefore, the power of a sinusoidal signal should be calculated by integrating the PSD at the signal frequency within the DFT bin spacing. For example, in the plot of Fig. 6.23a, the power of the IF signal (at $f_{IF} = 41.65$ MHz) can be calculated as $P_{IF,out} = -35.50 + 10\log_{10}(1.907e3) = -2.7$ dBFS. The SNR of the IF signal samples in Fig. 6.23a is about 71.3 dB, with the harmonics and the DC component excluded from the noise power calculation.

6.2.7 Digital-to-Analog Converter

6.2.7.1 DAC Noise Model

In digital LLRF systems, DACs are widely used for controlling the RF amplitude and phase or driving the piezo tuner. In this section, we will discuss the DAC noise transfer model and the noise added in the DAC. The discussion in this section has referred to the articles (Analog Devices 1999; Kester 2009; Wang 2015).

A DAC accepts a digital input $x_D(k)$, a clock $x_{CLK}(t)$, and generates an analog output. In LLRF systems, x_D often contains noise because it may be a result of the feedback control based on noisy measurements. The DAC updates its analog output using x_D at each zero-crossing time of x_{CLK} . To simplify the discussion, we assume that x_{CLK} is a sinusoidal signal and the DAC output is updated at the clock's rising edges. Each analog change is held on for a clock cycle until the next value arrives. This is the so-called zero-order-hold (ZOH). Typically, a low-pass filter (LPF) is used at the DAC output to clean the signal and filter out the images and harmonics at higher frequencies.

The expression of x_{CLK} has been given in (6.62). The digital input is written as

$$x_D(k) = A_D[1 + \alpha_D(k)] \sin(2\pi f_D k T_s + \Delta\varphi_D(k)), k = 0, 1, 2, \dots \quad (6.70)$$

where A_D and f_D are its amplitude and frequency, respectively. The terms α_D and $\Delta\varphi_D$ are the amplitude and phase noise, and T_s is the period of the clock. The data point $x_D(k)$ will be output at $t_k = kT_s - \Delta\varphi_{CLK}(t_k)/(2\pi f_{CLK})$. Ideally, $x_D(k)$ should be output at the time kT_s , but due to the noisy clock, it is output with a time error of $t_k - kT_s$. The DAC output (excluding the noise added by the DAC) is written as

$$\begin{aligned} x_{out0}(t_k) &= x_D(k) \\ &= A_D[1 + \alpha_D(k)] \sin\left(2\pi f_D t_k + \frac{f_D}{f_{CLK}} \Delta\varphi_{CLK}(t_k) + \Delta\varphi_D(k)\right). \end{aligned} \quad (6.71)$$

The value of $x_{out0}(t_k)$ is held on from t_k to t_{k+1} by the ZOH. Then, the amplitude and phase variations in the DAC output caused by the digital input and clock are

$$\begin{aligned} \alpha_{out0}(t_k) &= \alpha_D(k), \\ \Delta\varphi_{out0}(t_k) &= \Delta\varphi_D(k) + \frac{f_D}{f_{CLK}} \Delta\varphi_{CLK}(t_k), k = 0, 1, 2, \dots \end{aligned} \quad (6.72)$$

Similar to (6.65), we can derive the relationship of the DSB noise PSDs when there is no frequency aliasing of the noise. We assume that all items in (6.72) are continuous-time variables (replace t_k and k by t) and then obtain

$$\begin{aligned} S_{\alpha,out}(f_m) &= S_{\alpha,D}(f_m) + S_{\alpha,dac}(f_m), \\ S_{\varphi,out}(f_m) &= S_{\varphi,D}(f_m) + \left(\frac{f_D}{f_{CLK}}\right)^2 S_{\varphi,CLK}(f_m) \\ &\quad + 2 \frac{f_D}{f_{CLK}} S_{\varphi,D \times CLK}(f_m) + S_{\varphi,dac}(f_m), \end{aligned} \quad (6.73)$$

where $S_{\alpha,D}$ and $S_{\varphi,D}$ are the DSB amplitude and phase noise of the digital input, and $S_{\alpha,dac}$ and $S_{\varphi,dac}$ are added by the DAC. The CPSD between $\Delta\varphi_D$ and $\Delta\varphi_{CLK}$ is denoted as $S_{\varphi,D \times CLK}$. Figure 6.24 shows the phase noise transferring from the clock

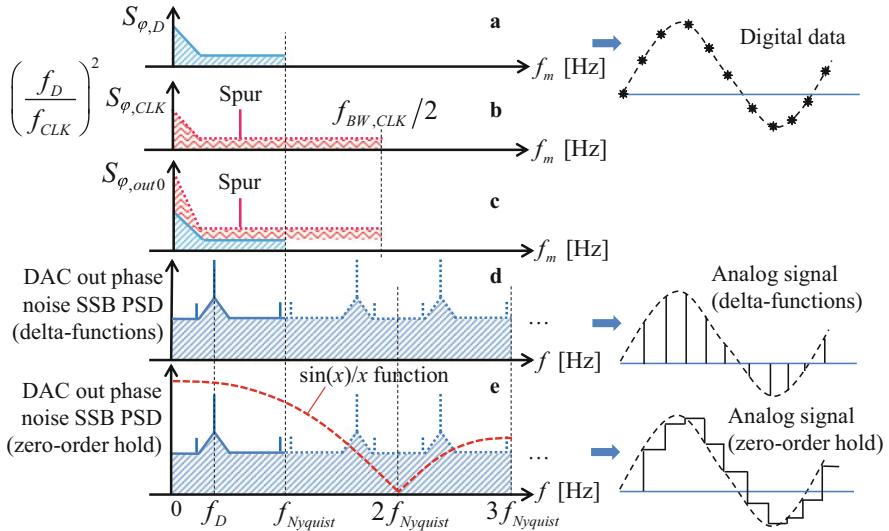


Fig. 6.24 Aliasing/imaging of the phase noise in the clock and digital input to the DAC output

and digital input to the DAC output. The DSB phase noise of x_D is defined in the Nyquist zone (Fig. 6.24a) since it is directly calculated with the phase of the discrete signal x_D . The clock phase noise is distributed in the analog bandwidth ($f_{BW,CLK}$) of the clock input circuit. Here we have normalized the DSB phase noise of the clock within $f_{BW,CLK}/2$ using the method in Fig. 6.20. The clock phase noise will affect the update time of the DAC output. Figure 6.24c shows the sum of $S_{\phi,D}$ and $(f_D/f_{CLK})^2 S_{\phi,CLK}$, assuming that $\Delta\varphi_D$ and $\Delta\varphi_{CLK}$ are uncorrelated. With a similar process as Fig. 6.19, we convert the DSB phase noise in Fig. 6.24c to SSB around the frequency f_D and alias all frequencies into the Nyquist zone. Then we obtain the SSB phase noise as in the first Nyquist zone of Fig. 6.24d. This process results in an equivalent sequence of samples (to be converted to analog output) containing noise from both the digital input and the clock. The DAC then produces analog outputs using this sequence of samples. In the frequency domain, the DAC output will image the spectrum in the first Nyquist zone to higher frequencies. Particularly, if there is no ZOH, the DAC analog output can be viewed as a series of delta-functions at t_k , then the spectrum at higher Nyquist zones will be the same (or image) as that in the first Nyquist zone (see Fig. 6.24d). However, most DACs use ZOH between two output points, which is equivalent to applying a filter on the spectrum of Fig. 6.24d. The magnitude response of the filter follows a pattern of $\sin(x)/x$ with $x = \pi f/f_{CLK}$, see Fig. 6.24e (the filtered spectrum is not shown). Typically, we select f_D to be smaller than $f_{Nyquist}$ and use a low-pass filter at the DAC output to filter out the images in higher Nyquist zones.

In LLRF systems, we use DACs to generate either baseband signals (for direct up-conversion) or IF signals (for IF up-conversion, up to tens of MHz).

When studying the DAC noise transfer model, we will assume that both the digital input and the clock are band-pass filtered. Therefore, the frequency aliasing and imaging can be neglected, and (6.73) can be directly used to calculate the DSB noise PSDs at the DAC output.

6.2.7.2 Noise Added by DAC

The DAC introduces additional internal noise via the clock input path or directly to the analog output.

The discussion about the noise added by the ADC clock input circuit (aperture jitter and spurs) is also applicable to the DAC. These additional noise terms in the DAC clock input circuit will also transfer to the DAC output following the mechanism in Fig. 6.24. The internal architecture of most DACs can attenuate the clock spurs, therefore, the clock spurs' contribution to the DAC output is often not well known. The quantization noise is determined by the vertical resolution (N) of the DAC, and the achievable maximum SNR is also given by $SNR_{qu} = 6.02 N + 1.76$ dB. Oversampling (with a higher f_{CLK}) is also beneficial to reduce the effects of the noise (e.g., the quantization noise and the clock aperture jitter) that is aliased in the Nyquist zone at the DAC's digital side.

Unlike ADCs, the DAC thermal noise will be added to the analog output directly. The thermal noise is an additive white noise generated by the resistance and buffer amplifiers in the DAC output circuit. Since it is not involved in the sampling process, the thermal noise is distributed within the full analog bandwidth of the DAC; therefore, its effects cannot be reduced by oversampling. When the DAC generates low-frequency signals (e.g., for direct up-conversion), the clock jitter is not relevant to the output noise. In this case, the thermal noise will dominate the overall noise, and the SNR can be improved by increasing the DAC output level.

In the DAC analog output, spurs coming from the higher-order harmonics, power supplies or glitches might also exist. The glitches often happen during the transient time when the DAC output changes (Garcia and LaJeunesse 1995).

As a summary, the DAC added amplitude and phase noise in terms DSB PSDs can also be described by (6.69), which is rewritten here as

$$\begin{aligned} S_{\alpha,dac}(f_m) &= F_{ZOH} \left\{ \zeta(f_m) D_{\alpha,th} + \sum_i D_{\alpha,outspur,i} \right\}, \\ S_{\varphi,dac}(f_m) &= F_{ZOH} \left\{ \zeta(f_m) (D_{\varphi,th} + D_{\varphi,apt}) + \sum_j D_{\varphi,clkspur,j} + \sum_k D_{\varphi,outspur,k} \right\}. \end{aligned} \quad (6.74)$$

We will only consider the noise power within the first Nyquist zone. Therefore, the offset frequency f_m should satisfy $0 < f_m \leq \max(f_D, f_{Nyquist} - f_D)$, as specified in (6.69).

Table 6.1 Noise of AD9780 (a DAC product of Analog Devices) specified in its datasheet

Parameter	One-tone Noise Spectral Density (NSD)
$f_{DAC} = 500 \text{ MSPS}, f_{OUT} = 40 \text{ MHz}$	-157 dBc/Hz
$f_{DAC} = 500 \text{ MSPS}, f_{OUT} = 120 \text{ MHz}$	-154.5 dBc/Hz
$f_{DAC} = 500 \text{ MSPS}, f_{OUT} = 380 \text{ MHz}$	-153 dBc/Hz
$f_{DAC} = 500 \text{ MSPS}, f_{OUT} = 480 \text{ MHz}$	-152 dBc/Hz

Here $D_{\alpha,th}$ and $D_{\varphi,th}$ are the contribution of the DAC thermal noise and $D_{\varphi,apt}$ is from the clock aperture jitter. The spurs added by the analog output circuits are denoted as $D_{\alpha,outspur}$ and $D_{\varphi,outspur}$, while the clock spurs will result in phase noise spurs denoted as $D_{\varphi,clkspur}$. Note that the noise is finally filtered by the ZOH (see Fig. 6.24e), which is described by a filter $F_{ZOH}\{\}$ in (6.74). The factor ζ has the same definition as in (6.69), with f_{IF} replaced by f_D .

The noise added by a DAC is typically specified in its datasheet for different output frequencies. We use the DAC product AD9780 of Analog Devices as an example (see the datasheet of AD9780/AD9781/AD9783, Analog Devices). Table 6.1 is the noise specification of AD9780 given by the datasheet. The specified noise PSD (i.e., NSD in the datasheet) is white noise at the DAC output, coming from the clock aperture jitter and the thermal noise (we neglect the quantization noise). We can estimate the noise added by a DAC with the following rules:

- (a) When f_D (i.e., f_{OUT} in the datasheet) is small, the noise PSD given by the datasheet is mainly the thermal noise. For example, from the datasheet, the noise PSD at $f_D = 40 \text{ MHz}$ is -157 dBc/Hz, which is an estimate of the SSB thermal noise PSD relative to the carrier signal power. Half of the thermal noise power contributes to the amplitude noise and another half to the phase noise. Therefore, the additional SSB amplitude (or phase) noise in the DAC output caused by the thermal noise will be $D_{\alpha,th} = D_{\varphi,th} \approx -160 \text{ dBc/Hz}$.
- (b) For a large f_D , the noise PSD given by the datasheet is mainly from the clock aperture jitter. For example, if we generate an output of $f_D = 480 \text{ MHz}$, the SSB phase noise induced by the aperture jitter can be estimated as $D_{\varphi,apt}|_{480\text{MHz}} \approx -152 \text{ dBc/Hz}$. Note that the aperture jitter-caused phase noise is dependent on the DAC output frequency. More accurately, $D_{\varphi,apt}$ is proportional to the square of f_D , as given by (6.66) with f_{IF} replaced by f_D . Therefore, we should calculate the value of $D_{\varphi,apt}$ for the desired f_D by scaling the value learned from the datasheet. For example, if our expected $f_D = 45 \text{ MHz}$, the phase noise added by the aperture jitter can be estimated as $D_{\varphi,apt}|_{45\text{MHz}} \approx -152 + 20 \times \lg(45/480) = -172.6 \text{ dBc/Hz}$.

We should keep in mind that the $D_{\alpha,th}$, $D_{\varphi,th}$, and $D_{\varphi,apt}$ are all slightly over estimated and their real values might be slightly smaller. The noise values in Table 6.1 all includes both the thermal noise and the aperture jitter. In our estimation, we have made the approximation neglecting one of them for the small or large f_D .

6.3 Noise Transfer in RF Control Loops

One of the primary goals of the LLRF system is to regulate and stabilize the RF field in the standing-wave cavities or traveling-wave structures. For RF systems operating in the CW mode or the pulsed mode with long pulses, feedback is usually adopted. However, the intra-pulse feedback is not applicable for short pulses (comparable to the loop delay); instead, pulse-to-pulse adaptive feedforward control is often used. The pulse-to-pulse control can only suppress the slow drifts because there is no control within the RF pulse. Figure 6.25 depicts a general RF control loop.

We are interested in how the noise in the RF reference signal, the RF driving chain, and the RF measurement chain transfers to the RF field for beam acceleration (Hoffmann 2008). The RF reference signal affects the phase noise in both the driving chain and the measurement chain, but its amplitude noise is not relevant because the reference input ports of the RF actuator and RF detector are usually saturated. The driving chain components (RF actuator, amplifier, klystron, transmission line and cavity or structure) add amplitude and phase noise directly to the cavity or structure input. The measurement chain components (RF pickup cables, RF detector) introduce RF measurement noise, which also affects the RF field when closing the control loop.

We will study the noise transfer relationship in the RF control loop for both the feedback (Ludwig et al. 2006) and the pulse-to-pulse control cases. It helps to understand the contributions of different noise sources to the RF field amplitude and phase fluctuations. The noise model can also be used to determine the noise requirements for the RF components based on the required RF field stability.

6.3.1 Noise Transfer in Feedback Control

Feedback control is often used for the control of standing-wave cavities. Figure 6.26 shows a typical RF feedback loop including noise. The blocks \mathbf{K} , \mathbf{G}_{cav} and \mathbf{G}_{det} are the phasor transfer functions of the controller, cavity, and RF detector, respectively. Here we have neglected the dynamics of the RF actuator, amplifier, and klystron, assuming that their bandwidths are much larger than the cavity bandwidth. The phasor \mathbf{r} is the set point of the cavity output \mathbf{y} that represents the cavity voltage experienced by the beam. The phase noise of the RF reference signal, $\Delta\varphi_{ref}$, is added

Fig. 6.25 General block diagram of an RF control loop

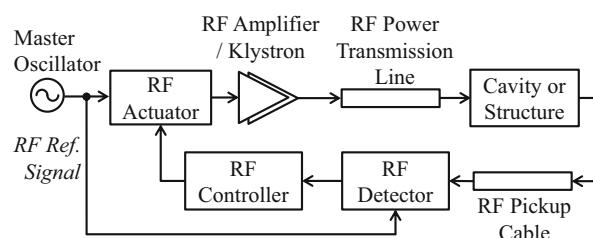
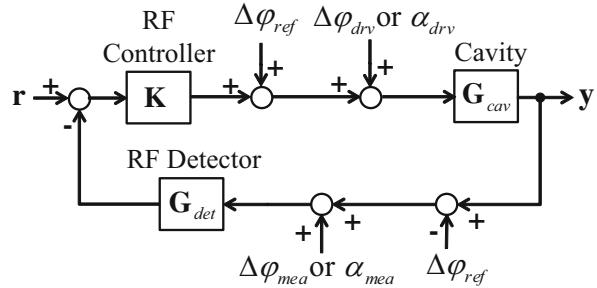


Fig. 6.26 Noise model of the feedback loop for a standing-wave cavity



in the RF driving chain and subtracted from the RF measurement chain. In the driving chain, the RF reference signal is modulated by the drive signal from the RF controller, so $\Delta\varphi_{ref}$ is added. In contrast, in the measurement chain, the cavity probe signal is demodulated using the RF reference signal, which subtracts the reference signal phase from the probe signal phase, resulting in the contribution of $-\Delta\varphi_{ref}$. The phase and amplitude noise added by the driving chain, $\Delta\varphi_{drv}$ and α_{drv} , also named as the *RF driving chain noise*, are applied to the cavity input. They are introduced by the driving chain components internally and the disturbances externally (e.g., temperature drifts in the power transmission line, microphonics in the cavity). The noise added by the measurement chain, $\Delta\varphi_{mea}$ and α_{mea} , also known as the *RF measurement chain noise*, comes from the RF pickup cable, down-conversion mixer, LO and clock generator, and ADCs.

To simplify the analysis, we assume \mathbf{K} is a proportional controller. We suppose

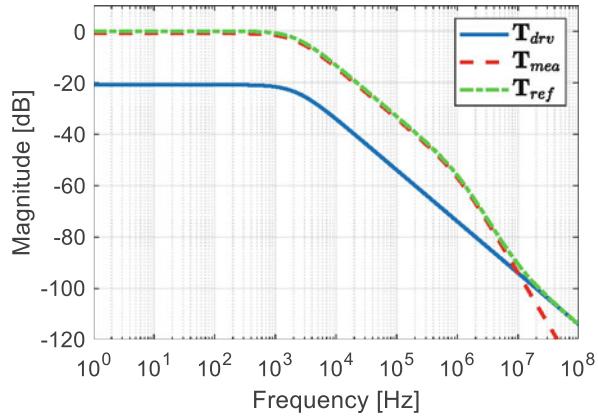
$$\mathbf{K}(\hat{s}) = g, \mathbf{G}_{cav}(\hat{s}) = \frac{\omega_{1/2}}{\hat{s} + \omega_{1/2}}, \mathbf{G}_{det}(\hat{s}) = \frac{\omega_{det}}{\hat{s} + \omega_{det}}, \quad (6.75)$$

where g is a constant and $\omega_{1/2}$ is the half-bandwidth of the cavity. The RF detector is modeled as a low-pass filter in the baseband with a cutoff frequency ω_{det} . Note that we use the angular frequency in transfer functions. Here we have neglected the detuning of the cavity. The closed-loop transfer functions from the noises of the driving chain, measurement chain and reference signal to the cavity field (i.e., cavity model output) can be written as

$$\begin{aligned} \mathbf{T}_{drv}(\hat{s}) &= \frac{\mathbf{G}_{cav}}{1 + \mathbf{K}\mathbf{G}_{cav}\mathbf{G}_{det}}, \mathbf{T}_{mea}(\hat{s}) = -\frac{\mathbf{K}\mathbf{G}_{cav}\mathbf{G}_{det}}{1 + \mathbf{K}\mathbf{G}_{cav}\mathbf{G}_{det}}, \\ \mathbf{T}_{ref}(\hat{s}) &= \mathbf{T}_{drv}(\hat{s}) - \mathbf{T}_{mea}(\hat{s}). \end{aligned} \quad (6.76)$$

The bode plots of the transfer functions above are depicted in Fig. 6.27 with the following parameters: $g = 10$, $\omega_{1/2} = 2\pi \times 200$ rad/s and $\omega_{det} = 2\pi \times 10^6$ rad/s. We can see from \mathbf{T}_{drv} that $\Delta\varphi_{drv}$ and α_{drv} will be suppressed by a factor of about g (assuming $g \gg 1$) corresponding to a closed-loop bandwidth $\omega_B \approx g\omega_{1/2}$. The bode plot of \mathbf{T}_{mea} indicates that $\Delta\varphi_{mea}$ and α_{mea} are transferred to y within ω_B and are filtered by the cavity and the RF detector at higher frequencies. Since $\Delta\varphi_{ref}$ transfers

Fig. 6.27 Bode plots of the closed-loop noise transfer functions of feedback control



to \mathbf{y} via both the driving chain and the measurement chain, \mathbf{T}_{ref} is the superposition of \mathbf{T}_{mea} and \mathbf{T}_{drv} . At lower frequencies, \mathbf{T}_{mea} is much larger than \mathbf{T}_{drv} , so \mathbf{T}_{ref} is dominated by \mathbf{T}_{mea} ; whereas at higher frequencies, the measurement path is further suppressed by the RF detector so \mathbf{T}_{drv} will dominate \mathbf{T}_{ref} . It indicates that the RF reference phase noise, which comes from the master oscillator and the synchronization system, will be transferred to the cavity output within the closed-loop bandwidth of the cavity. It also implies that the LLRF feedback cannot suppress the RF reference phase noise.

With a higher feedback gain, we will obtain better suppression of the RF driving chain noise, as well as a larger closed-loop bandwidth. However, more RF measurement chain noise will be transferred to the cavity output. Therefore, there exists an optimal value of g for minimizing the total noise in the cavity output (see Fig. 6.28). If we can reduce the measurement noise, the optimal g can be increased.

The total amplitude and phase noise PSDs ($S_{\alpha,tot}$ and $S_{\phi,tot}$) in the cavity output are contributed from all the noise sources that are assumed to be uncorrelated:

$$\begin{aligned} S_{\alpha,tot}(f_m) &= |\mathbf{T}_{drv}(j2\pi f_m)|^2 S_{\alpha,drv}(f_m) + |\mathbf{T}_{mea}(j2\pi f_m)|^2 S_{\alpha,mea}(f_m), \\ S_{\phi,tot}(f_m) &= |\mathbf{T}_{drv}(j2\pi f_m)|^2 S_{\phi,drv}(f_m) + |\mathbf{T}_{mea}(j2\pi f_m)|^2 S_{\phi,mea}(f_m) \\ &\quad + |\mathbf{T}_{ref}(j2\pi f_m)|^2 S_{\phi,ref}(f_m), \end{aligned} \quad (6.77)$$

where $S_{\phi,drv}$, $S_{\phi,mea}$ and $S_{\phi,ref}$ are the DSB phase noise PSDs added by the driving chain, measurement chain and in the RF reference signal, respectively; $S_{\alpha,drv}$ and $S_{\alpha,mea}$ are the PSDs of the corresponding amplitude noise. In some RF systems, the klystron operates in saturation, isolating the upstream amplitude noise and reducing the amplitude jitter in the cavity. In this case, the amplitude feedback cannot be applied. The cavity voltage amplitude drift needs to be controlled by other means, like adjusting the klystron modulator high voltage.

Fig. 6.28 Optimal gain to minimize the errors in the cavity output. Lower measurement noise helps to increase the optimal gain. Here g_{max} is the maximum allowed gain to keep the loop stable

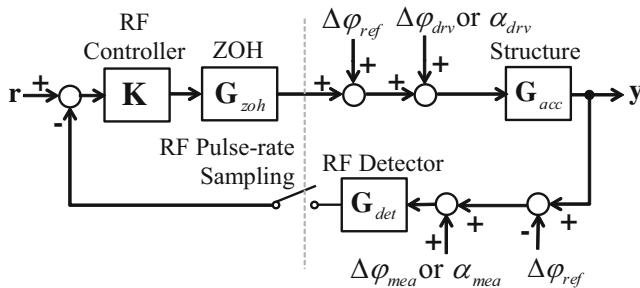
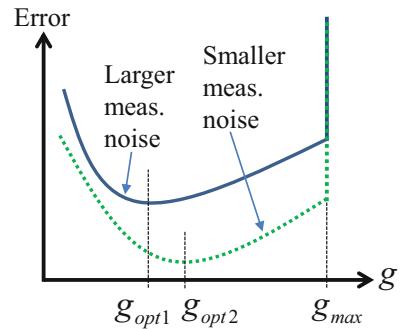


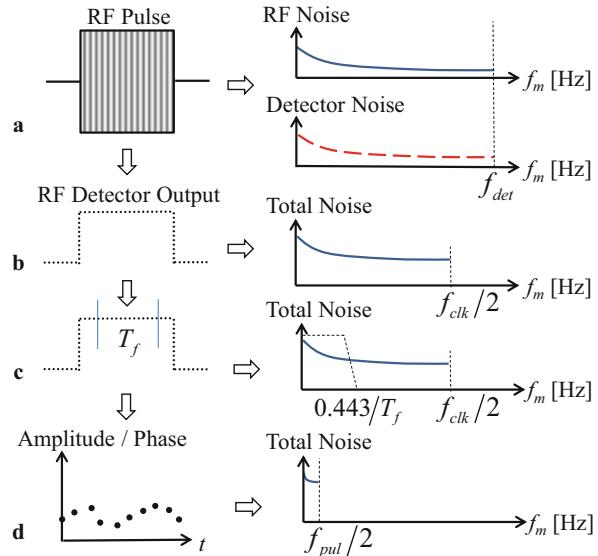
Fig. 6.29 Noise model of the pulse-to-pulse control loop for a traveling-wave structure

6.3.2 Noise Transfer in Pulse-to-Pulse Control

The RF stations with traveling-wave structures usually operate in the pulsed mode. The pulse width (i.e., structure filling time) is often too short to implement meaningful feedback control. Therefore, we often adopt pulse-to-pulse adaptive feed-forward control to suppress the RF field drifts. Figure 6.29 illustrates the noise model of a pulse-to-pulse control loop. The amplitude and phase waveforms of each RF pulse are measured to determine a pair of amplitude and phase values by averaging the samples within a time window. Then the controller generates necessary corrections to the RF drive amplitude and phase for the next pulse. The corrections will be applied before the next pulse and be held for the entire pulse. The RF measurement process of a pulse will alias the noise into the Nyquist zone of the pulse repetition rate, which is explained in Fig. 6.30, according to the traveling-wave structure model discussed in Sect. 3.5.

Figure 6.30a illustrates that the RF pulse is measured by an RF detector with an analog bandwidth f_{det} . Note that the RF detector also adds noise to the RF signal. The total noise in the RF signal will be filtered by the RF detector analog bandwidth, and when sampled at the frequency f_{clk} , it will be aliased to the Nyquist zone of f_{clk} (see plot b). We average the RF detector output within a time window T_f . It corresponds to filtering the noise using a moving-average filter with a cut-off frequency of about

Fig. 6.30 RF measurement of a pulse and the noise conversion in frequency domain



$0.443/T_f$ given by (3.63). See plot **c**. The time window is typically selected to be the filling time of the traveling-wave structure. Therefore, the measurement bandwidth is limited to the RF-beam interaction bandwidth in the structure. Finally, the sequence of the amplitude and phase values at the RF pulse rate f_{pul} will alias the filtered noise power into the Nyquist zone of f_{pul} , as in plot **d**.

The results of Fig. 6.30d are the inputs to the pulse-to-pulse controller. Due to the frequency aliasing, the noise in the measurements may have contributions from different Nyquist zones at higher frequencies. For example, if $f_{pul} = 100$ Hz, the noise at 1 Hz, 99 Hz and 101 Hz are all added to the 1 Hz noise power in the first Nyquist zone. In most LLRF systems, we select a sampling rate satisfying $f_{clk}/2 > f_{det}$, so there is no frequency aliasing in Fig. 6.30b.

The controller **K** performs pulse-to-pulse control, which is equivalent to a discrete-integral feedback controller updating at the RF pulse repetition rate. With the z-transform, we have $\mathbf{K}(z) = k / (z - 1)$, where k is the feedback gain. The corresponding phasor Laplace transform of **K** can be written below by substituting z with $z = e^{s\Delta T_{pp}}$ where $\Delta T_{pp} = 1/f_{pul}$ is the time interval between adjacent RF pulses:

$$\mathbf{K}(\hat{s}) = \frac{k}{e^{\hat{s}\Delta T_{pp}} - 1}. \quad (6.78)$$

The output of **K** is the amplitude and phase correction for the next RF pulse. It is updated at the pulse repetition rate and held during the next pulse. In Fig. 6.29, we used a zero-order hold to model the holding function. The holding time is the RF pulse width ΔT_{pw} , which is usually much smaller than the RF pulse interval ΔT_{pp} . Here we have assumed that ΔT_{pw} equals the RF-beam interaction time; that is, $\Delta T_{pw} = T_f$. The phasor transfer function of the ZOH is given by

$$\mathbf{G}_{zoh}(\hat{s}) = \frac{e^{\hat{s}\Delta T_{pw}/2} (1 - e^{-\hat{s}\Delta T_{pw}})}{\hat{s}\Delta T_{pw}}. \quad (6.79)$$

The phasor transfer function of the traveling-wave structure has been given by (3.62), which happens to be the same as that of the ZOH here, so $\mathbf{G}_{acc} = \mathbf{G}_{zoh}$. Similar to (6.76), the closed-loop noise transfer functions for Fig. 6.29 can be written as

$$\begin{aligned} \mathbf{T}_{drv}(\hat{s}) &= \frac{\mathbf{G}_{acc}}{1 + \mathbf{KG}_{zoh}\mathbf{G}_{acc}\mathbf{G}_{det}}, \quad \mathbf{T}_{mea}(\hat{s}) = -\frac{\mathbf{KG}_{zoh}\mathbf{G}_{acc}\mathbf{G}_{det}}{1 + \mathbf{KG}_{zoh}\mathbf{G}_{acc}\mathbf{G}_{det}}, \\ \mathbf{T}_{ref}(\hat{s}) &= \mathbf{T}_{drv}(\hat{s}) - \mathbf{T}_{mea}(\hat{s}). \end{aligned} \quad (6.80)$$

An example of the bode plots of the closed-loop transfer functions above is depicted in Fig. 6.31 with the parameters as $f_{pul} = 100$ Hz, $\Delta T_{pw} = 1$ μ s, and $k = 0.5$.

In the first Nyquist zone (0 Hz to $f_{pul}/2$), \mathbf{T}_{drv} and \mathbf{T}_{mea} are the same as that in the discrete-integral feedback loop (Geng and Hong 2016). In higher Nyquist zones, the transfer functions image their responses in the first Nyquist zone. Note that \mathbf{T}_{drv} will suppress the RF driving chain noise at frequencies around mf_{pul} ($m = 0, 1, 2, \dots$), whereas \mathbf{T}_{mea} will pass the measurement noise to the structure at these frequencies. We also notice that the RF reference phase noise will be fully transferred to the structure. At higher frequencies, \mathbf{T}_{drv} , \mathbf{T}_{mea} and \mathbf{T}_{ref} will be finally filtered by the bandwidth of the traveling-wave structure, which is about 443 kHz in this example calculated with (3.63).

The total amplitude and phase noise PSDs in the structure's RF field can still be calculated with (6.77) using the closed-loop transfer functions in (6.80).

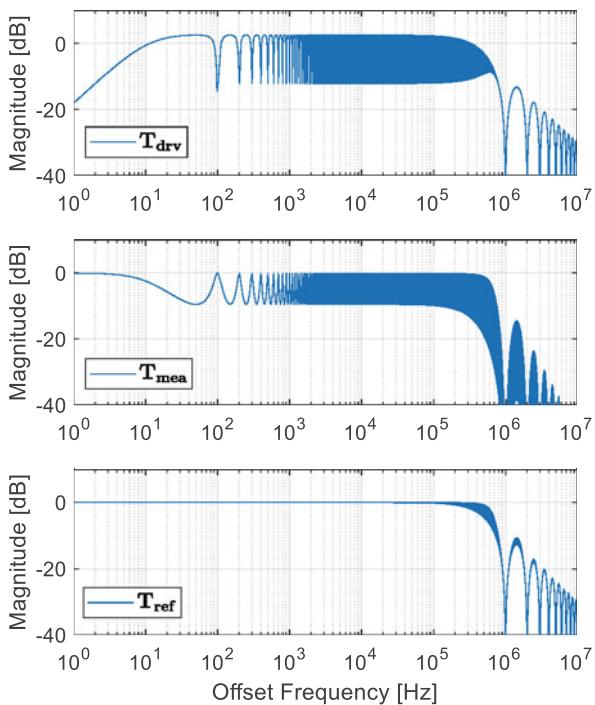
6.4 RF System Noise Specification

The *RF system noise model* describes how the various noise sources in an RF station determine the RF field fluctuation. On the other hand, the noise transfer relationship helps to determine the *noise budget* (i.e., the upper limit of the noise power) of each RF component satisfying the overall RF field stability requirement. Specifying the performance of HPRF and LLRF components is an important task when defining the requirements for the RF system.

6.4.1 Noise Specification Strategy

We will follow a top-to-down procedure to analyze the noise model and specify the performance requirements of the RF system:

Fig. 6.31 Bode plots of the closed-loop noise transfer functions of pulse-to-pulse control



(a) *Define a global noise model for the accelerator*

At the machine level of an accelerator, each RF station can be modeled as a black box providing RF fields for beam acceleration. The beam dynamics study describes the relationship between the beam parameter deviations and the RF amplitude and phase changes. This is referred to as a *global noise model*. It can be static (for Linacs) described by constant response matrices or dynamical (for storage rings) by differential equations.

(b) *Specify the RF station stability requirements*

The global noise model can be used to predict the resulting beam fluctuations from the given RF field fluctuations and the fluctuations of other sources (e.g., bunch charge or dipole magnetic field). This allows us to specify the stability requirements for the involved subsystems (including RF stations) for the desired beam stability. The procedure to determine the stability requirements for each subsystem may be iterative. We can define the fluctuations of various subsystems and apply them to the global noise model. If the resulting beam stability is satisfactory, the assumed fluctuations of the subsystems can define their tolerances. We should tune the required stability of different subsystems to compromise the technical difficulties and cost.

(c) *Define the frequency range for noise PSD integration*

The RF station stability requirements derived from the global noise model are typically RMS amplitude and phase fluctuations of the RF field. However, we

often measure or evaluate the amplitude and phase noise as PSDs in the frequency domain. The RMS values are calculated by integrating the PSDs in a frequency range. See (6.32) and (6.37). We need to define the integration frequency range.

The lower boundary of the frequency range, denoted as f_L , is typically determined by the closed-loop bandwidth of the beam-based feedback. The effects of the RF station's slow amplitude and phase fluctuations below f_L can usually be compensated for by the beam-based feedback. Therefore, we only consider the RF noise at frequencies above f_L .

The upper boundary frequency, f_U , is typically defined by the closed-loop bandwidth ω_B (defined in Sect. 6.3.1) of the RF control loop with $f_U = \omega_B/(2\pi)$. As seen from Fig. 6.27, ω_B is a critical parameter determining how the RF reference phase noise, measurement chain noise, and driving chain noise transfer to the cavity field. For RF stations operated by a pulse-to-pulse control loop, f_U is often chosen as the Nyquist frequency of the RF pulse repetition rate, that is, $f_U = f_{pul}/2$.

- (d) *Determine the upper limits of the RF driving chain, measurement chain noise and the RF reference phase noise*

Equation (6.77) shows that the RF driving chain noise is suppressed, and the RF measurement chain noise and RF reference phase noise are transferred to the cavity voltage within the closed-loop bandwidth of the RF control loop. Note that the noise here is specified as PSDs, and we only consider them within the frequency range from f_L to f_U . This allows us to specify the maximum noise added by different parts of the RF station. The resulting RF field fluctuation, considering the closed-loop transfer functions, should satisfy the RF station's overall stability requirements (as RMS values). In other words, we split the total noise budget of the RF station into three parts in the RF reference, the driving chain, and the measurement chain. Of course, trade-offs should be made considering the technical difficulties and cost. For example, we may allocate more noise budget in the driving chain because a low-jitter klystron might not be available, whereas a low-noise RF detector could be easier to obtain.

- (e) *Determine the noise budgets of the RF driving chain and measurement chain components*

This can be done following the same approach as above. That is, we make assumptions of the noise added by the RF components (in terms of PSDs) and estimate the overall noise added in the driving chain and measurement chain. By satisfying the driving chain and measurement chain stability requirements specified in the previous step, we find a combination of the allowed noise of different RF components. This combination, a compromise between the technical difficulties and cost, can define the RF components' tolerances.

We will present a detailed noise model of an RF station in Sect. 6.5. The RF reference signal is special, which adds correlated phase noise to multiple RF stations. We will discuss it in more detail in Sect. 6.4.3.

6.4.2 Accelerator Global Noise Model

In a particle accelerator, the amplitude and phase fluctuations of an RF station will affect the beam parameters, especially the longitudinal beam parameters. The global noise model describes the relationship between the RF fluctuations and the variations of beam parameters (Shih et al. 1994; Mastorides et al. 2010; Maalberg et al. 2020). The determination of the global noise model is usually based on beam dynamics simulations (mainly for machines under design) or measurements (mainly for machines already in operation). We assume that the relationship ($T_{rf-beam}$) between the RF deviation and beam deviation is linear. Then the effect of each noise source can be studied separately, and the overall beam deviation is the superposition of the effects of all noise sources. The transfer function $T_{rf-beam}$ can be derived by simulating or measuring the beam deviations in response to a small change in the RF amplitude or phase.

6.4.2.1 Noise Specification of Linacs

For electron linear accelerators, the concerned beam parameters are typically the beam energy, bunch length and bunch arrival time (Dinter 2018). The $T_{rf-beam}$ is usually static and described by a constant matrix denoted as the *sensitivity matrix* (or *response matrix*). Of course, the sensitivity matrix is only valid around a particular operating point (e.g., specific bunch charge and beam energy) because $T_{rf-beam}$ is highly nonlinear for large RF deviations. With such a static global noise model, the beam jitter is often specified as RMS values. Therefore, the maximum allowed amplitude and phase jitter of an RF station is also specified as RMS values without information on their frequency distributions.

Let us use SwissFEL (Fig. 6.32) as an example to demonstrate the global noise model (Beutner and Reiche 2010). We will also use it to specify the stability requirements for the RF stations. SwissFEL (Milne et al. 2017) is an FEL machine based on a normal-conducting Linac. It employs an S-band ($f_{SB} = 2998.8$ MHz) RF Gun with a 2.6-cell standing-wave cavity to generate electron bunches. The Booster1 consists of two S-band traveling-wave structures, each powered by a separate klystron. A laser heater (LH) is used to mitigate the micro-bunching instability in the bunch compressors. The Booster2 consists of two S-band and one

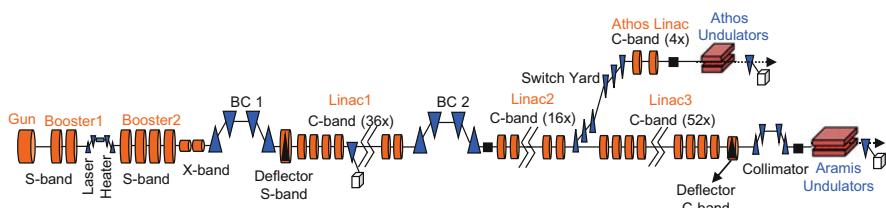


Fig. 6.32 Layout of SwissFEL linear accelerator

X-band ($f_{XB} = 11.424$ GHz) RF stations, each driving two traveling-wave structures. The Booster2 S-band RF stations operate off-crest to generate the required energy chirp for bunch compression in the first bunch compressor, BC1. The X-band RF station works at a decelerating phase to linearize the energy distribution along the electron bunch for optimal compression performance. After BC1, the Linac1 with nine C-band ($f_{CB} = 5712$ MHz) RF stations are used to ramp the beam energy and generate the necessary energy chirp for the second bunch compressor, BC2. Each C-band RF station drives four traveling-wave structures. The Linac2 consists of four C-band RF stations, which boost the beam energy to 3 GeV to feed the soft X-ray line “Athos” and the Linac3 of the hard X-ray line “Aramis”. In the Aramis beam line, 13 C-band RF stations are used to ramp the electron beam energy up to 5.8 GeV before injection in the undulators. The RF stations operate in the pulsed mode at a pulse repletion rate of 100 Hz.

The sensitivity matrix of the SwissFEL Linac (Reiche 2016; Geng et al. 2019) is derived with the beam-tracking code LiTrack (Bane and Emma 2005) and given by

$$\begin{bmatrix} (\Delta E/E)_{BC1} \\ \Delta\varphi_{b,BC1} \\ (\Delta L/L)_{BC1} \\ (\Delta E/E)_{BC2} \\ \Delta\varphi_{b,BC2} \\ (\Delta L/L)_{BC2} \\ (\Delta E/E)_{L3} \\ \Delta\varphi_{b,L3} \\ (\Delta L/L)_{L3} \end{bmatrix} = \begin{bmatrix} -0.001 & -0.288 & 0.474 & 0.599 & 0.278 & -0.071 & 0.003 & 0 & 0 & 0 & 0 \\ -0.004 & 0.006 & 1.643 & 2.072 & 0.970 & -0.248 & 0.006 & 0 & 0 & 0 & 0 \\ 0.362 & 13.501 & 15.624 & 8.199 & 34.035 & -2.831 & -11.882 & 0 & 0 & 0 & 0 \\ -0.001 & -0.044 & 0.703 & -0.889 & -0.411 & 0.106 & -0.005 & 0.845 & 0.249 & 0 & 0 \\ -0.005 & -0.052 & 0.728 & 0.916 & 0.436 & -0.110 & 0 & 1.104 & 0.323 & 0 & 0 \\ 5.733 & 68.079 & -36.768 & -100.583 & 94.446 & 3.774 & -56.055 & 20.083 & 32.506 & 0 & 0 \\ -0.002 & -0.011 & -0.256 & -0.327 & -0.141 & 0.038 & -0.006 & 0.313 & 0.094 & 0.639 & -0.002 \\ -0.005 & -0.052 & 0.728 & 0.916 & 0.436 & -0.110 & 0 & 1.104 & 0.323 & 0 & 0 \\ 5.733 & 68.079 & -36.768 & -100.583 & 94.446 & 3.774 & -56.055 & 20.083 & 32.506 & 0 & 0 \end{bmatrix} \begin{bmatrix} (\Delta Q/Q)_{LH} \\ \Delta\varphi_{b,LH} \\ (\Delta E/E)_{LH} \\ \alpha_{bsr2} \\ \Delta\varphi_{bsr2} \\ \alpha_{XB} \\ \Delta\varphi_{XB} \\ \alpha_{L1} \\ \Delta\varphi_{L1} \\ \alpha_{L23} \\ \Delta\varphi_{L23} \end{bmatrix} \quad (6.81)$$

The noise sources at the right side include:

- $(\Delta Q/Q)_{LH}$: initial bunch charge deviation at the exit of laser heater.
- $\Delta\varphi_{b,LH}$: initial bunch arrival time (converted to an S-band phase $\Delta\varphi_b = -2\pi f_{SB}\Delta t_b$, where Δt_b is the deviation of the bunch arrival time) at the exit of laser heater.
- $(\Delta E/E)_{LH}$: initial beam energy deviation at the exit of laser heater.
- $\alpha_{bsr2}, \Delta\varphi_{bsr2}$: total amplitude and phase deviations of Booster2.
- $\alpha_{XB}, \Delta\varphi_{XB}$: amplitude and phase deviations of X-band RF station.
- $\alpha_{L1}, \Delta\varphi_{L1}$: total amplitude and phase deviations of Linac1.
- $\alpha_{L23}, \Delta\varphi_{L23}$: total amplitude and phase deviations of Linac2 and Linac3.

The amplitude deviations are relative values with $\alpha := \Delta A/A$. In the global noise model, we assume that the noise of different RF stations is uncorrelated. Therefore, the RMS jitter of a Linac section is scaled from the RMS jitter of an individual RF station. For example, if the RMS jitter of the C-band stations is $\alpha_{CB,rms}$ and $\Delta\varphi_{CB,rms}$, then the overall RMS jitter of Linac2+Linac3 (with 17 C-band RF stations) should be calculated as $\alpha_{L23,rms} = \alpha_{CB,rms}/\sqrt{17}$, $\Delta\varphi_{L23,rms} = \Delta\varphi_{CB,rms}/\sqrt{17}$. As we have introduced before, the Booster2 consists of 2 RF stations and Linac1 consists of 9, so

their jitter in the equation above should be scaled by a factor of $1/\sqrt{2}$ and $1/3$ from the jitter of an individual RF station, respectively.

The sensitivity between the laser heater beam parameters and the RF Gun and Booster1 is studied separately. Upstream the laser heater, the beam energy is low and the space-charge effects in the bunches are not negligible. Therefore, different codes (e.g., ASTRA) (Floettmann 2017) should be used to make the simulation. In this example, we treat the laser heater beam parameter variations as inputs to the noise model so that the faster LiTrack code (with space charge effects neglected) can be used.

The resulting beam parameters at the left side of (6.81) are:

- $(\Delta E/E)_{BC1}$: beam energy deviation at the exit of BC1.
- $\Delta\varphi_{b,BC1}$: bunch arrival time (normalized to S-band phase) at the exist of BC1.
- $(\Delta L/L)_{BC1}$: bunch length deviation at the exit of BC1.
- $(\Delta E/E)_{BC2}$: beam energy deviation at the exit of BC2.
- $\Delta\varphi_{b,BC2}$: bunch arrival time (normalized to S-band phase) at the exist of BC2.
- $(\Delta L/L)_{BC2}$: bunch length deviation at the exit of BC2.
- $(\Delta E/E)_{L3}$: beam energy deviation at the exit of Linac3.
- $\Delta\varphi_{b,L3}$: bunch arrival time (normalized to S-band phase) at the exist of Linac3.
- $(\Delta L/L)_{L3}$: bunch length deviation at the exit of Linac3.

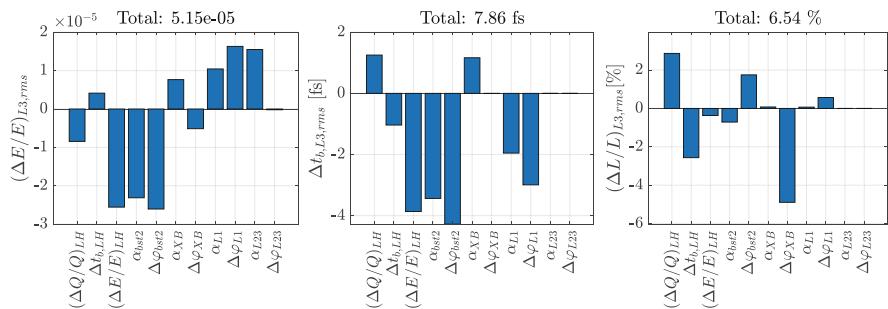
SwissFEL requires that at the end of Linac3, the beam jitter should satisfy $(\Delta E/E)_{L3,rms} < 5e-4$, $\Delta t_{b,L3,rms} < 20$ fs and $(\Delta L/L)_{L3,rms} < 5\%$.

To determine the required RF stability in terms of RMS jitter, we assume the noise sources at the right side of (6.81) are uncorrelated. We assign each time only one noise source a jitter (keep other noise sources zero) and calculate the resulting jitter of the beam parameters. Then the overall beam parameter jitter can be calculated by summing up the contributions from all noise sources. For example, for the i th noise source with a given jitter, we obtain an energy jitter $(\Delta E/E)_{L3,rms,i}$, then the overall energy jitter is $(\Delta E/E)_{L3,rms} = \sqrt{\sum_{i=1}^n (\Delta E/E)_{L3,rms,i}^2}$, where n is the number of the noise sources. In this example, we have chosen the tolerances of the noise sources as in Table 6.2.

Figure 6.33 illustrates the resulting RMS beam jitter at the exit of Linac3. The bars represent the contributions of different noise sources, and the total beam jitter is depicted above the plot, assuming all noise sources uncorrelated. The RMS jitter of each noise source is derived from Table 6.2 and the number of RF stations in the corresponding Linac section. Note that we also show the direction of the beam parameter variation, though the magnitudes are RMS values. A negative bar indicates that the deviations of the noise source and beam parameter change in opposite directions. The total beam jitter when all noise sources are present (uncorrelated) is also shown in Fig. 6.33. The specification in Table 6.2 satisfies the requirements of $(\Delta E/E)_{L3,rms}$ and $\Delta t_{b,L3,rms}$ very well, but not that of $(\Delta L/L)_{L3,rms}$. To satisfy the requirement of $(\Delta L/L)_{L3,rms}$, we need to reduce further the jitter of sensitive noise sources, including $(\Delta Q/Q)_{LH}$, $\Delta t_{b,LH}$, $\Delta\varphi_{SB}$ and $\Delta\varphi_{XB}$ implied by the third plot. Note that the specification in Table 6.2 is already stringent for state-of-the-art RF systems.

Table 6.2 Stability requirements for SwissFEL subsystems

Noise source	Notation	Tolerance (RMS)
Bunch charge at laser heater	$(\Delta Q/Q)_{LH}$	0.005
Bunch arrival time at laser heater	$\Delta t_{b,LH}$	20 fs
Beam energy at laser heater	$(\Delta E/E)_{LH}$	1e-4
S-band amplitude	α_{SB}	1e-4
S-band phase	$\Delta\varphi_{SB}$	0.015°
X-band amplitude	α_{XB}	2e-4
X-band phase	$\Delta\varphi_{XB}$	0.05°
C-band amplitude	α_{CB}	1e-4
C-band phase	$\Delta\varphi_{CB}$	0.03°

**Fig. 6.33** Estimate of beam jitter from the jitter of noise sources

Therefore, it implies that we may need more discussions with the beam dynamics experts about the bunch length stability requirement.

The SwissFEL example has assumed that the jitter of different RF stations is uncorrelated. This approximation is valid for most electron Linacs. However, two possible noise sources may cause correlated jitter in multiple RF stations. One is the master oscillator. Its phase noise will affect all RF stations via the reference signal distribution. The second source is the beam, which passes the cavities or structures of all RF stations and yields correlated error due to the beam loading effects. In the state-of-the-art RF systems, the phase noise of the master oscillator is usually minimal, and the beam loading effects can be well suppressed by the feedback or feedforward control. Therefore, we typically assume uncorrelated jitter in various RF stations when performing the global noise analysis.

6.4.2.2 Noise Specification of Synchrotrons

Synchrotrons are circular accelerators, in which the particles are accelerated or stored for many turns. The concerned beam parameters include the bunch arrival time (t_b) and energy (E_b) of the central particle in a bunch, and the standard deviations of t_b (bunch length, $\Delta t_{b,rms}$) and E_b (energy spread, $\Delta E_{b,rms}$) of all the

particles in the bunch. To simplify the discussion, we define a *beam current phase* $\varphi_{beam} := 2\pi f_{RF} t_b$ to represent the bunch arrival time with respect to the RF frequency. The phase of the RF field in the cavity is denoted as φ_{RF} . Note that φ_{RF} and φ_{beam} are defined with respect to a stable reference signal, whose frequency is a constant for electron synchrotrons, or time varying for synchrotrons accelerating nonrelativistic beams (e.g., protons).

We denote the RF phase noise as $\Delta\varphi_{RF}$ and the amplitude noise as α_{RF} defined by $\alpha_{RF} := \Delta V_{RF}/V_{RF}$, where V_{RF} and ΔV_{RF} are the cavity voltage and its deviation, respectively. The propagation from $\Delta\varphi_{RF}$ and α_{RF} to the deviation of beam current phase $\Delta\varphi_{beam}$ are dynamical and described by differential equations. This is because the RF fluctuations will affect the longitudinal synchrotron oscillation of the beam. The noise transfer functions (Mastorides et al. 2010; Lin et al. 2016) have a typical form of

$$\begin{aligned} T_{PN}(\hat{s}) &= \frac{\Delta\varphi_{beam}(\hat{s})}{\Delta\varphi_{RF}(\hat{s})} = \frac{\omega_{sync}^2}{\hat{s}^2 + 2\xi\hat{s} + \omega_{sync}^2}, \\ T_{AN}(\hat{s}) &= \frac{\Delta\varphi_{beam}(\hat{s})}{\alpha_{RF}(\hat{s})} = \frac{\omega_{sync}^2 \operatorname{tg}\varphi_{sync}}{\hat{s}^2 + 2\xi\hat{s} + \omega_{sync}^2}, \end{aligned} \quad (6.82)$$

where ω_{sync} is the *synchrotron oscillation* frequency (angular frequency, the corresponding ordinary frequency is denoted as f_{sync}), ξ is a damping factor related to the radiation energy loss of the particle, and φ_{sync} is the synchronous phase, the constant accelerating phase of the *synchronous particle* without longitudinal oscillations. Note that the accelerating phase is the difference between the RF and beam current phases ($\varphi_{RF} - \varphi_{beam}$). Here we have used the phasor Laplace transform because the transfer functions are for baseband signals. Figure 6.34 shows an example of the bode plots and step responses of the transfer functions in (6.82). The example assumes that $f_{sync} = 22$ Hz and the step changes of $\Delta\varphi_{RF}$ and α_{RF} are 0.01.

The step response of T_{PN} in Fig. 6.34b indicates that if we introduce a step change in $\Delta\varphi_{RF}$, the synchrotron oscillation will be stimulated. Due to the damping from

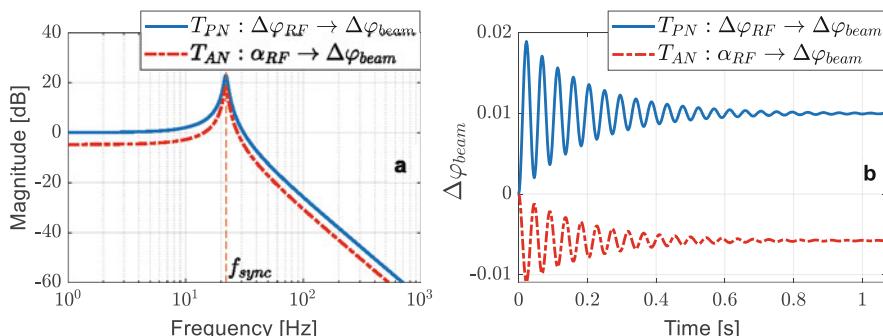


Fig. 6.34 Example of the bode plots (a) and step responses (b) of the transfer functions in (6.82)

synchrotron radiation, the particle will finally settle down in the same synchronous accelerating phase as before; therefore, $\Delta\varphi_{beam}$ must follow $\Delta\varphi_{RF}$ after the step response reaches a steady state gain. This behavior of T_{PN} is implied by the 0 dB gain at low frequencies in the bode plot. It also implies that when we change the RF phase by a constant offset, the beam arrival time at the cavity gap will follow it exactly after the oscillation damps, resulting in an unchanged φ_{sync} (determined by $\varphi_{RF} - \varphi_{beam}$).

As for T_{AN} , if we apply a step-change in α_{RF} , we will also obtain a change in $\Delta\varphi_{beam}$ after the oscillation damps. A typical step response is shown in Fig. 6.34b. This can be explained by the dependency of φ_{sync} to the cavity voltage V_{RF} . Even if φ_{RF} remains constant, φ_{beam} will change due to the φ_{sync} deviation caused by the V_{RF} step change.

The bode plots illustrate that the RF noise at frequencies around f_{sync} will be amplified, contributing to the beam fluctuations. The LLRF system must suppress the RF noise at these frequencies, or it will degrade the beam stability significantly. Here we have described T_{PN} and T_{AN} as continuous transfer functions. Nevertheless, a particle only experiences the RF noise at discrete time-instants when it passes through the cavity in each turn. That is, the particle experiences RF errors corresponding to sampling the RF noise at its *revolution frequency* f_{rev} , which aliases the RF noise power within the closed-loop bandwidth of the cavity to the frequency range between 0 Hz to $f_{rev}/2$. It implies that the RF noise at frequencies around $kf_{rev} \pm f_{sync}$ ($k = 0, 1, 2, \dots$) will all stimulate the synchrotron oscillation. It also implies that the LLRF system should maintain a strong suppression of the RF noise at all these frequencies. Therefore, in addition to the nominal feedback (i.e., direct feedback), a storage ring LLRF system may also employ the so-called *one-turn delay feedback control* (Blas and Garoby 1991; Mastoridis et al. 2012; Schmid et al. 2017). The one-turn delay feedback uses a comb-filter to increase the loop gain around kf_{rev} ($k = 0, 1, 2, \dots$), resulting in a stronger disturbance rejection at these frequencies. To keep the closed loop stable, we adjust the one-turn delay feedback loop delay to be a period of the particle revolution (i.e., traveling time of a synchronous particle for one turn). The one-turn delay feedback provides control at around kf_{rev} in a wide band. In contrast, the nominal feedback reduces low-frequency disturbances, and its closed-loop bandwidth is limited by the loop delay and the cavity pass-band modes.

The one-turn delay feedback pushes the RF control action over a much wider band than the nominal feedback. This imposes stricter requirements on the RF amplifiers driving the cavity (Fox et al. 2007; Fox et al. 2010). The amplifiers should have enough bandwidth with a flat frequency response over the control band. Due to the one-turn delay feedback, the RF amplifier produces power at the revolution harmonics at both sidebands of the central RF operating frequency. The intermodulation caused by the nonlinearity should be low to avoid too much interference between the control of different revolution harmonics.

The beam stability requirements are specified as the maximum allowed fluctuations of the arrival time t_b and the energy E_b of the particles. They are the basis for determining the required RF field stability. In the Linac case of (6.81), the relationship between the beam and RF field deviations is a static matrix; that is, the resulting beam fluctuation levels are independent of the RF field fluctuation frequencies.

