Introduction to LLRF

1. Abstract

In this lecture an introduction to Low Level Radio Frequency (LLRF) is given. The lecture is divided into two parts. First, after explaining the purpose of LLRF, an introduction to the theoretical background is given. This covers the modeling of accelerating RF cavities, the deduction of fundamental equations, and short peaks into signal processing and controller theory. This part also includes a demo with a cavity simulator, to give the listener a better feeling, on how a basic system behaves. The second part of the lecture focuses on the architecture of LLRF systems, actual applications, and real-world examples. At the end of the lecture, the listener should have a good idea what LLRF is about and what path to follow, if he or she wants to get involved with this topic.

2. Introduction

LLRF stands for Low Level Radio Frequency. The goal is to control RF fields withing cavities. Depending on the application, the requirements on the LLRF system vary drastically. E.g., in a single cavity vertical test stand, it is important to control the amplitude, but the phase basically does not matter. On the other hand, in a particle accelerator both, amplitude and phase have to be controlled precisely. Furthermore, there are different types of cavities, such as normal conducting (NC) cavities and superconducting radio frequencies (SRF) cavities. Depending on the application, the frequencies are different, ranging typically from several MHz to tens of GHz.

Especially in SRF cavities, voltages of several tenths of MV are common. Such high fields are impossible to detect directly. Thus, only a very small fraction of the field is coupled out via a pickup antenna. Its signal is then sent to the LLRF system for detection. Typically, the input power to the LLRF system is less than 1 W, hence Low Level RF.

Depending on the purpose of the facility, cavities are operated in a specific mode. There are continuous wave (CW) machines, in which a continuous RF field is maintained within the cavities. Its duty factor is 100% and, if the facility is a particle accelerator, beam can be accelerated all the time. On the other hand, cavities can be operated in pulsed mode. This can be distinguished into short pulse mode (SP) and long pulse mode (LP). In SP mode the duty factor is typically about 1%, meaning only during 1% of the time RF is fed into the cavity. Beam acceleration is possible only during a portion of this time. In LP mode the duty factor is typically 10% to 50%.

The most basic setup including an LLRF system is shown Figure 1. It comprises the LLRF system, which generates corresponding to a set setpoint a drive signal. This is fed into an amplifier, which amplifies it. It is then fed into a