However, for a storage ring, we must consider the dynamical behavior (6.82), which implies that the RF field fluctuations at different frequencies will cause different levels of beam fluctuations. Therefore, the following procedure can be followed to specify the desired RF field stability to satisfy the beam stability requirements:

- (a) Based on the RF system properties (e.g., RF frequency, cavity bandwidth, estimated disturbances including their magnitudes and frequencies, desired feedback structure, etc.), make assumptions of the RF field amplitude and phase noise PSDs at different frequencies.
- (b) Apply the RF field noise PSDs from the previous step to (6.82) and to the transfer functions from $\Delta\varphi_{RF}$ (and α_{RF}) to ΔE_b (energy deviation, not shown here), and estimate the resulting fluctuations (RMS values) of t_b and E_b . Here we should consider the RF noise PSDs in the frequency range from f_L to f_U as defined in Sect. 6.4.1.
- (c) Repeat the two steps above, revising the RF noise assumptions until the beam stability requirements are satisfied. The RF field noise tolerances can then be defined as the finally adopted RF amplitude and phase noise PSDs.

The procedure above yields the RF field stability requirements as allowed maximum amplitude and phase noise PSDs. They specify the tolerances of the RF field amplitude and phase fluctuations at each frequency. Note that the RMS amplitude and phase jitter are not enough to specify the RF field stability requirements due to the frequency-dependent noise transfer functions.

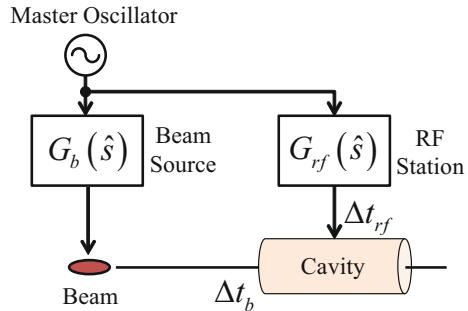
6.4.3 Specification of RF Reference Phase Noise

As mentioned before, the first step to specify the stability requirements for an RF station is to split the total noise budget to the RF reference signal, the driving chain, and the measurement chain. Here we assume that these three parts are uncorrelated. The RF reference phase noise comes from two sources: the master oscillator (MO) distributes correlated phase noise to all RF stations, and the RF reference receiver at each RF station adds uncorrelated phase noise. From the noise budget allocated to the RF reference signal, we need to specify the phase noise of the master oscillator and that added by the RF reference receiver.

First, let us study how the MO phase noise affects the beam acceleration (Hoffmann 2008; Hanaki et al. 2014; Gallo 2018). Typically, the MO generates an ultimate reference signal distributed to and synchronizing various subsystems, including the beam source and the RF stations. Therefore, both the time of bunch generation and the RF phase will be affected by the MO phase noise. It implies that when we consider the beam accelerating phase (i.e., beam phase), which is defined by (3.33), the MO phase noise effects will be partially canceled. We will use a simple model as Fig. 6.35 to explain this situation.

We define the MO phase noise as $\Delta\varphi_{MO}$. We use two first-order transfer functions, G_b and G_{rf} , to model how $\Delta\varphi_{MO}$ affects the beam-current and RF-field phase

Fig. 6.35 A simple model to study the MO phase noise effects on the beam phase



variations (denoted as $\Delta\varphi_{beam}$ and $\Delta\varphi_{rf}$, respectively). The beam will interact with the RF field in the cavity. We use Δt_b to describe the time delay for the MO phase variation to reach the cavity, corresponding to the bunch arrival time. Similarly, Δt_{rf} is the delay of the MO phase noise transferred to the cavity voltage. The model can be written as

$$G_b(\hat{s}) = \frac{\omega_b}{\hat{s} + \omega_b} e^{-\hat{s}\Delta t_b}, G_{rf}(\hat{s}) = \frac{\omega_B}{\hat{s} + \omega_B} e^{-\hat{s}\Delta t_{rf}}, \quad (6.83)$$

where ω_b and ω_B are the bandwidths of the beam source and the RF station in which they lock to the MO. Here ω_B is the closed-loop bandwidth of the RF control loop, as discussed in Sect. 6.3.1. Then, the beam phase variation can be calculated as

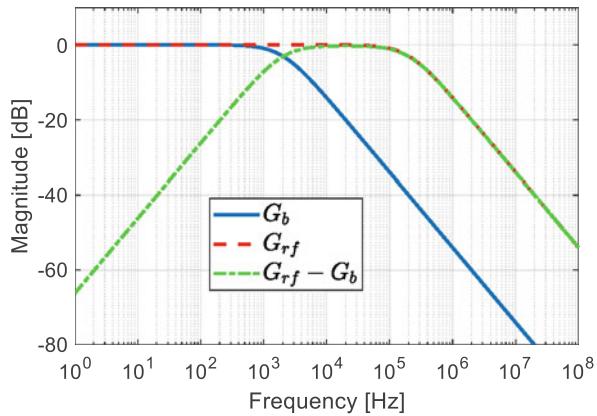
$$\Delta\varphi_b = \Delta\varphi_{rf} - \Delta\varphi_{beam} = (G_{rf} - G_b)\Delta\varphi_{MO}. \quad (6.84)$$

This is the relationship between the variations of the beam phase and the MO phase. Note that the variables in (6.84) should be the Laplace transforms of the phase variations. We can plot G_b , G_{rf} and $G_{rf} - G_b$ in the frequency domain in Fig. 6.36. Here we have used the following parameters: $\omega_b = 2\pi \times 2e3$ rad/s, $\omega_B = 2\pi \times 2e5$ rad/s, $\Delta t_b = 1 \mu s$ and $\Delta t_{rf} = 2 \mu s$.

The bode plot of $G_{rf} - G_b$ indicates that the MO phase noise within the *common locking bandwidth* (denoted as f_{lock} , below which all subsystems are synchronized (locked) to the MO), has no impact on beam acceleration. Therefore, we may adopt the following approach to specify the MO phase noise:

- First, we determine the MO phase-locking bandwidths of different subsystems of the accelerator. The minimum value of the subsystem locking bandwidths is defined as the common locking bandwidth f_{lock} . We assume that the MO phase noise at frequencies lower than f_{lock} is not relevant for the beam acceleration according to the example in Fig. 6.36.
- For each RF station, we specify the upper limit of the RMS jitter ($\Delta\varphi_{ref,rms}$) of the RF reference phase noise $\Delta\varphi_{ref}$ from the RF station's total noise budget.
- We split the tolerance of $\Delta\varphi_{ref}$ of each RF station into two parts: one for the MO phase noise contribution (denoted as $\Delta\varphi_{MO,res}$), and the other for the phase noise

Fig. 6.36 Bode plots of the transfer functions from the MO phase noise to the beam current (G_b), cavity voltage (G_{rf}) and the beam phase ($G_{rf} - G_b$)



added by the RF reference receiver (denoted as $\Delta\varphi_{ref,add}$). We assume that these two parts are uncorrelated, so their RMS values satisfy

$$\Delta\varphi_{ref,rms} = \sqrt{\Delta\varphi_{MO,res,rms}^2 + \Delta\varphi_{ref,add,rms}^2}. \quad (6.85)$$

We call $\Delta\varphi_{MO,res}$ the *residual MO phase noise*. Its RMS value includes the MO phase noise within a frequency range from f_{lock} to the RF station's closed-loop bandwidth ($f_U = \omega_B/(2\pi)$). In contrast, the RMS value of $\Delta\varphi_{ref,add}$ includes the phase noise introduced by the RF reference receiver in the frequency range from f_L to f_U as defined in Sect. 6.4.1. This is because the $\Delta\varphi_{ref,add}$ of different RF stations are uncorrelated.

- (d) If the $\Delta\varphi_{MO,res}$ specification is in the form of RMS value, we often convert it to an average PSD within the frequency range between f_{lock} and f_U :

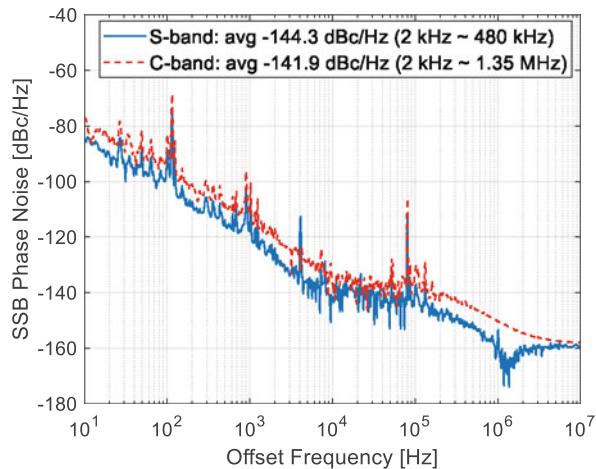
$$S_{\varphi,MO,res} = \Delta\varphi_{MO,res,rms}^2 / (f_U - f_{lock}). \quad (6.86)$$

Let us still use the SwissFEL as an example to demonstrate the specification of the MO phase noise. The phase stability requirements of the RF stations have been defined in Table 6.2 in terms of RMS jitter: $\Delta\varphi_{SB,rms} = 0.015^\circ$, $\Delta\varphi_{CB,rms} = 0.03^\circ$ and $\Delta\varphi_{XB,rms} = 0.05^\circ$. In the discussion below, we will remove _{SB}, _{CB} and _{XB} from the subscripts so that the results can be applied to all RF stations. As mentioned before, the RF station jitter budget is allocated to the RF reference signal, the RF driving chain and the RF measurement chain. We define that the allowed RF reference phase jitter is $\Delta\varphi_{ref,rms} = 0.3\Delta\varphi_{rms}$ ($\Delta\varphi_{rms}$ is the allowed phase jitter of the RF station). We further divide $\Delta\varphi_{ref,rms}$ into $\Delta\varphi_{ref,add,rms}$ and $\Delta\varphi_{MO,res,rms}$. Here we assume $\Delta\varphi_{ref,add,rms} = 0.3\Delta\varphi_{ref,rms}$ and $\Delta\varphi_{MO,res,rms}$ can be calculated with (6.85). Then we can calculate the desired MO phase noise PSD (DSB) between f_{lock} and f_U using (6.86). The results are summarized in Table 6.3. In the calculation, we have chosen $f_{lock} = 2$ kHz that is the bandwidth of the Gun laser oscillator locking

Table 6.3 Specification of the MO phase noise for SwissFEL

	S-band	C-band	X-band
Closed-loop bandwidth (f_U , kHz)	480	1350	4220
Allowed RF ref. phase jitter ($\Delta\varphi_{ref,rms}$, deg)	0.0045	0.009	0.015
Receiver added phase jitter ($\Delta\varphi_{ref,add,rms}$, deg)	0.0013	0.0027	0.0045
Residual MO phase jitter ($\Delta\varphi_{MO,res,rms}$, deg)	0.0043	0.0086	0.0143
Max MO phase noise ($S_{\varphi,MO,res}$, dB Br^2/Hz)	-139.3	-137.8	-138.3

Fig. 6.37 Phase noise of the S-band and C-band reference signals (measurement of the X-band reference is not available)



to the MO. The laser pulses are used in a photocathode RF Gun to generate electron bunches. The closed-loop bandwidths of the RF stations are also listed in Table 6.3.

Figure 6.37 shows the measurements of the S-band and C-band reference signals at the SwissFEL. The average phase noise PSDs between f_{lock} and f_U have been calculated. Note that the measurement results are SSB phase noise, so we need to add 3 dB to obtain the DSB phase noise values. Compared to the MO phase noise requirements in Table 6.3 (as DSB phase noise), we can conclude that the requirements are satisfied. For the MO phase noise below f_{lock} , the requirements may be relaxed because its effects on beam acceleration are canceled in different subsystems.

In this section, we introduced a simplified approach to specifying the MO phase noise. We have assumed that the MO-caused phase jitter for different RF stations is uncorrelated beyond the common locking bandwidth. This assumption simplifies our calculation but may be too optimistic. Consider again the SwissFEL. Above f_{lock} , all the RF stations of the same type (S-band or C-band) still suffer from correlated phase jitter resulting from the MO phase noise. That is, even if $\Delta\varphi_{beam}$ is no longer correlated with $\Delta\varphi_{rf}$ at frequencies above f_{lock} , the beam is still exposed to correlated phase jitter from multiple RF stations. Therefore, the MO phase noise effects between f_{lock} to f_U on beam acceleration with multiple RF stations may be stronger than the uncorrelated case. For example, we have obtained in Table 6.3 that the

$\Delta\varphi_{MO,res,rms}$ for a C-band RF station should be smaller than 0.0086° RMS. Let us consider Linac1 consisting of 9 C-band RF stations. If we follow our assumption that $\Delta\varphi_{MO,res}$ is uncorrelated for different RF stations, the overall beam phase jitter caused by the residual MO phase noise in Linac1 is $0.0086^\circ/3 = 0.0029^\circ$ RMS. However, if we consider the correlation between the residual MO phase noise in different RF stations, the resulting overall beam phase jitter should be larger (but less than 0.0086° RMS). Based on this consideration, we need tighter MO phase noise requirements compared to that shown in Table 6.3. However, the method introduced in this section will still provide a good estimate for the phase noise requirements for the master oscillator.

6.5 RF Station Noise Model

In the previous section, we have introduced a strategy to analyze and specify the RF system noise in an accelerator. A global noise model is used to derive the RF station noise specification based on the beam stability requirements. In this section, we will work out a detailed noise model for an RF station (Hoffmann 2008), which will be used to specify the requirements of each RF component in the RF station.

6.5.1 RF Station Noise Overview

The amplitude and phase noise added by an RF station originates from both the High-Power RF (HPRF) and LLRF components. The noise transfers to the RF field following different paths according to the location of its source in the RF station. Figure 1.13 shows a typical RF station controlled by an LLRF system, which detects the RF field in the cavity, regulates its amplitude and phase and generates drive signals to the HPRF system.

As mentioned before, we will discuss the noise added by an RF station in three aspects: the RF reference (phase) noise, the driving chain noise, and the measurement chain noise. Their contributions to the overall RF station noise are given by (6.77).

In Sect. 6.4.3, we have discussed the RF reference phase noise, which is a superposition of the residual MO phase noise and the phase noise added by the RF reference receiver. The RF reference phase noise within the RF station's closed bandwidth will be transferred to the RF field.

The RF driving chain noise consists of the contributions from:

- *RF actuator (DACs + I/Q modulator)*: the RF actuator transfers the RF reference phase noise to its output and adds amplitude and phase noise. Figure 1.13 uses the simple direct up-conversion scheme, where the DACs generate baseband signals and modulate the RF reference signal via an I/Q modulator. Both the DACs and

- the I/Q modulator (consisting of RF power splitter/combiner and mixers) will introduce additional amplitude and phase noise. See Sects. 6.2.2, 6.2.4 and 6.2.7.
- *RF Amplifier and Klystron*: the amplifier and klystron transfer the input noise within their bandwidths and add additional amplitude and phase noise. See Sect. 6.2.3.
 - *RF power transmission line*: the RF power transmission line, which is usually made from coaxial cables or waveguides, is passive and mainly introduces phase drift due to temperature changes.
 - *RF cavity or structure*: the standing-wave cavity or traveling-wave structure is passive and filters the amplitude and phase noise in the RF input. It also introduces amplitude and phase fluctuations from external disturbances like mechanical vibrations or cooling-water temperature fluctuations.

The RF driving chain noise will be suppressed by the LLRF feedback loop within the closed-loop bandwidth of the RF station.

The RF measurement chain noise consists of the contributions from:

- *RF pickup cable*: the RF pickup cable introduces slow phase and amplitude drifts.
- *RF detector (mixer + ADC)*: the mixer down-converts the RF frequency to IF with respect to the LO frequency. The IF signal is then sampled by an ADC. With the down-conversion and sampling process, the LO and clock phase noise is transferred to the ADC samples. Of course, the mixer and ADC also introduce uncorrelated amplitude and phase noise, reducing the RF detection resolution. See Sects. 6.2.4 and 6.2.6.
- *LO and clock generator*: the LO and clock signals are generated from the RF reference signal and inherit its phase noise within the corresponding bandwidths. Moreover, the LO and clock generator adds additional phase noise to the LO and clock signals. The amplitude noise of the LO and clock signals are usually not relevant for the RF stability because the LO inputs of mixers and the clock inputs of ADCs and DACs often operate in saturation. See Sects. 6.2.4 and 6.2.5.

The low-frequency RF measurement noise (i.e., amplitude and phase drifts) within the RF station's closed-bandwidth will be transferred to the RF field. See Fig. 6.27. Therefore, the drifts in the RF measurement chain are critical for the long-term stability of the RF station.

6.5.2 RF Reference Phase Noise

As discussed in Sect. 6.4.3, the RF reference phase noise, $\Delta\varphi_{ref}$, is a superposition of the residual MO phase noise, $\Delta\varphi_{MO,res}$, and the phase noise added by the RF reference receiver, $\Delta\varphi_{ref,add}$. More specifically, we use $\Delta\varphi_{ref,add}$ to model the added phase noise by the entire distribution chain from the MO to the RF reference signal, including the phase drift in the signal distribution line and the phase noise added by the transmitter and receiver of the synchronization signal. The phase drift

in the synchronization system will be discussed in Chap. 8. Assuming that these two parts are uncorrelated, we obtain the following equations for the time-domain values and the noise PSDs.

$$\begin{aligned}\Delta\varphi_{ref}(t) &= \Delta\varphi_{MO,res}(t) + \Delta\varphi_{ref,add}(t), \\ S_{\varphi,ref}(f_m) &= S_{\varphi,MO,res}(f_m) + S_{\varphi,ref,add}(f_m),\end{aligned}\quad (6.87)$$

where f_m is the offset frequency with respect to the carrier frequency. Later we will neglect the independent variables t and f_m once the variables' meanings are clear. Note that $S_{\varphi,MO,res}$ is defined from f_{lock} to f_U , and $S_{\varphi,ref,add}$ is defined from f_L to f_U . See Sect. 6.4.3. Typically, we have $f_{lock} > f_L$.

6.5.3 RF Driving Chain Noise

The RF driving chain components add uncorrelated amplitude and phase noise that can be described in the time domain as

$$\begin{aligned}n_{drv} = n_{dac,add} + n_{iqm,add} + n_{amp,add} + \\ n_{kly,add} + n_{tline,add} + n_{cav,add}, n \text{ for } \alpha \text{ or } \Delta\varphi.\end{aligned}\quad (6.88)$$

Here $n_{dac,add}$, $n_{iqm,add}$, $n_{amp,add}$, $n_{kly,add}$, $n_{tline,add}$ and $n_{cav,add}$ represent the amplitude or phase noise added by the DAC, I/Q modulator, amplifier, klystron, power transmission line and cavity or structure, respectively. We will use the subscript cav to denote both the standing-wave cavities and travelling-wave structures. If the terms above are not correlated, their PSDs will satisfy

$$\begin{aligned}S_{n,drv} = S_{n,dac,add} + S_{n,iqm,add} + S_{n,amp,add} + \\ S_{n,kly,add} + S_{n,tline,add} + S_{n,cav,add}, n \text{ for } \alpha \text{ or } \varphi.\end{aligned}\quad (6.89)$$

The terms above are defined at the offset frequency range $f_L < f_m < f_U$, where f_L and f_U are defined in Sect. 6.4.1.

In some RF stations, the amplifier or klystron operates in saturation. This configuration is quite often adopted in normal-conducting Linacs, such as the LCLS and SwissFEL. In this case, the amplitude noise in the driving chain does not follow the relations above because the saturated components will suppress the amplitude jitter in their inputs. However, we can still apply the above results to the phase noise, even with a saturated amplifier or klystron.

The overall RF driving chain noise is filtered by the closed-loop transfer function \mathbf{T}_{drv} and contributes to the cavity field according to (6.77). Note that the noise discussed above does not include the amplitude and phase errors caused by repetitive disturbances (e.g., Lorenz-force detuning and beam loading in a superconducting cavity operating in the pulsed mode), which can be suppressed with feedforward controls. Only random noise is considered here.

6.5.4 RF Measurement Chain Noise

The RF measurement chain components include the LO and clock generator, RF pickup cable, down-conversion mixer, and ADC. They all belong to LLRF.

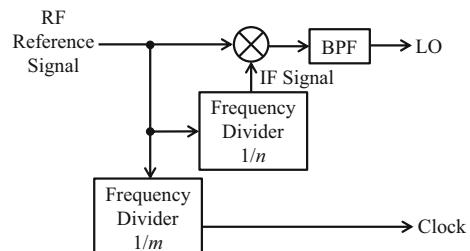
The LO and clock generator transfers the input RF reference phase noise within the locking bandwidths and adds additional phase noise to both the LO and clock signals. A typical LO and clock generator is shown in Fig. 6.38, where we use frequency dividers to derive the clock and IF frequencies. The IF frequency is then mixed with the reference signal, and the lower sideband of the mixing product is extracted as the LO signal. Usually, a narrow band-pass filter (e.g., cavity filter) is needed here, which introduces more low-frequency phase fluctuations due to its sensitivity to the environmental temperature. Using the noise transfer relations of the frequency divider and mixer, the time-domain phase noise in the LO and clock signals can be calculated as

$$\begin{aligned}\Delta\varphi_{LO}(t) &= \frac{f_{LO}}{f_{ref}} \Delta\varphi_{ref}(t - \Delta t_{LO}) + \Delta\varphi_{LO,add}(t), \\ \Delta\varphi_{CLK}(t) &= \frac{f_{CLK}}{f_{ref}} \Delta\varphi_{ref}(t - \Delta t_{CLK}) + \Delta\varphi_{CLK,add}(t),\end{aligned}\quad (6.90)$$

where f_{LO} , f_{CLK} and f_{ref} are the frequencies of the LO, clock and reference signals, respectively. It shows that $\Delta\varphi_{LO}$ and $\Delta\varphi_{CLK}$ have a correlated part coming from the common input $\Delta\varphi_{ref}$ with different delays (Δt_{LO} and Δt_{CLK}). The terms $\Delta\varphi_{LO,add}$ and $\Delta\varphi_{CLK,add}$ are the added phase noise by the LO and clock generator contributing from the frequency dividers, mixers, and amplifiers (not shown in Fig. 6.38). They are often partly correlated because the LO and clock paths may share some common frequency dividers or amplifiers. Nevertheless, we will assume them as uncorrelated to simplify the discussion. The overall phase noise introduced by the LO and clock signals when they are used for RF signal detection can be calculated as

$$\begin{aligned}\Delta\varphi_{LOCLK}(t) &= -\Delta\varphi_{LO}(t) - \frac{f_{IF}}{f_{CLK}} \Delta\varphi_{CLK}(t) \\ &\approx -\Delta\varphi_{ref}(t - \Delta t_{LO}) - \Delta\varphi_{LO,add}(t) - \frac{f_{IF}}{f_{CLK}} \Delta\varphi_{CLK,add}(t).\end{aligned}\quad (6.91)$$

Fig. 6.38 Block diagram of a simplified LO and clock generator



The minus sign comes from the RF detection principle. For the approximation, we have neglected the difference between the clock and LO delays. The frequency relation $f_{ref} = f_{LO} + f_{IF}$ has been used, where f_{IF} is the IF frequency. In the equation above, the term $-\Delta\varphi_{ref}(t - \Delta t_{LO})$ describes the coupling of the RF reference phase noise into the measurement chain via the LO and clock generator with a delay Δt_{LO} . This leads to a more precise closed-loop transfer function for the RF reference phase noise for (6.76) and (6.80):

$$\mathbf{T}_{ref}(\hat{s}) = \mathbf{T}_{drv}(\hat{s}) - \mathbf{T}_{mea}(\hat{s})e^{-\hat{s}\Delta t_{LO}}. \quad (6.92)$$

The delay Δt_{LO} is generally small compared to the time-constant of the standing-wave cavity or the filling time of the traveling-wave structure. That is, within the closed-loop bandwidth of the RF feedback loop, the phase lag caused by Δt_{LO} can usually be neglected. For example, if $\Delta t_{LO} = 10$ ns and the closed-loop bandwidth of the cavity is 500 kHz, then the maximum phase lag caused by the term $e^{-\hat{s}\Delta t_{LO}}$ is about -1.8° , which can be neglected when calculating \mathbf{T}_{ref} . Therefore, to simplify our following discussion, we will neglect the effects of Δt_{LO} .

The noise added by the RF measurement chain comes from the LO and clock generator, RF pickup cable, mixer, and ADC. Their time-domain expressions are

$$\begin{aligned} \Delta\varphi_{mea} &= -\Delta\varphi_{LO,add} - \frac{f_{IF}}{f_{CLK}}\Delta\varphi_{CLK,add} + \Delta\varphi_{cab,add} \\ &\quad + \Delta\varphi_{mix,add} + \Delta\varphi_{adc,add}, \\ \alpha_{mea} &= \alpha_{cab,add} + \alpha_{mix,add} + \alpha_{adc,add}, \end{aligned} \quad (6.93)$$

where the subscripts $_{cab}$, $_{mix}$ and $_{adc}$ denote the noise added by the RF pickup cable, mixer, and ADC, respectively. Note that we have excluded the term $\Delta\varphi_{ref}$ in (6.91) from the calculation of $\Delta\varphi_{mea}$ because $\Delta\varphi_{ref}$ has been considered separately in our noise model. With the direct up-conversion scheme, the LO and clock added phase noise ($\Delta\varphi_{LO,add}$ and $\Delta\varphi_{CLK,add}$) only contributes to $\Delta\varphi_{mea}$. However, if the IF or SSB up-conversion scheme is used, $\Delta\varphi_{LO,add}$ and $\Delta\varphi_{CLK,add}$ will also appear in the driving chain. The RF pickup cable contributes to the low-frequency phase drift ($\Delta\varphi_{cab,add}$) and (minor) gain drift ($\alpha_{cab,add}$) due to temperature changes. The ADC and mixer will mainly generate wide-band phase and amplitude noise ($\Delta\varphi_{mix/adc,add}$ and $\alpha_{mix/adc,add}$). The noise items in (6.93) are considered uncorrelated so the noise PSDs can be written as

$$\begin{aligned} S_{\varphi,mea} &= S_{\varphi,LO,add} + \left(\frac{f_{IF}}{f_{CLK}}\right)^2 S_{\varphi,CLK,add} + S_{\varphi,cab,add} \\ &\quad + S_{\varphi,mix,add} + S_{\varphi,adc,add}, \\ S_{\alpha,mea} &= S_{\alpha,cab,add} + S_{\alpha,mix,add} + S_{\alpha,adc,add}. \end{aligned} \quad (6.94)$$

The terms above are defined at the offset frequency range $f_L < f_m < f_U$, where f_L and f_U are defined in Sect. 6.4.1.

The overall noise added in the RF measurement chain will be transferred to the RF field via the closed-loop transfer function \mathbf{T}_{mea} , as in (6.77).

6.5.5 Estimation of RF Measurement Chain Noise

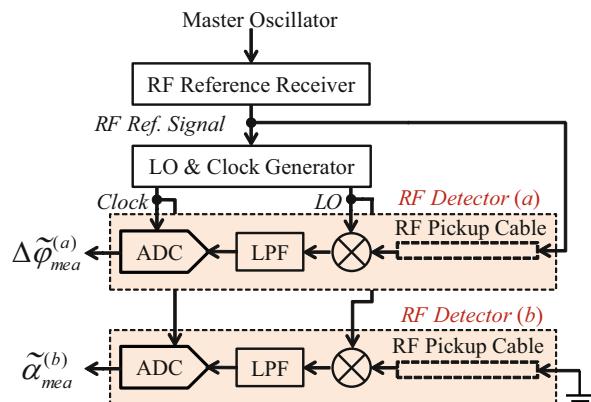
The phase noise $\Delta\varphi_{mea}$ added by the RF measurement chain can be estimated by measuring the RF reference signal with an independent RF detector channel. See the channel (a) in Fig. 6.39. Similarly, we can estimate α_{mea} by terminating the RF detector input with a matched load, as shown by the channel (b) in Fig. 6.39. We will use a symbol “~” above the variable to indicate that it is a measurement or estimate.

We may route the RF reference signal through the same type of cable (with the same length and in the same routing path with the same environmental conditions) that picks the RF field. It helps estimate the added phase noise by the RF detector electronics and the phase drift introduced by the RF pickup cables. The phase noise measurement of the RF reference signal can be written as

$$\begin{aligned}\Delta\tilde{\varphi}_{ref} = & -\Delta\varphi_{LO,add} - \frac{f_{IF}}{f_{CLK}} \Delta\varphi_{CLK,add} + \Delta\varphi_{cab,add}^{(a)} \\ & + \Delta\varphi_{mix,add}^{(a)} + \Delta\varphi_{adc,add}^{(a)} \\ \approx & \Delta\tilde{\varphi}_{mea}^{(a)}.\end{aligned}\quad (6.95)$$

Here the superscript ^(a) denotes that the noise is added by the RF detector channel (a). Compared to (6.93), the phase measurement $\Delta\tilde{\varphi}_{ref}$ is a good estimate of $\Delta\varphi_{mea}$. However, the measured amplitude noise of the RF reference signal, $\tilde{\alpha}_{ref} = \alpha_{ref} + \alpha_{cab,add}^{(a)} + \alpha_{mix,add}^{(a)} + \alpha_{adc,add}^{(a)}$, contains both the RF reference amplitude noise α_{ref} and that added by the RF detector α_{mea} . To estimate α_{mea} , we may use the

Fig. 6.39 Estimation of the noise added by the RF measurement chain



RF detector channel (*b*) as in Fig. 6.39, where we denote the measurement results as $x^{(b)}(t)$. Since the input port is terminated, $x^{(b)}$ is the white noise added by the entire RF measurement chain. From the discussion in Sect. 6.1.8, half of the power in $x^{(b)}$ will be allocated to the amplitude noise. Therefore, we can estimate α_{mea} as

$$\tilde{\alpha}_{mea}^{(b)} = \frac{x^{(b)}/\sqrt{2}}{A}, \quad (6.96)$$

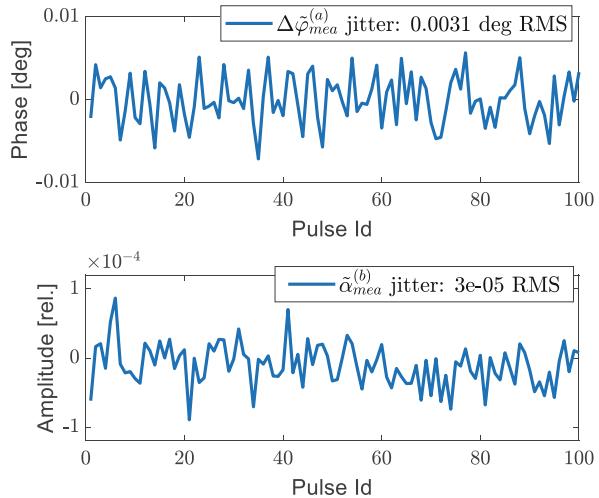
where A is the amplitude of the RF signal to be measured by the RF detector. Let us use an example to demonstrate the method above to estimate the noise added by the RF measurement chain.

In this example, we will use the data collected from the second S-band RF station in the SwissFEL injector. It operates in the pulsed mode at a repetition rate of $f_{pul} = 100$ Hz with a pulse width of about 920 ns (the filling time of the traveling-wave structure). See Sect. 6.4.2.1 for more details of the SwissFEL. To simplify the discussion, we only consider the amplitude and phase noise at offset frequencies higher than $f_L = 1$ Hz. That is, we focus on the short-term jitter of the RF field. The long-term drift (< 1 Hz) is more difficult to analyze because the drifts in the RF reference distribution system, RF pickup cables and RF detectors are not easy to measure. Furthermore, the effects of the RF noise slower than 1 Hz can be compensated for by the beam-based feedback system.

The amplitude and phase of each pulse are measured by averaging within a time window of 920 ns. It corresponds to limiting the measurement bandwidth below about 480 kHz. See the discussion in Sect. 6.3.2, Fig. 6.30. Furthermore, to achieve the lower noise-frequency bound of 1 Hz, we will calculate the amplitude and phase RMS jitter using the measurements of 50 RF pulses. We have determined this number of pulses for the calculation of RMS values according to the discussions in Sect. 6.1.9, where $f_s = f_{pul} = 100$ Hz in our case. The pulse-to-pulse amplitude and phase sequences will alias the RF noise power within the measurement bandwidth (480 kHz) into the Nyquist zone (0 Hz to 50 Hz) of the pulse repetition rate. Then the RMS value will include the aliased noise power between $f_L = 1$ Hz and $f_U = f_{pul}/2 = 50$ Hz (equivalently, the RMS value includes the unaliased noise power from 1 Hz to 480 kHz approximately).

The measurement results of $\Delta\tilde{\varphi}_{mea}^{(a)}$ and $\tilde{\alpha}_{mea}^{(b)}$ following the method in Fig. 6.39 are shown in Fig. 6.40. We have assumed that A is 85% of the ADC full scale to calculate the relative amplitude jitter. Although the measurements for 100 pulses are shown, the RMS values were calculated with 50 pulses. Note that an RMS value is calculated for every 50 pulses, and the results are averaged to derive the RMS values displayed in the plots. Compared to the S-band RF field stability requirements specified in Table 6.2 ($\alpha_{SB,rms} < 1e-4$ and $\Delta\varphi_{SB,rms} < 0.015^\circ$), the noise added by the RF measurement chain is negligible.

Fig. 6.40 Estimation of the noise added by the RF measurement chain for an S-band RF station of SwissFEL. The concerned frequencies are:
 $f_{ref} = f_{RF} = 2998.8$ MHz,
 $f_{LO} = 3040.5$ MHz,
 $f_{IF} = 41.65$ MHz and
 $f_{CLK} = 249.9$ MHz



6.5.6 Estimation of RF Driving Chain Noise

To estimate the noise added by the RF driving chain, a third independent RF detector channel (*c*) has been introduced to measure the RF field signal picked by a probe installed in the cavity. See Fig. 6.41. We keep the channel (*a*) to measure the RF reference signal and the channel (*b*) to measure the white noise. They are used to estimate the RF measurement noise, which should be excluded from the channel (*c*) results.

The channel (*c*) measures the RF field in the cavity or structure and the results can be written as

$$\begin{aligned}\Delta\tilde{\varphi}_{cav} &= \Delta\varphi_{act} + \Delta\varphi_{drv} + \Delta\varphi_{mea}^{(c)}, \\ \tilde{\alpha}_{cav} &= \alpha_{act} + \alpha_{drv} + \alpha_{mea}^{(c)},\end{aligned}\quad (6.97)$$

where $\Delta\varphi_{act}$ and α_{act} are the phase and amplitude actuation signals from the LLRF controller. We used the superscript ^(*c*) to denote the measurement noise in the RF detector channel (*c*). Note that $\Delta\varphi_{ref}$ affects both the driving chain and the measurement chain with opposite signs, so it does not show up in the measurement result. The reference input of the I/Q modulator is saturated; therefore, α_{ref} will not affect the amplitude noise in the RF field. Note that $\Delta\varphi_{act}$ and α_{act} are known, which can be read from the RF controller. We have also the estimates of $\Delta\tilde{\varphi}_{mea}^{(a)}$ and $\tilde{\alpha}_{mea}^{(b)}$ from the channel (*a*) and (*b*) according to the discussion in the previous section. Clearly, if the noise of all the RF detectors is identical, that is, $\Delta\varphi_{mea}^{(c)} = \Delta\tilde{\varphi}_{mea}^{(a)}$ and $\alpha_{mea}^{(c)} = \tilde{\alpha}_{mea}^{(b)}$, the exact values of $\Delta\varphi_{drv}$ and α_{drv} can be calculated from (6.97). However, in practice, the RF detector noise in different channels is only partially correlated. The correlation comes from the commonly added phase noise by the LO and clock

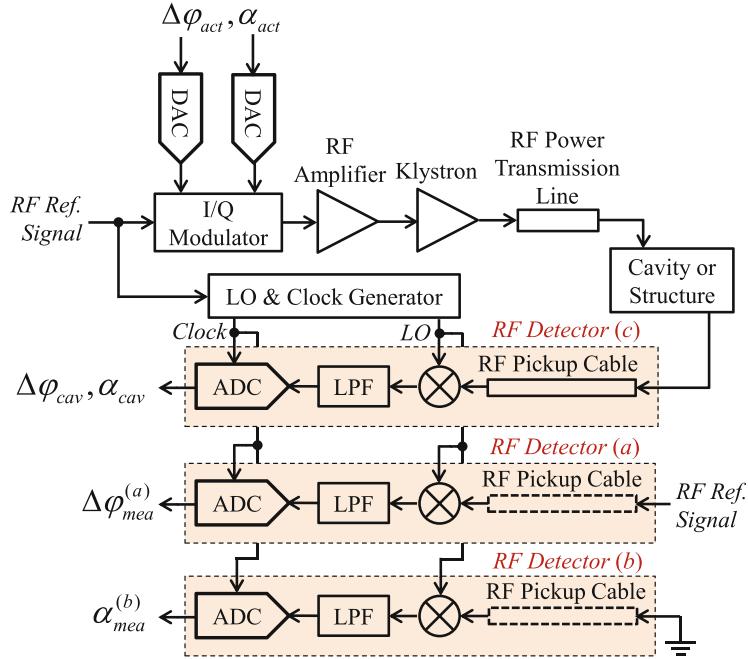


Fig. 6.41 Estimation of the noise added by the RF driving chain

generator and the temperature changes in the RF detector chassis. The white noise added by different channels is uncorrelated. Therefore, in principle, the phase noise added by the RF detectors are more correlated than the added amplitude noise. Typically, we use the following algorithm to estimate the noise added by the RF driving chain:

$$\begin{aligned}\Delta\tilde{\varphi}_{drv} &= \Delta\tilde{\varphi}_{cav} - \Delta\varphi_{act} - \text{LPF}\{\Delta\tilde{\varphi}_{mea}^{(a)}\}, \\ \tilde{\alpha}_{drv} &= \tilde{\alpha}_{cav} - \alpha_{act} - \text{LPF}\{\tilde{\alpha}_{mea}^{(b)}\},\end{aligned}\quad (6.98)$$

where $\text{LPF}\{\cdot\}$ represents the low pass filtering function. Compared to (6.97), we have replaced $\Delta\varphi_{mea}^{(c)}$ with $\Delta\tilde{\varphi}_{mea}^{(a)}$, and $\alpha_{mea}^{(c)}$ with $\tilde{\alpha}_{mea}^{(b)}$. The low-pass filtered noise reflects the slow drifts in the RF measurement chain and is typically correlated in all RF detector channels. Nevertheless, in some cases, e.g., when we only consider the high-frequency noise, the noise added by different RF detector channels is uncorrelated. Then (6.98) is no longer applicable; that is, we cannot subtract the estimated measurement noise of other channels from the time-domain measurement of the cavity pickup signal. However, in this case, we may still estimate the RMS jitter of the noise added by the RF driving chain as

$$\begin{aligned}\Delta\tilde{\varphi}_{drv,rms}^2 &= \text{RMS}^2\{\Delta\tilde{\varphi}_{cav} - \Delta\varphi_{act}\} - \text{RMS}^2\left\{\Delta\tilde{\varphi}_{mea}^{(a)}\right\}, \\ \tilde{\alpha}_{drv,rms}^2 &= \text{RMS}^2\{\tilde{\alpha}_{cav} - \alpha_{act}\} - \text{RMS}^2\left\{\tilde{\alpha}_{mea}^{(b)}\right\},\end{aligned}\quad (6.99)$$

where $\text{RMS}\{\}$ returns the RMS value of the item in the bracket. Figure 6.42 is an example of the estimate of the RF driving chain noise. We have used the same data of the SwissFEL S-band RF station as in Sect. 6.5.5. The RMS jitter is calculated with the data of 50 RF pulses, which contains the noise power (not aliased by the sampling at the RF pulse rate) from 1 Hz up to the bandwidth of the traveling-wave structure (480 kHz). Since above 1 Hz, the RF detector noise of different channels are almost uncorrelated, we have used (6.99) to estimate the RMS jitter of $\Delta\tilde{\varphi}_{drv}$ and $\tilde{\alpha}_{drv}$. The jitter added by the RF driving chain is comparable to but still smaller than the tolerances ($\alpha_{SB,rms} < 1\text{e-}4$ and $\Delta\varphi_{SB,rms} < 0.015^\circ$), so this RF station is promising to satisfy the requirements. We will verify this in the next section.

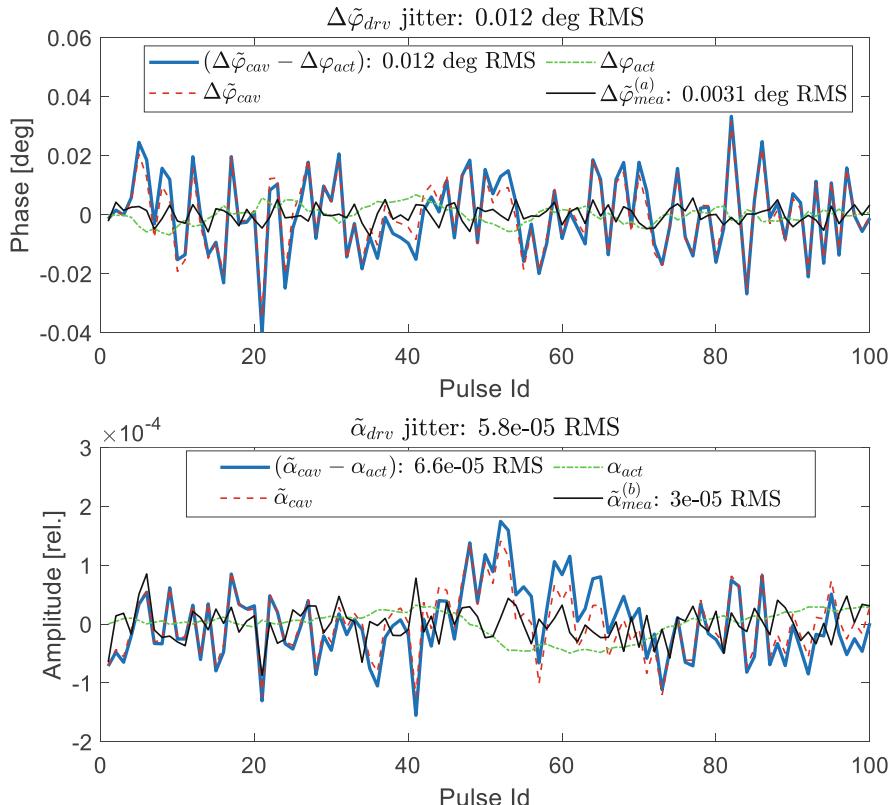


Fig. 6.42 Estimation of the noise added by the RF driving chain for the S-band RF station of SwissFEL

6.5.7 Estimation of RF Field Noise

In Fig. 6.41, we have used the RF detector channel (c) to measure the RF field in the cavity or structure. The question is, what is the relationship between the measurement results ($\Delta\tilde{\varphi}_{cav}, \tilde{\alpha}_{cav}$) and the actual phase and amplitude fluctuations ($\Delta\varphi_{field}, \alpha_{field}$) in the RF field?

The measurement results have been given in (6.97). Similarly, $\Delta\varphi_{field}$ and α_{field} can be calculated as

$$\Delta\varphi_{field} = \Delta\varphi_{ref} + \Delta\varphi_{act} + \Delta\varphi_{drv}, \alpha_{field} = \alpha_{act} + \alpha_{drv}. \quad (6.100)$$

Compared to (6.97), we obtain

$$\begin{aligned} \Delta\tilde{\varphi}_{cav} &= \Delta\varphi_{field} - \Delta\varphi_{ref} + \Delta\varphi_{mea}^{(c)}, \\ \tilde{\alpha}_{cav} &= \alpha_{field} + \alpha_{mea}^{(c)}, \end{aligned} \quad (6.101)$$

which implies that the phase measurement $\Delta\tilde{\varphi}_{cav}$ cannot detect the RF field phase error caused by $\Delta\varphi_{ref}$ via the RF driving chain. That is, $\Delta\tilde{\varphi}_{cav}$ measures the fluctuations of the relative phase between the RF field and the RF reference signal. Therefore, if the phase feedback loop is based on $\Delta\tilde{\varphi}_{cav}$, it cannot suppress $\Delta\varphi_{ref}$, which agrees with our closed-loop noise transfer eqs. (6.76) and (6.80). For convenience, we define a new variable

$$\delta\varphi_{field} = \Delta\varphi_{field} - \Delta\varphi_{ref} \quad (6.102)$$

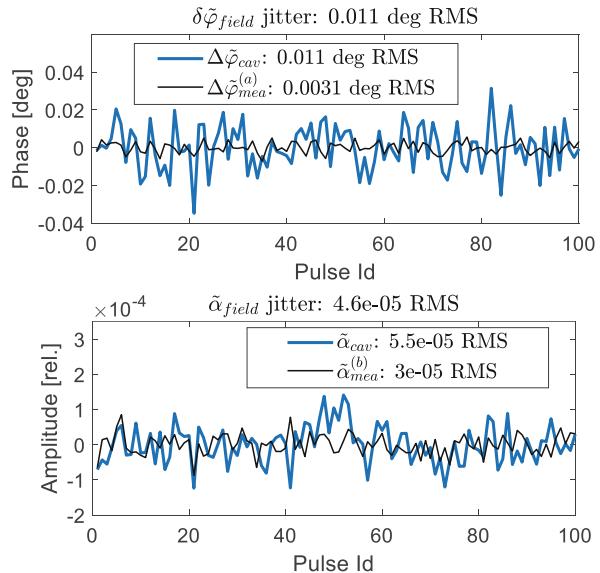
for the RF field phase noise relative to the RF reference phase noise. Like the previous section, we can use the following formula to estimate the RF field amplitude and phase variations in the cavity or structure:

$$\delta\tilde{\varphi}_{field} = \Delta\tilde{\varphi}_{cav} - \text{LPF}\left\{\Delta\tilde{\varphi}_{mea}^{(a)}\right\}, \tilde{\alpha}_{field} = \tilde{\alpha}_{cav} - \text{LPF}\left\{\tilde{\alpha}_{mea}^{(b)}\right\}. \quad (6.103)$$

$$\begin{aligned} \delta\tilde{\varphi}_{field,rms}^2 &= \Delta\tilde{\varphi}_{cav,rms}^2 - \text{RMS}^2\left\{\Delta\tilde{\varphi}_{mea}^{(a)}\right\}, \\ \tilde{\alpha}_{field,rms}^2 &= \tilde{\alpha}_{cav,rms}^2 - \text{RMS}^2\left\{\tilde{\alpha}_{mea}^{(b)}\right\}. \end{aligned} \quad (6.104)$$

Note that (6.103) is used to estimate the time-domain values of $\delta\varphi_{field}$ and α_{field} . The low-pass filters select the frequency range where the phase and amplitude noise of different RF detector channels are correlated. If they are uncorrelated, we can still use (6.104) to estimate the RMS jitter of the RF field in the cavity or structure.

Fig. 6.43 Estimation of RF field jitter for the S-band RF station of SwissFEL



We estimated the RF field jitter in the traveling-wave structure of the SwissFEL S-band RF station using the same data as shown in Fig. 6.42. The results are depicted in Fig. 6.43. The amplitude and phase jitter of the RF field is estimated to be 4.6e-5 RMS and 0.011° RMS, respectively. Same as before, we considered the noise power (not aliased by the sampling at the RF pulse rate) between 1 Hz and the bandwidth of the traveling-wave structure (480 kHz).

6.5.8 Validation of RF Station Noise Model

To validate the closed-loop noise transfer eq. (6.77), we calculate the PSDs of the noise added by the RF driving chain ($S_{\alpha,drv}$ and $S_{\varphi,drv}$) and the RF measurement chain ($S_{\alpha,mea}$ and $S_{\varphi,mea}$). We have used the same data as that has been used in Sects. 6.5.5, 6.5.6 and 6.5.7. See the results in Fig. 6.44. Because the data covers 30 minutes, we can see the amplitude and phase noise spectra at lower frequencies. The 100 Hz data results in PSDs defined between 0 Hz and 50 Hz. In our calculation, we have assumed a discrete integral feedback controller, described by (6.78), with $k = 0.15$. To calculate the noise PSDs transferred to the RF field in the structure, we apply the closed-loop transfer functions T_{mea} and T_{drv} to the noise added by the measurement chain and driving chain, respectively. Figure 6.31 has depicted the bode plots of T_{mea} and T_{drv} up to the bandwidth of the traveling-wave structure. When processing the 100 Hz data, we should use the values of T_{mea} and T_{drv} within the first Nyquist zone, i.e., from 0 Hz to 50 Hz. The resulting noise PSDs in the RF

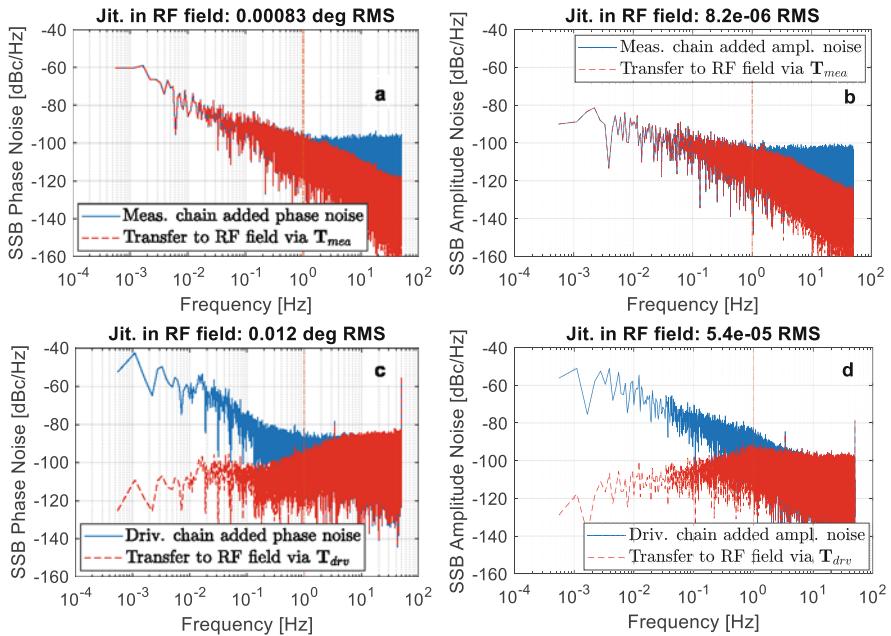


Fig. 6.44 Estimation of the RF field noise using the measurement and driving chain noise. The resulting RF field jitter is calculated by integrating the transferred noise power from 1 Hz to 50 Hz

field contributed from the measurement chain and driving chain are shown in Fig. 6.44.

We also calculate the resulting phase and amplitude jitter in the RF field caused by the noise added by the measurement chain and driving chain. As before, the jitter only considers the noise power above 1 Hz. The RF field jitter caused by the measurement chain noise (0.00083° RMS for phase and $8.2\text{e-}6$ RMS for amplitude) is much smaller than that caused by the driving chain noise (0.012° RMS for phase and $5.4\text{e-}5$ RMS for amplitude). Therefore, we can neglect the high-frequency (> 1 Hz) noise added by the RF measurement chain. Then the overall RF field jitter is approximately the same as that caused by the RF driving chain noise. Compared to the estimate in Fig. 6.43, we obtain slightly higher jitter with the frequency-domain noise transfer functions here. The difference in the amplitude jitter is more significant. It may be caused by the saturation of the klystron, which makes the RF field amplitude jitter less sensitive to the amplitude noise added by the RF driving chain. Therefore, the directly measured RF field amplitude jitter ($4.6\text{e-}5$ RMS) is smaller than that calculated by the linear transfer functions ($5.4\text{e-}5$ RMS).

Figure 6.44a is the phase noise PSD added by the RF detector. If the results of this RF detector are used as inputs to the feedback controller, quite some low-frequency phase noise will transfer to the RF field, as seen from the PSD curve plotted in dashed lines. The low-frequency phase noise in the RF detector comes mainly from the LO and clock generator. Without suppressing the low-frequency noise added by

the RF detector, the RF field stability cannot be guaranteed even if we implement perfect feedback control. In Sect. 6.6, we will introduce a few methodologies to mitigate the low-frequency phase noise, i.e., the drift, in RF detectors.

The RF field jitter estimated in Figs. 6.43 and 6.44 is within the tolerances defined in Table 6.2 ($\alpha_{SB,rms} < 1e-4$ and $\Delta\varphi_{SB,rms} < 0.015^\circ$). However, we still cannot conclude that the stability requirements are finally satisfied by this RF station. Below are the reasons:

- (a) The results in Figs. 6.43 and 6.44 only include the contributions from the noise added by the RF driving chain and measurement chain. The RF reference phase noise $\Delta\varphi_{ref}$ is still unknown. As mentioned before, we cannot measure $\Delta\varphi_{ref}$ with the RF detectors because they use the RF reference signal as the phase reference. Therefore, the actual RF field phase jitter experienced by the beam must be larger than what we have estimated.
- (b) Our jitter estimation only includes the noise power for frequencies higher than 1 Hz. If we consider a longer time, the RMS jitter of the RF field amplitude and phase may be larger. This situation can be seen from the PSDs plotted in dashed lines in Fig. 6.44, representing the noise power transferred to the RF field at different offset frequencies. Let us consider the time scale of the entire data set (around 30 minutes) for the PSD calculation. The resulting RF field jitter from the measurement chain noise becomes 0.0044° RMS for phase and $1.4e-5$ RMS for amplitude. The contribution from the driving chain noise does not change much because the low-frequency noise is suppressed by the feedback. Fortunately, in our case, the overall RF field jitter is still within the tolerances for this longer time scale. Of course, if the beam-based feedback is present, the low-frequency noise effects can be mitigated. In summary, when we specify the stability requirements for an RF station (or an RF component), the frequency range to evaluate the jitter should be mentioned explicitly.

6.5.9 Specification of RF Component Noise

As we have mentioned in Sect. 6.4.1, one of the major use cases of the RF station noise model is to specify the noise budget for each RF component.

Assume that we have been given the overall noise specification of the RF station. The specification can be upper limits of either PSDs or RMS values. An RMS value can also be converted to white noise PSDs within the relevant frequency band (e.g., $f_L < f_m < f_U$). With a particular feedback loop configuration, we can specify the noise allowed in the RF reference signal ($\Delta\varphi_{ref}$), the RF measurement chain ($\Delta\varphi_{mea}$ and α_{mea}) and the RF driving chain ($\Delta\varphi_{drv}$ and α_{drv}). These three noise sources will be transferred to the RF field in the cavity or structure via different closed-loop transfer functions, as (6.77). Since the transfer functions have different frequency responses, we may specify the noise as PSDs in the frequency domain. For example, we can specify the noise added by the RF measurement chain as upper limits of PSDs at different offset frequencies. It allows us to place tighter requirements on the

low-frequency measurement noise according to the frequency response of \mathbf{T}_{mea} in Figs. 6.27 and 6.31. In contrast, the high-frequency driving chain noise is more critical, especially in the pulsed machine. This is because the high-frequency driving chain noise will be amplified by the pulse-to-pulse control (see the bode plot of \mathbf{T}_{drv} in Fig. 6.31). In summary, from this step, we will derive the requirements for $S_{\varphi,ref}$, $S_{\varphi,mea}$, $S_{\alpha,mea}$, $S_{\varphi,drv}$ and $S_{\alpha,drv}$ as functions of frequency. All these noise PSDs are considered in the frequency range $f_L < f_m < f_U$, as defined in Sect. 6.4.1.

Finally, we should breakdown the noise PSDs derived above to their components following (6.87), (6.89) and (6.94). As mentioned before, when defining the noise budget for each RF component, we need to compromise the technical difficulties and cost. For example, to breakdown $S_{\varphi,drv}$ to the driving chain components, we should reserve a little more budget for the klystron because further reducing the klystron phase jitter may be more challenging.

6.6 RF Detector Drift Correction

The analysis in Sect. 6.5.8 illustrates that the low-frequency phase noise in the RF detector is harmful to the RF field's long-term stability. Techniques have been developed to compensate for the RF detector drifts, such as the reference tracking and the drift calibration discussed in this section. We will also briefly introduce the beam-based feedback that corrects the RF field drifts based on the direct observation of the beam parameter fluctuations.

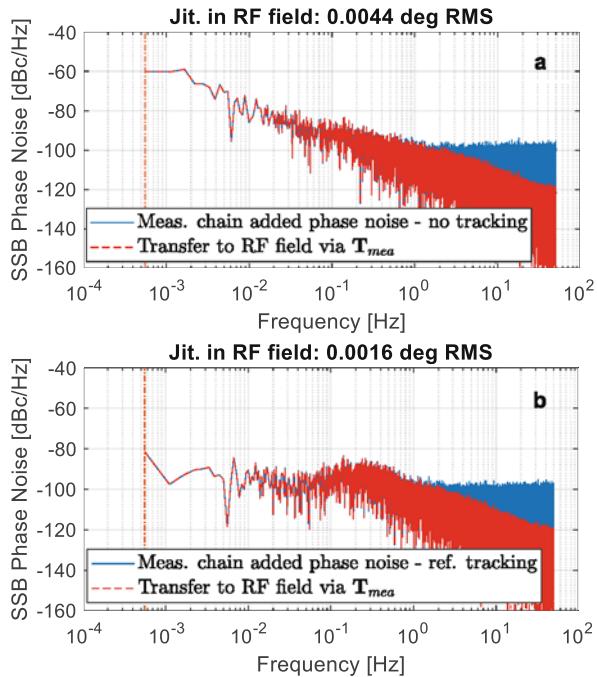
6.6.1 Reference Tracking

We have discussed the reference tracking in Chap. 5. It was used to deal with some RF detection difficulties, such as the time-varying IF frequency, the asynchronous LO or clock, and the phase uncertainties caused by race conditions. Here we will highlight the drift rejection capability of the reference tracking (Doolittle et al. 2006).

To stabilize the RF field in the cavity or structure, we must measure its fluctuations ($\delta\varphi_{field}$ and α_{field} , defined in Sect. 6.5.7) accurately. As shown in Fig. 6.41, if we directly use the output $\Delta\tilde{\varphi}_{cav}$ as the RF field phase measurement, we will suffer from the phase noise added by the RF detector since $\Delta\tilde{\varphi}_{cav} = \delta\varphi_{field} + \Delta\varphi_{mea}^{(c)}$, according to (6.101) and (6.102). Here $\Delta\varphi_{mea}^{(c)}$ is the measurement noise that will be transferred to the cavity field through \mathbf{T}_{mea} . See Fig. 6.45a.

With reference tracking, we use $\delta\tilde{\varphi}_{field}$ given by (6.103) as the RF field phase measurement. It can be rewritten as $\delta\tilde{\varphi}_{field} = \delta\varphi_{field} + \Delta\varphi_{mea}^{(c)} - \text{LPF}\{\Delta\tilde{\varphi}_{mea}^{(a)}\}$; that is, the measurement noise becomes $\Delta\varphi_{mea}^{(c)} - \text{LPF}\{\Delta\tilde{\varphi}_{mea}^{(a)}\}$. The low-pass filter ensures that the uncorrelated phase noise in channel a is removed. Therefore, the

Fig. 6.45 Contribution of the phase noise added by the RF measurement chain to the RF field phase jitter with/without reference tracking. The resulting RF field jitter is calculated by integrating the transferred noise power from 5.56e-04 Hz to 50 Hz



common-mode phase noise (e.g., the phase noise introduced by the LO and clock generator) in these two channels will be canceled. This situation is depicted in Fig. 6.45b. With reference tracking, we can reduce the low-frequency measurement noise and reduce the RF field phase drift. Note that we cannot simply discard the low-pass filter because a direct subtraction of two phases will double the power of the uncorrelated noise (e.g., white noise introduced by the mixers and ADCs).

6.6.2 Drift Calibration

Reference tracking assumes that two RF detector channels suffer from similar phase drifts (e.g., from the common LO and clock signals or similar temperature fluctuations if they are installed in the same chassis). Nevertheless, there always exists uncorrelated noise in two channels that will degrade the performance of the reference tracking. An alternative approach is the so-called *drift calibration* (see Fig. 6.46a) (Ludwig et al. 2010; Lin et al. 2020)

To calibrate the RF detector's phase drift, we inject a reference signal into the same channel measuring the cavity pickup signal. If the RF station operates in the pulsed mode, the reference signal should be injected between the RF pulses, so we need an RF switch to control the injection time. For a CW RF station, the reference signal should have a different frequency from the RF operating frequency. For the

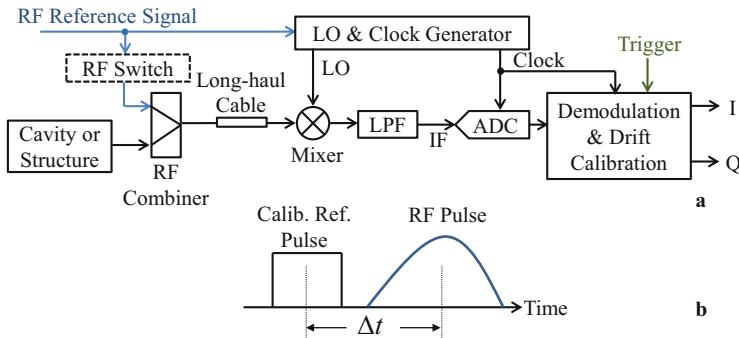


Fig. 6.46 (a). Block diagram of drift calibration; (b). Cabliration and RF field pulses for an RF station operating in the pulsed mode.

CW case, the RF detector drift can be evaluated from the measured reference phase in real-time and subtracted from the RF field phase measurement. However, for pulsed RF stations, the reference phase is measured at a different time from the RF pulse. Let us check this situation in detail in Fig. 6.46b.

Assume that the delay between the RF pulse and the calibration pulse is Δt . The RF detector phase drift is evaluated with the calibration pulse and then subtracted from the RF pulse after Δt . Therefore, the phase detection of RF pulses with drift calibration still suffers from an equivalent measurement noise given by

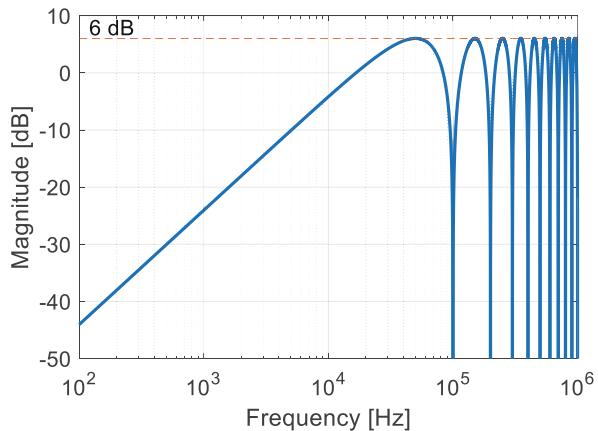
$$\Delta\varphi'_{mea}(t) = \Delta\varphi_{mea}(t) - \Delta\varphi_{mea}(t - \Delta t), \quad (6.105)$$

where $\Delta\varphi_{mea}$ is the phase noise introduced by the RF detector. If $\Delta t = 0$, we will obtain a perfect cancellation of the measurement noise, whereas, for a non-zero Δt , we better study (6.105) in the frequency domain using Laplace transforms:

$$\Delta\varphi'_{mea}(\tilde{s}) = \Delta\varphi_{mea}(\tilde{s}) \left(1 - e^{-\tilde{s}\Delta t} \right). \quad (6.106)$$

The gain between $\Delta\varphi_{mea}$ and $\Delta\varphi'_{mea}$ at each frequency is depicted in Fig. 6.47, where we have assumed $\Delta t = 10 \mu s$. The frequency response implies that at the frequencies $k/\Delta t$ ($k = 0, 1, 2, \dots$), we obtain a full cancellation of $\Delta\varphi_{mea}$, whereas, at $(k + 1/2)/\Delta t$, $\Delta\varphi_{mea}$ is doubled. Therefore, for the drift calibration operating in the pulsed mode, we must low-pass filter the phase measurement of the calibration pulse before subtracting it from the RF pulse phase. In other words, the drift calibration can only suppress the measurement noise with a frequency lower than $f_{dcal} = 1/(2\pi\Delta t)$ satisfying $|1 - e^{-j2\pi f_{dcal}\Delta t}| \approx 1$. Of course, reducing Δt helps to increase the maximum frequency of drift calibration.

Fig. 6.47 Suppression of $\Delta\varphi_{mea}$ at different frequencies for $\Delta t = 10 \mu\text{s}$



6.6.3 Beam-Based Feedback

Experience shows that the reference tracking can remove most of the RF detector phase drift. Its implementation is simple without requiring extra hardware but an additional RF detector channel, which makes the reference tracking popular for drift correction. Drift calibration can usually achieve better results. However, it requires extra hardware (e.g., RF switches, reference signals routing to each RF detector channel, and power combiners), increasing the complexity.

Both reference tracking and drift calibration can only correct the drifts in RF detectors. The RF field drifts experienced by the beam may also be caused by other sources, such as the beam arrival time drift, the RF reference phase drift, and the cavity probe's coupling phase drift. These drifts cannot be corrected by the LLRF system and will cause errors in beam acceleration. In practice, we use the so-called *beam-based feedback* (Himel et al. 1993; Koprek et al. 2010; Sugiyama et al. 2019) to stabilize the beam parameters directly. Figure 6.48 shows an example of the beam-based feedback system implemented at FLASH in DESY (Koprek et al. 2010). The beam parameters, such as the beam current (measured by the Toroid), bunch arrival time (measured by the BAM – Bunch Arrival Time Monitor) and the bunch length (measured by the BCM – Bunch Compression Monitor), are stabilized by actuating on the amplitude or phase set points of the LLRF system.

The beam-based feedback system has been proven to compensate for the RF field residual drifts mentioned above very efficiently. It forms cascaded control loops together with the LLRF control loops. In this book, we will not discuss the detail of the beam-based feedback. A separate book (*Intelligent Beam Control in Accelerators*) is planned to cover the beam-based feedback, as well as the optimization and the machine-learning-based control of the beam in particle accelerators.

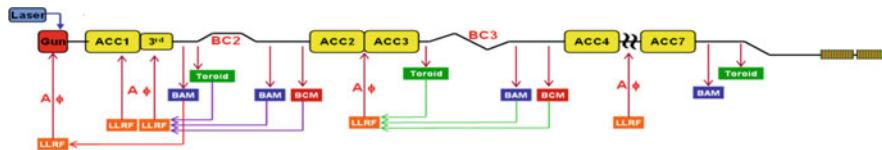


Fig. 6.48 Concept of the beam based feedback at FLASH (ACC represents the accelerating module, each consisting of 8 TESLA cavities) (Courtesy of W. Koprek et al)

6.7 Summary

This chapter focuses on the noise in accelerator RF systems. We first introduced several important concepts that are frequently used in the RF system noise analysis. Then the noise models of several basic RF components were discussed. The noise model of other RF components, such as the I/Q modulator/demodulator, can be built up by combining the models of multiple basic RF components (e.g., power splitter/combiner, mixer, amplifier, etc.). The noise transfer relations in an accelerator RF system distinguish the contributions from the RF reference signal, RF driving chain, and RF measurement chain to the cavity field stability. We also introduced a systematic top-to-down approach to specifying the stability requirements of the accelerator RF systems. The discussion of the RF station noise model provides guidelines on how to measure the noise in the RF driving chain, measurement chain, and cavity field. Finally, we briefly discussed several methods to mitigate the drifts in the RF detectors.

References

- Analog Devices (1999) A technical tutorial on digital signal synthesis. <http://www.ieee.li/pdf/essay/dds.pdf>. Accessed 27 May 2021
- Bane K L F, Emma P (2005) LiTrack: a fast longitudinal phase space tracking code with graphical user interface. In: Proceedings of the 2005 particle accelerator conference, Knoxville, TN, USA, 16–20 May 2005
- Beutner B, Reiche S (2010) Sensitivity and tolerances study for the SwissFEL. In: Proceedings of the 32nd international free electron laser conference, Malmö, Sweden, 23–27 August 2010
- Blas F, Garoby R (1991) Design and operational results of a “one-turn-delay feedback” for beam loading compensation of the CERN PS ferrite cavities. Paper presented at the IEEE 1991 particle accelerator conference. San Francisco, CA, USA, 6–9 May 1991
- Boudot R, Rubiola E (2012) Phase noise in RF and microwave amplifiers. <https://arxiv.org/abs/1001.2047>. Accessed 28 May 2021
- Brannon B (2004) Sampled systems and the effects of clock phase noise and jitter. Analog Devices application note (AN-756). <https://www.analog.com/media/en/technical-documentation/application-notes/AN-756.pdf>. Accessed 27 May 2021
- Candy J C, Temes G C (1991) Oversampling methods for data conversion. In: Proceedings of the 1991 IEEE pacific rim conference on communications, computers and signal processing, Victoria, BC, Canada, 9–10 May 1991

- Cárdenas-Olaya AC, Rubiola E, Friedt J-M et al (2017) Noise characterization of analog to digital converters for amplitude and phase noise measurement. *Rev Sci Instrum* 88:065108. <https://doi.org/10.1063/1.4984948>
- Dinter H (2018) Longitudinal diagnostics for beam-based intra bunch-train feedback at FLASH and the European XFEL. Ph.D. thesis, Hamburg University
- Doolittle L et al (2006) Digital low-level RF control using non-IQ sampling. In: Proceedings of the 23rd linear accelerator conference, Knoxville, TN, USA, 20–25 August 2006
- Floettmann K (2017) ASTRA: a space charge tracking algorithm. https://www.desy.de/~mpyflo/Astra_manual/Astra-Manual_V3.2.pdf. Accessed 29 May 2021
- Fox J, Mastorides T, Rivetta C et al (2007) Selecting RF amplifiers for impedance controlled LLRF systems – nonlinear effects and system implications. In: Proceedings of the 22ed particle accelerator conference, Albuquerque, NM, USA, 25–29 June 2007
- Fox J, Mastorides T, Rivetta C et al (2010) Lessons learned from positron-electron project low level RF and longitudinal feedback. *Phys Rev ST Accel Beams* 13(5):052802. <https://doi.org/10.1103/PhysRevSTAB.13.052802>
- Gallo A (2011) Basic of RF electronics. CERN accelerator school: RF for accelerators. <https://cds.cern.ch/record/1407402/files/p223.pdf>. Accessed 28 May 2021
- Gallo A (2018) Timing and synchronization. CERN accelerator school: beam instrumentation. <https://cas.web.cern.ch/schools/tuusula-2018>. Accessed 29 May 2021
- Garcia J, LaJeunesse S G (1995) Understanding glitch in a high speed D/A converter. Intersil technical brief (TB325.1). <https://www.renesas.com/us/en/document/tb325-understanding-glitch-high-speed-da-converter>. Accessed 27 May 2021
- Geng Z, Hong B (2016) Design and calibration of an RF actuator for low-level RF systems. *IEEE Trans Nucl Sci* 63(1):281–287. <https://doi.org/10.1109/TNS.2015.2507204>
- Geng Z, Craievich P, Kalt R et al (2019) RF jitter and electron beam stability in the SwissFEL Linac. In: Proceedings of the 39th free electron laser conference, Hamburg, Germany, 26–30 August 2019
- Hanaki H, Dewa H, Suzuki S et al (2014) Evaluation of beam energy fluctuation caused by phase noise. In: Proceedings of the 27th international linear accelerator conference, Geneva, Switzerland, 31 August–5 September 2014
- Himel T, Allison S, Grossberg P et al (1993) Adaptive cascaded beam-based feedback at the SLC. In: Proceedings of the 1993 IEEE particle accelerator conference, Washington, DC, USA, 17–20 May 1993
- Hoffmann M (2008) Development of a multichannel RF field detector for the low-level RF control of the free-electron laser at Hamburg. Ph.D. thesis, Technische Universität Hamburg
- Howard RM (2002) Principles of random signal analysis and low noise design: the power spectral density and its applications. Wiley, New York
- Analog Devices, Datasheet of dual 12-/14-/16-bit, LVDS interface, 500 MSPS DACs, AD9780/AD9781/AD9783. https://www.analog.com/media/en/technical-documentation/data-sheets/AD9780_9781_9783.pdf. Accessed 27 May 2021
- Kester W (2009) Evaluating high speed DAC performance. Analog Devices tutorial (MT-013). <https://www.analog.com/media/en/training-seminars/tutorials/MT-013.pdf>. Accessed 28 May 2021
- Koprek W, Behrens C, Bock M K et al (2010) Intra-train longitudinal feedback for beam stabilization at FLASH. In: Proceedings of the 32nd international free electron laser conference, Malmo, Sweden, 23–27 August 2010
- Lapointe L (2016) Optimal clock sources for GSPS ADCs design guide. Texas Instruments reference designs (TIDA-00479). <https://www.ti.com/tool/TIDA-00479>. Accessed 27 May 2021
- Lin H, Du B, Huang G et al (2016) A time domain analysis method for RF noise. In: Proceedings of the 7th international particle accelerator conference, Busan, Korea, 8–13 May 2016

- Lin Z, Du Y, Huang G et al (2020) Application of a drift compensation low-level radio frequency system based on time-multiplexing pick-up/reference signals. *Rev Sci Instrum* 91(12):124706. <https://doi.org/10.1063/5.0016549>
- Ludwig F, Hoffmann M, Schlarb H et al (2006) Phase stability of the next generation RF field control for VUV- and X-ray free electron laser. In: Proceedings of the 10th European particle accelerator conference, Edinburgh, Scotland, 26–30 June 2006
- Ludwig F, Gerth K, Hacker K et al (2010) Drift calibration techniques for future FELs. In: Proceedings of the 1st international accelerator conference, Kyoto, Japan, 23–28 May 2010
- Maalberg A, Kuntzsch M, Petlenkov E (2020) Simulation of RF noise propagation to relativistic electron beam properties in a linear accelerator. *IFAC PapersOnLine* 53(2):348–354. <https://doi.org/10.1016/j.ifacol.2020.12.184>
- Mastorides T, Rivetta C, Fox J et al (2010) RF system models for the CERN large hadron collider with application to longitudinal dynamics. *Phys Rev ST Accel Beams* 13(10):102801. <https://doi.org/10.1103/PhysRevSTAB.13.102801>
- Mastoridis T, Baudrenghien P, Molendijk J (2012) LHC one-turn delay feedback commissioning. In: Proceedings of the 3rd international particle accelerator conference, New Orleans, LA, USA, 20–25 May 2012
- Milne CJ, Schietinger T, Aiba M et al (2017) SwissFEL: the Swiss X-ray free electron laser. *Appl Sci* 7(7):720. <https://doi.org/10.3390/app7070720>
- Neu T (2010) Clock jitter analyzed in the time domain, Part 1. Texas Instruments analog application journal. <https://www.ti.com/jp/lit/pdf/slyt379>. Accessed 28 May 2021
- Neu T (2017) Clocking the RF ADC: should you worry about jitter or phase noise. Texas Instruments analog application journal. <https://www.ti.com/lit/pdf/slyt705>. Accessed 27 May 2021
- Reiche S (2016) SwissFEL tolerance study. PSI internal presentation
- Rubiola E (2006) Tutorial on the double balanced mixer. <https://arxiv.org/pdf/physics/0608211>. Accessed 29 May 2021
- Rubiola E (2008) Phase noise and frequency stability in oscillators. Cambridge University Press, Cambridge
- Rubiola E (2021) Enrico's chart of phase noise and two-sample variances, <http://rubiola.org/pdf-static/Enrico's-chart.pdf>. Accessed 15 June 2021
- Schmid L, Baudrenghien P, Hagmann G (2017) One-turn delay feedback with a fractional delay filter. <https://public.cells.es/workshops/www.llrf2017.org/pdf/Posters/P-77.pdf>. Accessed 29 May 2021
- Shih H-J, Ellison JA, Newberger BS et al (1994) Longitudinal beam dynamics with RF noise. *Particle Accelerators* 43(3):159–194
- Sugiyama Y, Tamura F, Yoshii M (2019) The feedback system for the longitudinal coupled-bunch instability in the J-PARC main ring. <https://arxiv.org/pdf/1910.07415.pdf>. Accessed 29 May 2021
- Szczepaniak ZR (2010) Spatial power combining techniques for semiconductor power amplifiers. In: Zhurbenko V (ed) Advanced microwave circuits and systems. InTech, Rijeka, pp p159–p176
- Teytelman D (2015) Ingredients for perfect RF transceiver. Tutorial presented at the LLRF2015 workshop, Shanghai, China, 2–6 November 2015
- Wang D (2015) Effects of clock spur on high-speed DAC performance. Texas Instruments application report (SLAA565A). <https://www.ti.com/lit/pdf/slaa565>. Accessed 28 May 2021

Chapter 7

Nonlinearity in RF Systems



The linearity of the high-power RF (HPRF) and low-level RF (LLRF) components is critical for the RF control performance. A nonlinear RF detector causes errors in the measurements of the amplitude and phase. Moreover, the RF driving chain nonlinearity produces higher-order harmonics in the output and results in gain and phase shift varying with the input power. In this chapter, the nonlinear effects in an accelerator RF system will be discussed. We also introduce several widely used approaches to linearize the RF amplifiers and to deal with the nonlinearity in feedback controllers.

7.1 Basic Concepts

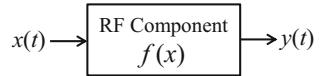
We have derived in Chap. 3 the linear model of various RF components, such as the transmission line, RF amplifier, and cavity. The linear models are described by linear (differential) equations or transfer functions, which are the basis for analyzing and designing the RF control systems. However, commercially available RF components like ADCs, DACs, mixers, RF amplifiers, and klystrons exhibit some level of nonlinearity. The nonlinearity may deteriorate the RF control performance or even cause instability. In this section, we will introduce some basic concepts about the RF system nonlinearity. In general, we model an RF component as a nonlinear network like Fig. 7.1.

The relationship between the output $y(t)$ and the input $x(t)$ can be described as

$$y(t) = a_0 + a_1 x(t - \Delta t_1) + a_2 [x(t - \Delta t_2)]^2 + a_3 [x(t - \Delta t_3)]^3 + \dots \quad (7.1)$$

where the coefficients a_k ($k = 0, 1, 2, \dots$) are real numbers (Cho 2005; Kumar et al. 2010). In the following, we limit our analysis to the third order in the polynomial function because the higher-order terms are usually negligible. Here Δt_k is the

Fig. 7.1 Nonlinear model of an RF component



transmission delay from x to the k th-order term in y . If we input a sinusoidal signal, $x(t) = A_{in}\sin(\omega t)$, the output can be written as

$$\begin{aligned} y(t) &= a_0 + a_1 A_{in} \sin(\omega t - \omega\Delta t_1) \\ &\quad + \frac{a_2 A_{in}^2}{2} [1 - \cos(2\omega t - 2\omega\Delta t_2)] \\ &\quad + \frac{a_3 A_{in}^3}{4} [3 \sin(\omega t - \omega\Delta t_3) - \sin(3\omega t - 3\omega\Delta t_3)]. \end{aligned} \quad (7.2)$$

It represents the steady-state response of the RF component for a single-frequency input. Here we call ω the *fundamental frequency* to distinguish with the other frequencies generated by the nonlinearity. Equation (7.2) implies that the steady-state output contains not only the fundamental-frequency signal but also its harmonics and a DC offset. We notice that the odd-number terms ($k = 1, 3, 5, \dots$) also generate output signals at ω , resulting in distortions to the linear term (the desired output). Let us select the fundamental-frequency terms from y and define:

$$y_\omega(t) := a_1 A_{in} \sin(\omega t - \omega\Delta t_1) + \frac{3}{4} a_3 A_{in}^3 \sin(\omega t - \omega\Delta t_3). \quad (7.3)$$

If we define the carrier frequency $\omega_c = \omega$, (7.3) can be written in terms phasors as

$$\mathbf{y}_\omega(t) = \left(\mathbf{a}_1 + \frac{3}{4} \mathbf{a}_3 A_{in}^2 \right) \mathbf{x}(t), \quad (7.4)$$

where \mathbf{x} and \mathbf{y}_ω are the phasors of the input x and the fundamental-frequency output y_ω . The complex coefficients \mathbf{a}_1 and \mathbf{a}_3 are defined as

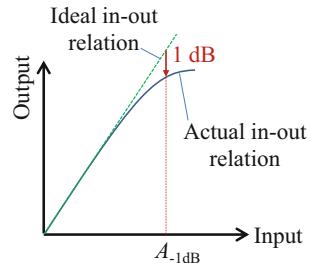
$$\mathbf{a}_1 := a_1 e^{-j\omega\Delta t_1}, \mathbf{a}_3 := a_3 e^{-j\omega\Delta t_3}. \quad (7.5)$$

Using the relation (7.4), we will explain some basic concepts like the 1-dB compression point, the third-order intercept point, and the AM-PM conversion.

7.1.1 1-dB Compression Point

According to (7.4), the fundamental-frequency component in the output of a nonlinear component comes from both the linear term and the higher odd-order terms.

Fig. 7.2 1 dB compression point of a nonlinear RF component



In practice, the complex coefficients \mathbf{a}_1 and \mathbf{a}_3 have inverse signs. This implies that the nonlinearity typically reduces the magnitude of \mathbf{y}_ω in the output. The higher the input amplitude A_{in} , the more the magnitude of \mathbf{y}_ω is reduced compared to the linear output term. This effect is known as the *gain compression* for the fundamental-frequency input. Figure 7.2 shows the gain compression on the input-output gain curve of a nonlinear RF component. The *1 dB compression point* is used to characterize the compression effect. It is defined as an operating point, at which the output power is 1 dB less compared to the ideal output power predicted by the linear gain of the RF component. The input amplitude of the 1 dB compression point, A_{-1dB} , satisfies

$$20 \log_{10} \left| 1 + \frac{3\mathbf{a}_3}{4\mathbf{a}_1} A_{-1dB}^2 \right| = -1 \text{ dB}. \quad (7.6)$$

7.1.2 Third-Order Intercept Point

The *third-order intercept point (IP3)* is another important characteristic of the nonlinearity in an RF component. The input IP3 (IIP3) is defined as an input amplitude/power, at which the fundamental-frequency signals generated by the linear term and by the third-order term have the same power. From (7.4), the IIP3 amplitude, A_{IIP3} , satisfies

$$|\mathbf{a}_1| = \left| \frac{3}{4} \mathbf{a}_3 A_{IIP3}^2 \right|. \quad (7.7)$$

Then the A_{IIP3} and the corresponding output amplitude, A_{OIP3} , can be calculated as

$$A_{IIP3} = \sqrt{\frac{4|\mathbf{a}_1|}{3|\mathbf{a}_3|}}, A_{OIP3} = |\mathbf{a}_1| A_{IIP3}. \quad (7.8)$$

Correspondingly, the third-order intercept point can also be defined in terms of the input and output powers (P_{IIP3} and P_{OIP3}).

7.1.3 AM-PM Conversion

The fundamental-frequency output of a nonlinear RF component is determined by (7.4). Because \mathbf{a}_1 and \mathbf{a}_3 are complex numbers, the phase of the complex gain \mathbf{y}_o/\mathbf{x} depends on A_{in} if the angles of \mathbf{a}_1 and \mathbf{a}_3 are not exact 180° out of phase. This results in the so-called *amplitude modulation (AM) to phase modulation (PM) conversion*. If we control the RF component using separated amplitude and phase loops, the AM-PM conversion will cause crosstalk between the two loops. This could result in a degradation of the control performance.

7.1.4 Nonlinearity Induced RF Detection Error

The RF detector measures the amplitude and phase (or I and Q components) of an RF signal. RF detectors often consist of amplifiers (e.g., for RF input or LO signal amplification), mixers and ADCs, and all of them may introduce nonlinearity. Here we briefly discuss the effects of nonlinearity on RF signal detection (Hoffmann 2008).

First, the RF detector nonlinearity generates higher-order harmonics of the input RF (or IF) signal to be measured. With the ADC sampling process, the harmonics may be aliased into the fundamental frequency resulting in measurement errors. The harmonic issue has been addressed in Chap. 6. Here we will focus on the RF detection errors caused by the compression effects in the fundamental-frequency output.

From (7.4), for an input $\mathbf{x}(t)$, the RF detector output consists of an ideal linear term $\mathbf{y}_{ideal}(t)$ and an error term $\mathbf{y}_{error}(t)$, which are given by

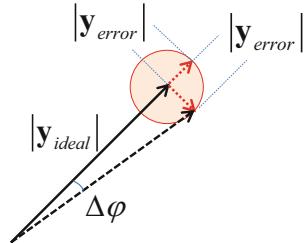
$$\mathbf{y}_{ideal}(t) = \mathbf{a}_1\mathbf{x}(t), \mathbf{y}_{error}(t) = \frac{3}{4}\mathbf{a}_3A_{in}^2\mathbf{x}(t). \quad (7.9)$$

These two output terms can be plotted in a vector diagram in Fig. 7.3. If \mathbf{y}_{error} has the same direction as \mathbf{y}_{ideal} , it generates a maximum amplitude error; while if the two vectors are perpendicular, it generates a maximum phase error. Then, the maximum amplitude and phase measurement errors caused by the nonlinearity can be calculated as

$$\left(\frac{\Delta A}{A}\right)_{NL} = \Delta\varphi_{NL} = \frac{3|\mathbf{a}_3|}{4|\mathbf{a}_1|}A_{in}^2 = \frac{A_{in}^2}{A_{IIP3}^2}. \quad (7.10)$$

It can also be expressed with the IIP3 or OIP3 powers as

Fig. 7.3 Vector diagram of the RF detection errors caused by nonlinearity



$$\left(\frac{\Delta A}{A}\right)_{NL} = \Delta\varphi_{NL} = \frac{P_{in}}{P_{IIP3}} = \frac{P_{out}}{P_{OIP3}}. \quad (7.11)$$

This means, we can estimate the RF detection errors by comparing the input (or output) power to the IIP3 (or OIP3) power of the RF detector. For example, if the RF detector input power is 0 dBm and the IIP3 power is 30 dBm, the resulting amplitude and phase measurement errors can reach 3.2% and 1.8°, respectively.

7.1.5 Nonlinearity Induced RF Driving Disturbances

Most RF systems of particle accelerators operate around a central frequency, i.e., the RF operating frequency. The input of the RF driving chain is typically a single-tone RF signal at the operating frequency. In this case, the gain drops caused by the fundamental-frequency components induced by the higher-order nonlinear terms, as discussed in Sects. 7.1.1, 7.1.2, 7.1.3 and 7.1.4, is a major concern. However, if the RF driving chain input contains multiple tones, such as in the LLRF systems of synchrotrons controlled by a one-turn delay feedback loop (Blas and Garoby 1991; Mastoridis et al. 2012; Schmid et al. 2017), the *intermodulation* between different tones will produce additional disturbances in the cavity input (Fox et al. 2007; Fox et al. 2010).

The intermodulation disturbances caused by a nonlinear RF amplifier with a two-tone input are depicted in Fig. 7.4. Here we assume that except for the fundamental frequency ($f_1 = f_{RF}$) input, another input with a frequency $f_2 = f_{RF} + \Delta f$ is amplified to correct a strong narrow-band ripple in the cavity field. We also assume that Δf is much smaller than f_{RF} . With the two-tone input, the intermodulation will produce frequency components at $nf_1 \pm mf_2$ (n and m are integers) in the amplifier output. Figure 7.4 shows only the intermodulation products caused by the third nonlinear term, including two frequencies at $2f_1 - f_2$ and $2f_2 - f_1$, which fall potentially in the cavity bandwidth. The intermodulation products are disturbances applied to the cavity input, causing cavity field errors. Their effects can partly be suppressed by the feedback controller, depending on the sensitivity function of the feedback loop at the disturbance frequencies.

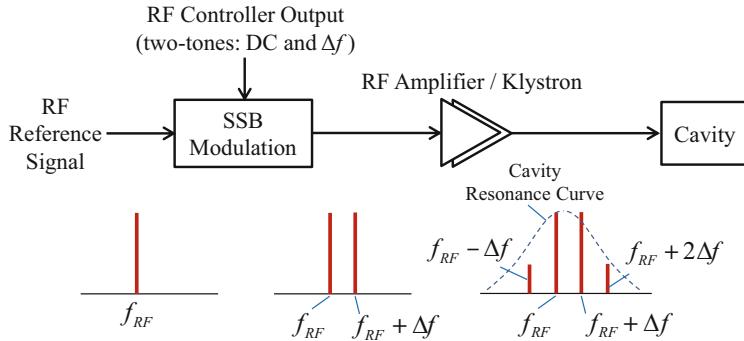


Fig. 7.4 Intermodulation disturbances caused by the RF driving chain nonlinearity

The intermodulation is an important source of errors for the one-turn delay feedbacks of the LLRF systems of synchrotrons. The one-turn delay feedback produces RF power at the frequencies of $f_{RF} \pm kf_{rev}$ ($k = 0, 1, 2, \dots$) to suppress the RF disturbances at these frequencies with strong impact to the beam stability, where f_{RF} is the RF operating frequency and f_{rev} is the particle revolution frequency (Fox et al. 2007). The intermodulation products of the RF operating frequency and one of the revolution harmonics will transfer energy to other revolution harmonics (similar to Fig. 7.4), degrading the performance of the one-turn delay feedback. Therefore, selecting amplifiers with high linearity is essential for controlling a synchrotron with a high beam current.

7.2 Nonlinearity of RF Amplifiers

RF amplifiers, including solid-state amplifiers and klystrons, are used to amplify the input RF power to supply the input power required by the next amplifier stage or the required drive power for the RF cavities or structures. Most amplifiers behave linearly only in a particular input power range. Nonlinearity becomes significant in the case of larger input power.

As described before, the nonlinearity yields harmonics and intermodulation in the amplifier output and reduces the gain of the fundamental frequency. When the amplifier amplifies an RF-frequency signal, the harmonics are typically out of the bandwidths of the downstream RF components, and therefore, are not visible in the system output. Of course, large harmonics may damage the downstream RF components. If the amplifier is used to amplify an IF signal, the harmonics may be captured by the RF detection process and cause measurement errors, which are discussed in detail in Chap. 5. Here we will focus on the gain compression of the fundamental-frequency signal by an amplifier.

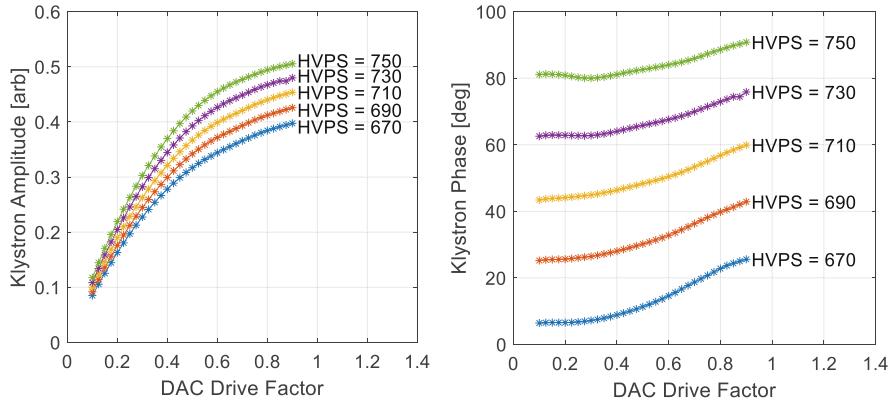


Fig. 7.5 Gain curves of a SwissFEL klystron. The input is the DAC drive factor determined by the RF controller with a value between 0 and 1.25, where 1.25 corresponds to the DAC full scale. The amplitude of the klystron output signal has been normalized, where 1 corresponds to the full scale of the RF detector ADC

The gain curves of a klystron at the SwissFEL for different modulator high voltages (voltage setting of the high-voltage power supplies (HVPS)) are depicted in Fig. 7.5. The gain compression (saturation) is visible in the amplitude curves. The phase curves illustrate the AM-PM conversion: at a particular HVPS, the phase shift of the klystron depends on the input power (determined by the DAC drive factor). Note that the nonlinearity is an overall effect of the entire RF driving chain, including the DAC, I/Q modulator, pre-amplifier, and klystron.

The nonlinearity of an amplifier can be modelled mathematically as a function between the output and input:

$$\mathbf{y} = F_a(V, x_a) e^{iF_p(V, x_a)} \mathbf{x}, \text{ where } x_a := |\mathbf{x}|. \quad (7.12)$$

The input \mathbf{x} and output \mathbf{y} of the amplifier are given as phasors referring to the fundamental frequency. We let x_a and y_a denote the input and output amplitudes, $x_a = |\mathbf{x}|$ and $y_a = |\mathbf{y}|$. The steady-state gain and phase shift of the amplifier are modelled by two (nonlinear) functions, F_a and F_p , which depend on x_a and V , where V is the power supply voltage of the amplifier. For some solid-state amplifiers, V is not adjustable, so F_a and F_p are only functions of x_a . For the example in Fig. 7.5, x_a is the DAC drive factor, V is the HVPS and y_a is the klystron output amplitude. The klystron output phase is the phase of the phasor \mathbf{y} .

If F_a and F_p are known, we can perform the following two tasks:

- (a) Given \mathbf{x} and V , predict the output \mathbf{y} . This is a direct application of the nonlinear model (7.12).

- (b) Given the desired output amplitude y_a and the desired headroom from saturation, $\eta = 1 - y_a/y_{sat}$ (where y_{sat} is the saturation output of the determined V), determine the power supply voltage setting V and the input amplitude x_a .

We call the task (a) as *output prediction*. The task (b) can be called *operating-point determination (OPD)*, which requires an inversion of the nonlinear model (7.12) (Rezaeizadeh 2014).

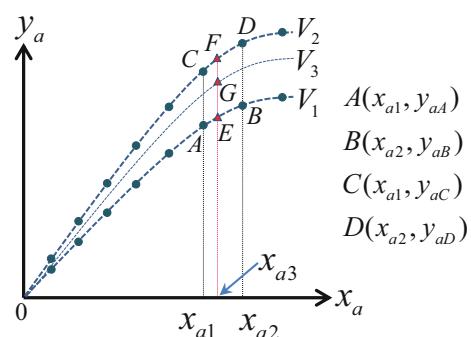
7.2.1 Look-Up Table Model

The simplest method to model F_a and F_p is to use the measurement results in Fig. 7.5 as a look-up table (LUT) directly. For output prediction, the algorithms like linear interpolation can be used to calculate the output y from a given x and V . As an example, Fig. 7.6 illustrates how to predict the output amplitude based on the LUT using the linear interpolation algorithm.

Figure 7.6 is a LUT representing the function F_a . Note that the LUT here is formed by the measurement results of y_a at discrete values of V and x_a . Suppose that we want to predict the output amplitude for an input $|x| = x_{a3}$ at a high voltage V_3 . From the LUT, we can find four measurement points (A , B , C , and D) around the desired operating point corresponding to V_3 and x_{a3} . We assume the point A and B are on the gain curve of V_1 , and C and D are on the gain curve of V_2 . The inputs and outputs of these points are shown in Fig. 7.6. Our goal is to estimate the output amplitude y_{aG} at the point G for the high voltage V_3 and input x_{a3} . We can perform the calculation following the steps below:

- On the curve of V_1 , calculate the output at point E , assuming the points A , E and B on a straight line: $y_{aE} \approx y_{aA} + (y_{aB} - y_{aA})(x_{a3} - x_{a1})/(x_{a2} - x_{a1})$.
- On the curve of V_2 , calculate the output at point F , assuming the points C , F and D on a straight line: $y_{aF} \approx y_{aC} + (y_{aD} - y_{aC})(x_{a3} - x_{a1})/(x_{a2} - x_{a1})$.

Fig. 7.6 Prediction of the output amplitude with LUT using linear interpolation. The dots are the measurement points that form a two-dimensional LUT



- (c) Assuming the output amplitude is proportional to V for a fixed input x_a , calculate the output at point G using the results from the last two steps:
 $y_{aG} \approx y_{aE} + (y_{aF} - y_{aE})(V_3 - V_1)/(V_2 - V_1)$.

Note that Fig. 7.6 only shows the LUT of amplitude curves. The output phase for the given x_{a3} and V_3 can be determined similarly using the phase LUT (not shown).

The LUT is easy to implement, so it is widely used in practice. Of course, we may have to consider the storage space occupied by a large LUT when implementing it in a small FPGA. To improve the LUT accuracy, we may make more measurements with reduced step sizes of V and x_a or use other interpolation algorithms, such as the polynomial interpolation or spline interpolation.

When using the LUT as an OPD, the desired output amplitude y_a is given and the goal is to determine the required V and x_a . As seen from Fig. 7.6, there are multiple combinations of V and x_a to achieve the desired y_a . Therefore, the headroom from saturation, η , should be also specified. To implement the OPD, we need one additional LUT (denoted as T_{V-ysat}), which describes the saturation output, y_{sat} , of each V (see Fig. 7.7). To simplify the discussion, we denote the two-dimensional LUT formed by the raw measurements as $T_{V,xa-y}$. The following procedure can be used to determine the operating point:

- From the desired output amplitude $y_{a,d}$ and the headroom η , calculate the corresponding saturation output $y_{sat,d} = y_{a,d}(1 - \eta)$.
- Using the LUT T_{V-ysat} , find out the required high voltage V_{OPD} for the calculated $y_{sat,d}$ above. The linear interpolation algorithm can be used here.
- From $T_{V,xa-y}$, calculate a gain curve for the required V_{OPD} . To obtain a gain curve for V_{OPD} , we select two measured gain curves at V_1 and V_2 that are closest to V_{OPD} . For each x_a value on the gain curves, we can interpolate the outputs at V_1 and V_2 to estimate the output at V_{OPD} .
- Using the calculated gain curve at V_{OPD} , calculate the required x_{aOPD} to achieve the desired $y_{a,d}$ with the linear interpolation algorithm.

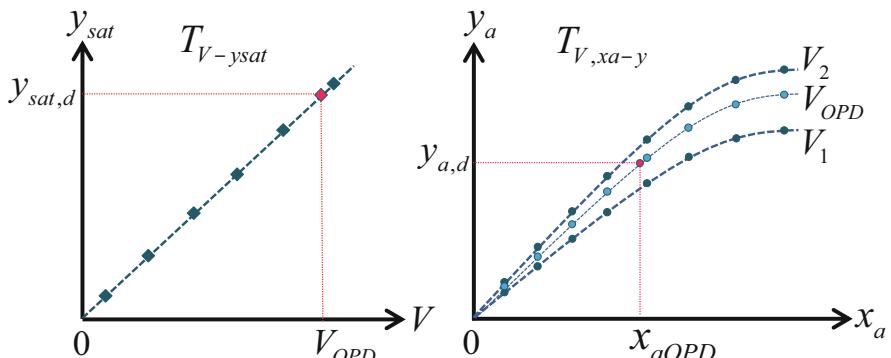


Fig. 7.7 LUT-based OPD

The LUT-based OPD has demonstrated excellent accuracy and is simple to implement. Typically, we only use the OPD to determine the operating point roughly. Accurate amplitude and phase settings of an RF station often rely on the feedback controls.

7.2.2 Analytical Model

An alternative approach to model F_a and F_p is to fit the measured data to some analytical functions. In this section, the *polynomial model*, *Saleh model* and *Ghorbani model* will be introduced (Jantunen 2004; Yadav et al. 2010). Note that we will apply these analytical models only for the nonlinear responses at a particular high voltage V .

The gain curves in Fig. 7.5 at a particular high voltage can be fitted as polynomial functions in the form of

$$F_a(x_a) = a_0 + a_1x_a + a_2x_a^2 + \dots, F_p(x_a) = b_0 + b_1x_a + b_2x_a^2 + \dots \quad (7.13)$$

Here we have neglected the independent variable V in (7.12). The coefficients a_k and b_k ($k = 0, 1, 2, \dots$) can be determined from the measurement data with the least-square algorithm. The order of the polynomial should be selected as a compromise between the fitting error and the prediction error. The fitting error is the difference between the data used for fitting and the values calculated by the fitted model; whereas the prediction error is the error for a new measurement compared to the value predicted by the fitted model. If the order of the polynomial is too high, the model has less fitting error but large prediction error, which is called *overfitting*. In this case, the model is sensitive to the measurement noise, leading to large errors compared to the system's real input-output relations. In contrast, if the order of the polynomial is too low, the model will lose information in the system's input-output relations, causing large fitting and prediction errors. This is the so-called *underfitting*. Typically, the order of the polynomial can be determined using the technique called *cross validation*, which is described briefly as follows:

- (a) Define a polynomial order.
- (b) Split the data set into two parts: one is used to fit the polynomial, which is called the *fitting set*, and the other (denoted as the *test set*) is used to validate the fitting results. The fitting performance is evaluated with the validation error using the data in the test set. The validation error is calculated as the RMS value of the difference between the measurement results and the values predicted by the polynomial model.
- (c) Repeat step b for different splits (e.g., split the original data set randomly) for the fitting and test sets. The validation error from each split can be averaged to obtain the overall validation error.
- (d) Repeat the steps a – c for different polynomial order. Then we can select a minimum order for the polynomial that yields a small validation error.

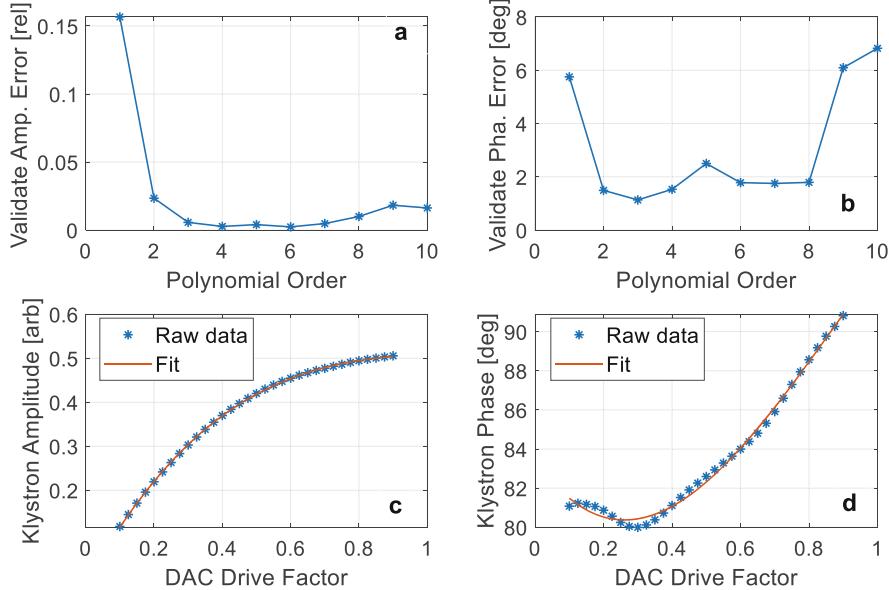


Fig. 7.8 (a, b) Selection of the order of polynomials with cross-validation; (c, d) The third-order polynomial fittings of the gain and phase curves

Fig. 7.8 shows the polynomial model of the gain and phase nonlinear functions for the data in Fig. 7.5 with HVPS = 750 V. The cross-validation illustrates that the third-order polynomials are suitable for both the gain and phase curves. The raw data and the fitted polynomials are also displayed for comparison.

The *Saleh model* can be used to model the traveling-wave tube (TWT) amplifiers or klystrons with F_a and F_p given by

$$F_a(x_a) = \frac{a_0 x_a}{1 + a_1 x_a^2}, F_p(x_a) = \frac{b_0 x_a}{1 + b_1 x_a^2}, \quad (7.14)$$

where a_0, a_1, b_0, b_1 are positive real numbers. The *Ghorbani model* is another widely used formula to describe the amplifier nonlinearity. It is typically used to model solid-state amplifiers, which do not have a large roll-off at saturation as TWT amplifiers. The Ghorbani model are described by

$$F_a(x_a) = \frac{a_0 x_a^{a_1}}{1 + a_2 x_a^{a_1}} + a_3 x_a, F_p(x_a) = \frac{b_0 x_a^{b_1}}{1 + b_2 x_a^{b_1}} + b_3 x_a. \quad (7.15)$$

The coefficients a_k and b_k in (7.14) and (7.15) can be determined from the measurement data with nonlinear fitting algorithms (e.g., `lsqcurvefit` function of Matlab). Figure 7.9 shows a comparison of the polynomial, Saleh and Ghorbani models applied to the same data as Fig. 7.8. The model predicted values are plotted over a

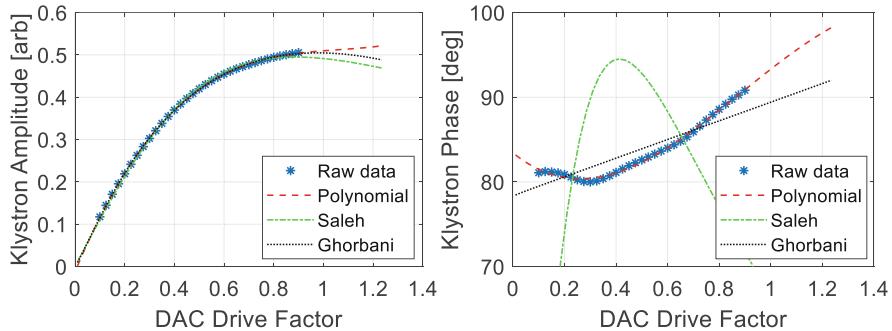


Fig. 7.9 Comparision of the polynomial, Saleh and Ghorbani models

wider input range than the measurement data to check whether the model can be applied to a wider range. All the three analytical models fit well to the gain curve, except that the polynomial model does not show an obvious saturation. For the phase curve, the Saleh model has a too large error and is not applicable.

There are also some other widely used models for amplifier nonlinearity such as the Rapp model and the White model. We will not discuss them in this book and the readers can refer to the articles (Jantunen 2004; Yadav et al. 2010).

The analytical models above are for a particular power supply voltage V . For most solid-state amplifiers, V is a constant, and then the analytical models can be directly used for both output prediction and OPD. It is straightforward to predict the output for a given x_a with the analytical models, as demonstrated in Fig. 7.9. When used as OPD, we need to determine the required input x_{aOPD} to achieve the desired output $y_{a,d}$ (we cannot change V here). This requires solving the analytical equation. For example, suppose we have obtained a third-order polynomial model, the required input can be derived by solving $y_{a,d} = a_0 + a_1x_{aOPD} + a_2x_{aOPD}^2 + a_3x_{aOPD}^3$, where x_{aOPD} should have a positive real value. For a klystron with adjustable high voltage, we may fit an analytical model for each V setting (see Fig. 7.5). To predict the output for a given input x_a and high voltage V , we may calculate the output at each high voltage using the corresponding analytical models. To obtain the output at the desired V , we can still use the interpolation algorithm described in the previous section. For klystrons, using the analytical models as OPD is usually over complicated. Therefore, we may use the LUT-based solution as Fig. 7.7 when the OPD functions are required for a klystron with an adjustable high voltage.

7.2.3 *Dynamical Model*

The nonlinear input-output relationship of an amplifier (7.12) describes the steady state. In Chap. 3, we have introduced the dynamical model of the amplifier in Eq. (3.65), where the amplifier is modeled as a first-order low-pass filter applied to the complex-envelope of the input RF signal. Comparing to (7.12), we can rewrite (3.65) as

$$\mathbf{G}_A(\hat{s}) = F_a(V, x_a) e^{jF_p(V, x_a)} e^{-\Delta t_A \hat{s}} \frac{\omega_{1/2,A}}{\hat{s} + \omega_{1/2,A}}. \quad (7.16)$$

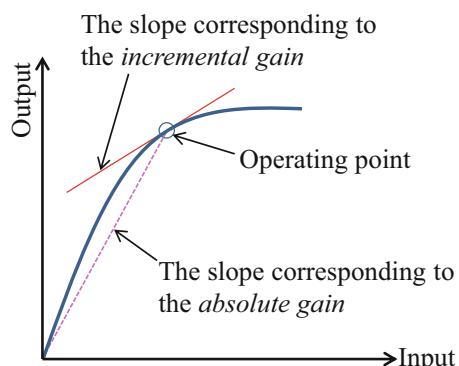
This is the dynamical model of the amplifier, including the nonlinearity, where F_a and F_p are either LUT-based models or analytical models described above. Note that the model (7.16) implies that the bandwidth of the amplifier is independent of the input level. Such an assumption is usually valid if the bandwidths of the amplifier and klystron are much larger than the closed-loop bandwidth of the cavity.

7.3 Handling of Amplifier Nonlinearity in RF Control

Nonlinear RF driving chain components, such as the pre-amplifier and klystron, may affect the stability and performance of the RF control loops. When the amplifier (or klystron) operates close to saturation, the loop gain of the RF feedback may be reduced. Around a particular operating point, the feedback loop gain is determined by the *incremental gain*, as depicted in Fig. 7.10. The incremental gain illustrates the ratio between the resulting output deviation and the corresponding small input change. In contrast, the ratio between the total output and input is denoted as the *absolute gain*.

We notice that at the saturation point, the incremental gain becomes zero and the feedback cannot affect the output by manipulating the input. If the amplifier is over saturated, the incremental gain becomes negative. The dependency of incremental gain to the operating point makes it difficult to guarantee the performance and stability of the RF feedback loop. For example, if the loop is optimized near saturation, the loop gain may exceed the gain margin when the operating point moves towards the linear region, resulting in instability. In contrast, when the operating point moves closer to saturation, the reduced loop gain will degrade the feedback performance in terms of command tracking and disturbance rejection. Particularly, if the operating point is over saturation, the negative incremental gain will change the RF control loop into positive feedback and therefore, make it unstable.

Fig. 7.10 Loop gain reduction when the operating point is close to saturation



To guarantee the RF control performance, we must consider the nonlinearity in the RF driving chain explicitly. First, the magnitude of the input signal (i.e., the DAC drive factor) should be limited to avoid operating the amplifier or klystron above the saturation point. Furthermore, several algorithms have been developed to improve the control stability and performance against nonlinearity. These algorithms are widely used in practical LLRF systems and are discussed in this section. One should keep in mind that these algorithms were developed mainly to compensate for the gain drops for a single-tone input at the RF operating frequency. They might not be optimized for suppressing the intermodulation effects if the amplifier input contains multiple tones (see Sect. 7.1.5). If the intermodulation of multiple tones is concerned, such as for the synchrotron RF systems equipped with one-turn delay feedbacks, selecting better amplifiers (e.g., with constant absolute and incremental gains within the control bandwidth, less intermodulation, and larger 1 dB compression and IP3 power) is essential.

7.3.1 *RF Amplitude Control with High Voltage*

The output power of a klystron (or an amplifier equipped with an adjustable power supply) can be adjusted by changing either the drive power or the power supply high voltage. As illustrated in Fig. 7.5, the amplitude of the klystron output changes with the DAC drive factor or the HVPS. If we fix the input level (e.g., select a DAC drive factor of 0.9), the output amplitude is approximately linear to the HVPS. This implies that we may regulate the klystron output power by manipulating the HVPS even if the klystron operates in saturation. Figure 7.11 shows an RF station controlled by two independent loops. The phase loop is regular, whereas the amplitude loop uses the modulator high voltage to control the accelerating voltage in the cavity or structure.

In normal-conducting Linacs operating in the pulsed mode, the RF stations not requiring fast amplitude regulations are often operated in saturation. At the saturation point, the klystron power efficiency is maximized and the amplitude jitter in the klystron drive signal can be suppressed. For such RF stations, we usually keep the klystron drive power fixed, which should saturate the klystron at the HVPS operating range. Then we adjust the HVPS to regulate the accelerating voltage. Typically, we cannot change the high voltage too fast in order to protect the high-voltage modulator. Therefore, the amplitude loop in Fig. 7.11 is usually slow and is only used to compensate for the slow accelerating voltage drift. At the SwissFEL, we have implemented the HVPS-based amplitude feedback loop, for which the step response is depicted in Fig. 7.12. The feedback controller employs simple discrete integral control. We have limited the ramping-up speed of the HVPS, so it takes longer time when increasing the accelerating voltage. The time constant of the closed loop can be estimated to be about several seconds from the response of a stepping down of the accelerating voltage set point.

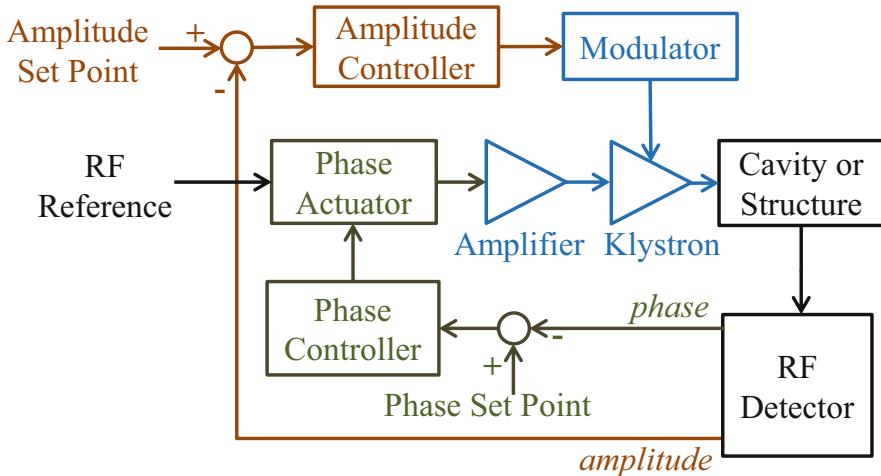


Fig. 7.11 An RF station controlled with a phase loop and an amplitude loop actuating on the modulator high voltage

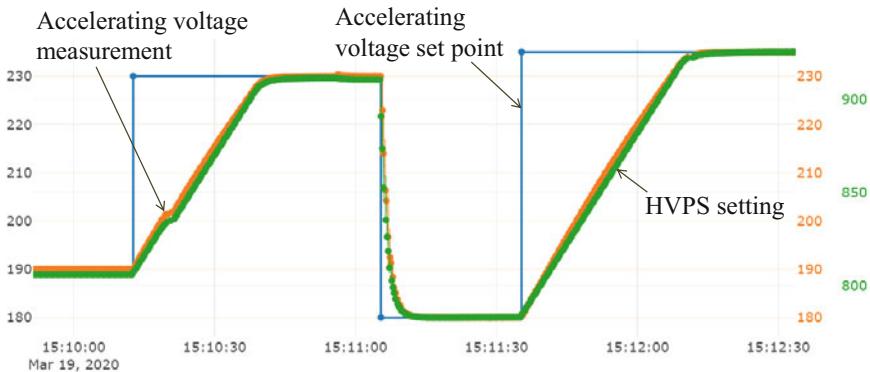


Fig. 7.12 Step response of the HVPS-based amplitude feedback loop at SwissFEL

Note that the fast phase regulation is still possible even if the klystron operates in saturation. Therefore, we may adjust the phases of multiple RF stations (without changing the amplitudes) to change the vector-sum amplitude and phase for beam acceleration. This approach enables the RF stations operating in saturation as fast actuators for the control of beam parameters (e.g., beam energy, bunch length, etc.). For example, assume that we have two RF stations, both operating in saturation with the same accelerating voltage 200 MV and the same phase 0° . The cavity voltage phasors induced by the two RF stations are denoted as $\mathbf{v}_{C1} = \mathbf{v}_{C2} = 200e^{j0}$ and the overall accelerating voltage phasor is $\mathbf{v} = \mathbf{v}_{C1} + \mathbf{v}_{C2} = 400e^{j0}$. Now we want to change the overall accelerating voltage to be 350 MV with a phase of 10° , that is,

$\mathbf{v}' = 350e^{j10\pi/180}$. Without changing the amplitudes of the two RF stations, we can choose $\mathbf{v}'_{C1} = 200e^{j\alpha}$ and $\mathbf{v}'_{C2} = 200e^{j\beta}$, where α and β can be determined by solving the equation $\mathbf{v}' = \mathbf{v}'_{C1} + \mathbf{v}'_{C2}$. Here we directly give the results: $\alpha = 38.96^\circ$, $\beta = -18.95^\circ$.

7.3.2 Gain Scheduling

The gain scheduling can be employed for the proportional (P) control or proportional-integral (PI) control to mitigate the effects of nonlinearity. The simplest implementation is to adopt different feedback gains for different operating points to compensate for the incremental gain changes depicted in Fig. 7.10. More generally, we may adopt different linear controllers (e.g., with different structures or parameters) when the operating point changes. This technique enables us to control a nonlinear system in a large range of operating points with a set of linear controllers (Ilka 2015).

In practice, the gain scheduling can also be used for linear systems to implement a trade-off between the required drive power and the control performance (Wibowo et al. 2018). A larger gain (assume the feedback loop is stable) usually results in better performance in terms of command tracking and disturbance rejection, but also requires larger drive power. Figure 7.13 shows the control of a TESLA cavity with both feedforward and feedback employing a P controller. The cavity parameters are as follows: resonance frequency $f_0 = 1.3$ GHz, $r/Q = 1036 \Omega$, $Q_L = 3e6$ and the loop delay is 5 μ s. As depicted in Fig. 7.13c, the cavity voltage ramps up before 500 μ s and a flattop follows where the beam is accelerated. To fill the cavity faster, a higher drive power is required, as shown in Fig. 7.13b. Typically, during the filling period, we do not need the cavity voltage to follow the set point strictly, because there is no beam in this period. Instead, we must minimize the peak drive power to relax the power requirements for the klystron. During the flattop of the cavity voltage, less drive power is required, and our primary goal is to reduce the cavity voltage error for stable beam acceleration. The different requirements imply that we can use different gain settings for the filling and flattop periods; that is, to use the gain scheduling. As shown in Fig. 7.13a, we ramp the gain setting during the filling period, which results in less peak power shown in Fig. 7.13b. The resulting cavity voltage (see Fig. 7.13d) does not change much for these two cases with constant gain or with gain scheduling. Note that the overshoot in the cavity voltage is not avoidable due to the large loop delay, which reduces the gain/phase margin of the cavity control loop.

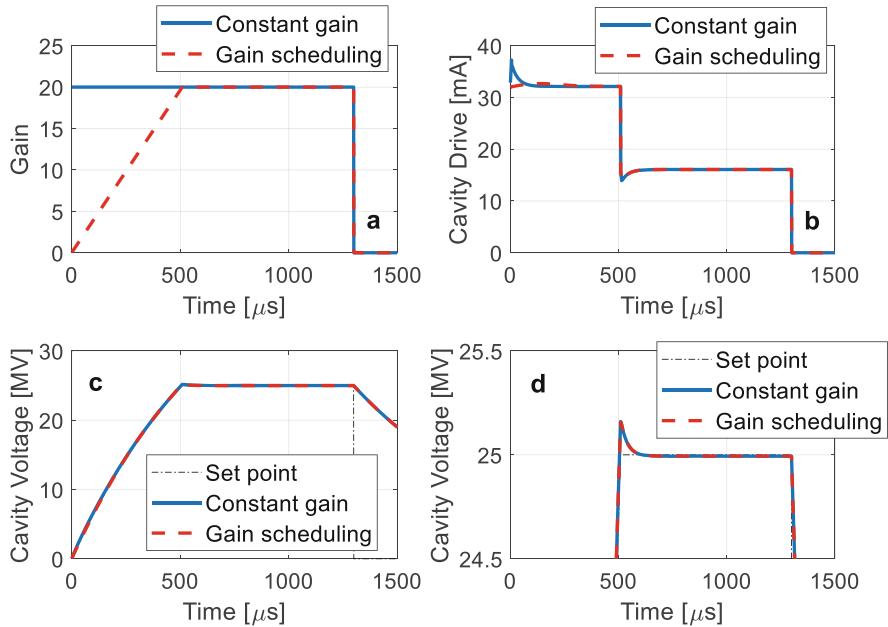


Fig. 7.13 Gain scheduling to reduce peak drive power for an RF station operating in pulsed mode. (a) gain setting; (b) cavity drive signal; (c) cavity voltage; (d) cavity voltage zoomed in around flattop

7.3.3 LUT-Based Linearization

Linear responses of the amplifier and klystron are preferred for the RF controller design. Benefiting from the digital implementation of RF controllers, we may pre-distort the output of the RF controller with a LUT, so that the output of the klystron appears as the output of a linear system (Cichalewski and Jalmuzna 2009; Omet et al. 2014; Sajedin et al. 2014; Bellandi et al. 2020). This idea is depicted in Fig. 7.14.

Assume that we have obtained the nonlinear responses of the RF driving chain for different DAC drive factors (i.e., the RF controller output as a DAC output level, which is the input of the RF driving chain). The desired amplitude response is a linear function from zero amplitude to the saturation point, and the desired phase response is a constant phase shift for all input levels. See Fig. 7.14. According to the desired amplitude response, if the RF driving chain input is $x_{a,d}$, the expected output is $y_{a,d}$. Due to the nonlinear amplitude response, we must use $x_{a,act}$ as the actual input to the RF driving chain to obtain $y_{a,d}$. Furthermore, we find out from the phase response that at $x_{a,act}$, we must rotate the phase of the RF driving chain input by $\Delta\theta$ to achieve the constant phase response. In summary, for the RF controller output $x_{a,d}$, if we change its amplitude to $x_{a,act}$ and rotate its phase by $\Delta\theta$ before inputting into the RF driving chain, we obtain an output from the klystron following the desired

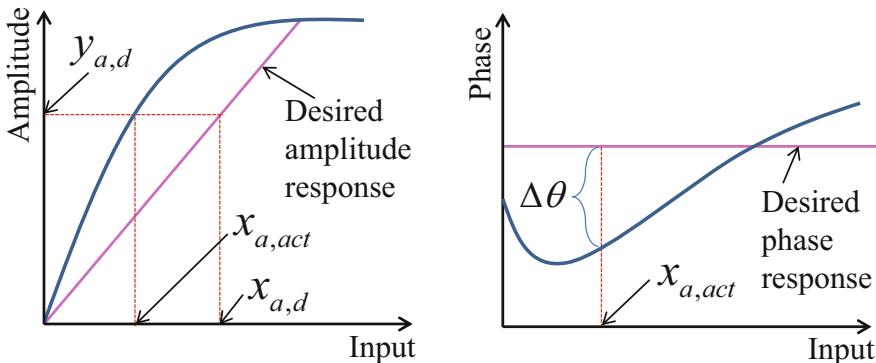


Fig. 7.14 Pre-distortion of the RF controller output to linearize the RF driving chain

responses above. This is a technique called *pre-distortion* to linearize the RF driving chain.

Typically, we can implement the pre-distortion algorithm using a LUT, which consists of two tables mapping $x_{a,d} \rightarrow x_{a,act}$ and $x_{a,d} \rightarrow \Delta\theta$. The LUT contains the mapping for different values of $x_{a,d}$ from zero to its maximum value at the saturation point. If a given $x_{a,d}$ does not match any exact points in the LUT, interpolation algorithms can be used to obtain the corresponding $x_{a,act}$ and $\Delta\theta$. The linear interpolation algorithm has been discussed in Sect. 7.2.1.

In practice, we can implement the LUT-based RF driving chain linearizer in the FPGA. To fit the FPGA calculation, we may vary slightly the mapping to $x_{a,d}^2 \rightarrow x_{a,act}/x_{a,d} \cdot \cos \Delta\theta$, $x_{a,d}^2 \rightarrow x_{a,act}/x_{a,d} \cdot \sin \Delta\theta$. Note that in FPGAs, the values are usually presented as fixed-point numbers. Figure 7.15 shows an example of the LUT-based linearizer that is used to pre-distort the RF controller output before inputting it into the RF driving chain via DACs.

7.3.4 Linearization Loop

The bandwidth of the RF driving chain, including the RF amplifier and klystron, is usually much larger than the closed-loop bandwidth of the cavity. Therefore, an additional inner loop can be implemented to regulate the RF driving chain output to compensate for the nonlinearity (Dawson and Lee 2004; Fox et al. 2005). The schematic of the RF driving chain linearizer loop can be found in Fig. 7.16. This method is widely used in RF systems operating in the CW mode or with long pulses.

The RF driving chain linearizer loop (inner loop) is cascaded with the cavity control loop (outer loop). When the response of the inner loop is much faster than that of the outer loop, the closed-inner loop appears as a linear component from the viewpoint of the outer loop. That is, the amplifier and klystron are effectively “linearized”.

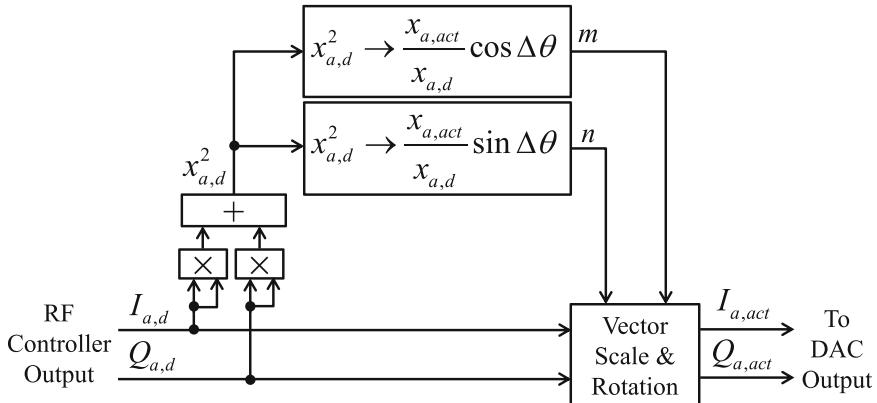


Fig. 7.15 An example of the LUT-based RF driving chain linearizer

Fig. 7.16 Feedback-based linearizer of the RF driving chain

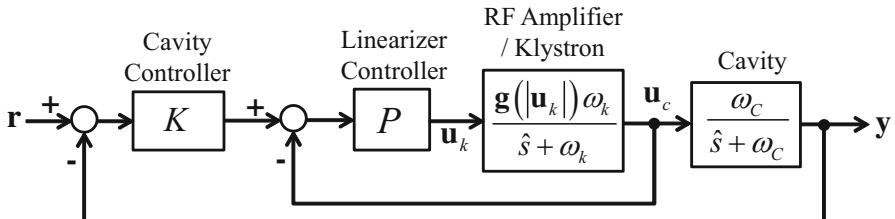
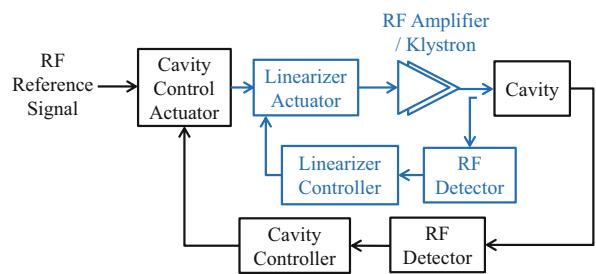
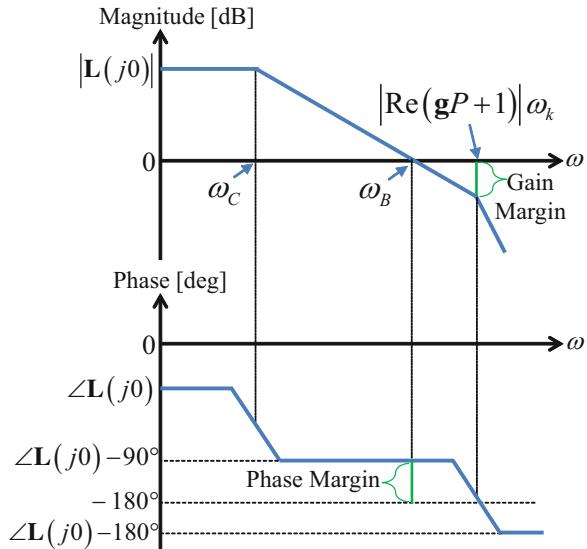


Fig. 7.17 Cascaded feedback loops for the RF driving chain linezrization and cavity control

The closed-loop bandwidths of the inner- and outer-loops should satisfy the stability conditions that are briefly discussed with the simplified model in Fig. 7.17. Here we assume that both loops are controlled with simple P controllers. The RF driving chain (RF amplifier and klystron) and the cavity are modeled by two first-order transfer functions with a half-bandwidth ω_k and ω_C , respectively. The nonlinearity of the RF driving chain is modeled as a complex gain that depends on the input level. The open-loop transfer function of the outer loop (including the closed-loop transfer function of the inner loop) can be calculated as

Fig. 7.18 Bode plot of the open-loop transfer function of the cascaded loops



$$L(\tilde{s}) = \frac{gKP\omega_k\omega_C}{(\tilde{s} + (\mathbf{g}P + 1)\omega_k)(\tilde{s} + \omega_C)}. \quad (7.17)$$

This is a phasor transfer function with a complex coefficient \mathbf{g} . The steady-state gain of the open-loop system is $L(j0) = gKP/(gP + 1)$. The Bode plot of \mathbf{L} is sketched in Fig. 7.18.

The open-loop transfer function \mathbf{L} has two poles at $p_1 = -\omega_C$ and $p_2 = -(gP + 1)\omega_k$. Since the bandwidth of the cavity is much smaller, we have $|\text{Re}(p_2)| \gg |\text{Re}(p_1)|$, where $\text{Re}(p_2)$ is the real part of p_2 . The Bode plot illustrates that the two poles yield a phase lag of -180° at the frequencies where $\omega > |\text{Re}(p_2)|$. To keep the closed-outer loop stable, we need a phase margin larger than zero; that is, the phase response at ω_B should be above -180° . Here ω_B is the zero-crossing frequency of \mathbf{L} , which approximates the closed-loop bandwidth of the outer loop. Therefore, to keep the cascaded loops stable, we require at least

$$\omega_B < |\text{Re}(gP + 1)|\omega_k. \quad (7.18)$$

Note that ω_B is usually larger than the cavity bandwidth ω_C to suppress the effects of the external disturbances coupled into the cavity. The relationship (7.18) implies that the closed-loop bandwidth of the inner loop should be larger than that of the outer loop. We notice that the phase of the complex gain \mathbf{g} in the klystron model affects the total phase response of the Bode plot via the phase of $L(j0)$. Therefore, \mathbf{g} may decrease the phase margin if the phase of \mathbf{g} is negative. Of course, when the gain of the inner loop is large (i.e., $|\mathbf{g}P| \gg 1$), the phase of $L(j0)$ is close to zero. The inner loop should prevent the operation of the klystron in saturation or over-saturation, where the loop gain becomes zero, or negative leading to instability.

7.4 Summary

This chapter provides an overview of the nonlinearity in accelerator RF systems. Nonlinear RF detectors cause RF measurement errors. The nonlinearity in the RF driving chain components, such as the RF amplifier or klystron, affects the RF control performance by either reducing the loop gain or producing intermodulation disturbances. We introduced the RF amplifier nonlinearity models that can be used to determine the operating point of the RF system. Finally, we discussed several widely used methods (amplitude feedback with modulator HV, gain scheduling, and amplifier linearization based on look-up tables or feedbacks) to mitigate the nonlinearity in accelerator RF systems.

References

- Bellandi A, Ayvazyan V, Butkowsk L et al (2020) Results on FPGA-based high-power tube amplifier linearization at DESY. *IEEE Trans Nucl Sci* 67(5):762–767. <https://doi.org/10.1109/TNS.2020.2980929>
- Blas F, Garoby R (1991) Design and operational results of a “one-turn-delay feedback” for beam loading compensation of the CERN PS ferrite cavities. Paper presented at the IEEE 1991 particle accelerator conference. San Francisco, CA, USA, 6–9 May 1991
- Cho C (2005) RF circuit nonlinearity characterization and modeling for embedded test. Ph.D. thesis, University of Florida
- Cichalewski W, Jalmuzna W (2009) High power amplifiers chain nonlinearity influence on the accelerating beam stability in free electron laser (FLASH). In: Proceedings of the 16th international conference mixed design of integrated circuits & systems, Lodz, Poland, 25–27 June 2009
- Dawson J L, Lee T H (2004) Cartesian feedback for RF power amplifier linearization. In: Proceedings of the 2004 American control conference, Boston, MA, USA, 30 June – 2 July 2004
- Fox J, Mastorides T, Teytelman D et al (2005) Klystron linearizer for use with 1.2 MW 476 MHz klystrons in PEP-II RF systems. In: Proceedings of 2005 particle accelerator conference, Knoxville, TN, USA, 16–20 May 2005
- Fox J, Mastorides T, Rivetta C et al (2007) Selecting RF amplifiers for impedance controlled LLRF systems – nonlinear effects and system implications. In: Proceedings of the 22ed particle accelerator conference, Albuquerque, NM, USA, 25–29 June 2007
- Fox J, Mastorides T, Rivetta C et al (2010) Lessons learned from positron-electron project low level RF and longitudinal feedback. *Phys Rev ST Accel Beams* 13(5):052802. <https://doi.org/10.1103/PhysRevSTAB.13.052802>
- Hoffmann M (2008) Development of a multichannel RF field detector for the low-level RF control of the free-electron laser at Hamburg. Ph.D. thesis, Technische Universität Hamburg
- Ilka A (2015) Gain-scheduled controller design. Ph.D. thesis, Slovak University of Technology in Bratislava
- Jantunen P (2004) Modelling of nonlinear power amplifiers for wireless communications. Dissertation, Helsinki University of Technology
- Kumar A, Reddy B, Reddy G (2010) Modeling RF power amplifier to study its nonlinear effects on RF communication system, with BER as a performance measure. *Int J Next-Gen Network* 2(4): 73–81. <https://doi.org/10.5121/ijngn.2010.2406>
- Mastoridis T, Baudrenghien P, Molendijk J (2012) LHC one-turn delay feedback commissioning. In: Proceedings of the 3rd international particle accelerator conference, New Orleans, LA, USA, 20–25 May 2012

- Omet M, Michizono S, Matsumoto T et al (2014) FPGA-based klystron linearization implementations in scope of ILC. *Nucl Instrum Methods Phys Res A* 768:69–76. <https://doi.org/10.1016/j.nima.2014.09.007>
- Rezaeizadeh A, Kalt R, Schilcher T et al (2014) Model-based klystron linearization in the SwissFEL test facility. In: Proceedings of the 36th international free electron laser conference, Basel, Switzerland, 25–29 August 2014
- Sajedin M, Ghorbani A, Davar H (2014) Nonlinearity compensation for high power amplifiers based on look-up table method for OFDM transmitters. *Int J Adv Comput Sci & Inf Technol* 3(4):354–367
- Schmid L, Baudrenghien P, Hagmann G (2017) One-turn delay feedback with a fractional delay filter. <https://public.cells.es/workshops/www.llrf2017.org/pdf/Posters/P-77.pdf>. Accessed 29 May 2021
- Wibowo SB, Matsumoto T, Michizono et al (2018) Digital low level RF control system for the international linear collider. *Phys Rev Accel Beams* 21(8):082004. <https://doi.org/10.1103/PhysRevAccelBeams.21.082004>
- Yadav A, Mazumdar D, Karthikeyan B R et al (2010) Linearization of Saleh, Ghorbani and Rapp amplifiers with Doherty technique. <http://www.sastechjournal.com/pdf/Journals/Sept2010/11.pdf>. Accessed 30 May 2021

Chapter 8

Timing and Synchronization



The subsystems of a particle accelerator must be turned on in a proper sequence to perform successful beam acceleration. Furthermore, the charged particles must interact with the RF fields (or the laser fields in case of laser-plasma acceleration is used) at the correct time for the desired accelerating phase. The timing and synchronization systems guarantee the required timing relations between the RF fields and the beam. The timing system defines the timing events and produces triggers if required for different subsystems to start or stop their operation. The synchronization system provides a common frequency and phase reference to all subsystems. We will introduce the basic concepts and architecture of the timing and synchronization systems in this chapter.

8.1 Overview

Two conditions for the timing relations should be satisfied for a successful acceleration of the charged particles. First, the RF fields in the cavities or structures must be established before the particles arrive. Second, the particles must be generated at a proper time, so that it travels to the cavity or structure and experiences the correct phase of the RF fields. The timing system satisfies the first condition, and the synchronization system guarantees the second (Loehl 2011; Schlarb 2013; Bellaveglia 2016; Damerau 2017; Gallo 2018).

A *timing system* provides triggers referenced to the timing events to different subsystems of an accelerator. The *timing event* is a logical concept consisting of information like the purpose of the event and defining a reference time by the time instant when it is received. On receiving a timing event, the timing system produces *triggers* in either electrical signals or firmware/software signals. The triggers for the different subsystems define the sequence in which they are activated. Let us use the linear accelerator (Dusatko et al. 2010; Kalantari and Biffiger 2017) in Fig. 1.14 as an example. The sequence of triggers to accelerate an electron bunch is as follows.

First, the Gun laser amplifier is turned on because it requires a longer time to reach a steady state. Then the RF pulses are started to feed RF power to the cavities or structures. At the same time, the beam diagnostic devices are triggered to start the measurement. Finally, the Gun laser pulse picker is triggered to release a laser pulse that produces an electron bunch in the photocathode RF Gun. The bunch is then accelerated by the downstream cavities or structures, which has established their desired accelerating voltages when the bunch arrives. The trigger sequence for a circular machine is more complicated. It includes the triggering of multiple kicker magnets to inject or extract particle bunches between the injector Linac and the rings (Chernousko et al. 2006; Zhao et al. 2008; Liu et al. 2021). The tolerance of the trigger timing jitter should be within several picoseconds or nanoseconds (i.e., a fraction of a period of the RF frequency), depending on the requirements of concrete applications. The timing system also provides timestamps to the data of different subsystems so that the data corresponding to the same beam acceleration cycle (e.g., amplitudes and phases of multiple RF stations accelerating the same bunch) can be correlated. The timestamp can be in the form of actual time, bunch index (ID), or pulse ID (for pulsed machines).

In contrast, the *synchronization system* provides precise timing control between different subsystems using continuous synchronization signals (e.g., RF signal). The timing jitter of a synchronization signal should be as low as several femtoseconds for a stable beam acceleration. As depicted in Fig. 1.14, a master oscillator (MO) provides the ultimate frequency and phase reference for the entire accelerator. The synchronization signals derived from the MO are distributed to different subsystems in terms of continuous-wave (CW) sinusoidal RF signals or a laser pulse train. At each RF station, as shown in Fig. 1.13, the synchronization signal is converted to an RF reference signal, providing frequency and phase reference for the RF station.

The low-level RF (LLRF) system is responsible for locking the frequency and phase of the RF fields in the cavities or structures to the RF reference signal. It is accomplished via LLRF feedbacks. For this purpose, the LLRF system generates other required frequencies like the local oscillator (LO) and clock signals using the RF reference signal. Furthermore, the LLRF system is also responsible for controlling the RF field phase to match the particle arrival time for a correct beam phase.

8.2 Master Oscillator

The primary concern of a master oscillator is the long-term frequency stability and the phase noise representing the short-term stability. We have discussed the MO phase noise in Sect. 6.4.3. The slow drifts of the MO central frequency are caused by the changes of temperature and humidity, or by the aging of the oscillator. For example, the SwissFEL MO employs a commercial RF signal generator (Rohde&Schwarz SMA100B) with a frequency drift below 1e-7 per year. To achieve a good frequency (phase) stability, we may combine multiple oscillators, which are either RF oscillators or laser oscillators. This topic will be discussed in Sect. 8.2.2.

Another critical requirement for the master oscillator is the high reliability and availability, which is fundamental for reliable beam operation. Therefore, redundancy is considered in the design of master oscillators, like the work at the European XFEL (Gasowski et al. 2018).

8.2.1 RF and Laser Oscillator

The *RF master oscillator (RMO)* often consists of an oven-controlled crystal oscillator (OCXO) with a high frequency-stability. The short-term frequency stability of OCXOs is typically $1\text{e-}12$ over a few seconds, while the long-term stability is limited to around $1\text{e-}8$ per year by the crystal aging. To improve the long-term stability, we may consider locking the OCXO to a Global Positioning System (GPS) disciplined rubidium oscillator (GPSD-Rb) (Zembala et al. 2014) that produces a 10 MHz reference output. The long-term frequency stability of a GPSD-Rb can be better than $1\text{e-}11$ per year. The typical frequency of an OCXO is below 200 MHz because a higher frequency requires a smaller crystal that is difficult to manufacture with a high frequency-accuracy. The RF frequency (e.g., 1.3 GHz for the RF systems driving TESLA cavities) is then derived from the OCXO output with a frequency multiplier or a phase-locked loop (PLL). A dielectric resonator oscillator (DRO) is often used in the PLL to synthesize the RF frequency at GHz level.

Nowadays, some commercial signal generators can produce RF signals with ultralow phase noise. For example, the above-mentioned Rohde&Schwarz SMA100B signal generator (with the enhanced phase noise option SMA-B22) can produce a 3 GHz RF signal with phase noise better than -133 dBc/Hz at 20 kHz offset from the carrier frequency. Such commercial RF signal generators can be directly used as the RMO if their performance satisfies the requirements. In this case, the 10 MHz reference signal generated by the GPSD-Rb can be directly connected to the commercial signal generator, often configurable to lock to an external reference. Figure 8.1 illustrates a typical master oscillator, including the RMO described above.

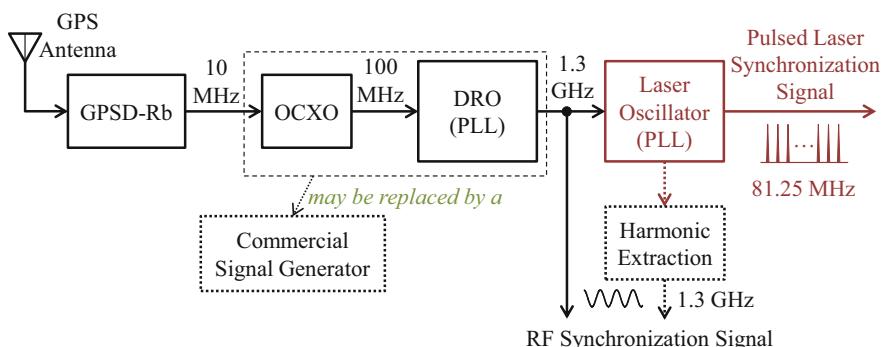


Fig. 8.1 Typical architecture of a master oscillator (the frequency values are examples)

A *laser master oscillator (LMO)* is typically formed by a mode-locked laser (e.g., Ti:sapphire laser) that generates laser pulses with pulse widths down to femtoseconds. The repetition rate of the laser pulses is often tens to hundreds of MHz. The high-frequency phase noise of the LMO is extremely low, whereas its low-frequency phase fluctuation may be large due to the thermal effects. Therefore, the LMO is always phase locked to an RMO (Titberidze 2017), which has much better low-frequency phase stability. We will provide more details about locking a laser oscillator to an RF reference signal in Sect. 8.4.3. In some applications (e.g., SwissFEL master oscillator, Hunziker et al. 2014), an RF signal is extracted from the pulsed laser signal generated by the LMO phase-locked to an RMO. This extracted RF signal has lower phase noise than that directly generated by the RMO, benefiting from the lower high-frequency phase noise of the LMO. The phase noise characteristics of two locked oscillators will be introduced in Sect. 8.2.2.

The synchronization system distributes either the *RF synchronization signal* or the *pulsed laser synchronization signal* to different accelerator subsystems. As mentioned above, the RF synchronization signal can directly be the RMO output or extracted from the pulsed laser synchronization signal. The synchronization system will be discussed in Sect. 8.4. Note that for small-scale accelerators, the LMO may be optional and a low-noise RMO is sufficient to provide frequency and phase references for the machine.

8.2.2 *Synchronization of Two Oscillators*

Phase-locked loops are used to synchronize two independent oscillators. The phase noise of the PLL output is a combination of the phase noise of both oscillators (Mehrotra 2002; Best 2007; Serrano et al. 2011). Therefore, with a PLL, we may achieve better phase stability than either oscillator running freely.

A general architecture of a PLL is shown in Fig. 8.2a. Here r is the reference signal produced by a reference oscillator, and y is the loop output generated by the PLL oscillator. The frequency of the PLL oscillator is controllable either by voltage (e.g., voltage-controlled oscillator – VCO) or by mechanical control using a motor or piezo translator (e.g., laser oscillator). In a physical PLL, the reference and PLL oscillators may be RF oscillators or laser oscillators; the reference and output signals may be CW RF signals or pulsed laser signals. Different combinations of these oscillators and signal types require different technologies to implement the PLL components like the phase detector and loop filter. However, the relationship between the phase noise of y and that of the two oscillators follows the same principle. Therefore, we assume that both oscillators are RF oscillators, and all the signals are CW RF signals to simplify the discussion. The frequency dividers with ratio M and N determine the relation between the frequencies of the two oscillators:

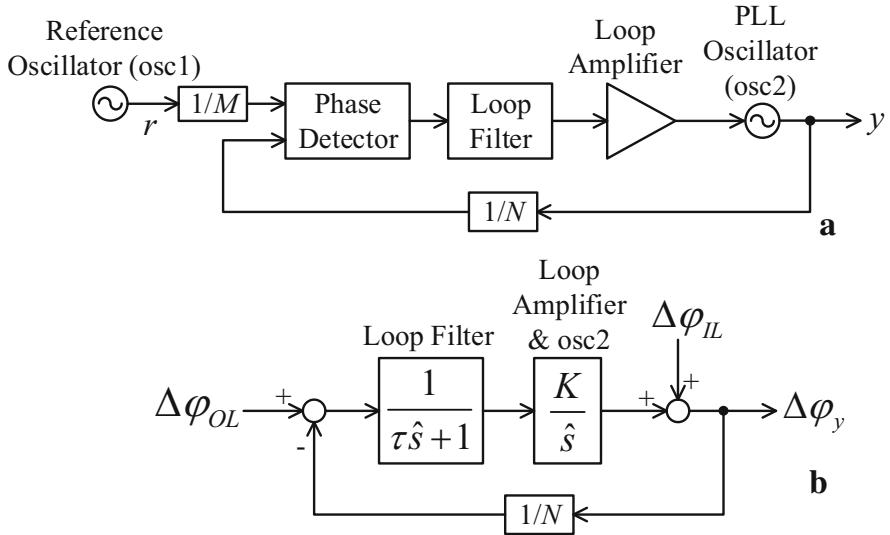


Fig. 8.2 General architecture of an PLL (a) and its phase noise model (b)

$$f_y = \frac{N}{M} f_r, \quad (8.1)$$

where f_y and f_r are the frequencies of y and r , respectively. The PLL output phase noise is a function of the phase noise in the reference and PLL oscillators, frequency dividers, phase detector, and loop amplifier. These noise sources contribute to the output phase noise following the PLL's closed-loop dynamics, see Fig. 8.2b. To simplify the discussion, we define the in-loop and out-of-loop phase deviations as

$$\begin{aligned} \Delta\varphi_{IL} &= \Delta\varphi_{amp} + \Delta\varphi_{osc2}, \\ \Delta\varphi_{OL} &= \Delta\varphi_{osc1}/M + \Delta\varphi_{divr} - \Delta\varphi_{PD} - \Delta\varphi_{divy}. \end{aligned} \quad (8.2)$$

Here $\Delta\varphi_{IL}$ is the in-loop phase noise, including the contributions from the loop amplifier ($\Delta\varphi_{amp}$) and the PLL oscillator ($\Delta\varphi_{osc2}$). As in Fig. 8.2b, we have defined that $\Delta\varphi_{IL}$ is added to the output of the PLL oscillator. Therefore, the term $\Delta\varphi_{amp}$ is the PLL oscillator output phase deviation caused by the amplifier noise. As a comparison, the term $\Delta\varphi_{osc2}$ is the phase noise of the PLL oscillator itself. The out-of-loop phase noise, $\Delta\varphi_{OL}$, includes the phase noise of the reference oscillator ($\Delta\varphi_{osc1}$), phase detector ($\Delta\varphi_{PD}$), and frequency dividers ($\Delta\varphi_{divr}$ and $\Delta\varphi_{divy}$ for the reference and output RF signals, respectively). We define that $\Delta\varphi_{OL}$ is added in the reference input of the closed loop. Note that the phase deviation of the reference oscillator is divided by M by the frequency divider. We have minus signs for $\Delta\varphi_{PD}$ and $\Delta\varphi_{divy}$ due to the phase subtraction in the phase detector.

Using the block transfer functions in Fig. 8.2b, where the loop filter is modeled as a first-order low-pass filter, the noise transfer functions can be written as

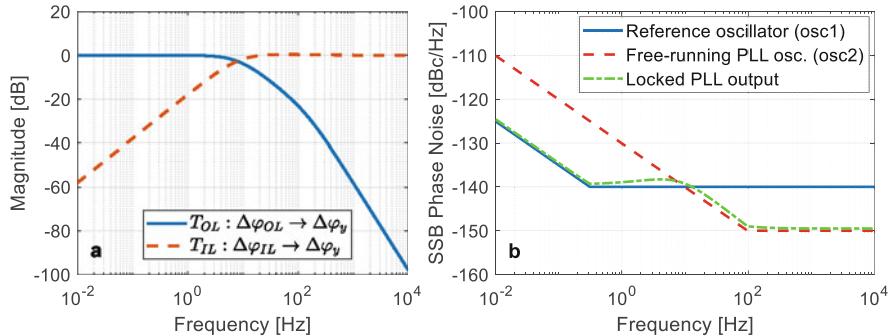


Fig. 8.3 (a). Frequency responses of the noise transfer functions of an PLL ($M = N = 1$); (b). Achieved phase noise with an PLL synchronizing two oscillators.

$$T_{IL}(\hat{s}) = \frac{N\tau\hat{s}^2 + N\hat{s}}{N\tau\hat{s}^2 + N\hat{s} + K}, T_{OL}(\hat{s}) = \frac{NK}{N\tau\hat{s}^2 + N\hat{s} + K}. \quad (8.3)$$

Here T_{IL} and T_{OL} are the transfer functions from $\Delta\varphi_{IL}$ and $\Delta\varphi_{OL}$ to $\Delta\varphi_y$, respectively. Their frequency responses (see Fig. 8.3a, with $M = N = 1$) are helpful for understanding the noise transfer relationship in the PLL. It is seen that within the closed-loop bandwidth of the PLL, the in-loop phase noise is suppressed but the out-of-loop phase noise is transferred to the output. Beyond the closed-loop bandwidth, the in-loop phase noise is transferred to the output whereas the output-of-loop noise is suppressed. Therefore, within the closed-loop bandwidth, the PLL output's phase noise is determined by the reference oscillator; in contrast, at higher frequencies, it is dominated by the PLL oscillator. Here we have neglected the phase noise added by the PLL components (phase detector, loop amplifier and frequency dividers). If the reference oscillator has lower low-frequency noise and the PLL oscillator has lower high-frequency noise, the PLL can combine them to achieve lower phase noise than either of them. See Fig. 8.3b.

8.3 Timing System

The timing system coordinates the actions of different accelerator subsystems, such as the RF, laser, beam diagnostics, kicker magnets, etc. Typical timing jitter is in the range of nanoseconds to picoseconds. A typical timing system (Korhonen 1999) for a large-scale accelerator is depicted in Fig. 8.4.

The *timing fiducial*, which defines the basic repetition rate of the machine, is often derived from the alternating current (AC) mains. The fiducial trigger is synchronized by a *common subharmonic signal* derived from the *master timing clock* that is distributed by the synchronization system (see Fig. 1.14). Then in the *event generator* (EVG), which is typically a digital device, a series of digital events are generated

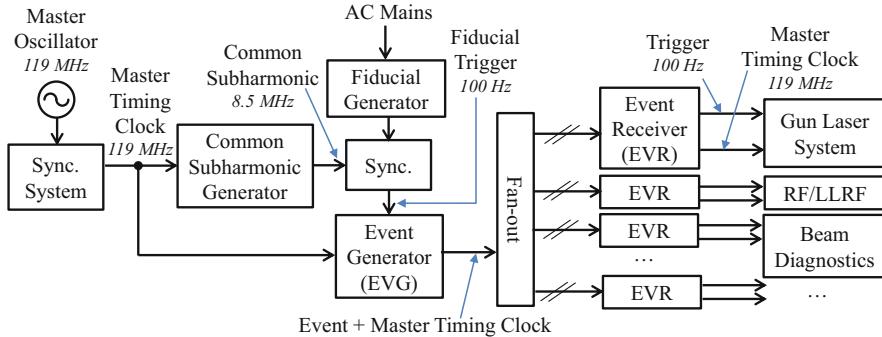


Fig. 8.4 Typical architecture of a timing system (the frequency values are examples)

on the arrival of each fiducial trigger. Synchronized by the master timing clock, the events are distributed to different locations of the accelerator. Note that the master timing clock is also distributed so that at the *event receiver (EVR)* side, triggers synchronized to the master timing clock can be generated. Optical fibers are typically adopted to distribute the events and master timing clock. In the EVR, the events are received and decoded. Each event has a specific purpose, such as starting the RF pulse, picking a laser pulse from a pulse train, or starting the kicker magnets. The EVR generates triggers for each subsystem when receiving the corresponding events and by adding specified delays. Besides generating and distributing events and master timing clock, the EVG also defines the time, bunch ID, or pulse ID (for a pulsed accelerator), and distributes them to the EVRs. They are used to timestamp the data of subsystems so that they can be correlated when performing data analysis.

8.3.1 Timing Fiducial Generation

A timing fiducial starts a sequence of timing events to perform a complete cycle of beam acceleration in a particle accelerator. In a pulsed Linac, the fiducial starts a bunch acceleration cycle, in which an electron bunch is generated, accelerated, and consumed (e.g., producing FEL). In a synchrotron light source, the fiducial may start an event sequence to generate an electron bunch with the injector Linac, inject the bunch into a booster ring to ramp its energy, extract the bunch from the booster, and inject it into the final storage ring.

The timing fiducial is typically derived from the zero-crossing time of the AC mains. See Fig. 8.5. The frequency of the AC mains is 50 Hz or 60 Hz, depending on the country where the machine is located. If we use all the three phases of the AC mains and the two zero-crossing points in a period of each phase, we can obtain a fiducial trigger with a rate of 6 times the AC frequency. For example, if the AC frequency is 60 Hz, a 360 Hz fiducial trigger can be obtained.

Fig. 8.5 Pulse generation from the zero-crossing times of AC mains (only one phase is displayed)

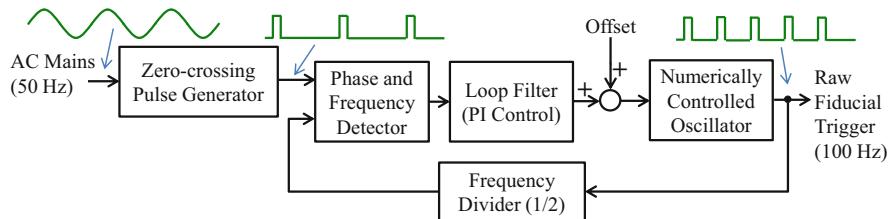
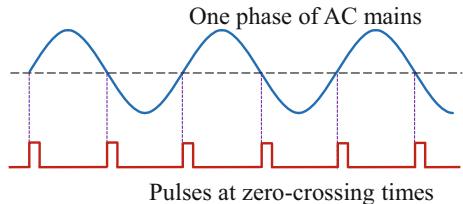


Fig. 8.6 Fiducial trigger generation with a digital PLL locking to the AC mains (the frequency values are examples)

Deriving the timing fiducial from the AC mains helps to sequence the triggers of various accelerator subsystems in a consistent manner so as not to perturb the AC line (Dusatko et al. 2010). Furthermore, running the beam acceleration cycle (e.g., accelerating a bunch in an RF pulse) synchronously to the AC mains can avoid the time-varying disturbances caused by the AC mains. For example, the 50 Hz AC mains may cause phase noise in the master oscillator at 50 Hz offset to the carrier frequency. If we run the 50 Hz RF pulses at the zero-crossing times of the AC mains, the RF pulses will experience a constant disturbance from the master oscillator phase noise caused by the AC mains. A constant disturbance is often much easier to be suppressed by the RF feedback loop. However, the AC-frequency fluctuations, which are not avoidable in the AC power system, will result in time interval changes between the fiducial trigger pulses. It may introduce fluctuations in some devices (e.g., the solid-state modulators of SwissFEL) due to different charging periods between the trigger pulses. Therefore, when generating the fiducial trigger, we must make a compromise between the tracking of AC mains and the equal trigger intervals.

A digital PLL, as shown in Fig. 8.6, can be adopted to lock the fiducial trigger to the pulses generated by the zero-crossing detection of the AC mains (Rybacyk and Shelley 1997; Rose et al. 2001). The loop filter, which is typically a PI controller, can be adjusted to tune the locking bandwidth to the AC mains. Larger feedback gains result in better tracking of the AC mains whereas smaller gains can improve the equal spacing of the fiducial trigger pulses. Another benefit of the PLL is to synthesize different frequencies other than the AC frequency. In the example of Fig. 8.6, we use one phase of the 50 Hz AC mains to generate a 100 Hz fiducial

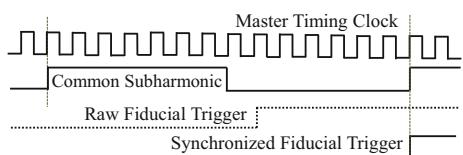
trigger. Such a digital PLL can be implemented in an FPGA. The phase and frequency detector (PFD) detects the time difference (e.g., as counting of FPGA clock cycles) between the rising edges of the two input square pulses. The numerically controlled oscillator (NCO) can be implemented as a pulse generator with its period controlled by the loop filter output.

8.3.2 Common Subharmonic

The raw fiducial trigger derived from the AC mains is not yet synchronous to the master oscillator. That is, the rising edges of the trigger pulses are fluctuating in time with respect to the rising edges of the master timing clock, which is phase locked to the MO. We must synchronize the trigger pulses with the machine clock to obtain a *synchronized fiducial trigger*. The goal is to make all the RF and clock frequencies derived from the MO ensure the same phases after each trigger (for a square-pulsed clock, we still use “phase” to represent the time difference to the rising edge). This is essential for repeatable beam qualities in each acceleration cycle started by a trigger pulse. It also implies that the time interval between two trigger pulses must cover full cycles of all the concerned frequencies. To accomplish this goal, we need to derive a *common subharmonic* of the machine frequencies and synchronize the raw fiducial trigger with it. The common subharmonic frequency is a common divisor of all the concerned frequencies. For example, the common subharmonic frequency of 2856 MHz, 119 MHz, 68 MHz, 102 MHz, and 25.5 MHz is 8.5 MHz. We may obtain the common subharmonic frequency from the master timing clock (119 MHz) with a frequency divider (1/14). See Fig. 8.4. The schematic to synchronize the raw fiducial trigger is depicted in Fig. 8.7. After synchronization, the trigger pulses’ rising edges will be aligned to the rising edges of the common subharmonic and the master timing clock.

Note that the common subharmonic frequency should be the common divisor of ALL the RF and clock frequencies in the machine, such as the RF operating frequency, laser oscillator frequency, LLRF LO and clock frequencies, etc. Therefore, we should use only a few different frequencies in the machine so that the common subharmonic frequency is not much lower than the lowest frequency. The time interval between two trigger pulses is a multiple of the common subharmonic period. Therefore, a lower common subharmonic frequency will cause a larger trigger interval variation when the raw fiducial trigger jumps to another cycle of the common subharmonic signal (see Fig. 8.7). When designing an accelerator, the

Fig. 8.7 Synchronization of the raw fiducial trigger with a common subharmonic signal



RF operating frequency is usually given, but the selection of the laser oscillator frequency, LLRF LO and clock frequencies are flexible. Therefore, we should keep in mind to simplify the relationships between different frequencies so that the common subharmonic frequency can be larger (e.g., several to tens of MHz).

8.3.3 Client Trigger Generation

On each synchronized fiducial trigger, the EVG broadcasts a sequence of events as well as the master timing clock to different subsystems of the accelerator. Typically, the beam acceleration cycle is started by a dedicated “sequence start” event and then driven by various events such as “laser fire”, “RF fire”, “BPM (beam position monitor)”, etc. See Fig. 8.8 as an example.

The EVR of each subsystem is configured to respond to different events. We use the LLRF system as an example. On receiving the “RF fire” event, the EVR generates a trigger synchronized to the recovered master timing clock with a user-configured delay. The trigger is input to the LLRF system (e.g., FPGA), where it is resynchronized to the LLRF clock. The LLRF system may use the resynchronized trigger to start the RF signal demodulation (see Sect. 5.2) or produce an RF pulse. If the LLRF clock phase changes with respect to the master timing clock, the resynchronized RF trigger may shift in time with respect to the trigger produced by the EVR. This situation may create uncertainties in the RF pulse starting time or in the RF phase measurement. We will discuss about this topic in more detail in Sect. 8.5.

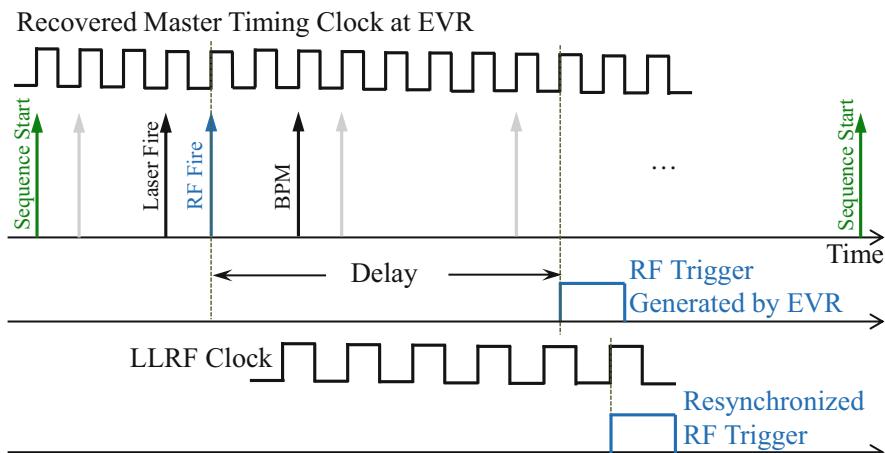


Fig. 8.8 Event sequence and trigger generation at the client devices (e.g., LLRF)

8.4 Synchronization System

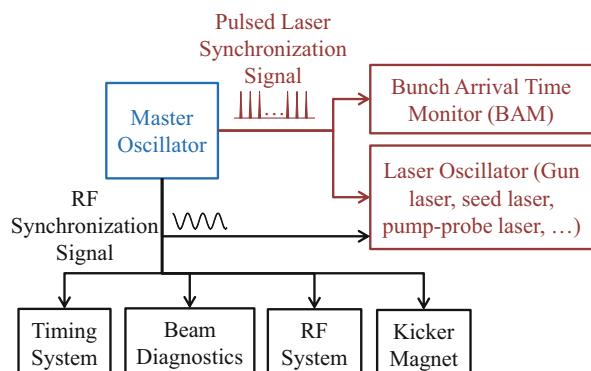
The synchronization system distributes the RF or pulsed laser synchronization signals in Fig. 8.1 to different accelerator subsystems (Byrd et al. 2010; Czuba et al. 2013; Lin et al. 2018a). These subsystems receiving the synchronization signals are called the *clients* of the synchronization system. Figure 8.9 sketches the conceptual architecture of a synchronization system with the clients that it synchronizes. At each client, the synchronization signals are used to lock local RF or laser oscillators, synthesize other frequencies, or directly as frequency and phase references. In a large-scale accelerator, the length of the distribution lines can be up to several kilometers, so the attenuation and phase drifts are the primary concerns. The high-frequency phase noise of the synchronization signals is dominated by the master oscillator. Furthermore, the synchronization system itself should not add much phase noise that worsens the synchronization precision.

8.4.1 Synchronization Signal Distribution

In practice, the synchronization signals are distributed via either coaxial cables or optical fibers.

Coaxial Cables are suitable for transmitting RF signals over a short distance. The distribution lines can be organized in a star topology when the number of clients is small. For a machine with many clients to synchronize, we often realize a reference coaxial distribution line and insert directional couplers at multiple points to pick up the synchronization signals. The coaxial-based design is reliable due to its simple architecture and mature technology. However, the attenuation of a coaxial cable becomes significant for a long distance, which requires amplification of the RF input signal and therefore increases the phase noise. The electrical length of a coaxial cable is sensitive to the environmental temperature and humidity, which may require

Fig. 8.9 Overview of the synchronization system and the synchronized clients



active stabilization for long-distance distribution. Besides, coaxial cables may also suffer from electromagnetic interferences (e.g., RF radiation) due to the limited shielding.

Optical Fibers are used to distribute pulsed laser signals. For long-distance distribution of a CW RF signal, we may also use optical fibers, which require modulating the electrical RF signal on a CW laser. The fiber-based distribution lines are usually organized in a star topology. Compared to the coaxial cables, the optical fibers have much smaller attenuation, so they can deliver synchronization signals to up to many kilometers. With the star topology, we can compensate for the drifts of a fiber with the reflected signal more easily. This method is described in Sect. 8.4.2. For coaxial cables connected with many directional couplers, drift compensation is more difficult due to the reflections from the directional coupler connectors. Fibers are also not sensitive to electromagnetic interferences. The primary disadvantage of the fiber-based design is its high complexity and cost. High complexity implies less reliability compared to the coaxial-based design. In addition, dispersion compensation is needed for the fibers that transmit pulsed laser synchronization signals to maintain ultrashort laser pulses (e.g., hundreds of femtoseconds level).

When designing a synchronization system, we must make tradeoffs between stability, reliability, and cost. For a small-scale synchronization system with less critical performance requirements, a coaxial-based design should be sufficient. Otherwise, a combination of both technologies, as shown in Fig. 8.10, can be used (Lamb et al. 2019). The coaxial distribution lines guarantee that the RF synchronization signals are always delivered to the clients reliably (e.g., to RF stations, BPMs, etc.), which is important for the machine's availability. The drifts of the coaxial distribution lines may be too large, but they can be compensated for with the fiber links. Each fiber link can be stabilized with the method depicted in Fig. 8.13. In the example of Fig. 8.10, we correct the amplifier output phase drift at each local distribution line using stabilized fiber links. Here we assume that the drifts in each local distribution line are acceptable (e.g., by limiting the length of the coaxial

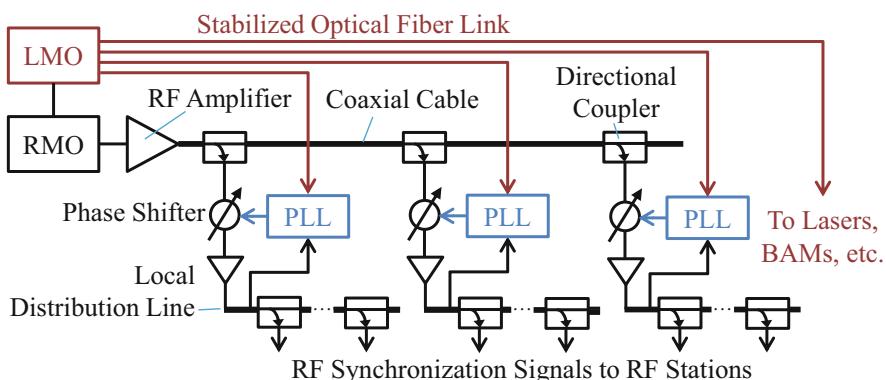


Fig. 8.10 An example with mixed coaxial- and fiber-based distributions

distribution line) or stabilized with other methods like Fig. 8.14. Though Fig. 8.10 is not a complete design of the synchronization system, it depicts the coaxial and fiber mixed design concept. In this example, we only require one stabilized fiber link for each local distribution line, and the cost can be significantly reduced compared to routing to each client a fiber link. For clients requiring stable reference signals, such as the photocathode laser, seeding laser, pump laser, and bunch arrival time monitor (BAM), we can use stabilized fiber links directly.

8.4.2 Phase Drift Mitigation

The phase drifts induced by the coaxial or fiber distribution lines must be mitigated to improve the accelerator's long-term stability. We have learned from Sect. 6.3 that the phase noise in the RF reference signal, derived from the synchronization signal, cannot be compensated for by the LLRF feedback. Therefore, we must control the distribution line drifts within the synchronization system before delivering the synchronization signals to the clients. In this section, we will introduce several passive or active methods to mitigate the phase drifts.

8.4.2.1 Phase-Stable Coaxial Cable

Since the phase drifts are caused mainly by temperature changes, we consider adopting phase-stable coaxial cables with smaller temperature coefficients. The temperature coefficient is specified as a constant in $\text{fs}/\text{m}/\text{K}$, representing the timing drift (in femtoseconds) in a unit length (1 meter) cable when the temperature changes by one Kelvin. For example, at the European XFEL, the coaxial cables for the long-distance distribution (with a smaller attenuation factor) have a temperature coefficient of about 20 to 30 $\text{fs}/\text{m}/\text{K}$ (Czuba and Sikora 2011; Czuba et al. 2013). The coaxial cables used for the local signal distribution (with a larger attenuation factor) are selected with temperature coefficients down to 0 $\text{fs}/\text{m}/\text{K}$ and even negative values to compensate for the long-distance cable drifts. The temperature coefficient of a coaxial cable is typically a nonlinear function with the temperature. We can control its operating temperature to select the lowest temperature sensitivity.

We should also use phase-stable coaxial cables to pick up the RF fields in the cavities or structures for better long-term RF stability. The discussion in Sect. 6.3 indicates that the drifts of the RF field pickup cables cannot be compensated for by the LLRF feedback either.

8.4.2.2 Temperature and Humidity Stabilization

Routing the coaxial or fiber distribution lines in an environment with smaller temperature and humidity fluctuations (e.g., in accelerator tunnel) is helpful to

reduce the phase drifts. Another benefit of routing the RF synchronization signal (with coaxial cables) in the tunnel is the capability of compensating for the cavity field pickup cable drifts (Czuba et al. 2018). See Fig. 8.11. From the directional coupler installed in the main distribution line in the tunnel, a reference cable delivers the RF synchronization signal to the LLRF system as the source of the RF reference signal for the RF station. Note that we distinguish the RF synchronization signal and the RF reference signal because the synchronization system may deliver a different frequency from that required by the RF station. In that case, the LLRF system should implement a frequency synthesizer to derive the RF reference signal with the correct frequency. The reference cable and the cavity field pickup cable should have the same type and length. We route these two cables in the same path so that they experience the same temperature and humidity changes. Therefore, the phase drifts of these two cables will be very close to each other. To simplify the discussion, we assume the phase drifts of these two cables are the same and denote it as $\Delta\varphi_{cable}$. With the noise transfer relationship (6.76), the contribution of the phase drifts of these two cables to the cavity field phase error $\Delta\varphi_{cav}$ can be calculated as

$$\Delta\varphi_{cav} = (\mathbf{T}_{ref} + \mathbf{T}_{mea})\Delta\varphi_{cable} = \mathbf{T}_{drv}\Delta\varphi_{cable}. \quad (8.4)$$

Note that the magnitude of \mathbf{T}_{drv} is small due to the RF feedback, which is depicted in the frequency response of \mathbf{T}_{drv} in Fig. 6.27. Therefore, with the cabling strategy in Fig. 8.11, the phase drifts in the cavity field pickup cable can be reduced according to (8.4). That is, the cavity field phase can lock better to the RF synchronization signal at the directional coupler output.

If we must route the coaxial or fiber distribution lines in an open area with large temperature fluctuations, we may consider stabilizing the cable with temperate-stable cooling water. See Fig. 8.12, which depicts two widely used strategies. At the LCLS in SLAC, the daily temperature change in the klystron gallery can exceed 15 °C, and both methods in Fig. 8.12 are adopted to stabilize the coaxial phase reference lines (Weaver and Hogg 1983; Jobe and Schwarz 1989).

At the synchronization clients, such as LLRF, we can use air-conditioned racks to reduce the local distribution drifts. The air-conditioned racks can stabilize the temperature and even humidity so that the synchronization of the clients (e.g., with frequency synthesizer) is stabilized. Of course, the cost is a concern if many

Fig. 8.11 Coaxial distribution line routing in the tunel and the cabling to compensate for the cavity pickup cable phase drifts

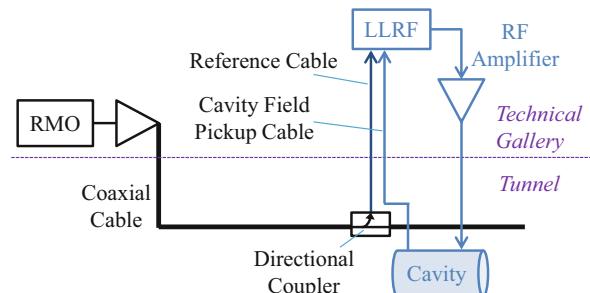


Fig. 8.12 Cable stabilization with temperature-stable cooling water. (a). Routing cables in a surrounding water pipe; (b). Routing cables together with a water hose

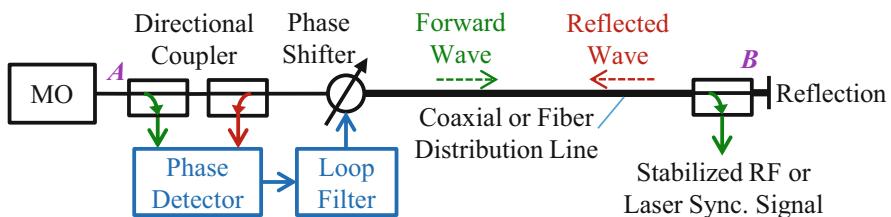
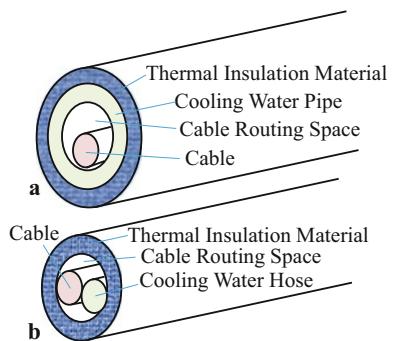


Fig. 8.13 Drift compensation for a star-distributed coaxial cable or optical fiber

such racks are required. Furthermore, if we route signals from the tunnel, as in Fig. 8.11, we can place the air-conditioned rack just above the cabling penetration. This arrangement helps to minimize the cabling distance in open areas.

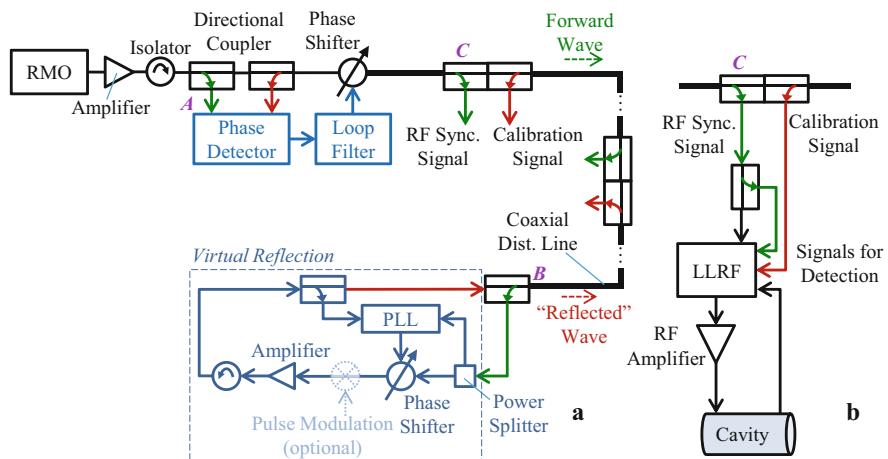
8.4.2.3 Active Drift Compensation

If the passive methods above with stable cables or environments are not adequate, we may measure the phase drifts in the coaxial or fiber distribution lines and correct them actively. Figure 8.13 is a widely used drift compensation strategy based on the interferometric principle (Frisch et al. 2000). This method works well for both coaxial cables and optical fibers if they are distributed in the star topology.

At the end of the distribution line, the forward wave is reflected. The reflected wave is coupled out at the starting point A of the cable and its phase is compared with the MO output. If the phase drift of the distribution line from A to B (see Fig. 8.13) is $\Delta\varphi_{AB}$, the phase detector measures a phase variation of $2\Delta\varphi_{AB}$. With the feedback loop, we can correct this phase drift by adjusting the phase shifter by $-\Delta\varphi_{AB}$. Note that the phase shifter applies to both the forward and reflected waves. For coaxial or optical distribution lines, the components in Fig. 8.13 are implemented differently, which are summarized in Table 8.1 for comparison. It has been demonstrated that the phase drift of an optical fiber of several kilometers can be controlled to be below 10 fs during a period of several days.

Table 8.1 Comparison of stabilized coaxial and fiber distribution lines

	Configuration 1	Configuration 2	Configuration 3
<i>Distributed signal</i>	CW RF sync. Signal	CW RF sync. Signal	Pulsed laser sync. Signal
<i>Distribution line</i>	Coaxial RF cable (e.g., 7/8" Heliax)	Optical fiber (CW laser with its amplitude modulated by the RF sync. signal)	Optical fiber (low dispersion)
<i>Directional coupler</i>	Coaxial RF directional coupler	Optical directional coupler	Optical directional coupler
<i>Phase shifter</i>	Coaxial RF phase shifter (e.g., variable phase shifter, I/Q modulator)	Piezo phase shifter (tuning of the fiber path length)	Piezo phase shifter (tuning of the fiber path length)
<i>Phase detector</i>	RF phase detector (e.g., mixer, I/Q demodulator)	Extract RF signal from the reflected laser with a photodiode and measure it with an RF phase detector	Optical cross correlator
<i>Reflection</i>	RF short circuit	Faraday rotator mirror	Faraday rotator mirror

**Fig. 8.14** Drift compensation for a coaxial distribution line with multiple coupling points

For a coaxial distribution line with multiple coupling points, as in Fig. 8.14a, we may face the following problems when directly applying the drift compensation algorithm above:

- (a) The coaxial distribution line is split into several segments with directional couplers inserted between them. The cable and directional coupler's connectors have slight mismatches. They all contribute to the reflected signal back to the point A, causing drift detection errors for the entire distribution line (from A to B).

- (b) The algorithm of Fig. 8.13 only stabilizes the phase at point B of the distribution line. We expect stable phases at all coupling points (e.g., point C in Fig. 8.14a), but their drifts are not well compensated.

To mitigate the drifts at each coupling point, we can use the strategy in Fig. 8.14. At the end of the coaxial distribution line (point B), we can amplify the forward signal and feed it back into the distribution line as an equivalent “reflected” wave. It is called *virtual reflection*. The amplified backward signal from the end will dominate the reflected signal at point A . It reduces the effects of the connector-induced reflections. Furthermore, we may modulate the backward signal in pulses so that the reflected signal from the end can be easily distinguished from the connector reflections that are CW signals. Note that we implement a phase-locked loop around the amplifier of the virtual reflection unit so that the backward signal at point B is phase locked to the forward signal.

At each coupling point, in addition to the directional coupler picking up the forward RF synchronization signal, we also insert a directional coupler to couple the reflected signal. The reflected signal is used as a calibration signal. To demonstrate the calibration algorithm, we assume that the distribution line (from A to B) has a phase drift of $\Delta\varphi_{AB}$. We also assume that the feedback loop at point A has successfully compensated for this phase drift by adjusting the phase shifter by $-\Delta\varphi_{AB}$. Let us use one of the coupling points (point C) as an example. The phase drift of the cable from A to C is denoted as $\Delta\varphi_{AC}$. Therefore, when the loop at A reaches a steady state, the phase drift in the RF synchronization signal at C can be calculated as

$$\Delta\varphi_{C,\text{sync}} = \Delta\varphi_{AC} - \Delta\varphi_{AB}, \quad (8.5)$$

where the term $-\Delta\varphi_{AB}$ comes from the phase shifter. Since the phase at B is locked to the phase of A (without relative drift), the calibration signal at C experiences a phase drift of

$$\Delta\varphi_{C,\text{calib}} = \Delta\varphi_{BC} = \Delta\varphi_{AB} - \Delta\varphi_{AC}, \quad (8.6)$$

where $\Delta\varphi_{BC}$ is the phase drift between the points B and C . Note that the phase drifts of the RF synchronization signal (forward) and the calibration signal (reflected) at each coupling point are identical in magnitude but opposite in sign. We can measure both signals with RF detectors (see Fig. 8.14b). Since the RF reference signal for the RF station is derived from the RF synchronization signal (forward), the LLRF system cannot measure $\Delta\varphi_{C,\text{sync}}$ directly (it only measures relative phases with respect to the RF reference signal). However, we can estimate $\Delta\varphi_{C,\text{sync}}$ by comparing the phases of the RF synchronization signal and the calibration signal:

$$\begin{aligned} \Delta\tilde{\varphi}_{C,\text{sync}} &= (\Delta\tilde{\varphi}_{C,\text{sync}} - \Delta\tilde{\varphi}_{C,\text{calib}})/2 \\ &= \frac{(\tilde{\varphi}_{C,\text{sync}} - \tilde{\varphi}_{C,\text{calib}}) - (\tilde{\varphi}_{C,\text{sync}0} - \tilde{\varphi}_{C,\text{calib}0})}{2}. \end{aligned} \quad (8.7)$$

Here $\Delta\hat{\varphi}_{C,\text{sync}}$ is the estimate of $\Delta\varphi_{C,\text{sync}}$, $\tilde{\varphi}_{C,\text{sync}0}$ and $\tilde{\varphi}_{C,\text{calib}0}$ are the initial measurements of the phases of the RF synchronization and calibration signals, $\tilde{\varphi}_{C,\text{sync}}$ and $\tilde{\varphi}_{C,\text{calib}}$ are the current measurements, and their drifts are calculated as $\Delta\tilde{\varphi}_{C,\text{sync}} := \tilde{\varphi}_{C,\text{sync}} - \tilde{\varphi}_{C,\text{sync}0}$ and $\Delta\tilde{\varphi}_{C,\text{calib}} := \tilde{\varphi}_{C,\text{calib}} - \tilde{\varphi}_{C,\text{calib}0}$. The initial measurements of the two phases should be recorded when the drift compensation is activated. The result (8.7) can be used to estimate the RF reference phase variation $\Delta\varphi_{\text{ref}}$ (as part of $\Delta\varphi_{\text{ref},\text{add}}$, see Sect. 6.5.2). As discussed in Sect. 6.5.7, the LLRF system only detects and stabilizes the RF field phase relative to the RF reference signal. Therefore, $\Delta\varphi_{\text{ref}}$ cannot be compensated for by the LLRF feedback. However, with the estimate of $\Delta\varphi_{\text{ref}}$, we can add an offset to the phase set point of the feedback to compensate for it, which improves the actual RF field stability experienced by the beam.

If the reflected wave is pulse-modulated, the synchronization and calibration signals at each coupling point can be combined and measured with a single RF detector channel. The two signals can be separated in software. The CW synchronization signal can be detected when the calibration pulse does not exist. During the calibration pulse, the measurement is a vector sum of the two signals. The calibration signal can be derived by subtracting the synchronization signal measurement from the vector sum. In practice, the pulsed reflected wave may cause problems due to the limited directivity of the directional couplers. The typical directivity of a coaxial directional coupler is no better than 40 dB, causing crosstalk from the reflected wave to the forward port into the RF synchronization signal. It can become a problem when we want to control the RF field stability to 10^{-4} level. However, if the RF station operates in the pulsed mode, we can generate the pulsed reflected wave between the RF pulses. Alternatively, we may replace the pulse modulation in Fig. 8.14a with frequency shifting. In this case, the reflected wave with a different frequency can be filtered out from the RF synchronization signal if crosstalk exists. The frequency-shifting scheme has been demonstrated in the work (Lin et al. 2018b).

8.4.2.4 Phase-Averaging Coaxial Line

The phase-averaging coaxial distribution is another widely used strategy to handle the cable drifts (Frisch et al. 2000; Doolittle et al. 2015; Lin et al. 2018b; Sikora et al. 2020). Figure 8.15 depicts a realization of the phase-averaging line with the following differences compared to Fig. 8.14:

- The distribution line is looped back to the RMO side, and the phase at the line endpoint B is locked to the RMO (point A). Here we double the cable length for the same distance of distribution but can lock the point A and B phases without using the reflected signal. Since the reflected signal is not needed for the phase locking, we can place the RF synchronization signal's power amplifier in the loop to have its drifts compensated by the PLL.

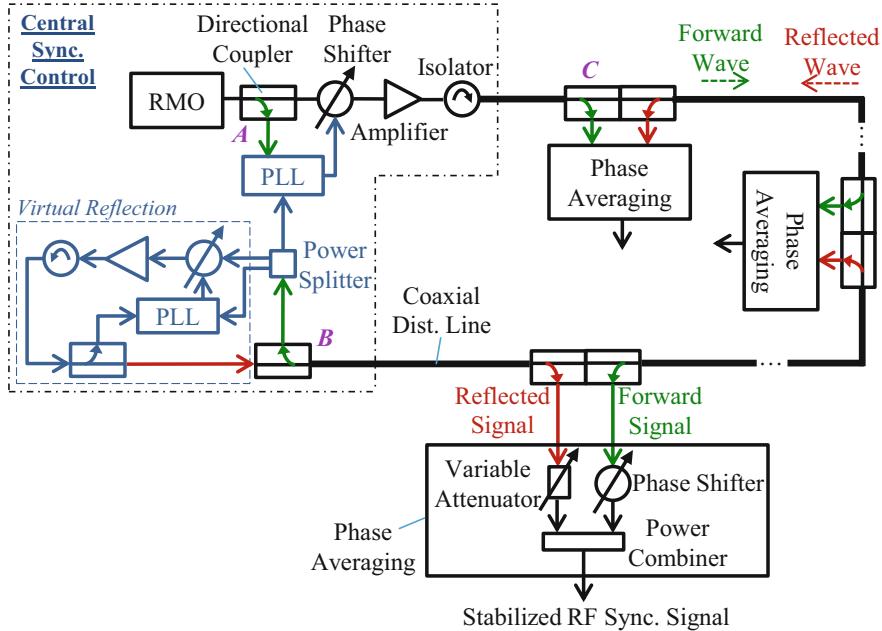


Fig. 8.15 A realization of the phase-averaging coaxial distribution line

- (b) At each coupling point, the forward (RF synchronization) and reflected (calibration) signals are combined with a phase-averaging unit. The output is a phase-stabilized RF synchronization signal.

Let us have a look at the principle of phase averaging. At each coupling point, the forward (x_{sync}) and reflected (x_{calib}) signals can be written as

$$\begin{aligned} x_{sync} &= A \cos(\omega_{RF}t + \varphi_{sync0} + \Delta\varphi_{sync}), \\ x_{calib} &= A \cos(\omega_{RF}t + \varphi_{calib0} + \Delta\varphi_{calib}), \end{aligned} \quad (8.8)$$

where ω_{RF} is the frequency of the RF synchronization signal, φ_{sync0} and φ_{calib0} are the initial phases of the two signals, and $\Delta\varphi_{sync}$ and $\Delta\varphi_{calib}$ are the phase drifts in the forward and reflected signals. The analysis (8.5) and (8.6) imply that

$$\Delta\varphi_{sync} + \Delta\varphi_{calib} = 0. \quad (8.9)$$

We have assumed that the amplitudes of the two signals are identical, which is necessary to cancel the drifts. Combining the two signals in (8.8), we get

$$\begin{aligned}
x_{phavg} &= x_{sync} + x_{calib} \\
&= 2A \cos \left(\frac{\varphi_{sync0} - \varphi_{calib0}}{2} + \frac{\Delta\varphi_{sync} - \Delta\varphi_{calib}}{2} \right). \\
&\quad \cos \left(\omega_{RF}t + \frac{\varphi_{sync0} + \varphi_{calib0}}{2} \right).
\end{aligned} \tag{8.10}$$

Here we have adopted the relationship (8.9). Equation (8.10) implies that the drift terms cancel in the RF output phase of the phase-averaging unit. We often adjust the phases of the forward and reflected signals so that $\varphi_{sync0} = \varphi_{calib0}$. Then, the amplitude of x_{phavg} becomes $2A\cos[(\Delta\varphi_{sync} - \Delta\varphi_{calib})/2]$, which is maximized but affected by the distribution line phase drifts. Therefore, by adjusting the phase shifter and variable attenuator in the phase-averaging unit (see Fig. 8.15), we obtain an RF output with a stable phase but an amplitude fluctuating with the phase drifts. In many applications, small amplitude drifts in the RF synchronization signal are not harmful. This is because most client systems accept RF synchronization signals with input ports operating in saturation.

In practice, obtaining the phase-averaged RF synchronization signal is not easy. Additional precise measurement equipment and sophisticated procedures are required to perform the adjustment (Sikora et al. 2020). It is more practical to use the forward and reflected signals at the coupling points in the same way as Fig. 8.14. That is, we directly use the forward signal as the RF synchronization signal and use the reflected signal to correct its phase drifts in the LLRF software. The phase-locking strategy for points A and B in Fig. 8.15 can be kept as an alternative to the approach in Fig. 8.14.

8.4.3 Client Synchronization

At the synchronization system clients, the RF or laser synchronization signals are used for various purposes, such as

- (a) Synthesizing the RF reference, LO or clock signals for the RF system, beam diagnostics, and fast kickers.
- (b) Synchronizing the laser oscillators for the photocathode RF Gun, seeding lasers, or pump lasers.
- (c) Acting as a timing reference for the bunch arrival time monitors.
- (d) Correcting the drifts of a local distribution line as in Fig. 8.10.

RF synchronization signals distributed by coaxial cables can be directly used to synthesize other frequencies. If the RF synchronization signal is modulated onto a CW laser and delivered via optical fibers, we must first extract the electrical RF signal from the modulated optical signal.

There are more choices in using a pulsed laser synchronization signal. We may generate RF signals from the laser pulse train by selecting a harmonic of the laser pulse repetition rate. The pulsed laser signal may also be directly used to synchronize a laser oscillator or as a reference for the BAMs.

8.4.3.1 RF Signal Extraction

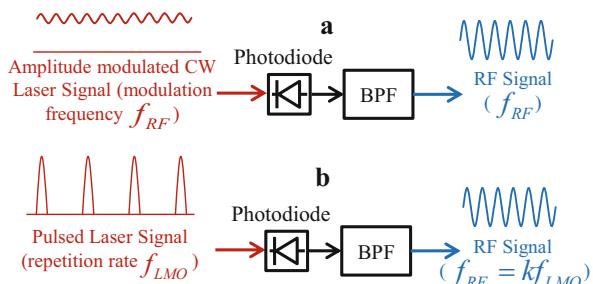
Fast photodiodes are the simplest device to extract electrical signals from optical signals (Lorbeer et al. 2007). Typical fast photodiodes can extract RF signals with frequencies up to tens of GHz, which are adequate for receiving typical synchronization signals in particle accelerators. This approach works for both the amplitude-modulated CW laser signals and the pulsed laser signals. See Fig. 8.16.

In Fig. 8.16a, the electrical output of the photodiode consists of a DC offset and a sinusoidal signal at the RF frequency, f_{RF} . A band-pass filter removes the DC offset. In Fig. 8.16b, the photodiode output is an electrical pulse train containing harmonics of the laser repetition rate, f_{LMO} . A narrow-band BPF is required to select the correct harmonic at the desired RF frequency. For example, if the laser pulse repetition rate is $f_{LMO} = 142.8$ MHz, the harmonic number of $f_{RF} = 2998.8$ MHz is $k = 21$. The BPF in Fig. 8.16b should be able to separate adjacent harmonics.

The photodiode is simple and effective for demodulating laser signals and extracting electrical RF signals from them. The high-frequency phase noise introduced by the photodiode is typically very low (Vargas et al. 2019). However, the phase of the output RF signal is sensitive to the temperature and laser power fluctuations. In practice, we must select photodiodes with small temperature coefficients (in $\text{ps}/^\circ\text{C}$) and operate them at optical power levels where the power sensitivity (in ps/mW) is minimized. However, if we can directly compare the phases of the optical input and the electrical output, a PLL can be built to stabilize the phase of the electrical RF output.

A device called balanced optical to microwave phase detector (BOM-PD) is available to detect the relative phase between a laser pulse train and an RF signal that is a harmonic of the laser repetition rate (Jung and Kim 2012).

Fig. 8.16 Extract electrical RF signals from optical signals via photodiodes



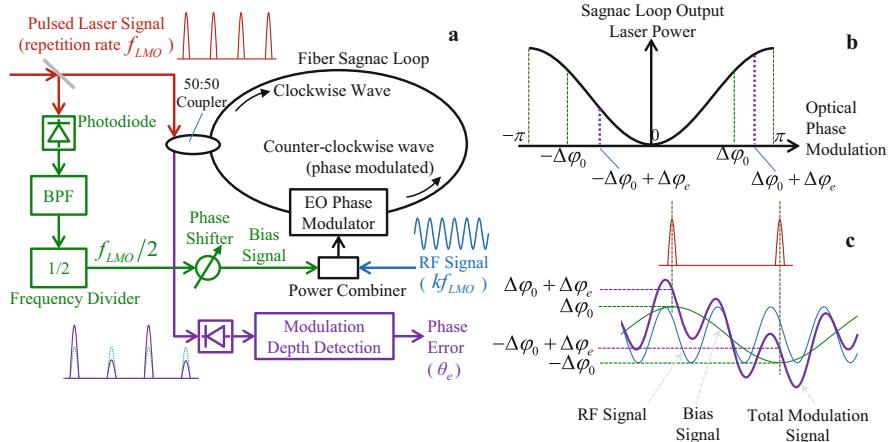


Fig. 8.17 Schematic of a BOM-PD and its principle of phase detection

The schematic of a BOM-PD is depicted in Fig. 8.17a, which is constructed based on a fiber Sagnac-loop interferometer. A directional electro-optical (EO) phase modulator modulates the optical carrier phase of the counterclockwise propagating wave. The output laser power of the interferometer is a function of the optical phase modulation, which produces phase imbalances between the two counter-propagating waves in the Sagnac loop. The relation between the optical phase modulation and the output laser power is shown in Fig. 8.17b. When the optical phase modulation is zero, i.e., no voltage applied to the EO phase modulator, the two counter-propagating waves interfere destructively at the output combiner, resulting in a minimum output laser power. When the optical phase modulation approaches $\pm\pi$, the output laser power increases to the maximum value.

As illustrated in Fig. 8.17a, we derive a sinusoidal signal of frequency $f_{LMO}/2$ from the pulsed laser signal via a photodiode and a frequency divider. Here the band-pass filter selects the fundamental frequency (f_{LMO}) in the photodiode output. The EO phase modulator of the Sagnac loop is biased by the $f_{LMO}/2$ sine wave. The sine wave's phase is adjusted via the phase shifter so that the laser pulses passing through the EO phase modulator experience the positive or negative peak voltages of the sine wave. Under this condition, the optical phase modulation for the laser pulses switches between $\pm\Delta\varphi_0$. See Fig. 8.17c. In this case, the output laser power is non-zero but without amplitude modulation since all laser pulses are equally attenuated.

We apply an RF signal of frequency $f_{RF} = kf_{LMO}$ to the modulation input of the EO phase modulator. The RF signal is power combined with the bias sine signal. Our goal is to detect the relative phase, θ_e , between the input pulsed laser signal and the input RF signal. We define $\theta_e = 0$ when the RF signal crosses zero while the laser pulses pass the EO phase modulator, causing no changes in the output laser power. In contrast, if $\theta_e \neq 0$, the input RF signal will generate an offset to the total

modulation voltage at the EO phase modulator's input, resulting in asymmetric optical phase modulations of $\pm\Delta\varphi_0 + \Delta\varphi_e$. See Fig. 8.17c and b. Therefore, a non-zero θ_e will cause an amplitude modulation on the laser pulses at the Sagnac loop output. The modulation depth can be detected as a measure of θ_e .

The BOM-PD is not sensitive to temperature fluctuations. The photodiodes in Fig. 8.17a extract the fundamental frequency (f_{LMO}) of the laser pulse train. Since f_{LMO} (tens of MHz) is much lower than the RF frequency (several GHz), the photodiode timing drifts will cause much smaller phase errors than the cases in Fig. 8.16, where the RF frequency is extracted directly from the laser signals. The phase error θ_e is represented by the amplitude modulation depth, which is a relative value and is not sensitive to the input laser power fluctuations. These features make the BOM-PD an excellent low-noise and low-drift phase detector.

A possible usage of the BOM-PD is driving a PLL, which locks a VCO of frequency kf_{LMO} ($k = 1, 2, \dots$) to the pulsed laser synchronization signal (Peng et al. 2014; Zhang et al. 2018). The VCO output is connected to the RF signal input port of the BOM-PD, and the phase error output is filtered to tune the VCO. This is another approach to extract RF signals from a pulsed laser synchronization signal besides the simple photodiode-based method in Fig. 8.16. We may also improve the setup of Fig. 8.16b to Fig. 8.18, which further stabilizes the RF signal extracted directly from a pulsed laser signal using a BOM-PD-based PLL. The BOM-PD can also be used in the PLL of Fig. 8.10 if the LMO distributes pulsed laser synchronization signals.

8.4.3.2 Frequency Synthesis

The RF synchronization signal (or the RF signal extracted from a pulsed laser synchronization signal) is used to synthesize other frequencies. The typical frequency synthesis methods are summarized in Table 8.2 and Fig. 8.19. These methods are combined for frequency conversion and generation, such as in the LO and clock generator, frequency down-converter and up-converter (see Chap. 5).

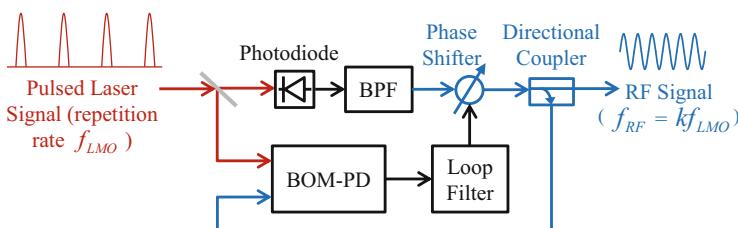
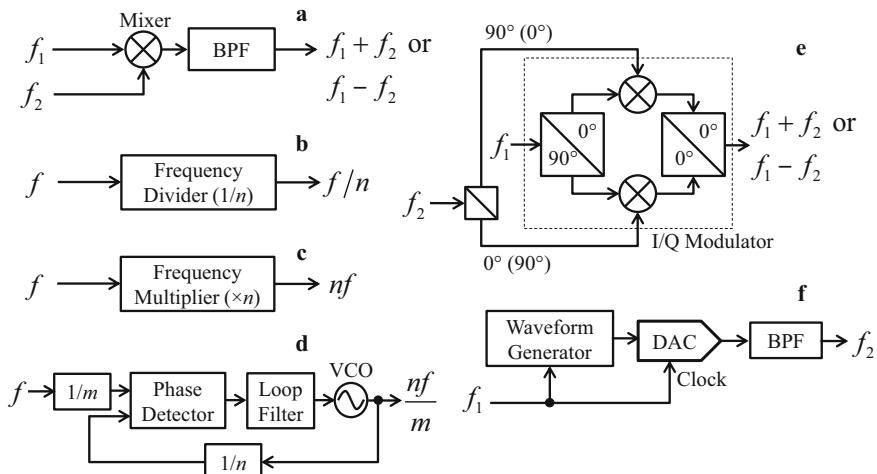


Fig. 8.18 Stabilization of the RF signal extracted by a photodiode

Table 8.2 Frequency synthesis methods

Methods	Comments
Mixer	A mixer (see Fig. 8.19a) produces the sum and difference of two input frequencies. The upper or lower sideband in the mixer output can be selected with a band-pass filter.
Frequency divider	A frequency divider (see Fig. 8.19b) produces an output frequency of a fraction ($1/n$) of the input frequency.
Frequency multiplier	A frequency multiplier (see Fig. 8.19c) generates an output frequency of integer times the input frequency. We only use it when multiplying the input frequency by a small factor (e.g., $\times n$ with $n < 10$). From Sect. 6.2.5, the wideband phase noise of the input signal will be increased by n times at the output. Therefore, if we need to multiply the input frequency by a large factor, a PLL is preferred.
Phase-locked loop	As discussed in Sect. 8.2.2, a PLL can generate a frequency determined by the frequency dividers in the reference input and the VCO output (see Fig. 8.19d). If we select $m = 1$, the PLL is equivalent to a frequency multiplier. With a properly selected VCO, the phase noise of the PLL can be minimized.
Single-sideband (SSB) modulation	SSB modulation uses an I/Q modulator (see Fig. 8.19e). By applying a 90° phase shift in the different branches of the power splitter for f_2 , we can generate either $f_1 + f_2$ or $f_1 - f_2$. Compared to a single mixer, the SSB modulation does not need the BPF.
Direct digital synthesis (DDS)	The DDS (see Fig. 8.19f) employs a fast DAC to produce sine waves. The DAC output frequency is configurable because the waveform generator is typically implemented in a digital processor like FPGA.

**Fig. 8.19** Schematics of typical frequency synthesis methods. (a). mixer; (b). frequency divider; (c). frequency multiplier; (d). PLL; (e). SSB modulation; (f). DDS

8.4.3.3 Synchronization of Laser Oscillator

In particle accelerators, we often need to synchronize a mode-locked laser oscillator to an external reference (Yang et al. 2015; Yang et al. 2017; Titberidze 2017). The LMO in Fig. 8.1 is a laser oscillator phase-locked to the RMO. In FEL machines, typical laser oscillators include the photocathode Gun lasers, FEL seeding lasers, and pump lasers for the pump-probe experiments.

A mode-locked laser oscillator is equipped with piezo and motor tuners. The motor tuner is used for a coarse tuning of the laser repetition rate, whereas the piezo tuner is used for fine tunings. There are two typical methods to synchronize a laser oscillator, see Fig. 8.20.

In the first setup, as Fig. 8.20a, the reference input is a CW RF signal. We extract an RF signal of the same frequency from the laser oscillator output. A PLL locks the extracted RF signal to the reference by acting on the piezo tuner. The RF reference frequency is usually a higher harmonic of the laser repetition rate to improve the locking precision. For example, an LMO producing a laser pulse train of 142.8 MHz can be phase-locked to an RF reference of 2998.8 MHz (21st harmonic of 142.8 MHz). Note that this scheme only performs locking over a small frequency range (fine-tuning), and the large-range (coarse) tuning of the laser repetition rate should be done by other means. The coarse tuning is required for a correct central frequency of the laser repetition rate. The fine-tuning PLL may lock to a different undesired subharmonic. In the above example, an LMO with a different repetition rate (e.g., 149.94 MHz, 1/20 of 2998.8 MHz) can also be locked to the RF reference via this fine-tuning PLL.

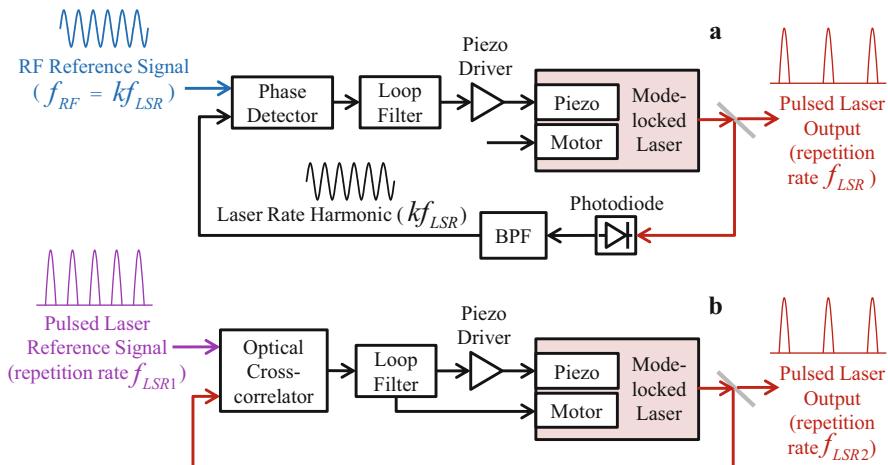


Fig. 8.20 Synchronization of a mode-locked laser oscillator. (a). locking to an RF reference; (b). locking to a pulsed laser reference

If the reference input is a pulsed laser signal, such as the pulsed laser synchronization signal, the scheme in Fig. 8.20b can be adopted. In this case, an optical cross-correlator is used to compare the time misalignment between the two laser pulse trains (Schulz et al. 2008). The repetition rates of the two laser pulse trains, f_{LSR1} and f_{LSR2} , should be a harmonic or subharmonic of either other. That is, they should stratify $f_{LSR1} = nf_{LSR2}$ or $f_{LSR1} = f_{LSR2}/n$, where n is a non-zero positive integer. Note that this method can perform both coarse- and fine-tunings of the f_{LSR2} laser oscillator. That is, the PLL can guarantee the correctness of the f_{LSR2} central frequency as well as the precise phase locking to the reference.

The locking bandwidth of a laser oscillator PLL is limited by the piezo bandwidth, which is typically up to several kilohertz.

8.5 Robust Timing Relations

One of the primary tasks operating a particle accelerator is establishing correct timing relations among different subsystems. The trigger delays must be adjusted to start different subsystems in correct sequences. The RF field phase must be aligned with the particle arrival time, which is called *phasing*. Furthermore, we expect that the timing relations can be kept or quickly restored after a reset or power cycle of any subsystems. It is important for improving the particle beam's availability.

This section focuses on the robustness of timing and synchronization systems (Geng 2020), such as the RF pulse timing uncertainty relative to the trigger, the phase uncertainty of frequency dividers after power cycles, and the race conditions between the trigger and clock. The possible solutions for robust timing and synchronization system design are also discussed. The discussion is based on an FEL machine. However, the concepts and methodologies can be applied to other accelerators like storage rings.

8.5.1 Timing Relation Highlight

Figure 8.22 highlights the timing relations between the master oscillator, timing, synchronization, photocathode Gun laser, and LLRF systems of an FEL machine. We assume that it operates in the pulsed mode with a repetition rate of 100 Hz. The displayed frequencies of different subsystems refer to the SwissFEL. These frequencies are typical for FELs based on S-band normal-conducting Linacs.

The frequencies and timing relations are summarized below as a basis for the robustness analysis later:

(a) *Synchronization frequency generation*

The master oscillator (MO) frequency is 142.8 MHz. It is directly used as the master timing clock for the timing system. The MO frequency is multiplied by 21 to obtain the RF reference frequency (2998.8 MHz) for the photocathode Gun laser and RF systems.

(b) *Master timing clock distribution*

The master timing clock is distributed via the EVG-EVR links and is always phase-locked to the MO. Therefore, it can be used as a global reference at different locations of the accelerator for diagnosing the phase errors of other signals (e.g., LLRF LO and clock).

(c) *Fiducial trigger synchronization*

The raw fiducial trigger derived from the AC mains is synchronized to the common subharmonic signal (1.9833 MHz). See Fig. 8.7. It results in a synchronized fiducial trigger with its rising edge aligned with the rising edge of the common subharmonic signal. Therefore, the time interval between two trigger pulses is an integer multiple of the common subharmonic period.

(d) *Photocathode laser synchronization*

The photocathode laser is phase-locked to the RF reference frequency (2998.8 MHz) that is the 42^{ed} harmonic of the laser repetition rate (71.4 MHz). The laser oscillator output then represents a frequency divider by 42. The divider input frequency is the RF reference frequency, and the output is the laser fundamental frequency (i.e., laser pulse repetition rate). A pulse picker is triggered to select 100 laser pulses per second to generate electron bunches in the photocathode RF Gun. Figure 8.21 illustrates the laser picker gate following a trigger pulse, selecting a laser pulse from the pulse train. The actual bunch generation time is determined by both the trigger time and the phase of the laser fundamental frequency signal.

(e) *LO and clock generation*

The LO and clock frequencies are derived from the RF reference frequency using the synthesis methods described in Sect. 8.4.3.2. Figure 8.22 shows a widely used design. Frequency dividers are used to derive the clock (249.9 MHz, 1/12 of the RF reference frequency) and the IF (41.65 MHz, 1/6 of the LLRF clock frequency) signals. The LO frequency (3040.45 MHz) is generated by mixing the reference and IF frequencies. Here a band-pass filter is required to select the correct sideband from the mixer output.

(f) *RF phase measurement*

We use digital RF detectors, as Fig. 5.3a, to measure the RF signal phase. The RF signal is down converted by the LO, and the resulting IF signal is sampled by an ADC at the LLRF clock frequency. In the LLRF controller, the non-I/Q demodulation algorithm ($n = 6, m = 1$) is used to calculate the I and Q components of the IF

signal. Here n and m are non-I/Q sampling parameters defined in Sect. 5.2.2. The non-I/Q demodulation process is (re)started by each trigger pulse. Therefore, the first ADC sample after the trigger time is defined as the phase-detection reference. Since the time intervals between trigger pulses are full periods of the common subharmonic frequency, the RF, LO, IF signals, and the LLRF clock will repeat their phases after each trigger time. Under this condition, we obtain an RF phase measurement that is repetitive for each RF pulse. The measured RF phase is determined by the relative timing between the RF, LO, clock signals, and the trigger. The timing relation between the LLRF trigger, clock, and the IF signal is depicted in Fig. 8.21.

(g) RF pulse generation

For RF stations operating in the pulsed mode, the RF pulses are started by the synchronized RF trigger, as depicted in Fig. 8.21. The LLRF trigger in the plot is generated by the LLRF EVR and is synchronized to the master timing clock. The trigger is resynchronized by the LLRF clock and defines the RF pulse starting time. At the first rising (or falling) edge of the LLRF clock after the trigger time, the macro

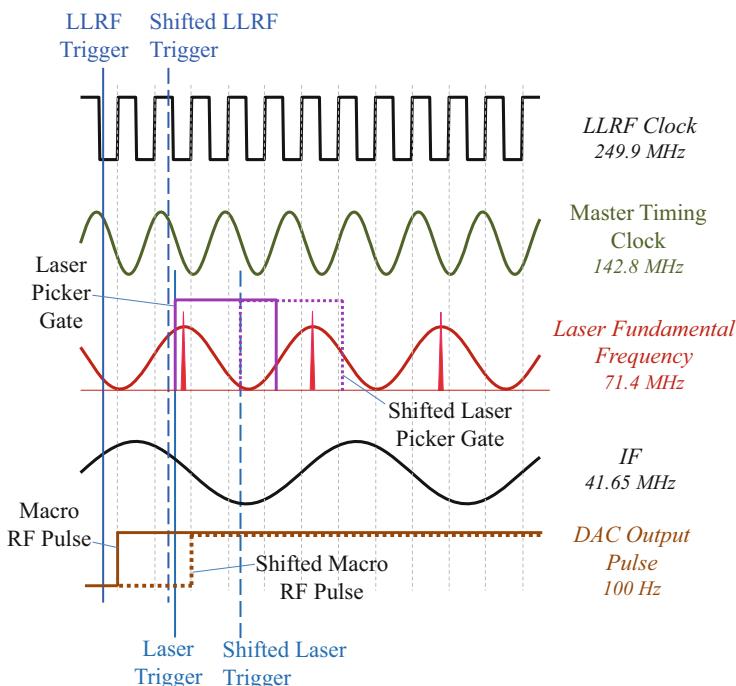


Fig. 8.21 Waveforms illustrating the timing relations between the trigger, LLRF clock, master timing clock, laser fundamental frequency, reference IF signal and the macro RF pulse produced by the DACs. The triggers for LLRF and laser are synchronous but with different delays to guarantee that the RF fields are established before the electron bunch arrives at the RF cavity or structure

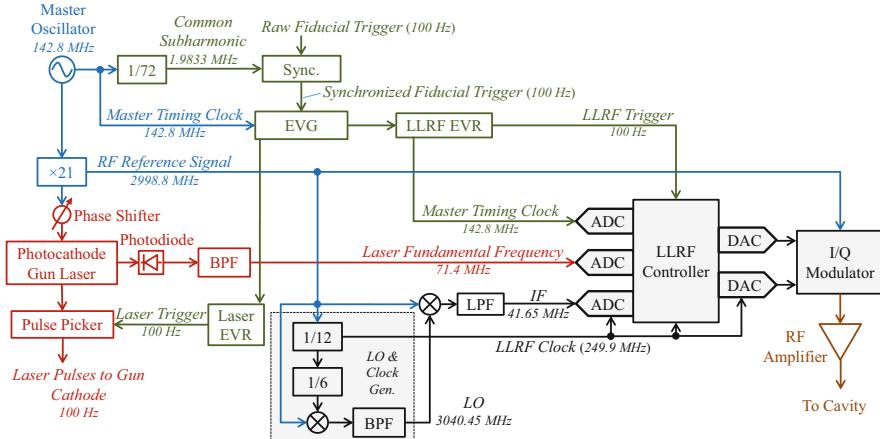


Fig. 8.22 Timing relations between the master oscillator, timing, synchronization, photocathode laser and LLRF systems

RF pulse is produced by the RF actuator DACs. This describes the relationship between the RF pulse starting time and the trigger time and LLRF clock phase.

8.5.2 Timing Relation Uncertainty

After rebooting or power cycling the components in Fig. 8.22, the timing relations discussed in the previous section may change, causing problems for beam acceleration.

The timing relation uncertainties come mainly from two sources:

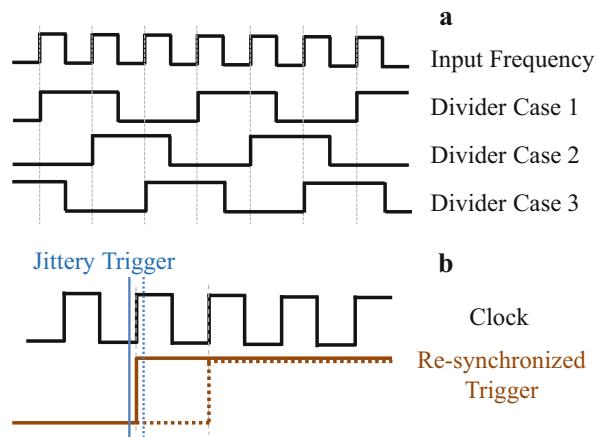
(a) Frequency divider reinitialization

A $1/n$ frequency divider is a counter that counts the number of input signal cycles and produces an output cycle every n counts. Reinitialization means either power cycling the frequency divider or interrupting its input for an arbitrary period. Due to the uncertain counting starting time after reinitialization, the frequency divider output may have n possible phases compared to the input frequency. This situation is depicted in Fig. 8.23a with a $1/3$ frequency divider as an example.

(b) Trigger-clock race condition

When the trigger and clock edges are close to each other, their relative jitter may cause the re-synchronized trigger jumping by one clock cycle randomly. See Fig. 8.23b, where the trigger is synchronized by the first rising edge of the clock after the input trigger time. The race condition happens when the input trigger time and the clock's rising edge coincide in the presence of a timing jitter.

Fig. 8.23 Sources of timing relation uncertainties. (a). reinitialization of a frequency divider (1/3 as example); (b). trigger-clock race condition (assuming a stable clock but jittery trigger)



The trigger-clock race condition affects the functions started by the trigger and synchronized by the clock, such as RF detection and macro RF pulse generation. If the trigger starts the non-I/Q demodulation, the race-condition will select a different sample as the demodulation starting point, causing a phase jump of $\pm\Delta\varphi$ (the IF phase covered by a clock cycle). When generating an RF pulse with the trigger synchronized by the LLRF clock, the race condition will cause a random jump in the RF pulse timing compared to the beam arrival time. If the RF pulse is not flat in amplitude or phase, the accelerating voltage and beam phase will be different when the beam interacts with different parts of the RF pulse.

The timing uncertainties caused by the frequency dividers in Fig. 8.22 are summarized as follows:

(a) *Reinitialization of the common subharmonic frequency divider*

After reinitialization, the timing of the common subharmonic's rising edge will change by multiple periods of the 142.8 MHz. This will cause the same timing changes in all triggers. The consequences of the trigger timing changes are explained as follows according to Fig. 8.21, where we assume that the triggers are shifted by one period (for example) of the 142.8 MHz:

1. The trigger-clock race condition, which is assumed not a problem before the trigger timing change, may happen because the trigger timing change is not full periods of the LLRF clock.
2. The laser pulse picker will select a different laser pulse with a different delay relative to the laser trigger. This is because the trigger timing change is half the laser pulse spacing. Since the LLRF and laser triggers have a constant relative delay, the electron bunch produced by the laser pulse will experience a different part of the RF pulse.
3. The measured RF phase changes by integer times of 60° . This is because the shifted LLRF trigger selects a different ADC sample of the IF signal as the non-I/Q demodulation starting point.

4. The macro RF pulse timing relative to the LLRF trigger changes up to one LLRF clock period. This is also because the trigger timing change is not full periods of the LLRF clock. It also causes the beam to interact with different parts of the RF pulse.

(b) *Reinitialization of the photocathode Gun laser*

As mentioned before, the Gun laser is equivalent to a frequency divider. If the laser oscillator is reinitialized (e.g., power cycled or re-locked), the laser timing will change by multiple periods of the reference frequency (2998.8 MHz). The consequence is that the laser-produced electron bunch will experience different parts of the RF pulse. The beam phase in RF cavities or structures will not change because the bunch arrival time changes by full RF (2998.8 MHz) periods. However, if the RF frequency is not a harmonic of the laser reference frequency, like the C-band RF stations operating at 5712 MHz (not shown in Fig. 8.22), the laser timing change will cause a beam phase error of multiple of 34.3° .

(c) *Reinitialization of the LO and clock generator*

After reinitialization, the LLRF clock timing shifts by multiple periods of the RF reference frequency (2998.8 MHz). With the LO generation scheme in Fig. 8.22, the measured phase of the IF signal (down-converted from the RF reference signal) will change by integer times of 60° . The reason is that the IF frequency is derived from the LLRF clock via a 1/6 frequency divider. According to the waveforms in Fig. 8.21, due to the LLRF clock phase changes, the trigger-clock race condition may happen, and the macro RF pulse timing relative to the LLRF trigger may change.

In summary, if the frequency dividers in Fig. 8.22 are reinitialized, then one or more of the following consequences may happen. The RF phase measurement may change by multiple of $\Delta\varphi$, i.e., the IF phase covered by an ADC clock period, which is 60° for the example of Fig. 8.21 and 8.22. The timing of the macro RF pulse may change by up to one DAC clock period. The electron bunch may arrive at a different time in the macro RF pulse and experience different accelerating voltage and phase if the RF pulse is not flat. The bunch may also experience a wrong accelerating phase if the RF frequency is not a harmonic of the Gun laser reference frequency.

These timing relation uncertainties create difficulties in operating the accelerator after a shutdown or maintenance. In the next section, we will discuss the possible methods to mitigate these uncertainties.

8.5.3 Strategies for Robust Timing Relations

We expect that the beam operation can be quickly restored for a particle accelerator after maintenance or troubleshooting, for which we may have to power cycle or reboot the devices of different subsystems, such as the timing, synchronization, Gun laser, and LLRF systems. The timing relations should be either unchanged (e.g., guaranteed by the hardware design) or be restored by automation procedures to

minimize the machine recovery time. In this section, a systematic strategy will be introduced to achieve robust timing relations.

8.5.3.1 Frequency Selection

In general, simple frequency relations between different subsystems are beneficial for robust timing relations. Let us still use Fig. 8.22 as an example. If we simplify the frequency relations by changing both the laser repetition rate and the LLRF clock frequency to 142.8 MHz, some issues described in Sect. 8.5.2 can be avoided. For example, when the triggers shift by one period of 142.8 MHz (see Fig. 8.21), the consequences (1), (2) and (4) caused by the common subharmonic frequency divider reinitialization can be avoided.

At the design stage of an accelerator, teams working on different subsystems should systematically coordinate the frequency selection. As a first principle, we should minimize the number of frequencies used in the machine.

Practically, the frequency selection is a compromise between the robust timing relations and other factors in different subsystems. For LLRF, the IF frequency should be higher than the environmental noise frequencies to be able to filter the noise. To detect fast RF signal envelope changes, we need a wideband RF detector, and the IF frequency should be higher than the desired RF detection bandwidth. In contrast, the RF detection noise is more sensitive to the ADC clock jitter at a higher IF frequency. When selecting the IF frequency, we must make a trade-off between the RF detection bandwidth, environmental noise separation, and clock jitter sensitivity. The LLRF clock frequency is mainly limited by the maximum speed of the ADCs, DACs, and FPGAs. High-speed ADCs, DACs, and FPGAs are often costly and more challenging to implement in hardware and firmware.

8.5.3.2 Timing Relation Diagnostics

If we measure the relative phases of the signals from various subsystems, we can gain insights into the timing relation changes. It is a basis for recovering the timing relations after a shutdown or maintenance of the accelerator subsystems.

Figure 8.22 illustrates the possible diagnostics of the concerned timing relations. We use the LLRF ADCs to measure the phases of the master timing clock and the laser fundamental frequency signal. The master timing clock recovered by the EVR is phase-locked to the master oscillator. The laser fundamental frequency signal is derived by a photodiode and a band-pass filter from the laser oscillator output. The ADC samples are demodulated using the non-I/Q demodulation algorithm. Considering the frequencies in the example of Fig. 8.22, the non-I/Q sampling parameters for the master timing clock are $n = 7$ and $m = 4$, and for the laser fundamental frequency are $n = 7$ and $m = 2$. They mean that n ADC samples cover m full periods of the input signal. The measured phases, including the IF signal phase, reflect the

relative timing among the trigger, LLRF clock, LO (i.e., IF), master timing clock (i.e., master oscillator) and Gun laser.

Since the master timing clock is always phase-locked to the master oscillator, it can be used as a reference to detect the timing changes in other signals. When the machine has a proper timing relation, we remember the phases of the signals measured by the three ADCs in Fig. 8.22 as references. We use φ_{MTC0} , φ_{LSR0} and φ_{IFO} to denote the reference phases of the master timing clock, laser fundamental frequency, and the IF signal. Afterward, we measure their phases iteratively and calculate the phase variations $\Delta\varphi_{MTC} = \varphi_{MTC} - \varphi_{MTC0}$, $\Delta\varphi_{LSR} = \varphi_{LSR} - \varphi_{LSR0}$ and $\Delta\varphi_{IF} = \varphi_{IF} - \varphi_{IFO}$. Note that the phases mentioned here are defined at the frequencies of the corresponding signals.

We use these results to diagnose the following timing relation changes:

(a) *Laser timing change relative to the master oscillator*

The waveforms in Fig. 8.21 imply that when the laser timing relative to the master oscillator (represented by the master timing clock) keeps unchanged, the terms $\Delta\varphi_{MTC}$ and $\Delta\varphi_{LSR}$ must satisfy

$$\frac{\Delta\varphi_{MTC}}{f_{MTC}} = \frac{\Delta\varphi_{LSR}}{f_{LSR}}, \quad (8.11)$$

where f_{MTC} is the master timing clock frequency and f_{LSR} is the laser fundamental frequency. Equation (8.11) means that $\Delta\varphi_{MTC}$ and $\Delta\varphi_{LSR}$ are only caused by the timing changes of the trigger or LLRF clock. Therefore, after the laser oscillator is reinitialized, the laser phase error (defined at the laser fundamental frequency) relative to the master timing clock (i.e., master oscillator) can be calculated as

$$\delta\varphi_{LSR} := \Delta\varphi_{LSR} - \frac{f_{LSR}\Delta\varphi_{MTC}}{f_{MTC}}. \quad (8.12)$$

Then the laser timing change in terms of the number of buckets (periods) of the laser reference frequency can be calculated as

$$k_{LSR,bucket} := \delta\varphi_{LSR} / \Delta\varphi_{LSR,bucket}. \quad (8.13)$$

Here $\Delta\varphi_{LSR,bucket}$ is the phase shift of the laser fundamental frequency signal when its timing changes by one bucket of the laser reference frequency. For our example in Fig. 8.22, $\Delta\varphi_{LSR,bucket} = 71.4/2998.8 \times 360^\circ = 8.57^\circ$. With the knowledge of the laser bucket changes, the laser timing can be restored following the approach introduced in Sect. 8.5.3.4.

(b) *LO (IF) timing change relative to the master oscillator*

In Fig. 8.22, the LO frequency is produced by mixing the IF and RF reference frequencies. We notice that the RF reference frequency has a synchronized and stable timing relation with the master oscillator due to the simple frequency

multiplication. Therefore, the LO phase error is determined by the IF phase change caused by the reinitialization of the frequency dividers in the LO and clock generator. We notice that the IF signal in Fig. 8.22 is similar to the laser fundamental frequency signal. That is, they are all equivalent to the outputs of frequency dividers using the RF reference frequency (2998.8 MHz) as input. Therefore, we can use the same algorithm (8.12) to calculate the IF phase error (defined at the IF frequency) relative to the master oscillator. We need to replace f_{LSR} with f_{IF} (the IF frequency) and replace $\Delta\varphi_{LSR}$ with $\Delta\varphi_{IF}$:

$$\delta\varphi_{IF} := \Delta\varphi_{IF} - \frac{f_{IF}\Delta\varphi_{MTC}}{f_{MTC}}. \quad (8.14)$$

Note that the LO phase error satisfies $\delta\varphi_{LO} = \delta\varphi_{IF}$, where $\delta\varphi_{LO}$ is defined at the LO frequency. Of course, we may also calculate the IF timing change in terms of the number of buckets (periods) of the RF reference frequency as

$$k_{IF,bucket} := \delta\varphi_{IF} / \Delta\varphi_{IF,bucket}. \quad (8.15)$$

Here $\Delta\varphi_{IF,bucket}$ is the phase shift of the IF signal when its timing changes by one bucket of the RF reference frequency. For our example in Fig. 8.22, $\Delta\varphi_{IF,bucket} = 41.65/2998.8 \times 360^\circ = 5^\circ$.

(c) LLRF clock timing change relative to the master oscillator

The measured phase variation $\Delta\varphi_{MTC}$ reflects the relative timing among the trigger, LLRF clock and the master timing clock (i.e., master oscillator). Therefore, if $\Delta\varphi_{MTC} \neq 0$, we can conclude that either the trigger timing or the LLRF clock timing changes relative to the master oscillator, but we cannot identify the exact sources. In this case, we need to judge the trigger timing first, following the method described in the next point.

(d) Trigger timing change relative to the master oscillator

According to Fig. 8.21, the trigger timing shift causes the non-I/Q demodulation starting points of all signals to change by the same number of samples. Therefore, we can calculate for each signal the ratio between its phase change and the phase shift in an ADC clock period. For the frequencies in Fig. 8.21, the time interval between subsequent samples corresponds to 205.71° , 102.86° and 60° for the master timing clock, laser fundamental frequency, and IF signal, respectively. When the three ratios are identical, we can conclude that the trigger timing has changed compared to the master oscillator.

The timing relation diagnostics above are for a single RF station. If the accelerator is equipped with multiple RF stations, the diagnostic results can be combined to gain further information. For example, if all RF stations report the same trigger timing change relative to the master oscillator, we can conclude that the common subharmonic frequency divider may have been reinitialized. Furthermore, if the

timing relations of all RF stations change randomly, a reinitialization of the master oscillator may have happened.

8.5.3.3 Reference Tracking

One of the primary tasks of a LLRF system is to maintain the RF field phase with respect to the RF reference phase, which is achieved by a phase feedback loop. As mentioned before, the timing relation irregularities between the trigger, LO, and clock will generate phase-detection errors that will be transferred to the RF field by the feedback. Fortunately, such phase measurement uncertainties can be mitigated by the reference tracking discussed in Sect. 5.2.5. Typically, the reference tracking based on PLL (Fig. 5.14) and Hilbert transform (Fig. 5.17) can only correct static phase-detection errors, whereas the direct phase tracking (Fig. 5.20) can also detect and compensate for the random phase jumps caused by the trigger-clock race conditions.

A valid reference tracking requires that both the RF reference signal and the RF signal to be measured are detected using the same trigger, LO, and clock. Therefore, the timing relation changes will cause the same phase errors in both signals. We subtract the reference signal phase from that of the RF signal to be measured to remove such common-mode phase errors. The subtraction represents the RF signal phase relative to the RF reference. If the RF signal to be measured is a pickup of the cavity field, the reference tracking guarantees that we measure the relative phase between the cavity field and the RF reference.

8.5.3.4 Frequency Divider Resynchronization

Resettable frequency dividers help to correct the phase uncertainties after reinitialization (Urbanski et al. 2018). The reset signal restarts the counter from zero and reinitializes the output phase. Therefore, if all frequency dividers are reset by a common signal, their relative phases can be maintained even after power cycles. The trigger synchronized by a common subharmonic (or the common subharmonic itself) is a good candidate to reset frequency dividers. Since the time interval between two trigger pulses covers full periods of all frequencies, the frequency divider reset can maintain the expected relative timing between different frequencies.

In practice, resettable frequency dividers for input frequencies over hundreds of MHz are not much available on the market. Besides, resetting a laser oscillator (equivalent to a frequency divider) at each trigger pulse is not practical. However, if we can measure the timing relation changes, we can correct them by inserting proper knobs, such as a delay stage or a phase shifter.

Figure 8.24 depicts an approach to resynchronize a laser oscillator using a phase shifter (Geng 2020). When the phase shifter is rotated by a full circle, the laser oscillator's output will be moved by one bucket (period) of the reference frequency.

Fig. 8.24 Laser phase resynchronization (the waveforms do not reflect the real signals)

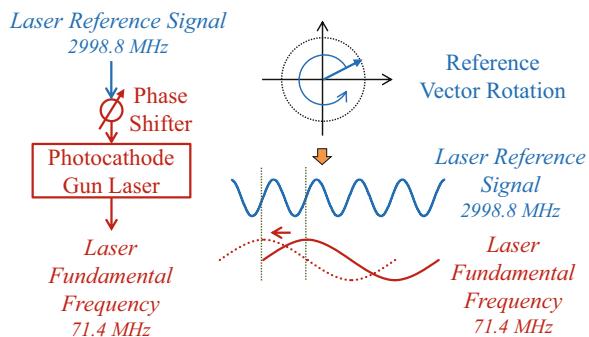
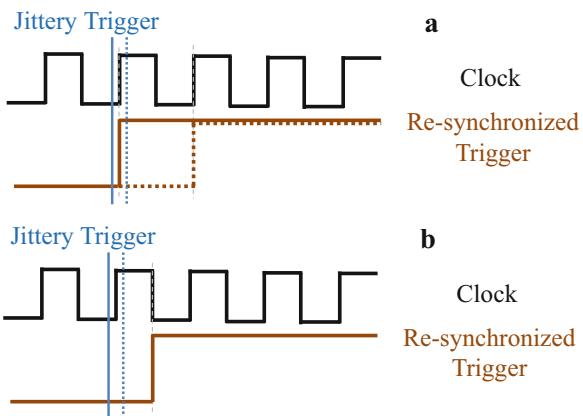


Fig. 8.25 Handling of trigger-clock race condition by synchronizing the trigger to the opposite clock edge



This method can be used to restore the relative timing between the laser oscillator and the master oscillator. Note that the rate of change of the reference phase should be limited to allow the laser PLL to follow. A continuous phase change in the laser reference signal is equivalent to changing the desired laser repetition rate ($d\omega = d\phi/dt$). Therefore, the laser PLL will act on the piezo tuner of the laser oscillator (see Fig. 8.20). A high-frequency drive to the piezo may stimulate its mechanical resonance modes and cause the laser PLL to lose locking.

The same method can be used to resynchronize other frequency dividers.

8.5.3.5 Race Condition Handling

To handle the trigger-clock race conditions, one can insert a piece of cable in the trigger or clock path. The goal is to introduce an additional delay to avoid aligning the trigger and clock edges. Alternatively, we can synchronize the trigger with the clock falling (or rising) edge if the race conditions happen for the clock rising (or falling) edge. See Fig. 8.25. Consequently, the trigger time will be moved by half of the clock period, resulting in a static change in the RF phase measurement or the macro RF pulse starting time.

With reference tracking, the static phase measurement error can be easily compensated. However, the uncertainty in the macro RF pulse starting time (up to an LLRF clock period) is difficult to compensate for. Fortunately, it is a minor influence on the beam acceleration and can be corrected by the beam-based feedbacks. One should be aware that the residual errors caused by the uncertain macro RF pulse timing make it difficult to restore the beam acceleration exactly, even if the machine's drifts are perfectly compensated.

8.6 Summary

This chapter offers an overview of the timing and synchronization systems of particle accelerators. These two systems are complex, and discussing all details is not practical. Therefore, we focused on the key points critical in the analysis and design. For the timing system, we used the EVG-EVR based architecture as the basis of discussion. Other timing system architectures, such as the White Rabbit, are not covered here. Of course, the basic concepts, like fiducial trigger generation, common subharmonic, and client trigger generation, are similar. For synchronization systems, we emphasized the phase drift compensation methods. The phase drift is one of the major problems to be solved by the designers of the synchronization system. Finally, we discussed the timing relation uncertainties when rebooting the timing, synchronization, or LLRF systems. We introduced several methods to diagnose and recover the timing relations after a boot. They are critical to the recovery of the beam operation quickly after a failure, maintenance, or shutdown of the machine.

References

- Bellaveglia M (2016) Timing and synchronization. In: Proceedings of the CERN accelerator school on FELs and ERLs, Hamburg, Germany, 31 May-10 June 2016
- Best RE (2007) Phase-locked loops: design, simulations, and applications, 6th edn. McGraw-Hill Education, New York
- Byrd J M, Doolittle L, Huang G et al (2010) Femtosecond synchronization of laser systems for the LCLS. In: Proceedings of the 1st international particle accelerator conference, Kyoto, Japan, 23-28 May 2010
- Chernousko Y, Gonias A, Heron M T et al (2006) The timing system for diamond light source. In: Proceedings of the 10th European particle accelerator conference, Edinburgh, Scotland, 26-30 June 2006
- Czuba K, Sikora D (2011) Temperature stability of coaxial cables. *Acta Physica Polonica A* 119(4): 333–337. <https://doi.org/10.12693/APhysPolA.119.553>
- Czuba K, Sikora D, Zembala L et al (2013) Overview of the RF synchronization system for the European XFEL. In: Proceedings of the 4th international particle accelerator conference, Shanghai, China, 12-17 May 2013
- Czuba K, Berlinski J, Czuba L et al (2018) Concept of the phase reference line for the European Spallation Source. In: Proceedings of the 22ed international microwave and radar conference, Poznan, Poland, 14-17 May 2018

- Damerau H (2017) Timing, synchronization & longitudinal aspects. In: Proceedings of the CERN accelerator school on beam injection, extraction and transfer, Erice, Italy, 10–19 March 2017
- Doolittle L, Huang G, Ratti A et al (2015) The LCLS-II LLRF system. In: Proceedings of the 6th international particle accelerator conference, Richmond, VA, USA, 3-8 May 2015
- Dusatko J, Allison S, Browne M et al (2010) The LCLS timing event system. In: Proceedings of the 14th beam instrumentation workshop , Santa Fe, NM, USA, 2–6 May 2010
- Frisch J, Brown D, Cisneros E (2000) Performance of the prototype NLC RF phase and timing distribution system. In: Proceedings of the 9th beam instrumentation workshop, Cambridge, MA, USA, 8-11 May 2000
- Gallo A (2018) Timing and synchronization. In: Proceedings of the CERN accelerator school beam instrumentation, Tuusula, Finland, 2-15 June 2018
- Gasowski B, Owczarek T, Czuba K et al (2018) Real-time redundancy for the 1.3 GHz master oscillator of the European XFEL. <https://arxiv.org/abs/1806.09281>. Accessed 03 June 2021
- Geng Z (2020) Robustness issues of timing and synchronization for free electron lasers. Nuclear Inst Methods Phys Res A 963:163738. <https://doi.org/10.1016/j.nima.2020.163738>
- https://en.wikipedia.org/wiki/Crystal_oven
- <https://white-rabbit.web.cern.ch>
- https://www.rohde-schwarz.com/ch/produkt/sma100b-produkt-startseite_63493-427776.html
- Hunziker S, Arsov V, Buechi F et al (2014) Reference distribution and synchronization system for SwissFEL: concept and first results. In: Proceedings of the 3rd international beam instrumentation conference, Monterey, CA, USA, 14-18 September 2014
- Jobe RK, Schwarz HD (1989) RF phase distribution systems at the SLC. In: Proceedings of the 13th particle accelerator conference, Chicago, IL, USA, 20-23 March 1989
- Jung K, Kim J (2012) Sub-femtosecond synchronization of microwave oscillators with mode-locked Er-fiber lasers. Opt Lett 37(14):2958–2960. <https://doi.org/10.1364/OL.37.002958>
- Kalantari B, Biffiger R (2017) SwissFEL timing system: first operational experience. In: Proceedings of the 16th int conf on accelerator and large experimental control systems, Barcelona, Spain, 8-13 October 2017
- Korhonen T (1999) Review of accelerator timing systems. In: Proceedings of the 7th international conference on accelerator and large experimental physics control systems, Trieste, Italy, 4-8 October 1999
- Lamb T, Felber M, Kozak T et al (2019) Femtosecond laser to RF synchronization and RF reference distribution at the European XFEL. In: Proceedings of the 39th international free-electron laser conference, Hamburg, Germany, 26-30 August 2019
- Lin Z, Du Y, Yang J et al (2018a) Development of sub-100 femtosecond timing and synchronization system. Rev Sci Instrum 89:014701. <https://doi.org/10.1063/1.5001768>
- Lin Z, Yang J, Du Y et al (2018b) A novel double sideband-based phase averaging line for phase reference distribution system. In: Proceedings of the 9th international particle accelerator conference, Vancouver, BC, Canada, 29 April-4 May 2018
- Liu F, Lei G, Duan Z et al (2021) The design of HEPS global timing system. Radiat Detect Technol Methods 2021. <https://doi.org/10.1007/s41605-021-00257-6>
- Loehl F (2011) Timing and synchronization. In: Proceedings of the CERN accelerator school on advanced accelerator physics, Chios, Greece, 18-30 September 2011
- Lorbeer B, Ludwig F, Schlarb H et al (2007) Noise and drift characterization of direct laser to RF conversion scheme for the laser based synchronization system for FLASH at DESY. In: Proceedings of the 22ed particle accelerator conference, Albuquerque, NM, USA, 25-29 June 2007
- Mehrotra A (2002) Noise analysis of phase-locked loops. IEEE Trans Circuits Syst I Fundam Theory Appl 49(9):1309–1316. <https://doi.org/10.1109/TCSI.2002.802347>
- Peng MY, Kalaydzhyan A, Kaertner FX (2014) Balanced optical-microwave phase detector for sub-femtosecond optical-RF synchronization. Opt Express 22(22):27103. <https://doi.org/10.1364/OE.22.027102>

- Rose C R, Lara P D, Merl R et al (2001) A timing-reference generator for power-grid-synchronized neutron-spallation facilities. In: Proceedings of the 2001 particle accelerator conference, Chicago, IL, USA, 18-22 June 2001
- Rybacyk LJ, Shelley FE (1997) Improvement to the LANSCE accelerator timing system. In: Proceedings of the 1997 particle accelerator conference, Vancouver, Canada, 12-16 May 1997
- Schlarp H (2013) Timing and synchronization. In: Proceedings of the CERN accelerator school on advanced accelerator physics, Trondheim, Norway, 18-29 August 2013
- Schulz S, Schmueser P, Zemella J et al (2008) An optical cross-correlation scheme to synchronize distributed laser systems at FLASH. In: Proceedings of 11th European particle accelerator conference, Genoa, Italy, 23-27 June 2008
- Serrano J, Alvarez P, Lipinski M et al (2011) Accelerator timing systems overview. In: Proceedings of the 2011 particle accelerator conference, New York, USA, 28 March-1 April 2011
- Sikora D, Czuba K, Jatczak P et al (2020) Phase drift compensating RF link for femtosecond synchronization of E-XFEL. *IEEE Trans Nucl Sci* 67(9):2136–2142. <https://doi.org/10.1109/TNS.2020.2966018>
- Titberidze M (2017) Pilot study of synchronization on a femtosecond scale between the electron gun REGAE and a laser-plasma accelerator. Ph.D. thesis, Hamburg University
- Urbanski M, Jatczak P, Sikora D et al (2018) Frequency divider module for the European XFEL phase reference signal distribution system. In: Proceedings of the 22nd international microwave and radar conference, Poznan, Poland, 14-17 May 2018
- Vargas EC, Kaertner FX, Berlin A et al (2019) Timing stability comparison study of RF synthesis techniques. In: Proceedings of the 39th international free-electron laser conference, Hamburg, Germany, 26-30 August 2019
- Weaver JN, Hogg HA (1983) Improving the phase stability of the SLAC RF driveline network for SLC operation. *IEEE Trans Nucl Sci* 30(4):2234–2236. <https://doi.org/10.1109/TNS.1983.4332772>
- Yang H, Jeon C-G, Jung K et al (2015) Femtosecond synchronization of 80-MHz Ti:Sapphire photocathode laser oscillator with S-band RF oscillator. In: Proceedings of the 37th international free electron laser conference, Daejeon, Korea, 23-28 August 2015
- Yang J, Du Y-C, Yan L-X et al (2017) Laser-RF synchronization based on digital phase detector. *Nucl Sci Tech* 28:57. <https://doi.org/10.1007/s41365-017-0205-z>
- Zembala L, Czuba K, Gasowski B et al (2014) Master oscillator for the European XFEL. In: Proceedings of the 5th international particle accelerator conference, Dresden, Germany, 15-20 June 2014
- Zhang W-Y, Liu X-Q, Feng L et al (2018) 2.856 GHz microwave signal extraction from mode-locked Er-fiber lasers with sub-100 femtosecond timing jitter. *Nucl Sci Tech* 29:91. <https://doi.org/10.1007/s41365-018-0419-8>
- Zhao LY, Yin CX, Liu DK (2008) The SSRF timing system. In: Proceedings of the 11th European particle accelerator conference, Genoa, Italy, 27-27 June 2008

Chapter 9

LLRF Applications



In addition to controlling the RF field for beam acceleration, the LLRF system also provides functions to facilitate the RF system operation. The operation-supporting tools are often called *high-level applications (HLAs)*. Typical HLAs implement the procedures like optimizing the RF system operational parameters, calibrating the RF measurements, identifying the system characteristics, and automating the RF system startup and conditioning. In this chapter, these HLA algorithms and procedures will be discussed.

9.1 Overview

In an LLRF system, RF detectors measure the RF signals picked from various locations of the RF system (e.g., cavity probe signal). RF actuators adjust the amplitude and phase of the cavity drive power. Digital LLRF systems often adopt FPGAs or DSPs to implement RF controllers to regulate the RF field. The RF controller input is the RF detector signal, and the output is the RF actuator drive signal. The RF detectors, controllers, and actuators need to be configured appropriately to perform the desired functions. Therefore, HLAs are needed to perform the following tasks (Brandt 2007):

(a) *Calibrate the RF measurements*

We need an accurate estimate of the RF field used for beam acceleration (i.e., accelerating voltage and beam phase) based on the RF detector outputs. In many cases, we also need to know the absolute power of the RF signals.

(b) *Optimize the operational parameters*

We need to optimize the RF system operating parameters, such as the loaded quality factor (Q_L) and detuning of the cavity, the pulse shape of the RF drive, the klystron

modulator high voltage, and the RF controller feedback gain. The goals are to improve the RF field stability and to minimize the required RF drive power.

(c) *Identify the RF system characteristics*

To obtain insights into the RF system, we estimate the system characteristics using the RF signal measurements. The characteristics include the RF plant system gain, the system phase, and the detuning and Q_L (or input coupling factor β) of the cavities. The identification is often performed in-situ, i.e., during system operation. The results can be used to detect exceptions (e.g., quench detection by observing the drop of Q_L) or as inputs to other HLAs for calibration or optimization.

(d) *Automation*

Automating the startup, interlock trip recovery, conditioning, and operation of the RF system is vital for large-scale accelerators consisting of multiple RF stations. It helps to reduce the required personnel for operation and to increase the RF system reliability. The automation is typically implemented as procedural sequencers or finite-state machines (FSM). Since this book focuses on the HLA algorithms, automation will not be discussed here.

LLRF HLA users are RF system operators and *physics applications* automating the beam operation, like the *beam energy manager* that automatically sets up the desired beam energy.

9.2 Parameter Optimization

In state-of-the-art digital LLRF systems, most control functions (e.g., RF demodulation, feedback/feedforward control, and RF actuation) are typically implemented in firmware or software. The firmware and software are configurable via many parameters and are flexible for different operation conditions. For example, we may use the same implementation to operate a cavity in either pulsed or CW mode by re-optimizing the parameters. On the other hand, we must set up and optimize these parameters, leading to the needs of HLA tools. Here, the term *optimization* is to find the “golden” values of the parameters aiming to improve the RF system performance. The required tools, which can be semi- or fully automated, assist the operators in optimizing the operational parameters. This section provides several examples that apply to most LLRF systems.

9.2.1 RF Pulse Shaping

For RF stations operating in the pulsed mode, the pulse shape is defined by a feedforward table stored in the LLRF firmware. The feedforward table is applied to the RF actuator via DACs to generate an RF pulse in response to a trigger.

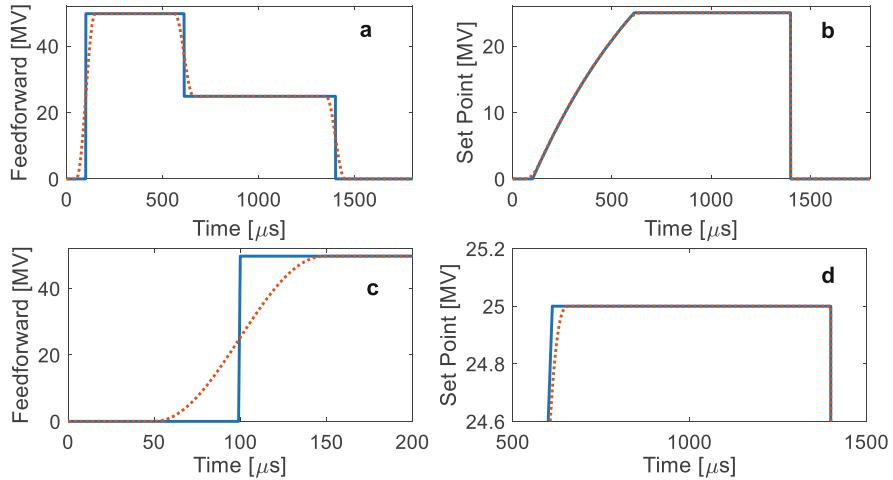


Fig. 9.1 Feedforward and setpoint tables for a superconducting cavity. The dotted lines are the smoothed tables. **(a)** Feedforward table; **(b)** Setpoint table; **(c)** Zoom-in of feedforward table at the rising edge; **(d)** Zoom-in of setpoint table at the flattop

The RF pulse shape is adapted to the dynamics and other requirements of different types of RF stations. The RF pulse for normal-conducting Linacs is typically a square pulse with a constant phase. If RF pulse compressors (e.g., SLED or BOC) are installed, a phase modulation by 180° is required (see Sect. 3.6.3). The typical pulse of a superconducting cavity operating in the pulse mode for the acceleration of a bunch train consists of an exponential rising part followed by a flattop. The exponential ramping of the cavity voltage can avoid excessive cavity drive power due to the large cavity time constant. The flattop is used for the bunch train acceleration. Therefore, the typical feedforward and setpoint tables of a superconducting cavity operating in the pulsed mode are depicted in Fig. 9.1. The feedforward signal is larger at the exponential rising stage of the cavity voltage for a faster filling; whereas at the flattop, the feedforward signal steps down to create the flattop. As we have demonstrated in Example 3.1 in Sect. 3.3.8, a different feedforward signal should be used if the beam is present.

The time-varying cavity voltage setpoint in Fig. 9.1 can be determined using the cavity equation introduced in Chap. 3 as

$$\begin{cases} \mathbf{v}_{SP}(t) = \frac{\mathbf{v}_{C0}(1 - e^{-\omega_{1/2}t})}{1 - e^{-\omega_{1/2}T_{fill}}}, & 0 \leq t < T_{fill} \\ \mathbf{v}_{SP}(t) = \mathbf{v}_{C0}, & T_{fill} \leq t \leq T_{pul} \end{cases} \quad (9.1)$$

where \mathbf{v}_{C0} is the desired cavity voltage at the flattop, $\omega_{1/2}$ is the half-bandwidth of the cavity, T_{fill} is the cavity filling time and T_{pul} is the overall pulse width. For example, in Fig. 9.1, $\omega_{1/2} = 2\pi \times 216$ Hz, $T_{fill} = 512$ μs and $T_{pul} = 1.3$ ms. The required

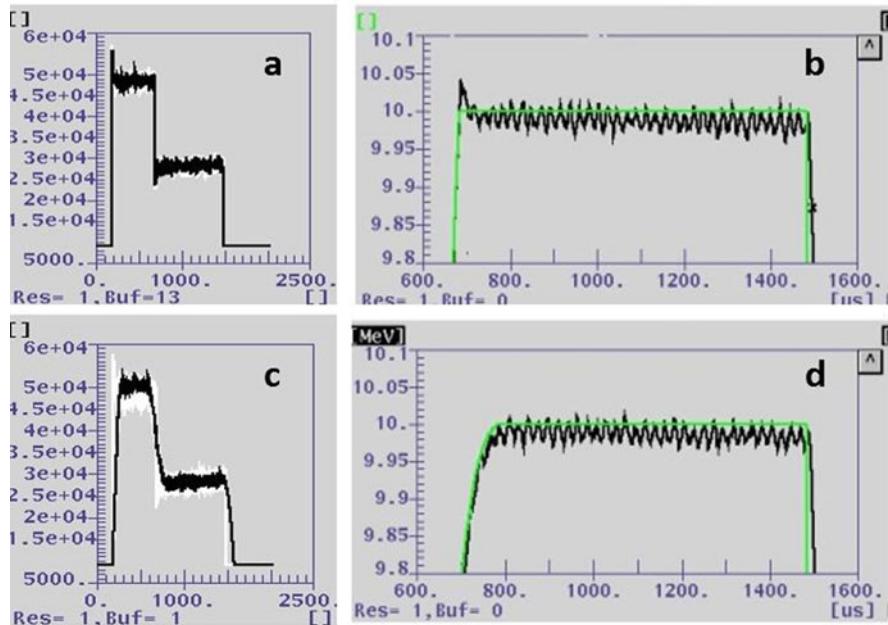


Fig. 9.2 Mitigation of overshoots. **(a)** Cavity drive without smoothing; **(b)** Cavity voltage setpoint and measurement without smoothing; **(c)** Cavity drive with smoothing; **(d)** Cavity voltage setpoint and measurement with smoothing

feedforward table can be calculated by inverting the cavity model. For the setpoint (9.1), the feedforward signal can be calculated as

$$\begin{cases} \mathbf{v}_{FF}(t) = \frac{\mathbf{v}_{C0}}{2(1 - e^{-\omega_{1/2}T_{fill}})}, & 0 \leq t < T_{fill} \\ \mathbf{v}_{FF}(t) = \mathbf{v}_{C0}/2, & T_{fill} \leq t \leq T_{pul} \end{cases} \quad (9.2)$$

where we have assumed the input coupling factor $\beta \gg 1$ for the superconducting cavity. The setpoint and feedforward tables (solid lines) in Fig. 9.1 contain discontinuities, resulting in overshoots when the cavity is controlled by a feedback loop. The overshoots can be reduced by smoothing the feedforward and setpoint tables. First, we smooth the feedforward table at the rising, stepping, and falling edges. Then we apply the smoothed feedforward table to the cavity equation to calculate the cavity response, which should be scaled to reach the desired cavity voltage at the flattop. This scaled cavity response can be used as the smoothed setpoint table (see the dotted lines in Fig. 9.1). Here we have used a half-sine function to smooth the rising and falling edges of the feedforward table. Other smoothing functions, such as a linear or an exponential function, may also be used. Figure 9.2 depicts the overshoot mitigation in the cavity drive and cavity voltage with the smoothed tables.

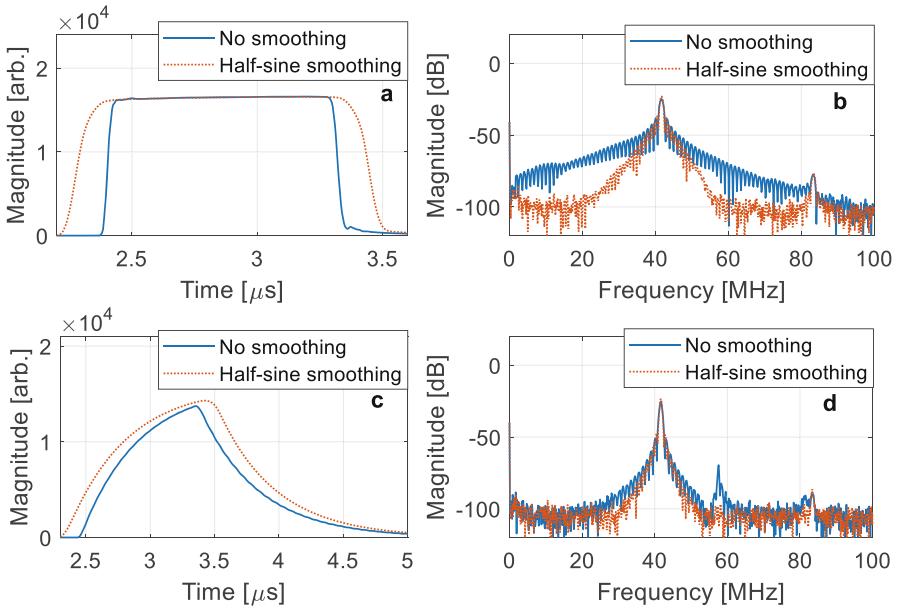


Fig. 9.3 Amplitude waveforms and spectra of the klystron output and cavity probe signals. The signals are down-converted to 41.65 MHz and sampled at 249.9 MHz. (a) Klystron output amplitude; (b) Klystron output IF spectra; (c) Cavity probe amplitude; (d) Cavity probe spectra

We notice that the smoothing reduces the flattop duration, which should be compensated for by increasing the pulse width.

Smoothing of the feedforward table reduces the bandwidth of the RF drive power, which can avoid exciting unwanted passband modes of the cavity. Let us use the SwissFEL RF Gun as an example to demonstrate the effects of a smoothed RF drive pulse (Geng 2017). The SwissFEL RF Gun is a 2.6-cell standing-wave cavity operating at 2998.8 MHz. It works in the π -mode with a pulse width shorter than 1 μ s. The nearest passband mode ($\pi/2$ -mode) is about 16 MHz below the π -mode frequency. A square RF pulse output by the klystron has a wideband that stimulates the $\pi/2$ -mode significantly. The $\pi/2$ -mode signal may corrupt the RF field measurement and cause errors in the closed-loop operation. The RF Gun is controlled by two pulse-to-pulse feedback loops to regulate its amplitude and phase, respectively. To reduce the bandwidth of the RF drive power, we smooth the rising and falling edges of the RF pulse with a half-sine function. Experiment shows that the smoothing can effectively suppress the passband modes in the RF Gun cavity, which can be seen in Fig. 9.3, where the $\pi/2$ -mode signal is aliased to around 58 MHz.

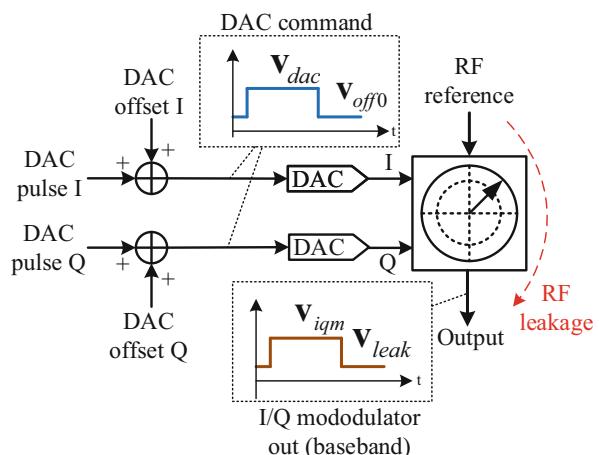
9.2.2 DAC Offset Correction

An RF actuator for direct up-conversion consists of one I/Q modulator and two DACs. See Fig. 9.4. The two DACs generate baseband signals to modulate the RF reference signal with the I/Q modulator. Due to the I/Q modulator's imperfection, the RF reference signal leaks to the output even with zero DAC outputs. The leakage may cause significant phase and amplitude actuation errors (Geng and Hong 2016). To compensate for the RF leakage, we can introduce constant offsets in DAC outputs. The DAC offsets can generate an additional RF output with the same amplitude but 180° out of phase to cancel the RF leakage signal.

The following procedure (Geng 2017) can be followed to determine the DAC offsets:

- Generate a test pulse and measure the resulting I/Q modulator output with an RF detector. We record the waveforms of the DAC commands and the I/Q modulator output. Both of them are baseband signals described by a vector waveform consisting of an in-phase (I) and a quadrature (Q) component.
- From the DAC command waveforms, we calculate the complex values of the DAC initial offset (\mathbf{v}_{off0}) and the DAC pulse (\mathbf{v}_{dac}). They can be calculated by averaging the corresponding parts of the DAC command vector waveforms. See Fig. 9.4.
- From the I/Q modulator output waveforms, we calculate the RF leakage (\mathbf{v}_{leak}) and the generated pulse (\mathbf{v}_{iqm}). They can be calculated by averaging the corresponding parts of the vector waveforms of the I/Q modulator output.
- Calculate the complex gain from the DAC command to the I/Q modulator output as

Fig. 9.4 RF leakage in a direct up-conversion RF actuator



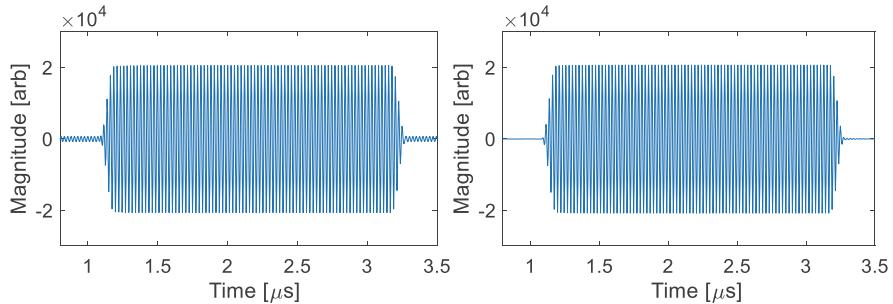


Fig. 9.5 IF signal of the I/Q modulator output without (left) and with (right) DAC offset correction

$$\mathbf{G} = \frac{\mathbf{v}_{igm} - \mathbf{v}_{leak}}{\mathbf{v}_{dac} - \mathbf{v}_{off0}}. \quad (9.3)$$

(e) Calculate the DAC offset required to compensate for the RF leakage as

$$\mathbf{v}_{off} = \mathbf{v}_{off0} - \mathbf{G}^{-1} \mathbf{v}_{leak}. \quad (9.4)$$

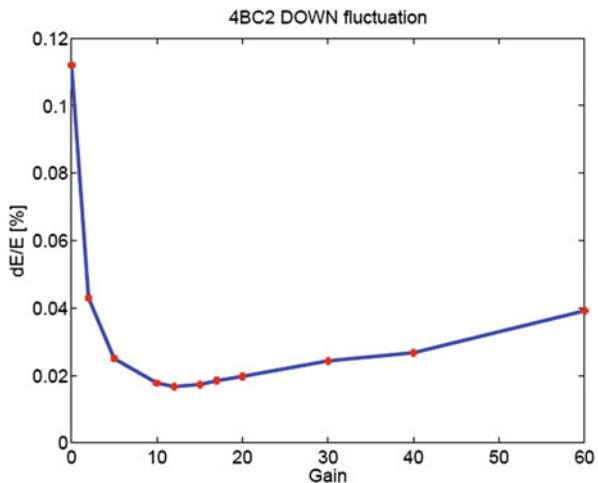
The algorithm above has considered the initial DAC offsets. Therefore, it can be executed iteratively. Figure 9.5 is an example at the SwissFEL, which illustrates that the algorithm can compensate for the RF leakage with one or two iterations.

9.2.3 Parameter Scanning

Many operational parameters of the RF system can be optimized experimentally. The simplest approach is to scan the parameter values in a range and find a “golden” value for optimizing the system output affected by the parameter. This method is suitable for single parameter optimization. Typically, the scanning can be first performed by simulation, especially when the experiment on a real system is not practical or not easy. Simulation can narrow down the optimal parameter value searching range, which defines the safe scanning range on a real system to avoid tripping the interlock.

When setting up the RF controller, we often need to find an optimal feedback gain. The goal is to minimize either the RF field or the beam parameter (e.g., beam energy) fluctuations. Figure 9.6 shows the proportional gain scanning for the feedback loop of a superconducting cavity. As discussed in Sect. 6.3.1, an optimal gain exists for minimizing the RF field error as a tradeoff between the disturbance rejection and the measurement noise contribution. Here we directly use the beam energy jitter as the optimization criteria. It can be seen that the optimal gain is around 12, with which a beam energy jitter of 0.018% is achieved. If we reduce the RF

Fig. 9.6 Gain scanning of an RF station at FLASH and the resulting beam energy jitter



measurement noise, the optimal gain can be increased, resulting in improved beam stability.

Many RF system parameters can be optimized via scanning. Here more examples are listed:

- Optimize the cavity loaded quality factor (or input coupling factor) to minimize the reflected RF power. See Sect. 3.3.9.
- Optimize the cavity detuning with the beam in the presence to minimize the reflected RF power. See Sect. 3.3.9.
- Optimize the cavity filling time and loaded quality factor to compromise the RF pulse width and the RF drive power. See Sect. 3.3.9.
- Optimize the RF pulse compressor's phase inversion time to compromise the RF pulse width and the RF peak power. See Sect. 3.6.3.
- Optimize the loop delay to mitigate the feedback instability caused by the passband modes. See Sect. 4.2.2.

The above examples are mainly for optimizing the parameters of an individual RF station. When scanning each of these parameters, we can typically find an observable quality as the optimization criteria. This feature makes the scanning-based optimization feasible. However, for a particle accelerator, we may need to optimize the setpoints of multiple RF stations to achieve the desired beam quality. For a light source (e.g., FEL or synchrotron radiation), the photon quality is often the optimization criteria. The relationship between the RF station setpoints and the beam (or photon) quality may be too complex for the scanning-based optimization. Furthermore, this relationship is usually nonlinear. In this case, modern optimization algorithms, such as the genetic algorithm (GA), the particle swarm optimization (PSO) algorithm, or the machine-learning algorithms, can be considered. These advanced optimization algorithms will be discussed in a newly planned book *Intelligent Beam Control in Accelerators* that will come soon.

9.3 RF Calibration

In general, *calibration* is to transform an uncalibrated detector signal into a calibrated measurement. For accelerator RF stations, we need to know the physical meanings of the RF detector outputs. For example, the cavity probe signal measurement yields amplitude and phase values, from which we need to know the *accelerating voltage* (V_{acc}) and the *beam phase* (φ_b) experienced by the beam. In Chap. 1, we have defined the accelerating voltage of a single cavity as the magnitude of the cavity voltage phasor ($V_{acc} = |\mathbf{v}_C|$). The amplitude of the RF detector output has a different unit, but V_{acc} has a unit with physical meanings like MV. Similarly, the RF detector output phase refers to an unknown reference phase, while φ_b refers to the beam arrival time. The calibration is necessary for regulating the RF station to provide the desired V_{acc} and φ_b . Therefore, the *RF calibration* is an important step to apply the RF station to beam operation.

Figure 9.7 depicts the RF measurements in an RF station with multiple cavities driven by the same RF amplifier. The required calibrations are listed as follows:

- Calibrate the V_{acc} and φ_b of each cavity;*
- Calibrate the total accelerating voltage and beam phase of the RF station;*
- Calibrate the forward and reflected RF power of each cavity.*

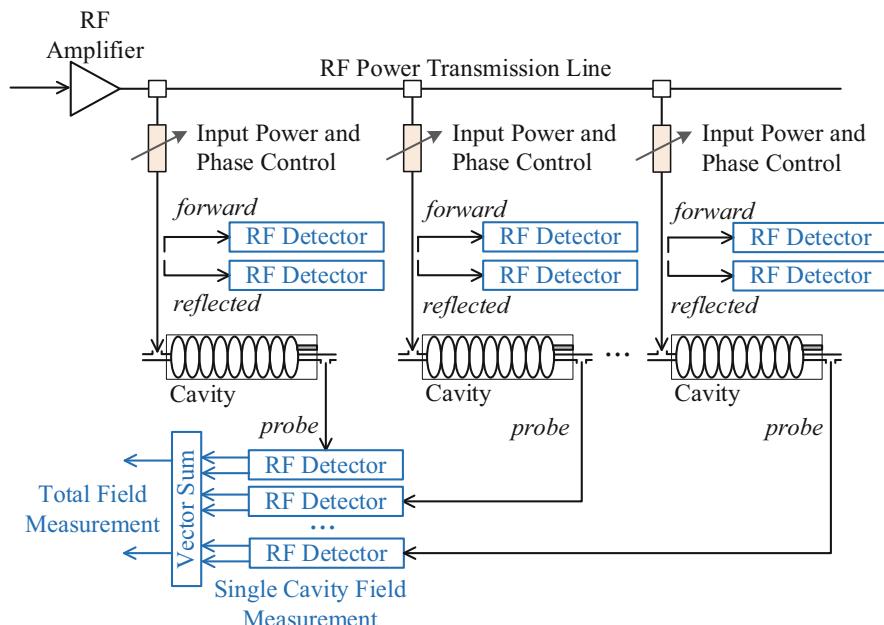


Fig. 9.7 RF measurements in an RF station with multiple cavities

The calibration under *a* above is to scale the amplitude and rotate the phase of each cavity's probe signal. When a particle passes through all cavities of the RF station, it experiences an integrated RF field, which is a *vector sum* of the cavity voltage phasors. The total acceleration to the beam is described by the *vector-sum accelerating voltage* ($V_{acc,vs}$) and the *vector-sum beam phase* ($\varphi_{b,vs}$). To perform proper beam acceleration using such an RF station with multiple cavities, we need to balance the inputs of different cavities. The *cavity input balancing* consists of two typical steps aiming at maximizing the accelerating voltage. First, we need to tune the input phases of different cavities so that the particle beam gets continuous acceleration from one cavity to another. Second, the input RF power to each cavity should be adjusted so that all cavities yield similar gradients. If some cavities obtain higher input powers, they may breakdown or quench before other cavities reach their maximum gradients. The calibration under *a* above provides inputs to the cavity input balancing procedure.

There are different methods to perform the calibration under *b*. The $V_{acc,vs}$ and $\varphi_{b,vs}$ are the quantities measured and regulated by the RF station controller. If the V_{acc} and φ_b of each cavity is known from the calibration under *a*, the $V_{acc,vs}$ and $\varphi_{b,vs}$ can be calculated directly as a vector sum

$$V_{acc,vs} e^{j\varphi_{b,vs}} = \sum_{n=1}^N V_{acc}^{(n)} e^{j\varphi_b^{(n)}}, \quad (9.5)$$

where N is the number of cavities. Here we use a superscript (n) to denote the quantities of the n th cavity. In many cases, we are interested only in the vector-sum accelerating voltage and beam phase, but not that of individual cavities. Therefore, we may directly perform a *vector-sum calibration* on the raw measurements of the cavity probe signals. Then we calibrate the vector-sum amplitude and phase into physical units as $V_{acc,vs}$ and $\varphi_{b,vs}$, respectively. This approach is widely used in practical LLRF systems and is discussed in Sect. 9.3.4. For high power RF experts, the cavity forward and reflected powers are of interest. We will discuss the *RF power calibration* algorithms in Sect. 9.3.6. In addition, to correct for the poor directivity of the directional couplers used to pick up the forward and reflected signals, we also need a *forward and reflected signal calibration* (see Sect. 9.3.5).

9.3.1 Beam-Induced Transient

A bunch train passing through a cavity will excite RF fields in the cavity. The cavity voltage induced by a short bunch train (with a duration much shorter than the cavity time constant) is called a *beam-induced transient*. For the discussions below, we will neglect the beam loss in the cavities of an RF station. The beam-induced transients in different cavities are correlated because they are excited by the same bunch train.

Therefore, if we can measure the beam-induced transients, we may use them as references to perform several calibrations mentioned above.

Note that we only consider the fundamental mode (e.g., the TM010 π -mode of a TESLA cavity) signal induced by the beam because it can be measured by a standard LLRF system. The beam phase can also be derived from the higher-order mode (HOM) signals of the cavity using RF detectors working at the HOM frequencies (Shi et al. 2018).

9.3.1.1 Physical Meaning of Beam-Induced Transient

First, let us use the cavity equation to analyze the physical meaning of the beam-induced transient. We assume that a cavity has no RF drive and its initial cavity voltage is zero. At $t = 0$, we inject a bunch train with an average current of \mathbf{i}_{b0} , described as a phasor. The beam current phasor magnitude is given by $|\mathbf{i}_{b0}| = Q_b/T_b$, where Q_b is the bunch charge and T_b is the bunch spacing in time. The phase of \mathbf{i}_{b0} reflects the bunch arrival time relative to the RF reference signal phase. The beam-induced cavity voltage \mathbf{v}_{Cb} is calculated with the equation below

$$\dot{\mathbf{v}}_{Cb} + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_{Cb} = 2\omega_{1/2}R_L\mathbf{i}_{b0}, \quad (9.6)$$

where $\omega_{1/2}$, $\Delta\omega$ and R_L are the cavity's half-bandwidth, detuning and loaded resistance, respectively. See Chap. 3 for details of these parameters. Here we only consider a short bunch train, which starts at $t = 0$ and ends at $t = \Delta T$ with $\Delta T \ll \tau_{cav} = 1/\omega_{1/2}$, where τ_{cav} is the time constant of the cavity. For example, the time constant of a TESLA cavity is around $\tau_{cav} = 735 \mu\text{s}$, and we use a bunch train of $\Delta T = 30 \mu\text{s}$ to measure the beam-induced transient. In this case, the bunch train-induced cavity voltage is very small in magnitude compared to its changing rate; that is, $|(\omega_{1/2} - j\Delta\omega)\mathbf{v}_{Cb}| \ll |\dot{\mathbf{v}}_{Cb}|$. Therefore, (9.6) can be approximated as

$$\dot{\mathbf{v}}_{Cb} \approx 2\omega_{1/2}R_L\mathbf{i}_{b0}, \quad (9.7)$$

then we obtain $\mathbf{v}_{Cb}(t) \approx 2\omega_{1/2}R_L\mathbf{i}_{b0}t$. We define the beam-induced transient (denoted as \mathbf{v}_{Cbt}) as the maximum cavity voltage induced by the bunch train, which is the value of $\mathbf{v}_{Cb}(t)$ at $t = \Delta T$:

$$\mathbf{v}_{Cbt} := \mathbf{v}_{Cb}|_{t=\Delta T} \approx 2\omega_{1/2}R_L\mathbf{i}_{b0}\Delta T. \quad (9.8)$$

The magnitude of \mathbf{v}_{Cbt} is independent of $\Delta\omega$ and is proportional to the bunch train's integrated charge ($|\mathbf{i}_{b0}|\Delta T$). Typically, the bunch charge can be measured accurately. Then we can calculate the beam-induced cavity voltage in physical units (e.g., MV) if $\omega_{1/2}$ and R_L are known. Furthermore, we also notice that the phase of \mathbf{v}_{Cbt} is approximately the same as that of the beam current phasor, \mathbf{i}_{b0} . These features of \mathbf{v}_{Cbt} enable us to perform various calibrations if we can measure the beam-induced

transients from the cavity probe signals. Of course, the measurement is only feasible if the magnitude of \mathbf{v}_{Cbtr} is large enough to be detectable by the resolution of the RF detectors. For a superconducting cavity (with a large R_L) and a large bunch charge, the beam-induced transient is often measurable. However, in many normal-conducting cavities accelerating a beam with a very low bunch charge, the beam-induced transient may be too small to be measured.

9.3.1.2 Measurement of Beam-Induced Transient

Ideally, the beam-induced transient should be measured while there is no RF drive and no initial RF field inside the cavity. It allows clear identification of the contribution from the beam. In practice, this approach is only feasible for the high-energy part of an accelerator. At the low-energy part, the beam loss may be significant since the beam energy loss in the cavity is relatively too large for efficient beam transmission. Usually, we measure the beam-induced transient with the RF drive on and the initial RF field in the presence. In this case, we inject a short bunch train during the flattop of the cavity voltage. The beam-induced cavity voltage is then calculated by subtracting the RF-induced cavity voltage from the measurement of the cavity probe signal. This relation is described as

$$\mathbf{v}_{Cb,mea}(t) = \mathbf{v}_{C,mea}(t) - \mathbf{v}_{Cr,mea}(t), \quad t \in [t_{bs}, t_{be}], \quad (9.9)$$

where $\mathbf{v}_{C,mea}$ is the measurement of the cavity probe signal. Here we use the subscript *mea* to indicate that the amplitude and phase of the signals are not calibrated. The cavity voltages (not calibrated) resulting from the RF drive and the beam are denoted as $\mathbf{v}_{Cr,mea}$ and $\mathbf{v}_{Cb,mea}$, respectively. The duration of the bunch train is between t_{bs} and t_{be} . To simplify the calculation, we usually apply an RF drive with a constant amplitude and phase and operate the cavity in an open loop. Furthermore, the bunch train should have a constant bunch charge and a constant bunch repetition rate, corresponding to a square pulse of \mathbf{i}_{bo} . In this case, the beam-induced cavity voltage variation is approximately a linear function of time. As an example, Fig. 9.8 shows the amplitude and phase waveforms of the probe signals of eight TESLA cavities driven by the same klystron. During the measurement, we have injected 30 bunches with a bunch charge of 3 nC and a bunch repetition rate of 1 MHz. The starting time of the bunch train is at 510 μs .

As mentioned before, the cavity voltage changes approximately linearly in a time interval much shorter than the cavity time constant. Therefore, we can estimate the RF-induced cavity voltage $\mathbf{v}_{Cr,mea}$ during the bunch train by applying linear interpolations to $\mathbf{v}_{C,mea}$. This algorithm is depicted in Fig. 9.9, which displays the linear fits of $\mathbf{v}_{C,mea}$ shortly before and after the bunch train. Note that $\mathbf{v}_{C,mea}$ has complex values and the linear fits should be applied to the complex waveform. Figure 9.9 shows only the amplitude waveform to demonstrate the principle of the algorithm.

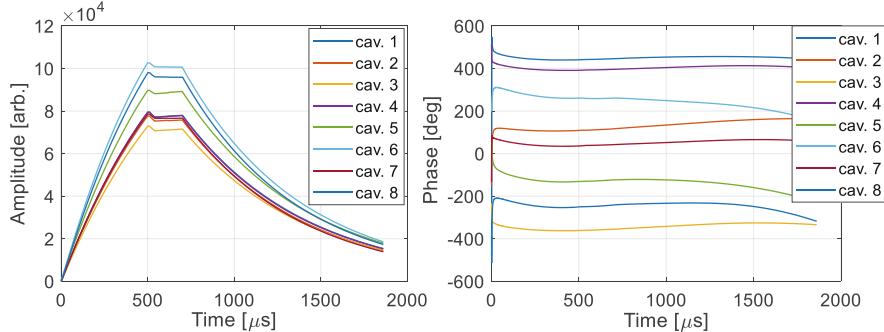
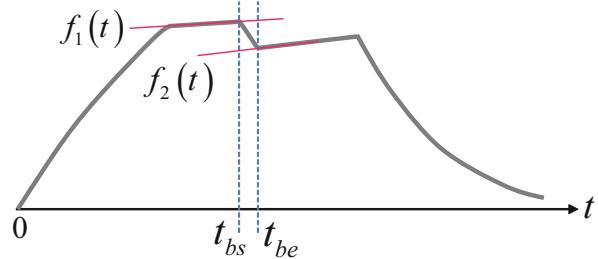


Fig. 9.8 Amplitude (left) and phase (right) waveforms of 8 cavity probe signals. The unit of amplitudes is in bits of a 16-bit ADC. The phases contain unknown offsets from the RF measurement paths

Fig. 9.9 Estimation of RF-induced cavity voltage during the bunch train



We fit the cavity voltage before and after the bunch train as two complex linear functions $f_1(t)$ and $f_2(t)$, respectively. Between t_{bs} and t_{be} , the beam is present, and the cavity voltage is reduced linearly due to the beam loading. Note that during the bunch train, only the overall cavity voltage $\mathbf{v}_{C,\text{mea}}$ is known from the cavity probe signal measurement. To calculate $\mathbf{v}_{Cb,\text{mea}}$, we need first to estimate the $\mathbf{v}_{Cr,\text{mea}}$ during the bunch train. Figure 9.9 implies that we can estimate the values of $\mathbf{v}_{Cr,\text{mea}}$ between t_{bs} and t_{be} with either $f_1(t)$ or $f_2(t)$, assuming that the RF-induced cavity voltage changes linearly around the bunch train duration. Therefore, we can use either of the following formulas to calculate $\mathbf{v}_{Cb,\text{mea}}$:

$$\begin{aligned} \mathbf{v}_{Cb,\text{mea}1}(t) &= \mathbf{v}_{C,\text{mea}}(t) - f_1(t) \\ \mathbf{v}_{Cb,\text{mea}2}(t) &= \mathbf{v}_{C,\text{mea}}(t) - f_2(t) - (f_1(t_{bs}) - f_2(t_{be})) \end{aligned} \quad t \in [t_{bs}, t_{be}]. \quad (9.10)$$

We often estimate $\mathbf{v}_{Cb,\text{mea}}$ using their average:

$$\begin{aligned} \mathbf{v}_{Cb,\text{mea}}(t) &= (\mathbf{v}_{Cb,\text{mea}1}(t) + \mathbf{v}_{Cb,\text{mea}2}(t))/2 \\ &= \mathbf{v}_{C,\text{mea}}(t) - \frac{f_1(t) + f_2(t)}{2} - \frac{f_1(t_{bs}) - f_2(t_{be})}{2}, \end{aligned} \quad t \in [t_{bs}, t_{be}]. \quad (9.11)$$

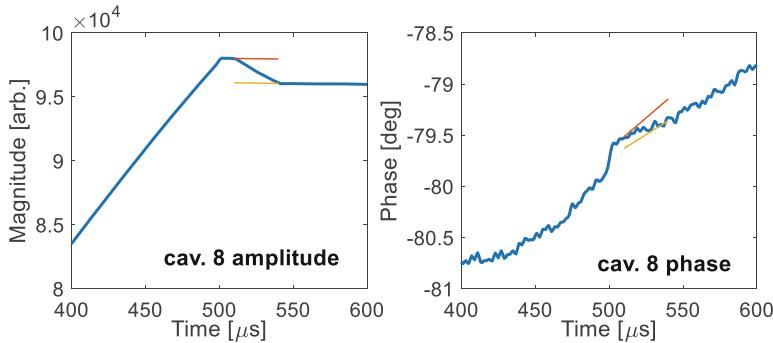
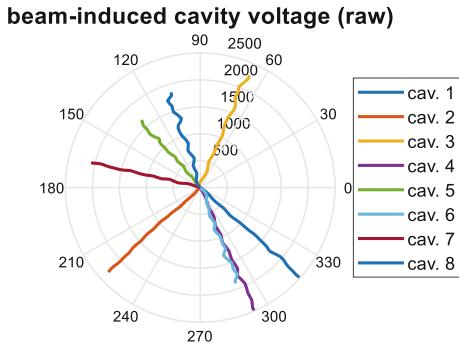


Fig. 9.10 Estimation of the beam-induced cavity voltage for the eighth cavity

Fig. 9.11 Vector plots of the beam-induced cavity voltages. The origin and the far ends of the plots correspond to the values of $\mathbf{v}_{Cb,mea}$ at t_{bs} and t_{be} , respectively



The algorithm (9.11) is applied to the data of Fig. 9.8. As an example, the fits of the amplitude and phase waveforms of the cavity 8 is shown in Fig. 9.10. The vector plots of the calculated $\mathbf{v}_{Cb,mea}$ for all the eight cavities are displayed in Fig. 9.11. Note that the amplitude and phase of $\mathbf{v}_{Cb,mea}$ are not calibrated yet. From the beam-induced transient definition (9.8), we can obtain for each cavity that $\mathbf{v}_{Cbtr,mea} = \mathbf{v}_{Cb,mea}(t_{be})$. Typically, $\mathbf{v}_{Cbtr,mea}$ is calculated with the linear fit of $\mathbf{v}_{Cb,mea}$, taking the value of the fitted function at t_{be} .

9.3.2 Accelerating Voltage and Beam Phase Calibration

The measurement of the cavity probe signal, $\mathbf{v}_{C,mea}$, is via a path with an unknown attenuation and phase shift. The amplitude of $\mathbf{v}_{C,mea}$ is proportional to the accelerating voltage, and the phase of $\mathbf{v}_{C,mea}$ has a constant offset to the beam phase. The beam phase describes the relative phase between the cavity voltage and the beam current. See Sect. 3.3.8 for details. Therefore, the *accelerating voltage and*

beam phase calibration is to find out the scale factor F_V and the phase offset Φ_{off} to obtain a calibrated cavity voltage $\mathbf{v}_{C,cal}$ from the raw measurement $\mathbf{v}_{C,mea}$:

$$\mathbf{v}_{C,cal} := V_{acc} e^{j\varphi_b} = F_V e^{j\Phi_{off}} \mathbf{v}_{C,mea}. \quad (9.12)$$

Note that the amplitude of $\mathbf{v}_{C,cal}$ is in a physical unit like MV, and the phase of $\mathbf{v}_{C,cal}$ equals to the beam phase φ_b . It implies that we define the phase of \mathbf{i}_{b0} to be π according to the definition of φ_b in (3.33). Such a definition is practical because when operating the RF station, we usually specify the desired beam phase as the phase setpoint. Therefore, if we can obtain the beam phase directly from the cavity field measurement, the control of beam phase can be done by controlling the phase of $\mathbf{v}_{C,cal}$. Here we have assumed that the phase of \mathbf{i}_{b0} (i.e., the bunch arrival time compared to the RF reference) is a constant. When the bunch arrival time changes, for example for a different setting of the chicane, the phase calibration Φ_{off} should be updated correspondingly.

9.3.2.1 Accelerating Voltage Calibration with RF Drive Power

When commissioning a particle accelerator, a rough estimate of the beam energy at various locations is necessary to set up the magnets for successful beam transmission. That is, we should estimate the accelerating voltage before the beam is available. In this case, the accelerating voltage cannot be calibrated with the beam. Instead, the RF drive power of the cavity or structure, P_{for} , should be used (Bravo et al. 2015). Here we assume that the power of the RF drive signal has been calibrated using the method described in Sect. 9.3.6.

For a standing-wave cavity, the relationship between the RF drive power and the accelerating (cavity) voltage is given by (3.23) and (3.30). The steady-state accelerating voltage induced by the RF drive power in a well-tuned cavity ($\Delta\omega = 0$) can be calculated as

$$V_{acc} = |\mathbf{v}_C| = 2R_L |\mathbf{i}_{for}| = 2\sqrt{\frac{2\beta R_L P_{for}}{\beta + 1}}, \quad (9.13)$$

where β is the input coupling factor and R_L is the loaded resistance of the cavity. The phasor \mathbf{i}_{for} is the equivalent cavity drive current corresponding to P_{for} . See Sect. 3.3.4 for details. For example, a TESLA cavity has the following parameters: $r/Q = 1036 \Omega$, $Q_L = 3e6$ and $\beta \gg 1$. For an RF drive power $P_{for} = 50$ kW, the expected accelerating voltage in the steady state is about 24.9 MV. For a traveling-wave structure, the accelerating (structure) voltage induced by the RF driver power can be calculated according to (3.60) as

$$V_{acc} = |\mathbf{v}_{ACC}| = E_0 L = \sqrt{LP_{for}r_{s,tw}(1 - e^{-2\tau_{tw}})}, \quad (9.14)$$

where L is the structure length, $r_{s,tw}$ is the shunt impedance per unit length and τ_{tw} is the total power attenuation factor of the structure. Let us use the SwissFEL S-band (2998.8 MHz) traveling-wave structure as an example. Its parameters are: $L = 4.15$ m, $r_{s,tw} = 56.5$ MΩ/m, quality factor $Q = 1.28\text{e}4$, filling time $T_f = 1$ μs, and $P_{for} = 24$ MW. From (3.61), we can calculate $\tau_{tw} = \omega T_f/(2Q) = 0.736$. Then the expected accelerating voltage is calculated to be about 65.85 MV, and the gradient is about 15.88 MV/m. Note that with the value of P_{for} , we can only estimate the accelerating voltage but not the beam phase. Then the voltage calibration factor can be calculated as $F_V = V_{acc}/V_{mea}$, where V_{mea} represents the measured magnitude of $|\mathbf{v}_{C,mea}|$ or $|\mathbf{v}_{ACC,mea}|$ in the steady state.

9.3.2.2 Accelerating Voltage and Beam Phase Calibration with Beam-Induced Transient

As discussed in Sect. 9.3.1.2, the beam-induced transient $\mathbf{v}_{Cbtr,mea}$ can be estimated from the cavity probe signal measurement $\mathbf{v}_{C,mea}$. Since $\mathbf{v}_{Cbtr,mea}$ is derived from $\mathbf{v}_{C,mea}$, the beam phase can be calculated from the relative phases between them. If the cavity parameters ($\omega_{1/2}$ and R_L) and the average beam current (\mathbf{i}_{b0l}) are known, the beam-induced transient in the physical unit, \mathbf{v}_{Cbtr} , can be calculated with (9.8). Therefore, the voltage calibration factor can be calculated as

$$F_V = |\mathbf{v}_{Cbtr}| / |\mathbf{v}_{Cbtr,mea}|. \quad (9.15)$$

From (9.12), we have $\angle \mathbf{v}_{C,cal} = \Phi_{off} + \angle \mathbf{v}_{C,mea}$. As mentioned before, the goal of the beam phase calibration is $\angle \mathbf{v}_{C,cal} = \varphi_b$, where φ_b is the beam phase, which can be calculated as follows according to (3.33):

$$\varphi_b(t_{be}) = \pi + \angle \mathbf{v}_{C,mea}(t_{be}) - \angle \mathbf{v}_{Cbtr,mea}. \quad (9.16)$$

Here we used the result of (9.8), stating that the phase of \mathbf{v}_{Cbtr} is the same as that of \mathbf{i}_{b0} . Note that we have used the values of φ_b and $\mathbf{v}_{C,mea}$ at the end of the bunch train $t = t_{be}$. This is because the beam-induced transient is defined as the beam-induced cavity voltage at t_{be} . With the discussion above, we can derive the phase calibration as

$$\Phi_{off} = \pi - \angle \mathbf{v}_{Cbtr,mea}. \quad (9.17)$$

By adding Φ_{off} to the phase of $\mathbf{v}_{C,mea}$, we also obtain the “beam phase” of the cavity voltage when there is no beam present. This “beam phase” implies that if a bunch with the same phase (relative to a stable RF reference) as the bunch train used for the calibration is injected into the cavity, it will experience the “beam phase” at the bunch injection time.

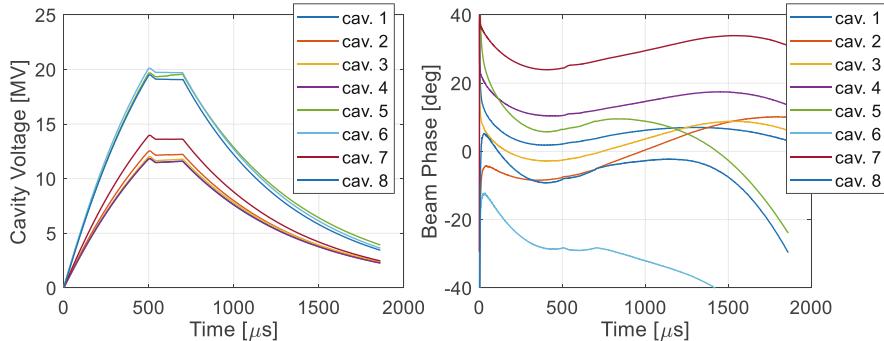


Fig. 9.12 Amplitude and phase waveforms of calibrated accelerating voltage and beam phase

The calibration (9.15)–(9.17) is applied to the data of Fig. 9.8. The resulting accelerating voltage and beam phase waveforms are shown in Fig. 9.12. When performing the accelerating voltage calibration, we have used the following parameters for all the eight cavities: $\omega_{1/2} = 2\pi \times 216.7$ rad/s, $r/Q = 1036 \Omega$ and $Q_L = 3e6$. The accelerating voltages in some cavities are lower because they have lower quench limits in gradient (i.e., when the cavity gradient is above the limit, the cavity has a large probability of quenching). Their drive powers are lowered so that all cavities can be operated at their gradient limits to maximize the acceleration efficiency. Another problem is the large beam phase deviations in different cavities, which implies that the cavity input phases should be optimized to improve the acceleration efficiency. We typically expect the same beam phase in all cavities of an RF station.

9.3.2.3 Accelerating Voltage and Beam Phase Calibration with Beam Energy

In practice, we often need to know the integrated (vector-sum) accelerating voltage and beam phase of an RF station with multiple cavities. As indicated by (9.5), one approach is to calibrate the individual cavities and calculate their vector sum. Comparing (9.5) to the calibrated cavity voltage of an individual cavity (9.12), we have $\mathbf{v}_{C,cal}^{(n)} = V_{acc}^{(n)} e^{j\varphi_b^{(n)}}$ for the n th cavity. Alternatively, we may calibrate the vector-sum accelerating voltage and beam phase by changing the RF phase and measuring the beam energy changes (Geng 2018). This method is also applicable for the RF stations without measurable beam-induced transients.

The beam energy is typically measured by a spectrometer consisting of a dipole magnet and a beam position monitor (BPM). See Fig. 9.13. A beam energy variation causes a different bending angle in the dipole magnet, resulting in a beam position change at the BPM. The relationship between the beam position x and the beam energy change ΔE is described as

$$\Delta E/E_0 = x/\eta. \quad (9.18)$$

Fig. 9.13 Beam energy measurement with a spectrometer

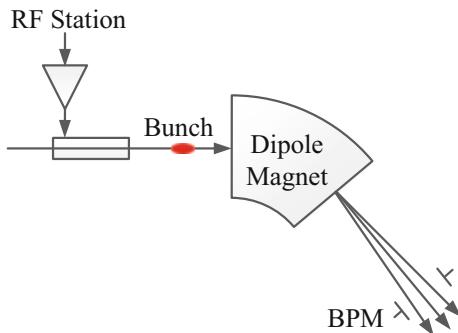
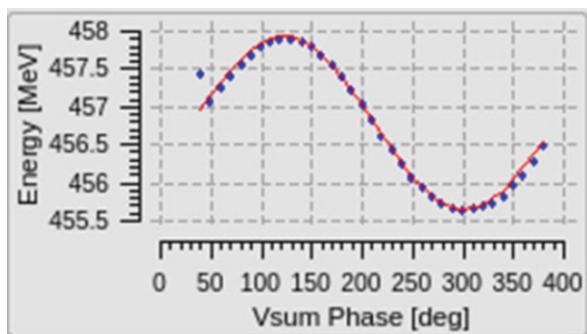


Fig. 9.14 Vector-sum phase scan and the resulting beam energy



Here E_0 is the energy of the beam when it passes through the center ($x = 0$) of the BPM, and η is the dispersion of the spectrometer.

To calibrate the vector-sum accelerating voltage and beam phase, we scan the RF station phase and measure the downstream beam energy simultaneously. The phase scan range is determined by the beam loss. Typically, a full-cycle (360°) scan is not feasible due to the beam loss caused by the massive beam energy change. If a full-cycle phase scan is required, such as when applying the RF station on beam for the first time, the accelerating voltage can be reduced to avoid beam loss. Figure 9.14 shows a full-cycle phase scan of an RF station of SwissFEL with a reduced accelerating voltage of around 1.2 MV.

Assume that a vector-sum phasor, $\mathbf{v}_{vs,mea}$, has been obtained to represent the integrated RF field experienced by the beam when it passes through the cavities of the RF station. Here $\mathbf{v}_{vs,mea}$ is a raw measurement that is not calibrated. Similar to (9.12), we introduce a scale factor $F_{V,vs}$ and a phase offset $\Phi_{off,vs}$ to derive the calibrated vector-sum phasor representing the integrated accelerating voltage and beam phase:

$$\mathbf{v}_{vs,cal} := V_{acc,vs} e^{j\phi_{b,vs}} = F_{V,vs} e^{j\Phi_{off,vs}} \mathbf{v}_{vs,mea}. \quad (9.19)$$

The goal here is to determine the values of $F_{V,vs}$ and $\Phi_{off,vs}$ from the phase scan.

The relationship between the beam energy and the RF phase is given by

$$\begin{aligned}\Delta E &= V_{acc,vs} \cos \varphi_{b,vs} \\ &= V_{acc,vs} \cos (\varphi_{vs,mea} + \Phi_{off,vs}),\end{aligned}\quad (9.20)$$

where $\varphi_{vs,mea}$ is the phase of $\mathbf{v}_{vs,mea}$. Note that the vector-sum beam phase $\varphi_{b,vs} = \varphi_{vs,mea} + \Phi_{off,vs}$. For the k th point of the phase scan, (9.20) can be rewritten as

$$\Delta E(k) = a \cos (\varphi_{vs,mea}(k)) + b \sin (\varphi_{vs,mea}(k)), \quad (9.21)$$

where

$$a = V_{acc,vs} \cos \Phi_{off,vs}, b = -V_{acc,vs} \sin \Phi_{off,vs}. \quad (9.22)$$

The parameters a and b can be calculated with the least-square algorithm, then

$$\begin{aligned}F_{V,vs} &= V_{acc,vs}/|\mathbf{v}_{vs,mea}| = \sqrt{a^2 + b^2}/|\mathbf{v}_{vs,mea}|, \\ \Phi_{off,vs} &= -\arctan(b/a).\end{aligned}\quad (9.23)$$

Note that the beam phase describes the relative timing between the RF field and the beam arrival time. If the beam arrival time is changed, the phase calibration should be updated correspondingly, usually done by simply repeating the above calibration procedure. However, if the changes of beam arrival time can be predicted accurately, such as when adjusting a magnetic chicane, the phase calibration can be updated by correcting the relative timing changes:

$$\Phi_{off,vs} = \Phi_{off0,vs} + 2\pi f_0 \Delta t_f. \quad (9.24)$$

Here $\Phi_{off0,vs}$ is the original phase calibration, f_0 is the RF frequency and Δt_f is the beam flight time change. The correction (9.24) should also be applied to the beam phase calibration of individual cavities.

9.3.3 Cavity Input Power and Phase Adjustment

The power and phase balances of the cavity inputs are essential for maximizing the beam acceleration efficiency for an RF station consisting of multiple cavities.

The power balancing is performed by adjusting the relative power of the cavity inputs so that each cavity can operate at its maximum gradient. For example, a cavity may start to breakdown or quench (for superconducting cavities) at a lower gradient than the nominal gradients of other cavities. Then we may reduce the input power of this cavity so that it does not limit the reachable gradients of other cavities.

The phases are balanced by adjusting the beam phases in different cavities until they are the same. It is important to provide continuous acceleration for the beam. If the beam-induced transients are available, the calibration results from Sect. 9.3.2.2 can be used to balance the phases. Let us analyze the example of Fig. 9.12. If the desired beam phase is zero, we must shift the input phases of the cavity 6 and 7 by about 30° and -30° , respectively. These two cavities have the most significant phase errors. The input phases of other cavities should be adjusted correspondingly.

If the beam-induced transients are not measurable, the power and phase balancing should be done when installing the cavities and the power transmission lines (e.g., waveguides). The attenuation and phase shift of each RF component of the power transmission lines can be measured offline with a network analyzer. Then the input power and phase of each cavity can be calculated based on the model. When the beam is available, a fine-tuning of the cavity drive power and phase is possible by observing the downstream beam energy. For example, with balanced phases in all cavities, the energy gain of the RF station is maximized.

9.3.4 Vector-Sum Calibration

Section 9.3.2 provides an approach to derive the calibrated vector sum using the calibrated cavity voltages of individual cavities. The procedure is as follows

$$\text{probe signal} \xrightarrow{\text{meas.}} \mathbf{v}_{C,mea} \xrightarrow{\text{calib.}} \mathbf{v}_{C,cal} \xrightarrow{\text{sum}} \mathbf{v}_{vs,cal} = \sum_{n=1}^N \mathbf{v}_{C,cal}^{(n)}. \quad (9.25)$$

That is, we calibrate the cavity voltage of each cavity and sum them up to determine the calibrated vector-sum phasor. The last equality above is equivalent to (9.5). Alternatively, we may first calculate the vector sum phasor $\mathbf{v}_{vs,mea}$ using the raw measurements of the cavity probe signals. Then $\mathbf{v}_{vs,cal}$ can be calibrated following the method described in Sect. 9.3.2.3. This approach corresponds to a procedure like

$$\text{probe signal} \xrightarrow{\text{meas.}} \mathbf{v}_{C,mea} \xrightarrow{\text{vector-sum}} \mathbf{v}_{vs,mea} = \sum_{n=1}^N \mathbf{C}_{vs}^{(n)} \mathbf{v}_{C,mea}^{(n)} \xrightarrow{\text{calib.}} \mathbf{v}_{vs,cal}. \quad (9.26)$$

In this section, we will introduce the algorithm to calibrate the vector sum, which determines the complex coefficient \mathbf{C}_{vs} for each cavity.

9.3.4.1 Vector-Sum Calibration Algorithm

The cavity probe signal is picked up via an RF cable and detected by an RF detector, both with unknown attenuation (or gains) and phase shifts. Therefore, the relative amplitudes and phases of the measurements ($\mathbf{v}_{C,mea}$) do not reflect the relative accelerating voltages and beam phases in different cavities. The vector-sum calibration is needed to scale and rotate the $\mathbf{v}_{C,mea}$ of each cavity to compensate for the unknown attenuation (or gain) and phase shift. As shown in (9.26), the vector-sum calibration coefficient is denoted as a complex number \mathbf{C}_{vs} .

Since the vector-sum calibration corrects only the relative amplitudes and phases of different cavities, we can use the first cavity as a reference. That is, we convert the measurements of all other cavities to the same reference plane as the measurement of the first cavity.

If the beam-induced transient is available, it can be used as a reference for the vector-sum calibration. To simplify the discussion, we assume that there is no beam loss when measuring the beam-induced transient. We also assume that the parameters ($\omega_{1/2}$ and R_L) of each cavity are known. Therefore, the beam-induced transient in each cavity can be calculated as

$$\mathbf{v}_{Cbtr}^{(n)} = 2\omega_{1/2}^{(n)} R_L^{(n)} \mathbf{i}_{b0} \Delta T, n = 1, 2, \dots, N. \quad (9.27)$$

Here N is the number of cavities and the superscript $^{(n)}$ denotes the n th cavity. In Sect. 9.3.1.2, we have derived the beam-induced transient ($\mathbf{v}_{Cbtr,mea}$) from $\mathbf{v}_{C,mea}$. Therefore, $\mathbf{v}_{Cbtr,mea}$ contains the same unknown attenuation (gain) and phase shift as $\mathbf{v}_{C,mea}$. To calibrate $\mathbf{v}_{Cbtr,mea}$ to reflect the same relative amplitude and phase as the calculated beam-induced transient (9.27), the coefficient \mathbf{C}_{vs} should satisfy

$$\frac{\mathbf{v}_{Cbtr}^{(n)}}{\mathbf{v}_{Cbtr}^{(1)}} = \mathbf{C}_{vs}^{(n)} \cdot \frac{\mathbf{v}_{Cbtr,mea}^{(n)}}{\mathbf{v}_{Cbtr,mea}^{(1)}}, n = 1, 2, \dots, N. \quad (9.28)$$

Note that $\mathbf{C}_{vs}^{(1)} = 1$ because we have used the first cavity as a reference. The coefficient \mathbf{C}_{vs} should be multiplied to the waveform of $\mathbf{v}_{C,mea}$ and the results should be summed up to determine the waveform of the vector-sum, as illustrated in (9.26).

The vector-sum calibration is applied to the data of Fig. 9.8 using the derived beam-induced transients in Fig. 9.11. For simplicity, we assume all the cavities have the same $\omega_{1/2}$ and R_L . The resulting waveforms of the calibrated amplitude and phase of each cavity and of the vector sum are depicted in Fig. 9.15. Note that the vector-sum amplitude has been scaled by 1/8 to fit in the plot. We can compare the waveforms in Fig. 9.8 ($\mathbf{v}_{C,mea}$, not calibrated), Fig. 9.15 ($\mathbf{C}_{vs}\mathbf{v}_{C,mea}$, vector-sum calibrated) and Fig. 9.12 ($\mathbf{v}_{C,cal}$, accelerating voltage and beam phase calibrated). It can be seen that the relative amplitudes and phases of different cavities in Fig. 9.15 have reflected the actual relative accelerating voltages and beam phases, as in Fig. 9.12.

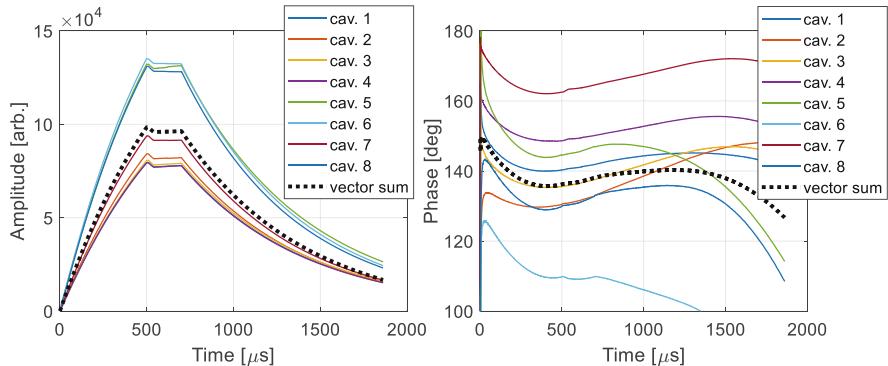


Fig. 9.15 Amplitude and phase waveforms with vector-sum calibration

The vector-sum calibration algorithm above relies on the measurements of the beam-induced transients. If the beam-induced transients are not measurable, like they are too small and below the RF detector resolution, the algorithm above is not applicable. Instead, a rough calibration can be performed by estimating the accelerating voltage and beam phase in each cavity using the cavity model. For example, assume we have two cavities (*A* and *B*), and we want to calibrate the vector sum using the cavity *A* as a reference. Assume that the accelerating voltage of cavity *B* is 90% of cavity *A*, and the beam phase of cavity *B* is 10° off from cavity *A*. We measure the two cavity probe signals and obtain the raw $\mathbf{v}_{C,mea}^{(A)}$ and $\mathbf{v}_{C,mea}^{(B)}$. Then, the vector-sum calibration coefficient for the cavity *B* can be calculated as $\mathbf{C}_{vs}^{(B)} = 0.9e^{j10\pi/180}\mathbf{v}_{C,mea}^{(A)}/\mathbf{v}_{C,mea}^{(B)}$. This method has been widely used in normal-conducting Linacs (e.g., SwissFEL), where one klystron drives multiple traveling-wave structures.

9.3.4.2 Vector-Sum Calibration Error

A feedback loop is often used in an RF controller to regulate the vector-sum by controlling the klystron input. Since a single klystron drives multiple cavities, the disturbance in one particular cavity will cause changes in the RF drives of all cavities due to the feedback. The goal of the RF controller is to stabilize the measured vector sum. However, if the vector-sum calibration has errors, the integrated RF field experienced by the beam may be slightly different from the measured vector sum. That is, even if the RF controller stabilizes the measured vector sum perfectly, the beam may still experience RF field errors in the presence of microphonics (Schilcher 1998). Microphonics are random and are uncorrelated in the different cavities. The vector-sum control using a single actuator (klystron) cannot compensate for such cavity-dependent disturbances completely.

The effects of vector-sum calibration errors are usually studied with numerical simulations. In the simulation, we assume that the vector-sum calibration errors and the microphonics are Gaussian-distributed random variables. To simplify the discussion, we denote the vector sum experienced by the beam as $\mathbf{v}_{vs,real}$ and the measured vector sum as $\mathbf{v}_{vs,mea}$. At the beginning of the simulation, we assume that the beam is correctly accelerated with no microphonics present, and the two vector-sum values are $\mathbf{v}_{vs,real0}$ and $\mathbf{v}_{vs,mea0}$, respectively. They are used as the zero-microphonics references. Due to the calibration errors, these two values are not equal. Then, for different combinations of vector-sum calibration errors and microphonics, we calculate the new vector-sum values $\mathbf{v}_{vs,real1}$ and $\mathbf{v}_{vs,mea1}$. The RF controller always corrects the errors in the measured vector sum, so it generates a correction of $\mathbf{v}_{corr} = \mathbf{v}_{vs,mea0} - \mathbf{v}_{vs,mea1}$. This correction will also affect the vector sum experienced by the beam that changes to $\mathbf{v}'_{vs,real1} = \mathbf{v}_{vs,real1} + \mathbf{v}_{corr}$, resulting in an error of $\Delta\mathbf{v}_{vs,real} = \mathbf{v}_{vs,real0} - \mathbf{v}'_{vs,real1}$.

Figure 9.16 shows the simulation results of the vector-sum control errors of 16 cavities for different scenarios. We assume a cavity half-bandwidth of 216 Hz and a cavity detuning fluctuation (microphonics) of 10 Hz RMS in the simulation. To simplify the description, we denote the vector-sum phase and amplitude calibration errors as $\Delta\varphi_{vscalib}$ and $\Delta\alpha_{vscalib}$, respectively. They are the phase and relative amplitude errors of the complex calibration coefficient \mathbf{C}_{vs} . The results are explained as follows:

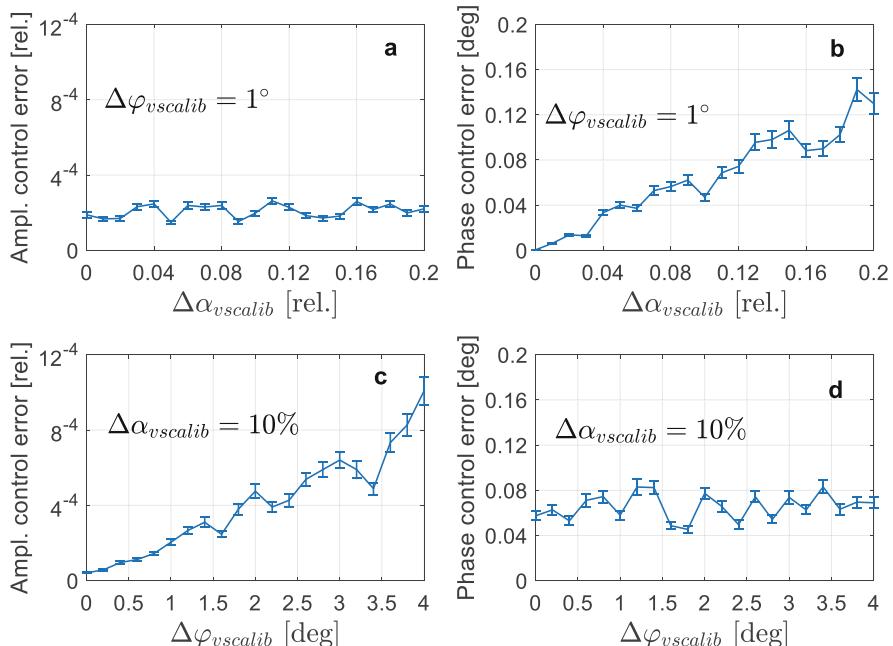


Fig. 9.16 Simulation of vector-sum control errors with different vector-sum calibration errors

- *Plot a:* $\Delta\varphi_{vscalib}$ is of 1° RMS; the vector-sum amplitude control error (relative RMS) as a function of $\Delta\alpha_{vscalib}$ from 0 to 20% RMS.
- *Plot b:* $\Delta\varphi_{vscalib}$ is of 1° RMS; the vector-sum phase control error ($^\circ$ RMS) as a function of $\Delta\alpha_{vscalib}$ from 0 to 20% RMS.
- *Plot c:* $\Delta\alpha_{vscalib}$ is of 10% RMS; the vector-sum amplitude control error (relative RMS) as a function of $\Delta\varphi_{vscalib}$ from 0° to 4° RMS.
- *Plot d:* $\Delta\alpha_{vscalib}$ is of 10% RMS; the vector-sum phase control error ($^\circ$ RMS) as a function of $\Delta\varphi_{vscalib}$ from 0° to 4° RMS.

The simulation shows that the phase control error is more sensitive to $\Delta\alpha_{vscalib}$, whereas the amplitude control error is more sensitive to $\Delta\varphi_{vscalib}$. To improve the vector-sum control performance, we should either reduce the vector-sum calibration errors or mitigate the microphonics (e.g., implementing active control of the cavity resonance frequency).

9.3.5 Cavity Forward and Reflected Signals Calibration

Using the calibrated cavity voltage, the RF powers fed into the cavity and reflected from the cavity can also be estimated. To do this, we need to convert the measurements of the forward and reflected signals to the same reference plane as the cavity probe signal.

The forward and reflected signals are picked up with directional couplers installed upstream of the cavity input coupler. The measured phasors of the forward and reflected signals are denoted as $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$, respectively. See the plot **b** and **c** of Fig. 9.17. Note that the measurements contain unknown attenuation (gains) or

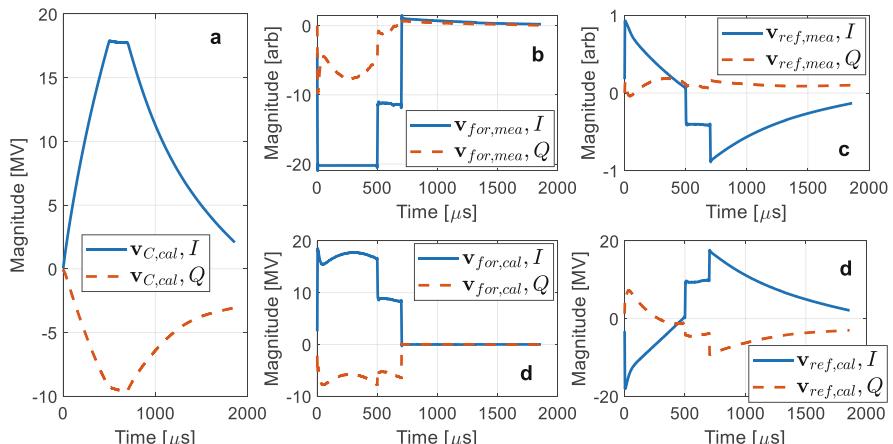


Fig. 9.17 Calibration of the forward and reflected signals. The waveform of each signal is in I/Q format representing the real and imaginary parts of the phasor

phase shifts from the signal pick-up and detection channels. In this section, we will calibrate $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$ using the calibrated cavity voltage ($\mathbf{v}_{C,cal}$, see Fig. 9.17a) as a reference. Here we assume that the relative group delays between different signals have been compensated. The waveforms are shifted to align in time, and we define the RF pulse starts at $t = 0$.

The directional couplers picking up the forward and reflected signals have limited directivities resulting in crosstalk between the two signals. From the viewpoint of the RF power source, the cavity represents the termination of the power transmission line. Considering the transmission line termination condition, we obtain

$$\begin{aligned}\mathbf{v}_{C,cal} &= \mathbf{v}_{for,cal} + \mathbf{v}_{ref,cal}, \\ \mathbf{v}_{for,cal} &= \mathbf{a}\mathbf{v}_{for,mea} + \mathbf{b}\mathbf{v}_{ref,mea}, \\ \mathbf{v}_{ref,cal} &= \mathbf{c}\mathbf{v}_{for,mea} + \mathbf{d}\mathbf{v}_{ref,mea}.\end{aligned}\quad (9.29)$$

Here $\mathbf{v}_{for,cal}$ and $\mathbf{v}_{ref,cal}$ are the calibrated forward and reflected signals, and the complex numbers \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{d} are the calibration coefficients. From (9.29), we have

$$\begin{aligned}\mathbf{v}_{C,cal} &= \mathbf{m}\mathbf{v}_{for,mea} + \mathbf{n}\mathbf{v}_{ref,mea}, \text{ where} \\ \mathbf{m} &:= \mathbf{a} + \mathbf{c}, \mathbf{n} := \mathbf{b} + \mathbf{d}.\end{aligned}\quad (9.30)$$

Then \mathbf{m} and \mathbf{n} can be calculated with linear fitting using the waveforms of $\mathbf{v}_{C,cal}$, $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$. During the cavity field decay after the end of the RF drive pulse (e.g., after around 700 μ s on the waveforms in Fig. 9.17), the following relation exists:

$$\begin{aligned}(\mathbf{a}\mathbf{v}_{for,mea} + \mathbf{b}\mathbf{v}_{ref,mea})|_{decay} &= 0, \text{ then we define} \\ \mathbf{z} &:= \frac{\mathbf{b}}{\mathbf{a}} = \text{mean}\left(-\frac{\mathbf{v}_{for,mea}}{\mathbf{v}_{ref,mea}}\Big|_{decay}\right).\end{aligned}\quad (9.31)$$

We can calculate \mathbf{z} with the values of $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$ during the decay stage. Substitute the definition of \mathbf{z} to (9.29), we obtain

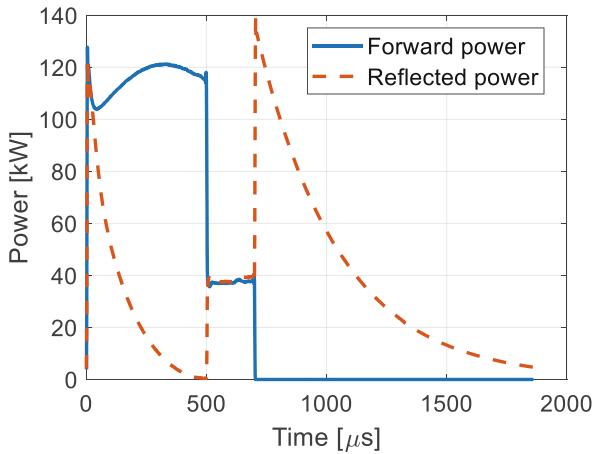
$$\mathbf{v}_{for,cal} = \mathbf{a}(\mathbf{v}_{for,mea} + \mathbf{z}\mathbf{v}_{ref,mea}).\quad (9.32)$$

If the cavity is only driven by the forward RF power (without beam), its equation can be written as

$$\dot{\mathbf{v}}_{C,cal} + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_{C,cal} = \frac{2\beta}{\beta + 1}\omega_{1/2}\mathbf{v}_{for,cal},\quad (9.33)$$

where $\omega_{1/2}$ is the half bandwidth, $\Delta\omega$ is the detuning, and β is the input coupling factor of the cavity. Near the starting time ($t = 0$) of the RF pulse, the cavity voltage

Fig. 9.18 Powers of the forward and reflected signals



is very small, but its derivative is large. Therefore, we can approximate the above equation around the zero time as

$$\begin{aligned} \dot{\mathbf{v}}_{C,cal}|_{t=0} &\approx \frac{2\beta}{\beta+1} \omega_{1/2} \mathbf{v}_{for,cal}|_{t=0} \\ &= \frac{2\beta}{\beta+1} \omega_{1/2} \mathbf{a} \left(\mathbf{v}_{for,mea}|_{t=0} + \mathbf{z} \mathbf{v}_{ref,mea}|_{t=0} \right). \end{aligned} \quad (9.34)$$

The value of $\omega_{1/2}$ can be calculated from the RF pulse decay stage. See Sect. 9.4.3.2. The derivative term in (9.34) can be calculated from the waveform of $\mathbf{v}_{C,cal}$ around $t = 0$. The determination of β can follow the method discussed in Sect. 9.4.1. Specifically, superconducting cavities have $\beta \gg 1$, so $2\beta/(\beta+1) \approx 2$. Since \mathbf{z} has been obtained from (9.31), we can calculate the value of \mathbf{a} using (9.34). Furthermore, we calculate the value of $\mathbf{b} = \mathbf{za}$ and derive the values of \mathbf{c} and \mathbf{d} according to the results of \mathbf{m} and \mathbf{n} obtained from (9.30). The calibrated forward and reflected signals are illustrated in plots **d** and **e** in Fig. 9.17.

Figure 9.17 shows that the calibrated forward and reflected signals also have the physical unit MV when using the calibrated cavity voltage as reference. Therefore, we can calculate the forward and reflected powers as

$$P_{for} = \frac{|\mathbf{v}_{for,cal}|^2}{2R_L}, \quad P_{ref} = \frac{|\mathbf{v}_{ref,cal}|^2}{2R_L}, \quad (9.35)$$

where R_L is the loaded resistance of the cavity. The corresponding powers of the forward and reflected signals of Fig. 9.17d and e are depicted in Fig. 9.18.

9.3.6 RF Signal Power Calibration

Except for the cavity forward and reflected signals, we often need to know the power of other RF signals, such as the klystron input and output. Therefore, the power calibration for the RF signal measurement is required.

Figure 9.19 is a typical configuration to pick up and measure an RF signal. A directional coupler picks up the RF signal from a high-power RF component (e.g., waveguide). The RF signal is transmitted via RF cables, and its level is optimized with attenuators at the input of the RF detector. The RF detector yields an uncalibrated amplitude and phase of the RF signal. The amplitude will be calibrated to estimate the RF power at the source of the signal.

The calibration is carried out with the following steps:

1. LLRF power calibration

At this step, we find the relation between the RF detector input power (P_{det}) and the measured amplitude (A_{RF}). The calibration can be done by inputting an RF signal with a known P_{det} to the RF detector. With the resulting A_{RF} , we can find a coefficient C_{det} satisfying

$$P_{det} = C_{det}A_{RF}^2. \quad (9.36)$$

2. Path attenuation calibration

At this step, the total attenuation from the directional coupler output to the RF detector input needs to be calibrated. To do this, we disconnect the RF cable from the directional coupler and inject there an RF signal with a known power P_{cpl} . Then we disconnect the RF detector input and measure the output power (P_{det}) of the attenuators. From this measurement, the path attenuation factor C_{path} can be calculated using the following relation

$$P_{cpl} = C_{path}P_{det}. \quad (9.37)$$

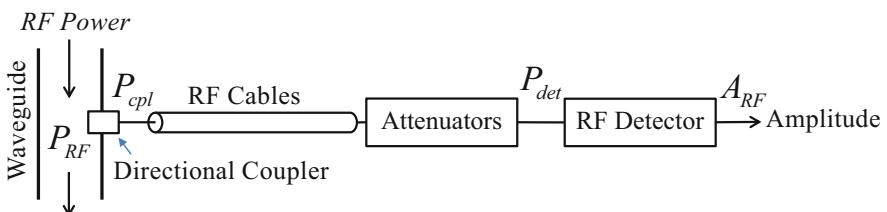


Fig. 9.19 Power calibration of an RF signal

3. Coupling factor calibration

At this step, we measure the coupling factor of the directional coupler. It should be measured with a network analyzer in the laboratory before installing the directional coupler. The coupling factor determines a coefficient C_{cpl} for

$$P_{RF} = C_{cpl} P_{cpl}. \quad (9.38)$$

The power P_{RF} at the high-power RF component can be then calculated as

$$P_{RF} = C_{cpl} C_{path} C_{det} A_{RF}^2. \quad (9.39)$$

With the calibration (9.39), the RF power can be calculated from the RF signal amplitude measured by the RF detector. A new power calibration is required if the RF measurement path is changed. The power calibration error of a few percent can be achieved. The error comes mainly from the coupling factor uncertainty and the thermal drifts in cables and attenuators.

9.4 RF System Identification

In this section, we will identify the RF system characteristics based on the RF signal measurements. The characteristics include the cavity input coupling factor, loaded quality factor, and detuning, providing insights into the RF system (Brandt 2007; Rybaniec et al. 2014; Jonsson 2015; Bellandi et al. 2021). These characteristics are used as inputs for the exception detection and parameter optimizations. For example, a sudden drop in the loaded quality factor of a superconducting cavity typically indicates a quench; the time-varying detuning in an RF pulse can be used to control the piezo tuner for feedforward compensation (see Sect. 4.6.3.3). The algorithms introduced here are based on the measurements from an RF station during operation. The cavity parameter measurements in the laboratory (e.g., with a network analyzer) will not be discussed.

9.4.1 Cavity Input Coupling Factor Identification

The cavity input coupling factor β describes the ratio between the RF power emitted from the cavity via the input coupler and the RF power dissipated in the cavity wall. The power ratio should be calculated after the cavity reaches the steady state. To use the cavity eq. (3.31) with voltage drives, we need to know the value of β , which can be obtained by one of the following methods:

- (a) For some normal-conducting cavities, β is a fixed value as a design parameter. In this case, we determine β offline before the cavity is installed.
- (b) For superconducting cavities, $\beta \gg 1$. Therefore, in the cavity equation, we can make the approximation as $\beta/(\beta + 1) \approx 1$.
- (c) We may estimate β with the measurements of the forward and reflected powers of the cavity in the steady state.
- (d) With the measured steady-state phasors of the cavity forward and reflected signals, β can be calculated from their ratios under different detuning $\Delta\omega$.

In this section, we will discuss the methods *c* and *d*, which are typically applied to normal-conducting cavities.

When a cavity reaches a steady state, the forward and reflected powers are described by (3.37) and (3.38). If there is no beam in the cavity and the cavity is well-tuned, which means, $i_{b0} = 0$ and $\Delta\omega = 0$, the power equations can be simplified to

$$P_{for} = \frac{\beta + 1}{\beta} \frac{v_{C0}^2}{8R_L}, P_{ref} = \frac{(\beta - 1)^2}{\beta(\beta + 1)} \frac{v_{C0}^2}{8R_L}. \quad (9.40)$$

Then we have

$$\alpha^2 := \frac{P_{ref}}{P_{for}} = \frac{(\beta - 1)^2}{(\beta + 1)^2}, \quad (9.41)$$

which results in

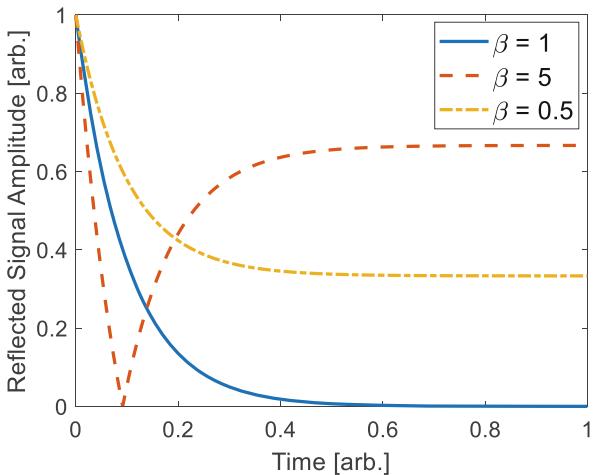
$$\beta = \frac{1 + \alpha}{1 - \alpha} \text{ or } \frac{1 - \alpha}{1 + \alpha}. \quad (9.42)$$

If we only use the steady-state forward and reflected powers, we cannot determine the exact value of β . To find out if β is larger or smaller than 1, the transient response of the cavity reflected signal can be observed. See Fig. 9.20. If we switch on a constant RF drive power at time zero, the amplitude of the cavity reflected signal follows different patterns for different β values. For example, for an overcoupled cavity with $\beta > 1$, the reflected signal has a clear undershoot.

Alternatively, β can also be determined by the phasors of the forward and reflected voltages denoted as \mathbf{v}_{for} and \mathbf{v}_{ref} , respectively (Kajfez 2003; Goryashko et al. 2015). Their ratio has been defined in (3.32) as a complex reflection factor of the cavity, which is rewritten as

$$\gamma_C := \frac{\mathbf{v}_{ref}}{\mathbf{v}_{for}} = \frac{2\omega_{1/2}}{\omega_{1/2} - j\Delta\omega} \frac{\beta}{\beta + 1} - 1. \quad (9.43)$$

Fig. 9.20 Transient responses of the reflected signal amplitude for different β values



From the discussion in Sect. 3.3.7, γ_C forms a circle on the complex plane when changing the cavity detuning. We are interested in two values of γ_C for $\Delta\omega = 0$ and $\Delta\omega = \infty$ given by

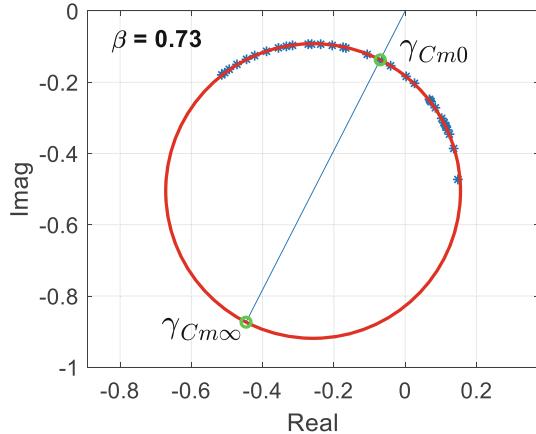
$$\gamma_{C0} := \gamma_C|_{\Delta\omega=0} = \frac{\beta - 1}{\beta + 1}, \quad \gamma_{C\infty} := \gamma_C|_{\Delta\omega=\infty} = -1. \quad (9.44)$$

In practice, the cavity forward and reflected signals are measured by RF detectors with unknown gains and phase shifts. Therefore, the ratio of the measured phasors of the forward and reflected signals, $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$, contains unknown complex coefficients compared to (9.43). We define the measured reflection factor as

$$\gamma_{Cm} := \frac{\mathbf{v}_{ref,mea}}{\mathbf{v}_{for,mea}} = \frac{\mathbf{b}\mathbf{v}_{ref}}{\mathbf{a}\mathbf{v}_{for}} = \mathbf{z}\gamma_C, \quad (9.45)$$

where **b** and **a** are unknown complex coefficients and their ratio are defined as $\mathbf{z} = \mathbf{b}/\mathbf{a}$. Here we assume that there is no crosstalk between the measurements of the forward and reflected signals. To avoid the influence of crosstalk, we should pick up the forward and reflected signals using a directional coupler with excellent directivity. Alternatively, if a circulator is installed between the RF power source and the cavity, we may consider picking up the forward signal at a location upstream the circulator and picking up the reflected signal from the load of the circulator. Equation (9.45) implies that when changing the cavity detuning, γ_{Cm} also forms a circle derived from the circle of γ_C . Figure 9.21 shows an example of the γ_{Cm} circle for the RF gun cavity of FLASH. On the γ_{Cm} circle, the points with the minimum and maximum magnitudes correspond to $\Delta\omega = 0$ and $\Delta\omega = \infty$, respectively. As shown in Fig. 9.21, we define these two points as

Fig. 9.21 Calibration of the coupling factor of the FLASH RF gun



$$\gamma_{Cm0} = \mathbf{z}\gamma_{C0} = \mathbf{z}\frac{\beta - 1}{\beta + 1}, \gamma_{Cm\infty} = \mathbf{z}\gamma_{C\infty} = -\mathbf{z}. \quad (9.46)$$

The unknown coefficient \mathbf{z} can be eliminated by calculating their ratio as

$$m := \frac{\gamma_{Cm0}}{\gamma_{Cm\infty}} = \frac{1 - \beta}{1 + \beta}. \quad (9.47)$$

Note that m is a real number because the phases of γ_{Cm0} and $\gamma_{Cm\infty}$ are identical. Then we obtain

$$\beta = \frac{1 - m}{1 + m}. \quad (9.48)$$

Therefore, if we can measure the circle of γ_{Cm} , the value of γ_{Cm0} and $\gamma_{Cm\infty}$ can be identified, from which we can calculate the value of β using (9.47) and (9.48).

To measure the γ_{Cm} circle, we need to detune the cavity. Due to the limited tuning range of the cavity tuner, it is impossible to detune the cavity more than several bandwidths. Therefore, only a part of the γ_{Cm} circle can be measured. The full circle can be derived by fitting the data in the complex plane. See Fig. 9.21. Some normal-conducting cavities do not have mechanical tuners equipped. These cavities can be detuned by varying the cooling water temperature. Another approach is to shift the frequency of the RF drive power using the method introduced in Sect. 4.6.1.1. It changes the detuning equivalently since $\Delta\omega$ is a relative value between the RF drive frequency and the cavity resonance frequency. Note that when changing the RF drive frequency, we introduce errors in the relative phases between the measured $\mathbf{v}_{for,mea}$ and $\mathbf{v}_{ref,mea}$. The errors are caused by the frequency-dependent phase shifts of the transmission lines between the forward and reflected signals' pickup points and the cavity. However, when the frequency change of the RF drive power is small compared to the RF operating frequency, these errors are negligible.

Figure 9.21 is measured at the FLASH RF gun. The cavity is detuned by changing the cooling water temperature, and the identified coupling factor is about 0.73.

The crosstalk between the measurements of the forward and reflected signals is not corrected in the above algorithm. Therefore, if there exists crosstalk, the calibration of β will have errors. For example, we assume that $\mathbf{v}_{ref,mea}$ contains a crosstalk from the forward signal. That is, $\mathbf{v}_{ref,mea} = \mathbf{b}\mathbf{v}_{ref} + \mathbf{c}\mathbf{v}_{for}$, the resulting γ_{Cm} circle will have the same shape but with an offset (\mathbf{c}/\mathbf{a}) compared to that without crosstalk. The offset will cause errors in identifying γ_{Cm0} and $\gamma_{Cm\infty}$, and lead to errors in calculating β . In the example of Fig. 9.21, if we assume that the crosstalk from the forward signal to the measurement of the reflected signal is $|\mathbf{c}|/|\mathbf{b}| = 0.1$ (i.e., 20 dB directivity), the resulting error in the value of β can be up to ± 0.14 .

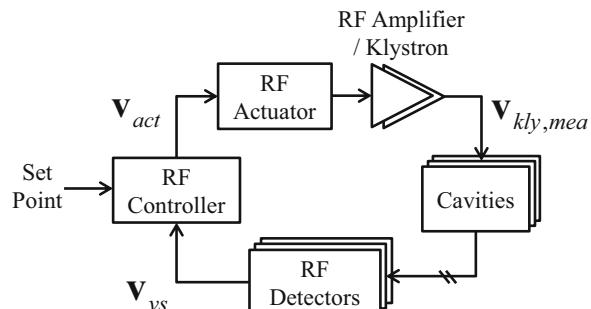
9.4.2 System Gain and System Phase Identification

The system gain and system phase play important roles in the RF control loop stability. The knowledge of system gain is required to determine the actual loop gain so that the feedback controller setting preserves the gain margin. As discussed in Sect. 4.2, the system phase must be in the correct range for the stability of an I/Q feedback loop. Here we will revisit this topic and introduce a method to measure the system gain and system phase in a typical RF system.

Figure 9.22 depicts a simplified RF control loop, where the *system* consists of all the blocks except for the RF controller. To simplify the discussion, we consider only the dynamics of the cavity when describing the system transfer function. The system input is the control signal \mathbf{v}_{act} produced by the RF controller, whereas the vector-sum phasor, \mathbf{v}_{vs} , is the system output, which can be viewed as the output of an equivalent single cavity. The input-output relationship of the system (without beam) is described as

$$\dot{\mathbf{v}}_{vs} + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_{vs} = g_{sys}e^{-j\varphi_{sys}}\omega_{1/2}\mathbf{v}_{act}, \quad (9.49)$$

Fig. 9.22 Simplified block diagram of an RF control loop



where $\omega_{1/2}$ and $\Delta\omega$ are the half-bandwidth and detuning of the equivalent cavity. Equation (9.49) defines the system gain and system phase. They are derived from the ratio between the steady-state output and input phasors of the system when $\Delta\omega = 0$. Note that φ_{sys} has a minus sign because it is a phase-lag. Typically, the waveform of \mathbf{v}_{act} can be directly read from the digital RF controller, and \mathbf{v}_{vs} is available from the measurement. Note that the \mathbf{v}_{vs} should be the vector-sum phasor used in the RF controller to compare to the setpoint. It can be either $\mathbf{v}_{vs,mea}$ or $\mathbf{v}_{vs,cal}$ (see Sect. 9.3.4), depending on the RF controller design. We often implement the feedback controller based on the raw measurement. That is, both the setpoint and the measurement are not calibrated to physical units. The physical-unit calibration is only used to translate the results to be displayed or convert the user desired accelerating voltage and beam phase to the setpoint of feedback control. In this case, the \mathbf{v}_{vs} corresponds to $\mathbf{v}_{vs,mea}$. Since the RF amplifier chain might be strongly nonlinear, we also measure the klystron output, $\mathbf{v}_{kly,mea}$, to facilitate the system gain and system phase identification. It helps to split the nonlinearity and the dynamics of the system, which leads to

$$\begin{aligned}\dot{\mathbf{v}}_{vs} + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_{vs} &= \omega_{1/2}\mathbf{C}_{kly}\mathbf{v}_{kly,mea}, \\ \mathbf{v}_{kly,mea} &= \mathbf{C}_{act}\mathbf{v}_{act}.\end{aligned}\quad (9.50)$$

The complex coefficient \mathbf{C}_{kly} converts the measurement of the klystron output signal to the same reference plane of \mathbf{v}_{vs} . The relationship between \mathbf{v}_{act} and $\mathbf{v}_{kly,mea}$ is described by a static complex coefficient \mathbf{C}_{act} , which depends on the magnitude of \mathbf{v}_{act} . Therefore, if we can estimate the values of \mathbf{C}_{kly} and \mathbf{C}_{act} , the system gain and system phase can be calculated using the relationship

$$g_{sys}e^{-j\varphi_{sys}} = \mathbf{C}_{kly}\mathbf{C}_{act}. \quad (9.51)$$

Here we assume that the relative group delays between different signals (\mathbf{v}_{act} , $\mathbf{v}_{kly,mea}$, and \mathbf{v}_{vs}) have been removed. We define the RF pulse starts at $t = 0$. The calculation of \mathbf{C}_{act} is straightforward since the waveforms of $\mathbf{v}_{kly,mea}$ and \mathbf{v}_{act} are known:

$$\mathbf{C}_{act}(t) = \mathbf{v}_{kly,mea}(t)/\mathbf{v}_{act}(t), t \in [0, T_{pul}]. \quad (9.52)$$

Here we define the RF pulse width as T_{pul} . Note that \mathbf{C}_{act} also forms a waveform, which should be calculated by the waveforms of $\mathbf{v}_{kly,mea}$ and \mathbf{v}_{act} .

The coefficient \mathbf{C}_{kly} is a constant, which can be estimated in the same way of calibrating the forward and reflected signals. Similar to (9.34), we have $\dot{\mathbf{v}}_{vs}|_{t \rightarrow 0} \approx \omega_{1/2}\mathbf{C}_{kly}\mathbf{v}_{kly,mea}|_{t \rightarrow 0}$, where $\omega_{1/2}$ is derived from the decay stage of the \mathbf{v}_{vs} pulse. Then \mathbf{C}_{kly} is calculated with the values of $\mathbf{v}_{kly,mea}$ and \mathbf{v}_{vs} at the RF pulse start

$$\mathbf{C}_{kly} \approx \frac{\dot{\mathbf{v}}_{vs}|_{t \rightarrow 0}}{\omega_{1/2}\mathbf{v}_{kly,mea}|_{t \rightarrow 0}}. \quad (9.53)$$

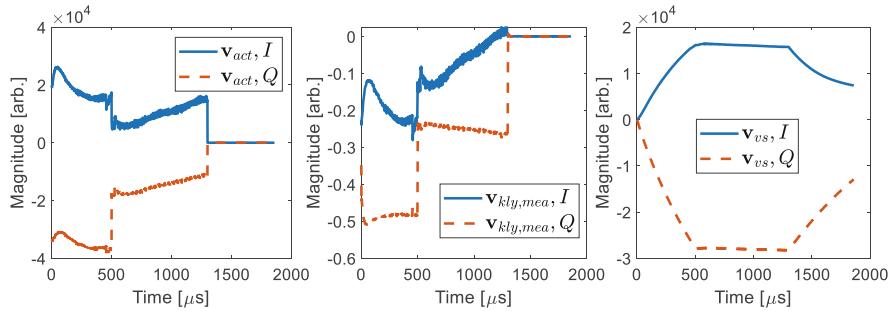


Fig. 9.23 Waveforms used to identify the system gain and system phase

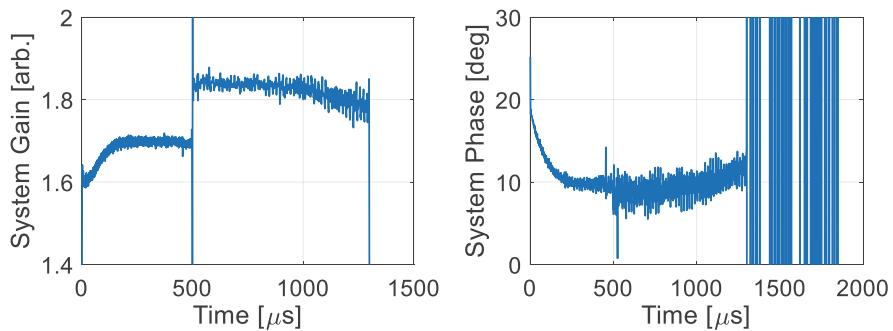


Fig. 9.24 Results of the system gain and system phase identification

With the results (9.52) and (9.53), we can derive the time-varying system gain and system phase according to (9.51).

Figure 9.23 shows an example of the waveforms of \mathbf{v}_{act} , $\mathbf{v}_{kly,mea}$, and \mathbf{v}_{vs} , all in the I/Q format. The identified system gain and system phase are illustrated in Fig. 9.24. The system gain has different values in the filling (0 to 500 μ s) and flattop (500 to 1300 μ s) stages of the RF pulse, caused by the amplifier chain nonlinearity. The identified system phase is about 10° during the RF pulse flattop. At the starting part of the pulse, the identified system phase changes significantly due to the large phase variation of \mathbf{v}_{act} caused by the feedback. The time-varying phase causes effective frequency changes, resulting in phase shift variations in the RF amplifier chain. When collecting the data of Fig. 9.23, the feedback loop is closed to regulate the vector sum. Operating the cavities in an open loop can improve the accuracy of the system gain and system phase identification.

In an RF controller, we often scale and rotate the input \mathbf{v}_{vs} (with a scale factor g_{ctrl} and a phase offset φ_{ctrl}) to normalize the system gain to 1 and the system phase to 0. That is, we require

$$g_{sys} g_{ctrl} = 1, -\varphi_{sys} + \varphi_{ctrl} = 0.$$

In this case, the loop gain and loop phase of the RF control loop (see Fig. 9.22) will be determined only by the dynamics of the cavity and the controller. Recall that the *loop gain* and *loop phase* are defined as the phase shift and gain of the entire loop, including all the blocks in Fig. 9.22, with the loop opened. Compared to the system gain and system phase concepts, the loop gain and loop phase include the system dynamics and are frequency dependent.

We notice that the system gain calculated with (9.51)–(9.53) is the *absolute gain*, as defined in Fig. 7.10. The system gain normalization above should refer to the *incremental gain* at the operating point. To calibrate the incremental system gain, we can collect two sets of data. One set is at the specified operating point consisting of \mathbf{v}_{act} , $\mathbf{v}_{kly,mea}$, and \mathbf{v}_{vs} , and the other is at a slightly shifted operating point described by $\mathbf{v}_{act} + \Delta\mathbf{v}_{act}$, $\mathbf{v}_{kly,mea} + \Delta\mathbf{v}_{kly,mea}$, and $\mathbf{v}_{vs} + \Delta\mathbf{v}_{vs}$. Then, (9.52) and (9.53) can be rewritten as

$$\begin{aligned}\mathbf{C}'_{act}(t) &= \Delta\mathbf{v}_{kly,mea}(t)/\Delta\mathbf{v}_{act}(t), t \in [0, T_{pul}], \\ \mathbf{C}'_{kly} &\approx \frac{\Delta\dot{\mathbf{v}}_{vs}|_{t \rightarrow 0}}{\omega_{1/2}\Delta\mathbf{v}_{kly,mea}|_{t \rightarrow 0}}.\end{aligned}\quad (9.54)$$

Then the incremental system gain and system phase (denoted as g'_{sys} and φ'_{sys}) can be calculated from

$$g'_{sys} e^{-j\varphi'_{sys}} = \mathbf{C}'_{kly} \mathbf{C}'_{act}. \quad (9.55)$$

Therefore, the vector scale and rotation block in the controller can be set as follows using the identified incremental system gain and system phase at the flattop:

$$g_{ctrl} = 1/g'_{sys}, \varphi_{ctrl} = \varphi'_{sys}. \quad (9.56)$$

In a practical RF controller, the feedback gain is typically optimized experimentally, as discussed in Sect. 9.2.3. Therefore, the system gain normalization in (9.56) is less critical compared to the system phase. A system phase near zero is important to keep the I/Q feedback loop stable. Typically, the system phases calculated by (9.55) and (9.51) are nearly the same; that is, $\varphi_{sys} \approx \varphi'_{sys}$. It implies that the incremental system gain and phase calibration (9.54)–(9.55) is often not needed. The simpler calibration (9.51)–(9.53) can be sufficient for most practical cases, especially when the RF driving chain linearization is implemented.

9.4.3 Cavity Parameters Identification

9.4.3.1 Basic Equations

The loaded quality factor (Q_L) and detuning ($\Delta\omega$) describe the dynamics of a cavity. The differential equation of the cavity is rewritten as follows

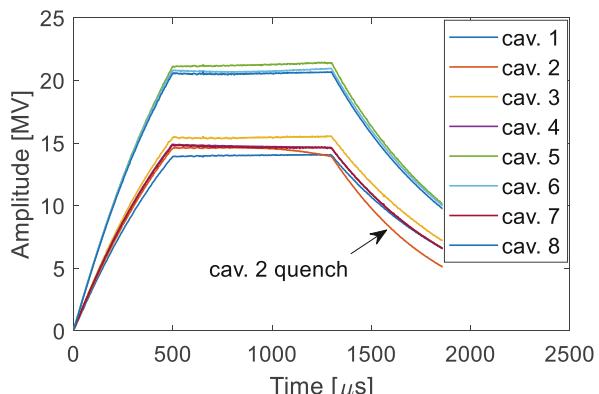
$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega)\mathbf{v}_C = \frac{2\beta}{\beta + 1}\omega_{1/2}\mathbf{v}_{for} + 2\omega_{1/2}R_L\mathbf{i}_{b0}. \quad (9.57)$$

The half-bandwidth is related to Q_L with $\omega_{1/2} = \omega_0/(2Q_L)$, where ω_0 is the resonance frequency of the cavity. From the measurements of the cavity forward, reflected and probe signals, the values of $\omega_{1/2}$ and $\Delta\omega$ can be calculated. The results are used to diagnose the cavity status or as the inputs for other applications (e.g., cavity frequency tuning based on $\Delta\omega$). For superconducting cavities, quench is a limiting factor for the maximum gradient. A cavity quench results in a higher RF power dissipation in the cavity walls and reduces the unloaded quality factor (Q_0). The relationship between Q_0 and Q_L is given by

$$Q_L^{-1} = Q_0^{-1} + Q_{ext}^{-1}, \quad (9.58)$$

where Q_{ext} is a constant once the geometries of the cavity and the input coupler are fixed. A quench typically causes a drop in the cavity voltage (see Fig. 9.25) and a reduction of Q_L . Note that the cavity voltage changes may also be induced by the RF drive power or detuning. Therefore, we use Q_L for the quench detection. We see from (9.57) that $\omega_{1/2}$ appears at both sides of the equation. On the right side, $\omega_{1/2}$ describes the coupling of the drive terms (\mathbf{v}_{for} and \mathbf{i}_{b0}). On the left side, it determines the cavity dynamics. As mentioned above, a quench changes only the cavity dynamics but not the coupling between the RF drive power or beam and the cavity. The cavity stored energy and the RF power coupled into and out of the cavity are only determined by the geometries of the cavity and the input coupler. That is, Q_{ext} is

Fig. 9.25 Cavity voltage drop in cav. 2 due to a quench



not affected by the cavity quench. To describe a quench in the cavity equation, we define the equivalent half-bandwidth and loaded resistance with respect to Q_{ext} as

$$\begin{aligned}\omega_{1/2}^{(ext)} &:= \frac{\omega_0}{2Q_{ext}} = \frac{\beta}{\beta + 1} \omega_{1/2}, \\ R_L^{(ext)} &:= \frac{1}{2} (r/Q) Q_{ext} = \frac{\beta + 1}{\beta} R_L.\end{aligned}\quad (9.59)$$

Note that the input coupling factor β is defined for the case when there is no quench.

Then the cavity equation can be rewritten as

$$\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega) \mathbf{v}_C = 2\omega_{1/2}^{(ext)} \mathbf{v}_{for} + 2\omega_{1/2}^{(ext)} R_L^{(ext)} \mathbf{i}_{b0}. \quad (9.60)$$

Therefore, to identify $\omega_{1/2}$ during the RF pulse, we can follow two steps. First, we calculate $\omega_{1/2}^{(ext)}$ and $R_L^{(ext)}$ following (9.59) when there is no quench. Then during the RF pulse, we solve the eq. (9.60) to determine the value of $\omega_{1/2}$, from which the Q_L can be calculated. A drop of Q_L is an indication for a quench.

9.4.3.2 Q_L and $\Delta\omega$ Identification at RF Pulse Decay

If a cavity operates in the pulsed mode, the cavity voltage decays after the RF drive power and the beam are turned off. During the RF pulse decay stage, the amplitude of the cavity voltage decreases exponentially, and the phase changes linearly with a slope equal to $\Delta\omega$. If we write the cavity voltage as $\mathbf{v}_C = v_C e^{j\varphi_C}$, the amplitude and phase during the decay stage can be written as

$$v_C(t) = v_{C0} e^{-\omega_{1/2} t}, \varphi_C(t) = \varphi_{C0} + \Delta\omega t, \quad (9.61)$$

where v_{C0} and φ_{C0} are the amplitude and phase at $t = 0$ when the field decay starts. Therefore, if the measurements of $v_C(t)$ and $\varphi_C(t)$ are available, we can calculate $\omega_{1/2}$ and $\Delta\omega$ by fitting the functions in (9.61). To simplify the exponential fitting, one can do a linear fit on the logarithm scale. We need to modify the expression of $v_C(t)$ to $\ln(v_C(t)) = \ln(v_{C0}) - \omega_{1/2} t$ before performing the linear fit. As an example, Fig. 9.26 shows the Q_L (derived from $\omega_{1/2}$) and $\Delta\omega$ identification of a TESLA cavity at the RF pulse decay starting at 1300 μ s. The fitted amplitude has been converted back to the exponential function. When there is no quench, we will use these results to define the $\omega_{1/2}^{(ext)}$ and $R_L^{(ext)}$ according to (9.59).

The resulting Q_L can be used for the quench detection. If a quench is detected from an RF pulse decay, the drive power of the next pulse can be stopped or reduced to protect the cavity. In most cases, this pulse-to-pulse quench detection and protection should be sufficient for a safe operation. However, if a serious quench happens, we must stop the RF drive power before the pulse is finished. Then we must

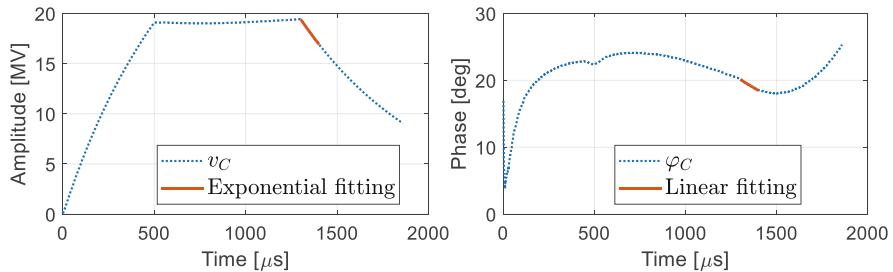
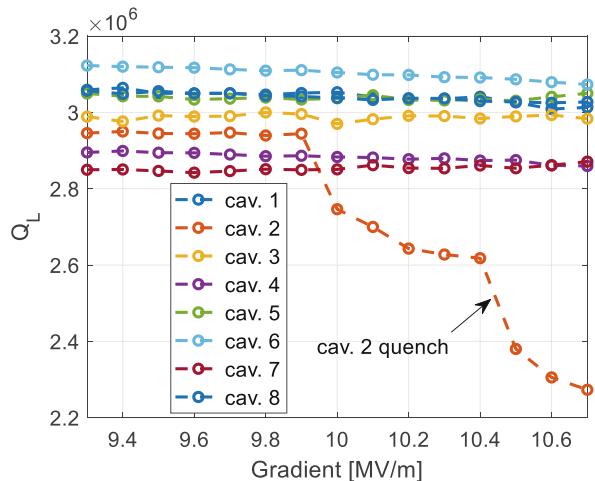


Fig. 9.26 Q_L ($2.97\text{e}6$) and $\Delta\omega$ (-45.6 Hz) identification at the RF pulse decay

Fig. 9.27 Identified Q_L at the RF pulse decay of eight cavities under different gradients. The gradient is determined during the flattop of the RF pulse, and Q_L is calculated for each pulse



identify the Q_L during the RF pulse in the presence of RF drive power or beam (see Sects. 9.4.3.3 and 9.4.3.4). Figure 9.27 illustrates the Q_L of eight TESLA cavities as functions of the cavity gradients. It can be seen that a quench happens in cavity 2 when the gradient reaches 10 MV/m.

For a cavity operating in the CW mode, the decay stage does not exist. To use the algorithm above, a small square-wave amplitude modulation can be introduced to the RF drive power, producing a small transient in the cavity voltage. Then we can calculate Q_L at the decay stage of the cavity voltage transient.

9.4.3.3 Q_L and $\Delta\omega$ Identification During RF Pulse

If the cavity is driven only by the RF power and there is NO beam, the Q_L and $\Delta\omega$ during the RF pulse can be calculated by solving (9.60) as

$$\omega_{1/2} - j\Delta\omega = \frac{2\omega_{1/2}^{(ext)} \mathbf{v}_{for} - \dot{\mathbf{v}}_C}{\mathbf{v}_C}. \quad (9.62)$$

Here $\omega_{1/2}^{(ext)}$ should be calculated from the cavity voltage decay when there is no quench. The \mathbf{v}_{for} waveform is the calibrated forward signal, as derived in Sect. 9.3.5. The term \mathbf{v}_C is either the calibrated cavity voltage waveform or the raw measurement of the cavity probe signal. The important point is that \mathbf{v}_C must be the reference when calibrating the forward signal for \mathbf{v}_{for} . Note that we need to calculate the first-order derivative of \mathbf{v}_C . If it is performed on the saved data after the RF pulse, we can use the central difference formula

$$\dot{\mathbf{v}}_C(k) = \frac{\mathbf{v}_C(k+1) - \mathbf{v}_C(k-1)}{2T_s}, \quad (9.63)$$

where k is the index of the samples and T_s is the time interval of sampling. Typically, the calculation of the derivative is sensitive to the noise of \mathbf{v}_C . This is because T_s is very short, during which the small \mathbf{v}_C transient is comparable to the noise. Furthermore, in (9.62), \mathbf{v}_C also appears at the denominator, so the calculation error can be large when the magnitude of \mathbf{v}_C is close to zero. Figure 9.28 illustrates the calculated $\omega_{1/2}$, Q_L and $\Delta\omega$ during the RF pulse for the same TESLA cavity as Fig. 9.26. It can be seen that the results are noisy even they are smoothed by a moving-average filter. The Q_L is a constant within the pulse when the cavity does not quench. The detuning varies with time due to the Lorentz force, which is dependent of the cavity voltage. In Sect. 4.6.1.3, another method has been introduced to derive \mathbf{v}_{for} from the forward signal without considering the crosstalk from the reflected signal. It works if the forward signal is picked up on the klystron side of the circulator where there is no reflected power.

To check the feasibility for quench detection using the intra-pulse Q_L identification, we apply the algorithm above to the cavity 2 of Fig. 9.27. The results for two gradients, 9.3 MV/m and 10.7 MV/m, are displayed in Fig. 9.29. It can be seen that the drop of Q_L is detectable within the flattop stage (from 500 to 1300 μ s). Therefore, a fast quench detection allows protecting the cavity by cutting the RF drive power immediately when the quench is detected.

9.4.3.4 Identification of Beam Drive

When there is no beam, we have identified the time-varying cavity parameters $\omega_{1/2}$ and $\Delta\omega$ using the measurements of the cavity probe, forward, and reflected signals. Since these three signals are not independent (i.e., $\mathbf{v}_C = \mathbf{v}_{for} + \mathbf{v}_{ref}$), it represents an underdetermined problem statement to calculate the time-varying $\omega_{1/2}$, $\Delta\omega$ and i_{b0} simultaneously. Therefore, if the beam is present, we cannot determine the time-varying $\omega_{1/2}$ and $\Delta\omega$ without knowing the beam current and beam phase. This may cause problems for the quench detection and the cavity detuning compensation.

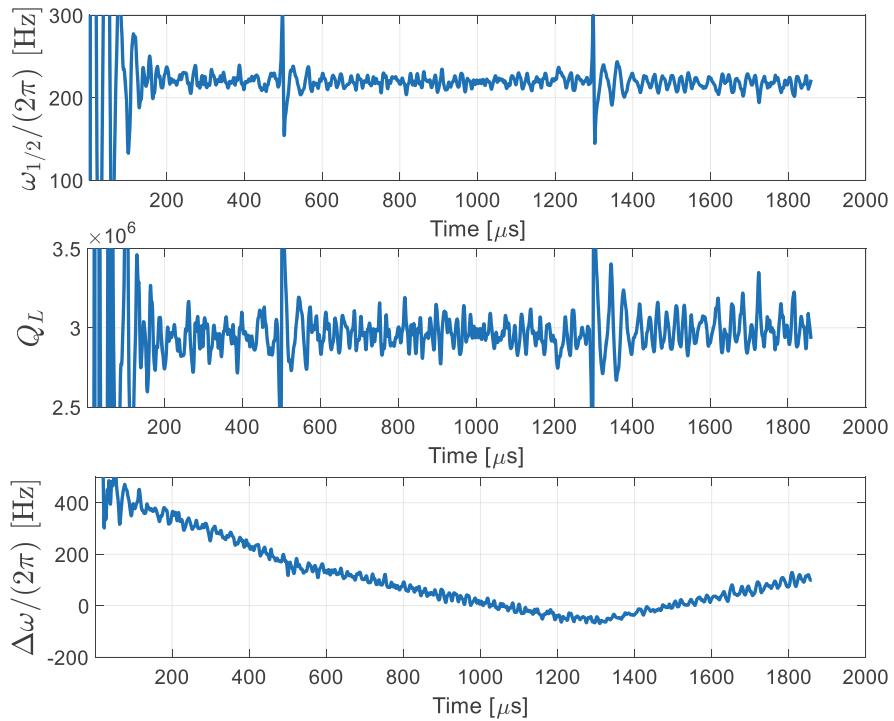
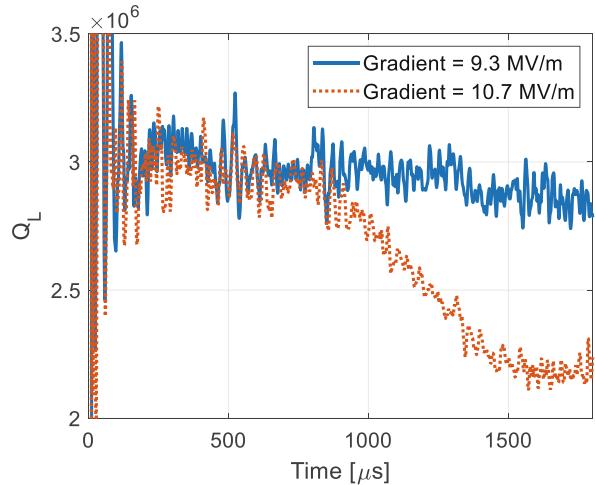


Fig. 9.28 Calculated $\omega_{1/2}$, Q_L and $\Delta\omega$ during the RF pulse

Fig. 9.29 Intra-pulse Q_L of the cavity 2 in Fig. 9.27



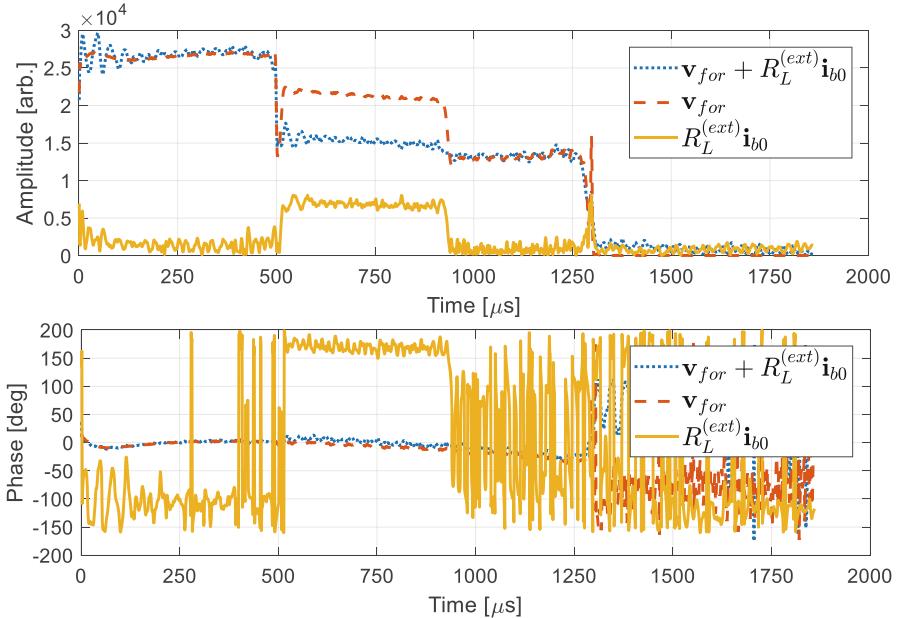


Fig. 9.30 Identification of the beam drive term

However, we can still obtain useful information if other conditions are known, such as:

- (a) If $\omega_{1/2}$ and $\Delta\omega$ are known, we can identify the time-varying beam drive \mathbf{i}_{b0} .
- (b) If \mathbf{i}_{b0} is known, we can calculate the time-varying $\omega_{1/2}$ and $\Delta\omega$.

We first discuss the case *a*. If the beam loading needs to be compensated (see Sect. 9.5), we need to identify the beam drive term \mathbf{i}_{b0} . It requires to know the time-varying $\omega_{1/2}$ and $\Delta\omega$ in the RF pulse where the beam is present. While the beam is off, we can use (9.62) to identify the time-varying $\omega_{1/2}$ and $\Delta\omega$ during the RF pulse. Assume that the cavity voltage does not change after the beam is turned on with a perfect feedback, the $\omega_{1/2}$ and $\Delta\omega$ values should be very close to that identified without beam. Therefore, we can calculate the time-varying beam drive term using the previously identified time-varying $\omega_{1/2}$ and $\Delta\omega$ as

$$R_L^{(ext)} \mathbf{i}_{b0} = \frac{\dot{\mathbf{v}}_C + (\omega_{1/2} - j\Delta\omega) \mathbf{v}_C}{2\omega_{1/2}^{(ext)}} - \mathbf{v}_{for}. \quad (9.64)$$

The values of $\omega_{1/2}^{(ext)}$, \mathbf{v}_{for} and the derivative of \mathbf{v}_C should be derived in the same way as in Sect. 9.4.3.3. As an example, Fig. 9.30 shows the time-varying beam drive term in a TESLA cavity calculated by (9.64). Here we have used the results of the time-

varying $\omega_{1/2}$ and $\Delta\omega$ derived from a previous RF pulse without beam. In this example, the beam phase is close to 0° (on-crest acceleration), so the phase of the beam drive term is about 180° off from the RF drive phase.

In practice, the time-varying $\omega_{1/2}$ and $\Delta\omega$ of the pulses with beam may be slightly different from that of the pulses without beam. The microphonics is a major source of the cavity detuning fluctuations from pulse to pulse. We often average the identified time-varying beam drive term of many pulses to reduce the errors caused by the microphonics.

The case *b* is simple. If we can measure the time-varying beam current and beam phase, the cavity eq. (9.60) can be solved directly to obtain the time-varying $\omega_{1/2}$ and $\Delta\omega$:

$$\omega_{1/2} - j\Delta\omega = \frac{2\omega_{1/2}^{(ext)} \left(\mathbf{v}_{for} + R_L^{(ext)} \mathbf{i}_{b0} \right) - \dot{\mathbf{v}}_C}{\mathbf{v}_C}. \quad (9.65)$$

One possibility is to adapt the pulsed drive of a piezo tuner to compensate for the Lorenz force detuning when the beam is on. In this case, the beam drive term does not change but $\Delta\omega$ changes significantly when changing the piezo drive. Therefore, we can first identify the time-varying beam drive \mathbf{i}_{b0} before adapting the piezo drive, and then use (9.65) to check the detuning changes during the adaption process.

9.5 Beam Loading Compensation

The beam loading, which is the beam-induced RF field, is one of the primary sources of cavity field errors. To suppress the errors caused by a heavy beam loading, we need a large loop gain if the cavity is controlled by a feedback controller. As we know, the maximum loop gain is limited by the gain margin and the measurement noise (see Fig. 6.28). However, if the beam loading is repetitive, we can introduce a feedforward signal to compensate for it.

With the method introduced in Sect. 9.4.3.4, we can identify the beam drive term $R_L^{(ext)} \mathbf{i}_{b0}$ using the vector sum \mathbf{v}_{vs} instead of \mathbf{v}_C . The quantities on the right side of (9.64) should be derived from \mathbf{v}_{vs} , which represents the output of an equivalent cavity. To compensate for the beam loading, the RF controller should generate an additional output $\Delta\mathbf{v}_{act}$. From (9.60) (with \mathbf{v}_C replaced by \mathbf{v}_{vs}) and (9.49), it is found that $\Delta\mathbf{v}_{act}$ should satisfy the following relationship to cancel the beam drive term

$$g_{sys} e^{-j\varphi_{sys}} \omega_{1/2} \Delta\mathbf{v}_{act} = -2\omega_{1/2}^{(ext)} R_L^{(ext)} \mathbf{i}_{b0}. \quad (9.66)$$

It leads to

$$\Delta \mathbf{v}_{act} = -\frac{2\beta}{\beta + 1} \frac{R_L^{(ext)} \mathbf{i}_{b0}}{g_{sys} e^{-j\varphi_{sys}}} = -\frac{2R_L \mathbf{i}_{b0}}{g_{sys} e^{-j\varphi_{sys}}}, \quad (9.67)$$

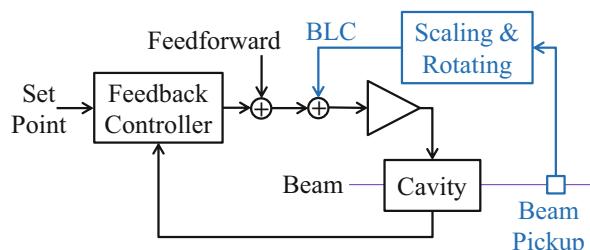
which is often implemented as a feedforward signal in the RF controller. The feedforward signal $\Delta \mathbf{v}_{act}$ must be synchronized with the beam. That is, it should be enabled or disabled at the exact time when the beam is turned on or off. Particularly, when the beam is cut off by the machine protection system (MPS), $\Delta \mathbf{v}_{act}$ should be disabled immediately. Therefore, a fast communication link between the MPS and the RF controller is required. Alternatively, the on-off state of the beam can be detected from the beam current measurement to enable or disable $\Delta \mathbf{v}_{act}$. We may follow the procedure below to apply the beam loading compensation:

- (a) Before switching on the beam, achieve the desired cavity voltage with feedforward and feedback controls.
- (b) Identify the system gain and system phase with the algorithm discussed in Sect. 9.4.2.
- (c) Determine approximately the feedforward signal $\Delta \mathbf{v}_{act0}$ for the beam loading compensation using (9.67). The beam current phasor \mathbf{i}_{b0} can be estimated from the desired beam parameters, such as the bunch charge, bunch rate, number of bunches, and the desired beam phase.
- (d) At the same time when the beam is turned on, apply $\Delta \mathbf{v}_{act0}$ in the RF controller. Good synchronization is needed to avoid large transients in the cavity voltage. For example, compared to the beam switch-on time, $\Delta \mathbf{v}_{act0}$ should be enabled no earlier or later than a clock cycle of the FPGA, in which the RF controller is implemented. This step helps to avoid beam loss due to the transients.
- (e) Identify the beam drive term using the data of the first RF pulse with the beam on and refine the feedforward signal $\Delta \mathbf{v}_{act}$ with (9.67). This step is only feasible if we have a good identification of the beam drive term.
- (f) Use the adaptive feedforward algorithm to compensate for the remaining beam loading disturbances.

The procedure above describes the switching-on sequence of the beam. The step *d* is very important for proton accelerators because we must avoid any beam loss.

For a long bunch train or a CW beam, we can also directly measure the beam current with a pickup (e.g., a toroid). After scaling the measured beam current and assigning it a proper phase, we can obtain a beam loading compensation signal to track the beam current changes. See Fig. 9.31. The benefit of this solution is that the

Fig. 9.31 Beam loading compensation (BLC) with the measurement of beam current



beam loading compensation is automatically turned on or off by the beam and can track the beam current fluctuations (Vogel et al. 2007; Kobayashi et al. 2010; Butkowski et al. 2018).

9.6 Summary

This chapter gives an overview of the high-level applications required to operate the LLRF and HPRF systems properly for the beam acceleration. Most examples are based on TESLA cavities operating in the pulsed mode. However, the algorithms can be easily expanded to RF systems with different configurations, such as with normal-conducting cavities, traveling-wave structures, or operating in the CW mode.

References

- Bellandi A, Butkowski L, Dursun B et al (2021) Online detuning computation and quench detection for superconducting resonators. *IEEE Trans Nucl Sci* 68(4):385–393. <https://doi.org/10.1109/TNS.2021.3067598>
- Brandt A (2007) Development of a finite state machine for the automated operation of the LLRF control at FLASH. Ph.D. thesis, Hamburg University
- Bravo B, Salom A, Ocampo J et al (2015) Calibration of the acceleration voltage of six normal conducting cavities at ALBA. In: Proceedings of the 6th international particle accelerator conference, Richmond, VA, USA, 3–5 May 2015
- Butkowski L, Branlard J, Omet M et al (2018) Implementation of the beam loading compensation algorithm in the LLRF system of the European XFEL. In: Proceedings of the 29th linear accelerator conference, Beijing, China, 16–21 September 2018
- Geng Z (2017) RF control optimization and automation for normal conducting linear accelerators. *IEEE Trans Nucl Sci* 64(8):2361–2368. <https://doi.org/10.1109/TNS.2017.2723508>
- Geng Z (2018) Beam-based optimization of SwissFEL low-level RF system. *Nucl Sci Tech* 29:128. <https://doi.org/10.1007/s41365-018-0460-7>
- Geng Z, Hong B (2016) Design and calibration of an RF actuator for low-level RF systems. *IEEE Trans Nucl Sci* 63(1):281–287. <https://doi.org/10.1109/TNS.2015.2507204>
- Goryashko V, Li H, Nicander H et al (2015) High-precision measurements of the quality factor of superconducting cavities at the FREIA laboratory. In: Proceedings of the 17th international conference on RF superconductivity, Whistler, BC, Canada, 13–18 September 2015
- Jonsson P (2015) High precision RF measurement for ESS cavity parameters. Dissertation, Lund University
- Kajfez D (2003) Random and systematic uncertainties of reflection-type Q-factor measurement with network analyzer. *IEEE Trans Microw Theory Techn* 51(2):512–519. <https://doi.org/10.1109/TMTT.2002.807831>
- Kobayashi T, Ikegami M (2010) Beam test of chopped beam loading compensation for the J-PARC linac 400-MeV upgrade. In: Proceedings of the XXV linear accelerator conference, Tsukuba, Japan, 12–17 September 2010

- Rybaniec R, Ayvazyan V, Branlard J et al (2014) Real-time estimation of superconducting cavities parameters. In: Proceedings of the 5th international particle accelerator conference, Dresden, Germany, 15–20 June 2014
- Schilcher T (1998) Vector sum control of pulsed accelerating fields in Lorentz force detuned superconducting cavities. Ph.D. thesis, Hamburg University
- Shi L, Baboi N, Jones RM (2018) Beam phase retrieval based on higher order modes in cylindrical superconducting radio frequency cavities. Rev Sci Instrum 89:105105. <https://doi.org/10.1063/1.5040191>
- Vogel E, Gerth C, Koprek W et al (2007) Beam loading compensation using real time bunch charge information from a toroid monitor at FLASH. In: Proceedings of the 22ed particle accelerator conference, Albuquerque, NM, USA, 25–29 June 2007

Index

A

Absolute gain, 277, 361
Absolute temperature, 195
Accelerating voltage, 7–10, 17, 21, 71, 72, 74, 78, 79, 83, 278, 279, 327, 335, 336, 340–344, 347, 348, 359
Accelerating voltage and beam phase calibration, 340–345
Accuracy, 96, 149–151, 153, 176, 178, 273, 274, 360
AC frequency, 293, 294
AC main, 293–295, 313
AC power system, 294
Active disturbance rejection control (ADRC), 85, 90–94, 96, 146
Adaptive control, 96–97
Adaptive feedforward, 26, 28, 85, 113–122, 146, 226, 369
Adaptive noise cancellation (ANC), 137–141
Agilent E5052B, 199
Air-conditioned rack, 301
Amplitude and phase control, 28
Amplitude modulation (AM) to phase modulation (PM) conversion, 268
Amplitude noise, 11, 170, 174, 175, 180, 197–201, 204, 205, 207, 210, 211, 218–221, 225–228, 238, 245, 246, 248–252, 254, 256
Analog-to-digital converter (ADC), 12, 18, 28, 31, 32, 149, 152–160, 162, 171, 175, 180, 183, 186, 187, 195, 201, 203, 213–221, 224, 245, 247, 248, 250, 268, 271, 313, 314, 316–318, 320, 339
Angular frequency, 150, 227, 238

Aperture jitter, 216–218, 220, 221, 224, 225

Aramis, 235
A Space Charge Tracking Algorithm (ASTRA), 236
Athos, 235
Attenuation length, 73
Autocorrelation, 184, 185, 191, 193
Automation, 17, 317, 328

B

Balanced optical to microwave detector (BOM-PD), 307–309
Bandwidth, 15, 28, 31, 36, 42, 43, 45, 56, 75, 76, 84–86, 89, 90, 102, 103, 106, 109, 119, 122, 123, 127, 129, 134–136, 144–146, 149, 154, 156, 159, 162, 163, 174–176, 185, 186, 188, 189, 196, 198, 200, 203, 211, 213–217, 223, 224, 226, 229–231, 239–245, 247, 250, 253, 255, 269, 270, 277, 278, 282–284, 294, 312, 318, 331, 357
Barrel open cavity (BOC), 2, 77, 329
Baseband signal, 36, 106, 149, 152, 165, 168, 174, 176–179, 223, 238, 244, 332
Beam acceleration cycle, 288, 294, 296
Beam-based calibration, 50
Beam-based feedback, 20, 233, 250, 257, 258, 261, 262
Beam diagnostics, 20, 50, 288, 292, 306
Beam drive term, 91, 367–369
Beam energy, 16, 167, 234–237, 279, 328, 333, 334, 338, 341, 343–346
Beam energy manager, 328

- Beam-induced cavity voltage, 53, 337, 338, 342
 Beam-induced drive current, 44
 Beam-induced transient, 336–340, 342–343, 347
 Beam loading, 5, 16, 21, 22, 25, 28, 35, 36, 43, 44, 54–58, 78, 79, 85, 91, 93, 94, 113, 117, 125, 126, 128, 237, 246, 339, 367–369
 Beam loading compensation (BLC), 28, 368–370
 Beam phase, 7, 9, 10, 17, 22, 50, 55, 57–59, 79, 80, 128, 195, 240–242, 244, 288, 316, 317, 327, 335–337, 340–348, 359, 365, 368, 369
 Beam position monitor (BPM), 296, 343, 344
 Bode plot, 70, 87–90, 111, 112, 115, 136, 137, 227, 228, 231, 232, 238, 239, 241, 242, 255, 258, 284
 Boltzmann’s constant, 195
 Booster1, 234, 236
 Booster2, 234, 235
 Bunch arrival time, 16, 50, 55, 234–238, 241, 261, 306, 317, 337, 341
 Bunch arrival time monitor (BAM), 261, 299
 Bunch compressor, 234, 235
 Bunch ID, 293
 Bunch length, 16, 50, 51, 234, 236, 237, 261, 279
 Bunch train, 16, 51–53, 113, 329, 336–339, 342, 369
- C**
 Carrier frequency, 36–38, 40–42, 51, 52, 66, 67, 69, 70, 75, 86, 99, 106, 149, 187, 195, 197–199, 204, 205, 207, 216, 246, 266, 289, 294
 Cascaded integrator-comb (CIC) filter, 167
 Cascaded loop, 20, 101, 104, 131, 284
 Cavity detuning, 45, 51, 91, 100, 105, 108–111, 122–124, 126, 127, 129–132, 134, 137, 138, 140–142, 334, 349, 356, 365, 368
 Cavity field, 1, 3, 14, 15, 18, 36, 89, 92, 93, 128, 146, 157, 183, 227, 233, 246, 258, 262, 269, 300, 321, 341, 351, 368
 Cavity forward signal, 350–356, 362
 Cavity input balancing, 336
 Cavity input coupling factor, 354–358
- Cavity probe signal, 18, 29, 91, 169, 227, 327, 331, 335, 336, 338–340, 342, 346–348, 350, 365
 Cavity reflected signal, 355
 Cavity resonance frequency, 8, 14, 15, 29–31, 47, 84, 98, 100, 102, 103, 105–107, 109, 122–124, 127, 129, 132, 134, 141, 143–146, 350, 357
 Cavity simulator, 81
 Cavity time constant, 48, 50, 61, 87, 150, 329, 336, 338
 Cavity voltage, 7, 8, 15, 22, 26, 29, 44, 46–48, 50–53, 55–61, 63, 78, 80, 86, 91–94, 97–99, 102–104, 106, 107, 122, 124–126, 128, 143–146, 202, 203, 226, 228, 233, 238, 239, 241, 242, 279–281, 329, 330, 335–343, 346, 350–352, 362–365, 367, 369
 Characteristic impedance, 44, 205
 Characteristic polynomial, 92
 Circulator, 2, 3, 44, 356, 365
 Clock, 18, 36, 104, 152, 154–157, 159–161, 163–169, 171, 174, 175, 177, 179, 180, 187, 194, 203, 213–225, 227, 245, 247, 248, 251, 256, 258, 259, 288, 295, 296, 306, 309, 312–323, 369
 Closed-loop bandwidth, 28, 32, 85, 102, 103, 110, 111, 115, 116, 130, 131, 134–136, 150, 170, 203, 211, 227, 228, 233, 239, 241–243, 245, 248, 277, 282, 284, 292
 Closed-loop pole, 87, 92, 93
 Closed-loop transfer function, 98, 99, 133, 134, 246, 248, 249, 283
 Coaxial cable, 13, 75, 245, 297–301, 306
 Coaxial distribution line, 297–300, 302, 303, 305
 Coherent sampling, 187, 188, 190
 Comb-filter, 239
 Command tracking, 25, 83, 87, 88, 96, 110, 277, 280
 Common locking bandwidth, 241, 243
 Common subharmonic signal, 292, 295, 313
 Complementary sensitivity function, 87
 Complex envelope, 38, 41, 149, 152, 153, 155, 276
 Complex reflection factor, 54, 355
 Complex signal, 26, 27, 37, 39
 Conjugate transpose, 119
 Constant-gradient traveling-wave structure, 73

Constant-impedance traveling-wave structure, 9
Control system, 3, 19, 21, 22, 35, 69, 76, 90, 94, 105, 126, 265
Coordinate rotation digital computer (CORDIC), 28, 166
Correlated noise, 175, 190, 194
Correlation coefficient, 191, 192
Coupling factor, 44, 45, 49, 54, 59, 65, 67, 77–79, 328, 330, 334, 341, 351, 354, 357, 358, 363
Coupling gain, 143, 145
Cross power spectral density (CPSD), 192, 193, 210, 214, 222
Cross validation, 274, 275
Cross-correlation, 191–193
Current source, 43

D

DAC drive factor, 271, 278, 281
1-dB compression point, 266–267
dBFS, 219–221
3-dB Hybrid, 150
DC-DC converter, 217
DC motor, 129, 133
DC offset, 152, 153, 155, 156, 171, 266, 307
Decay stage, 61, 124, 125, 351, 352, 359, 363, 364
Delta-function, 50
Detuning, 14, 15, 21, 30, 31, 47, 54, 56, 59, 61, 63, 77–79, 91, 97, 102, 108, 115, 119, 121–128, 132–134, 139–142, 227, 246, 327, 328, 337, 351, 354, 355, 357, 359, 362, 365, 368
Detuning angle, 47, 55, 100, 126, 128
Deutsches Elektronen-Synchrotron (DESY), 23, 261
DFT bin, 187–190, 202, 221
DFT bin spacing, 187, 189, 220, 221
Dielectric resonator oscillator (DRO), 289
Digital down-conversion (DDC), 155, 166–168
Digital signal processor (DSP), 31, 165, 166
Digital-to-analog converter (DAC), 98, 109, 123, 130, 177–180, 183, 221–225, 246, 271, 281, 310, 317, 332–333
Diode detector, 32, 150, 151
Dipole magnet, 343
Direct digital synthesis (DDS), 21, 310
Direct down-conversion, 152, 153, 177
Direct feedback, 239
Direct reference phase tracking, 175–176

Direct RF sampling, 154, 157
Direct up-conversion, 149, 177–178, 223, 224, 244, 248, 332
Directivity, 304, 336, 351, 356, 358
Discrete Fourier series, 160
Discrete Fourier transform (DFT), 185–190, 219, 221
Disturbance, 1, 10–17, 23, 25, 43, 78, 83, 85, 88, 90–96, 98, 102, 108–111, 113, 122, 126–128, 131–141, 183, 203, 227, 239, 240, 245, 246, 269–270, 284, 285, 294, 348, 369
observer, 90, 91, 93
rejection, 25, 83, 87, 90, 91, 93, 96, 110, 133, 146, 239, 277, 280, 333
Double sideband (DSB), 199–201, 204, 207, 214–216, 218, 220, 222, 224, 242
Down-conversion, 151, 154, 166, 180, 210, 211, 227, 245, 247
Drift, 4, 12, 13, 17, 20, 21, 25, 35, 83, 96, 129, 169, 175, 176, 202–203, 226–229, 245, 248–250, 252, 257–262, 278, 288, 297–306, 309, 323, 354
calibration, 13, 258–261
compensation, 13, 298, 301–304, 323
DSB phase noise, 200, 217, 223, 228, 243

E

Effective number of bit (ENOB), 213
Effective shunt impedance, 8
Electromagnetic interference (EMI), 178
Energy multiplication factor, 78
Energy signal, 184
European XFEL, 83, 289, 299
Event generator (EVG), 292, 293, 296
Event receiver (EVR), 293, 296, 314, 318
External impedance, 44, 48, 64
External quality factor, 45, 65
Extraction, 307–309

F

Fast Fourier transform (FFT), 138, 169, 172, 187
Feedback, 11–13, 16, 18, 20–22, 25–32, 63, 69, 79, 83–90, 92, 93, 95–98, 101–105, 107–118, 122, 127, 130–138, 143, 145, 146, 150, 153, 157, 170, 171, 175, 176, 178, 183, 203, 222, 226–231, 237, 239, 240, 245, 248, 254–257, 265, 269, 270, 274, 277, 278, 280, 283, 285, 288, 294, 299–301, 303, 304, 321, 323, 328, 330, 331, 333, 334, 348, 358–361, 367–369

- Feedforward, 16, 18, 21, 25–26, 28, 29, 32, 57, 79, 83, 84, 106, 108, 113–117, 119–122, 124, 141–142, 146, 169, 170, 237, 246, 280, 328–331, 354, 368, 369
- Fiducial trigger, 292–295, 313, 323
- Field programmable gate array (FPGA), 22, 23, 28, 78, 79, 155, 165, 166, 168, 178, 273, 282, 295, 296, 310, 369
- Filling stage, 57, 60, 61
- Filling time, 3, 61, 62, 72, 74, 75, 229, 230, 248, 250, 329, 334, 342
- Finite impulse response (FIR) filter, 167
- Finite-state machine (FSM), 328
- FIR model, 118–119
- Fitting set, 274
- Flattop stage, 57, 61, 365
- Flicker noise, 207–209, 220
- 1/f noise, 195–197, 207, 208
- Forward and reflected signal calibration, 336
- Forward RF power, 49, 54, 59–61, 351
- Fourier transform, 40, 185, 193
- Free-electron laser (FEL), 1, 2, 17, 20, 234, 293, 311, 312, 334
- Free-Electron Laser in Hamburg
(FLASH), 117, 123, 261, 262, 334, 356–358
- Free-running SEL, 98–99, 102, 103
- Frequency aliasing, 156–158, 214–216, 222, 224, 230
- Frequency and phase reference, 287, 288, 290, 297
- Frequency divider, 169, 211–213, 220, 247, 290–292, 295, 308, 310, 312, 313, 315–318, 320–322
- Frequency domain, 37, 39, 40, 70, 177, 179, 181, 202, 256
- Frequency multiplier, 212, 289, 310
- Frequency offset, 47, 67, 69, 153
- Frequency synthesis, 22, 104, 309, 310
- Friis formula, 197
- Full scale (FS), 195, 219, 220, 250, 271
- Fundamental frequency, 157–159, 162, 266–271, 308, 309, 313, 314, 318–320
- G**
- Gain compression, 267, 270, 271
- Gain margin, 79, 85, 87, 111, 135–137, 277, 358, 368
- Gain scheduling, 97, 280, 281, 285
- Generator driven resonator (GDR), 25, 28–30, 83–86, 98, 99, 101–104, 146
- Genetic algorithm, 20, 334
- Ghorbani model, 274–276
- Global noise model, 231, 232, 234–240, 244
- GPS disciplined Rubidium oscillator
(GPSD-Rb), 289
- Group velocity, 3, 72, 76
- Gun laser, 17, 19, 20, 242, 288, 311–313, 317, 319
- H**
- Half-bandwidth, 45, 57, 66, 67, 75–77, 89, 91, 97, 102, 111, 115, 119, 121, 124, 134, 136, 149, 211, 227, 283, 329, 337, 349, 351, 359, 362, 363
- Higher-order mode (HOM), 64, 337
- High-level application (HLA), 327, 328
- High-power radio frequency (HPRF), 1, 3, 5, 11, 16, 18, 19, 23, 27, 167–169, 183, 231, 244, 265, 370
- High-voltage power supply
(HVPS), 11, 12, 271, 275, 278
- Hilbert transform, 171–175, 321
- HVPS-based amplitude feedback, 278, 279
- I**
- IF up-conversion, 178, 180–181, 223
- IIR system, 119
- Impulse response, 118–121
- Incremental gain, 277, 278, 280, 361
- Infinite impulse response (IIR) filter, 167
- In-loop phase noise, 291, 292
- In-phase/quadrature control, 26–28
- Input coupler, 3, 43, 44, 48, 54, 60, 61, 63–67, 72, 350, 354, 362
- Input IP3 (IIP3), 267–269
- Input-output controllability, 108, 110
- Interferometric principle, 301
- Intermediate frequency (IF), 18, 28, 31, 36, 38, 149, 151–166, 168–175, 178–180, 195, 210, 211, 213–221, 223, 245, 247, 248, 258, 268, 270, 313, 314, 316–320, 331, 333
- Intermodulation, 239, 269, 270, 278, 285
- Inverse Fourier transform, 40
- Inverse Laplace transform, 43, 50
- Inverse z-transform, 119
- I/Q demodulation, 155–157, 161
- I/Q demodulator, 38, 39, 152–153, 155, 166, 302
- I/Q modulation, 22

I/Q modulator, 11, 18, 31, 86, 177–180, 244–246, 251, 262, 271, 302, 310, 332, 333

I/Q modulator imbalance, 178

I/Q sampling, 155–157, 162

Iterative learning control (ILC), 117–122, 146

J

Jitter, 10, 12, 17, 20, 83, 85, 152, 154, 169, 175, 177, 179, 180, 183, 198–203, 206, 210, 212, 219, 220, 224, 228, 234–237, 240–244, 246, 250, 252–259, 278, 288, 292, 315, 318, 333, 334

K

Kalman filter, 95, 96

L

Laplace transform, 37, 40, 41, 43, 46, 50, 51, 65, 106, 114, 241, 260

Laser heater, 234–237

Laser master oscillator (LMO), 290, 309, 311

Laser oscillator, 242, 288–290, 295–297, 306, 307, 311–313, 317–319, 321, 322

Laser repetition rate, 307, 311, 313, 318, 322

Latency, 28, 32, 149, 150, 168, 176

Least mean square (LMS), 138–140

Left half-plane (LHP), 87

Linac1, 235, 244

Linac2, 235

Linac3, 235, 236

Linac Coherent Light Source (LCLS), 12, 27, 246, 300

Linear interpolation, 272, 273, 282, 338

Linear model, 35, 76, 105, 117, 265

Linear quadratic Gaussian

(LQG), 95–97, 122

Linear quadratic regulator (LQR), 95

Linearity, 149, 150, 176, 265, 270

Linearization, 281–285, 361

LiTrack, 235, 236

LLRF node, 3, 17, 18, 20, 33

Loaded impedance, 48

Loaded quality factor, 15, 44, 45, 49, 56, 59, 61, 62, 65, 67, 79, 92, 327, 334, 354, 362

Loaded resistance, 45, 49, 57, 65, 70, 337,

341, 352, 363

Local oscillator (LO), 18, 36, 39, 151–154,

166–169, 171, 175, 178–180, 194, 210, 211, 227, 245, 247, 248, 251, 256, 258, 259, 288, 295, 296, 306, 309, 313, 314, 317, 319–321

LO & clock generator, 18, 154, 168, 180, 227, 245, 247, 248, 251, 256, 259, 309, 317, 320

Look-up table (LUT), 74, 272–274, 281, 282

Loop delay, 11, 26, 32, 35, 85–87, 89, 90, 99, 102, 103, 106, 110, 111, 115, 136, 150, 226, 239, 280, 334

Loop gain, 29, 35, 87, 89, 90, 98–100, 102, 103, 110, 122, 135–137, 239, 277, 284, 285, 358, 361, 368

Loop phase, 29, 30, 35, 89, 100, 136, 361

Lorentz force, 15, 16, 91, 96, 122, 124, 125, 133, 141, 142, 144, 365

Lorentz force detuning (LFD), 15, 16, 85, 93, 94, 109, 113, 124, 129, 141–146

LO signal, 39, 151–153, 166, 178–180, 210–212, 247, 268

Lower sideband, 151, 153, 166, 178, 179, 199, 200, 210–212, 215, 247, 310

Low-level radio frequency (LLRF), 1, 3, 10, 13, 16–23, 25, 27, 28, 31–35, 37, 38, 78, 79, 83, 90, 94, 113, 123, 125, 129, 150, 151, 153–156, 166, 167, 170, 176, 179, 186, 193, 197, 203, 204, 210, 211, 220–223, 226, 228, 230, 231, 239, 244, 245, 247, 251, 261, 265, 269, 270, 278, 288, 295, 296, 299, 300, 303, 304, 306, 312–321, 323, 327–370

LQG controller, 95, 96

lsqcurvefit, 275

M

Machine learning, 20, 334

Machine protection system (MPS), 16, 19, 369

Maintainability, 17

Master oscillator (MO), 11, 18, 19, 29, 30, 167, 228, 237, 240–245, 288–292, 294, 295, 297, 301, 312, 313, 315, 318–322

Master timing clock, 292, 293, 295, 296, 313, 314, 318–320

Matched case, 59, 60

Mathematical model, 35

- MATLAB, 43, 103, 108, 131, 172, 174, 190, 275
 Mean value, 184, 191, 202
 Mechanical mode, 62, 63, 131, 136, 137, 141, 144
 Mega-Samples Per Second (MSPS), 32, 154, 213, 225
 Microphonics, 14, 15, 22, 25, 31, 33, 85, 91, 92, 94, 109–111, 122, 124, 129, 132, 140, 141, 143, 145, 157, 203, 227, 348–350, 368
 Mixer, 12, 18, 32, 150–152, 156, 162, 166, 175, 178, 180, 183, 196, 209–212, 227, 245, 247, 248, 259, 262, 265, 268, 302, 310, 313
 $8\pi/9$ -mode, 79, 89, 90, 123
 Mode-locked laser, 290, 311
 Model predictive control (MPC), 32, 97
 Model uncertainty, 88, 91, 96, 97
 Modulated noise, 216
 Motor tuner, 129, 133, 135, 145, 311
 Moving-average filter, 162, 163, 165, 167, 229, 365
 Multi-cell cavity, 8, 43, 63–71
 Multiple input multiple output (MIMO), 22, 27, 43
- N**
- Noise, 11–12, 21, 31–33, 35, 50, 85, 88, 91, 94, 95, 98, 109, 113, 122, 138, 139, 150–152, 154, 165, 170, 174–176, 178, 180, 181, 183–204, 206–240, 244–260, 262, 274, 291, 292, 300, 318, 333, 334, 365, 368
 bin, 190
 budget, 183, 231, 233, 240, 241, 257, 258
 factors, 196–197, 207–210
 model, 85, 183, 204–227, 229, 231, 233, 236, 244–258, 262, 291
 specification, 225, 231–244, 257
 Noise figure (NF), 196, 197
 Non-I/Q demodulation, 149, 155, 157–166, 168, 169, 175, 176, 216, 218, 220, 313, 314, 316, 318, 320
 Non-I/Q sampling, 157–159, 162, 314, 318
 Nonlinearity, 151, 156, 159, 176, 239, 265–285, 359, 360
 Normalized effective shunt impedance, 8, 92
 Normalized shunt impedance per unit length, 10
 Notch filter, 90, 136–138, 163, 165
- Numerically controlled oscillator (NCO), 139, 166, 168–171, 177–179, 295
 Nyquist band, 154, 156, 157, 161, 162
 Nyquist frequency, 156, 187, 203, 233
 Nyquist plot, 87, 89, 90
 Nyquist zone, 178, 187, 189, 202, 203, 214–218, 223, 224, 229–231, 250, 255
- O**
- Offset frequency, 70, 87, 99, 106, 144, 196, 198, 200, 201, 204, 214, 216–218, 224, 246, 249, 250, 257
 Offset frequency domain, 41
 On-crest acceleration, 7, 368
 One-sided PSD, 185, 189
 One-turn delay feedback, 239, 269, 270
 Open-loop transfer function, 86, 87, 103, 111, 135–137, 283, 284
 Operating point, 12, 25, 29, 35, 76, 78, 100, 122, 145, 234, 267, 272–274, 277, 280, 285, 361
 Operating-point determination (OPD), 272–274, 276
 Optical cross-correlator, 302, 312
 Optical distribution line, 301
 Optical fiber, 293, 297, 298, 301, 302, 306
 Optimal control, 85, 95–96
 Out-of-loop phase noise, 291, 292
 Output IP3 (OIP3), 268, 269
 Output prediction, 272, 276
 Oven-controlled Crystal Oscillator (OCXO), 289
 Overfitting, 274
 Oversampling, 218, 224
- P**
- Parallel RLC circuit, 43–45
 Particle revolution frequency, 29, 122, 270
 Particle swarm optimization, 20, 334
 Passband mode, 8, 30, 43, 63, 67–71, 89, 90, 124, 331, 334
 Periodogram method, 190–192
 Personnel protection system (PPS), 19
 Perturbation, 11, 21, 96, 121, 122, 128, 137
 Phase and frequency detector (PFD), 295
 Phase-averaging coaxial line, 304–306
 Phase detector, 21, 23, 30, 31, 104, 105, 150–151, 170, 290–292, 301, 302, 307, 309

- Phase-lag, 9, 75, 76, 88, 89, 110, 115, 135, 136, 173, 248, 284, 359
- Phase-lead, 88, 89, 115, 116
- Phase-locked loop (PLL), 25, 28, 30–31, 83–85, 103–113, 126, 146, 169–171, 179, 180, 212, 289–292, 294, 295, 303, 304, 307, 309–312, 321, 322
- Phase margin, 79, 85, 88, 102, 110, 111, 115, 136, 137, 280, 284
- Phase noise, 11, 170, 174, 175, 177, 180, 184, 195–201, 204–207, 209–226, 228, 231, 233, 237, 238, 240–252, 254–260, 288–292, 294, 297, 299, 307, 310
- Phase noise analyzer, 199
- Phase-stable coaxial cable, 299
- Phase velocity, 3, 9, 72, 76
- Phasing, 312
- Phasor, 26, 37–41, 43–46, 48, 50–52, 55, 56, 58, 59, 65–67, 70, 86, 93, 106, 114, 118, 149, 152, 153, 155, 157, 226, 266, 271, 279, 335–337, 341, 344, 346, 350, 355, 356, 358, 359, 369
- Laplace transform, 40–43, 45, 46, 50, 51, 65, 66, 68, 230, 238
- transfer function, 41–47, 50, 52, 63, 66, 68–71, 74–77, 79, 86, 89, 226, 230, 231, 284
- Photocathode laser, 299, 313, 315
- Photodiode, 302, 307–309, 318
- Physics application, 19, 328
- Piezo tuner, 129–131, 135–137, 140–142, 221, 311, 322, 354, 368
- π -mode, 8, 49, 63, 69–71, 79, 89, 90, 123, 331, 337
- Polynomial model, 274–276
- Ponderomotive effect, 142–146
- Ponderomotive instability, 15, 21, 143–146
- Positive feedback loop, 29, 98, 143
- Power combiner, 205, 206, 209, 261
- Power signal, 184
- Power spectral density (PSD), 14, 185–201, 204, 206–210, 212, 214, 215, 217, 219–221, 225, 232, 242, 256, 257
- Power spectrum, 189, 202
- Power splitter, 150, 178, 205–206, 209, 245, 262, 310
- Pre-distortion, 282
- Proportional (P) control, 280
- Proportional-derivative (PD) controller, 133, 135
- Proportional-integral (PI) control, 86, 280
- Proportional-integral-derivative (PID) control, 21
- Pulse compressor, 2, 18, 77
- Pulsed laser synchronization signal, 290, 297, 298, 307, 309, 312
- Pulse ID, 288, 293
- Pulse shaping, 328–331
- Pulse-to-pulse control, 26, 226, 229–233, 258
- Pump laser, 299, 306, 311
- Q**
- Quality factor, 10, 63, 66, 72, 84, 342
- Quantization noise, 216, 217, 224, 225
- Quench detection, 328, 362, 363, 365
- R**
- Race condition, 175–177, 258, 312, 315–317, 321–323
- Radio frequency (RF), 1–23, 25, 26, 28–44, 46, 48–50, 52–61, 63, 64, 69–72, 74–80, 83–92, 94–98, 103–105, 108, 113–115, 117–126, 128, 131, 132, 135, 141–143, 146, 149–181, 183–262, 265–285, 288–300, 302, 304–309, 311–318, 321–323, 327–339, 341–343, 345–348, 350–370
- actuator, 11, 18, 25, 27, 32, 38, 86, 149, 176, 178, 226, 244, 315, 327, 328, 332
- amplifier, 1, 11, 12, 16, 18, 28–31, 33, 35, 57, 59, 61, 76, 91, 98, 106, 122, 185, 197, 207–209, 239, 245, 265, 269–277, 282, 283, 285, 335, 359, 360
- calibration, 335–354
- conditioning, 98
- controller, 1, 11, 18, 25–28, 30, 31, 35, 38, 69, 74, 79, 80, 83–86, 91, 96, 101, 120, 123, 150, 167, 176, 178, 183, 227, 251, 271, 281, 282, 327, 328, 333, 348, 349, 358–361, 368, 369
- detector, 11–13, 18, 27, 33, 38, 39, 50, 53, 86, 88–90, 123, 125, 149, 150, 152–157, 159, 167–170, 175, 176, 180, 183, 186, 193–195, 213, 226–229, 233, 245, 249–254, 256–262, 265, 268, 269, 271, 285, 303, 304, 313, 318, 327, 332, 335, 337, 338, 347, 348, 353, 354, 356
- drive current, 44, 57–59
- drive power, 1, 3, 5, 11, 16, 29, 31, 35, 36, 43, 48, 55–59, 70, 72, 79, 98, 102, 122, 124, 125, 128, 145, 328, 331, 334, 341–342, 355, 357, 362–365

- Radio frequency (RF) (*cont.*)
 driving chain, 11, 13, 18, 86, 176, 226,
 227, 233, 242, 246, 251–257, 262,
 265, 269–271, 277, 278, 281–283,
 285, 361
 driving chain noise, 94, 227, 228, 231,
 233, 244–246, 251–253, 256
 field, 1, 3, 6, 7, 10–17, 19, 20, 23, 25–27,
 29–31, 33, 35, 38, 55, 62, 69, 71, 72,
 75, 78, 80, 83–146, 150, 153, 169, 175,
 176, 181, 183, 203, 226, 229, 231–233,
 238–241, 244, 245, 249–251, 254–261,
 287, 288, 299, 304, 312, 314, 321, 327,
 328, 331, 333, 336, 338, 344, 345,
 348, 368
 measurement chain, 13, 18, 226, 227, 242,
 245, 247–252, 255–257, 259, 262
 measurement chain noise, 227, 228, 233,
 245, 247–250
 mixer, 11, 150, 151
 operating frequency, 30, 33, 36, 47, 52, 71,
 86, 89, 103, 123, 143, 167, 168, 239,
 259, 269, 270, 278, 295, 296, 357
 pickup cable, 18, 20, 226, 227, 245,
 247–250
 power calibration, 336
 power-induced drive current, 44, 49
 pulse compressor, 77–78, 329, 334
 reference frequency, 36, 154, 167, 177,
 311, 313, 317, 319, 320
 reference phase noise, 211, 228, 231, 233,
 240–248, 254, 257
 reference receiver, 11, 18, 240, 242,
 244, 245
 reference signal, 11–13, 18, 27, 29, 36,
 166–170, 174–180, 226–228, 233, 240,
 242, 244, 245, 249, 251, 254, 257, 262,
 288, 290, 299, 300, 303, 304, 317, 321,
 332, 337
 signal envelope, 36, 41, 42, 46, 75, 318
 station, 1–3, 11–13, 18, 19, 22, 26, 27, 29,
 36, 77, 167, 183, 203, 229, 231–237,
 240–260, 262, 274, 278–281, 288, 298,
 300, 303, 304, 314, 317, 320, 321, 328,
 329, 334–336, 341, 343–346, 354
 synchronization signal, 290, 298, 300,
 303–306, 309
 system noise model, 231
 Rapp model, 276
 Real signal, 37, 39, 322
 Redundancy, 17, 289
 Reference tracking, 13, 149, 166, 168–176,
 179–181, 258–259, 261, 321, 323
 Reflected RF power, 2, 59–61, 80, 334, 335
 Reflected signal, 54, 58, 79, 298, 302–306, 336,
 350–353, 355–359, 365
 Regular frequency, 150
 Reinitialization, 315–318, 320, 321
 Reliability, 17, 32, 177, 289, 298, 328
 Reproducibility, 17
 Residual MO phase noise, 242, 244, 245
 Resonance curve, 47, 100, 108, 143–145
 Resonance frequency, 12, 14, 21, 22, 29, 31,
 44, 45, 49, 51, 62, 65–67, 69–71, 98,
 107, 123, 124, 129, 131, 133, 135,
 141–143, 146, 280, 362
 Resonance mode, 66, 130, 131, 135, 136,
 144, 322
 Response matrix, 232, 234
 Resynchronization, 321–322
 Revolution frequency, 132, 167, 239
 RF-induced cavity voltage, 338, 339
 RF master oscillator (RMO), 289, 290,
 304, 311
 RHP pole, 102, 116
 Right half-plan (RHP), 102
 Rise time, 149, 176
 Robust control, 85, 96–97
 Root-mean-square (RMS) value, 184
- S**
 Sagnac-loop interferometer, 308
 Saleh model, 274–276
 Schottky diode detector, 27
 s -domain (frequency domain), 65, 172, 185,
 203, 223, 230, 233, 241, 257, 260
 δ -domain (offset frequency domain), 41
 Seeding laser, 299, 306, 311
 Self-excited loop (SEL), 21, 25, 28–31, 83–85,
 98–104, 146
 operating frequency, 29, 30, 99, 100, 102
 starting condition, 98, 99
 Sensitivity function, 87, 269
 Sensitivity matrix, 234, 235
 Shunt impedance, 10, 45
 Shunt impedance per unit length, 10, 73,
 74, 342
 Signal bin, 190
 Signal-to-noise ratio (SNR), 32, 53, 149–151,
 176, 186–191, 195, 196, 199–201, 204,
 213, 217, 220, 221, 224

Simulink, 103, 111, 112
 Single bunch, 16, 50–53
 Single-cavity control, 33
 Single-cell cavity, 6, 43–63, 69, 70
 Single input single output (SISO), 22, 27, 41
 Single sideband (SSB), 70, 149, 178–180, 198, 215, 216, 218, 223, 225, 310 modulation, 310
 phase noise, 199, 217, 223, 225, 243 up-conversion, 178–180, 248
 SLAC, 12, 300
 SLAC energy doubler (SLED), 77, 78, 329
 Solid-state amplifier, 4, 5, 33
 Solid-state modulator, 12, 294
 Spectral leakage, 187, 190
 Spectrometer, 343, 344
 s plane, 37, 41, 42
 \hat{s} plane, 41, 42
 Spur, 214, 215, 217–221, 224, 225
 Spurious frequency, 152, 216, 217
 Spurious signal, 162, 214
 Standing-wave cavity, 2, 3, 6, 7, 10, 71, 124, 227, 234, 245, 248, 331, 341
 Star topology, 297, 298, 301
 State-space equation, 46, 66, 67, 92, 95
 Stationary stochastic process, 183
 Steady state, 15, 21, 47, 48, 54–57, 59, 61, 63, 100, 103, 107, 126–128, 131, 144, 150, 170, 176, 239, 266, 271, 276, 284, 288, 303, 341, 342, 354, 355, 359
 Stepping motor, 129, 133
 Structure voltage, 10, 74
 Subsampling, 154, 155
 SwissFEL, 2, 3, 5, 11–13, 26, 27, 183, 234–237, 242, 243, 246, 250, 251, 253, 255, 271, 278, 279, 288, 290, 294, 312, 331, 333, 342, 344, 348
 Synchronization client, 300
 Synchronization system, 11, 18–20, 36, 228, 246, 287, 288, 290, 292, 297–312, 323
 Synchronized fiducial trigger, 295, 296, 313
 Synchronous particle, 238, 239
 Synchrotron, 122, 237–240, 269, 270, 278, 293
 oscillation, 238, 239
 radiation, 1, 238, 334
 System gain, 79, 86, 89, 97, 102, 328, 358–361, 369
 System identification, 79, 85, 91, 97, 354–368
 System phase, 28, 79, 86–90, 96–102, 328, 358–361, 369

T

Test set, 274
 TeV-Energy Superconducting Linear Accelerator (TESLA), 4, 14, 15, 49–51, 53, 54, 57, 62–64, 67–71, 89, 90, 99, 100, 111, 120, 121, 125, 126, 131, 142, 262, 280, 289, 337, 338, 341, 363–365, 367, 370
 Thermal noise, 186, 194–197, 204, 206, 209, 216–221, 224, 225
 Third-order intercept point (IP3), 266, 267
 Time domain, 37, 41, 43, 46, 50–52, 66, 68, 114, 165, 172, 184, 185, 195, 246–248, 252, 254
 Time-reversed low-pass filter, 115–117
 Timestamp, 288, 293
 Timing event, 287, 293
 Timing fiducial, 292–295
 Timing relation, 287, 312–323
 Timing system, 19, 287, 288, 292–296, 313, 323
 Ti: sapphire laser, 290
 TM010 mode, 6, 15
 Tolerance, 17, 83, 203, 232, 233, 236, 237, 240, 241, 253, 257, 288
 Toroid, 261, 369
 Transfer function, 11, 35, 37, 40–43, 69, 70, 96, 102, 106–109, 111, 112, 114–116, 119, 122, 127, 129, 130, 133–136, 142, 186, 227, 231, 233, 234, 238–240, 242, 255–257, 265, 283, 291, 292, 358
 Transformer, 44, 48, 63
 Transit-time factor, 7, 10
 Transmission line, 3, 44, 48, 49, 54, 59, 61, 64, 75–76, 204, 205, 226, 227, 245, 246, 265, 346, 351, 357
 Traveling-wave structure, 3, 9, 10, 71–75, 229–231, 245, 248, 250, 253, 255, 341, 342
 Trigger, 19, 154, 169, 175, 177, 287, 288, 293–296, 312–323, 328
 Tuning control loop, 127–129, 132
 Two-sided PSD, 185, 188, 189

U

Uncorrelated noise, 32, 175, 193, 194, 206, 259
 Underfitting, 274
 Undersampling, 154
 Unloaded impedance, 48
 Unloaded quality factor, 8, 45, 65, 67, 78, 362

Up-conversion, 149, 151, 178–180, 210

Upper sideband, 151, 178, 210, 212

V

Vector sum, 33, 80, 126, 279, 304, 336, 343, 344, 346–350, 358–360, 368

Vector-sum accelerating voltage, 336, 343, 344

Vector-sum beam phase, 336, 345

Vector-sum calibration, 33, 336, 346–350

Vector-sum control, 25, 33, 348–350

Virtual reflection, 303

Voltage-controlled oscillator (VCO), 30, 31, 104–111, 113, 290, 309, 310

W

White model, 276

White noise, 96, 137, 190, 195–197, 201, 204, 207–210, 212, 217, 218, 224, 225, 250–252, 257, 259

Wiener-Khinchin theorem, 185

Wilkinson divider, 150

Windowing, 187, 188

Z

Zero order hold (ZOH), 222, 223, 225, 230, 231

z-transform, 119, 230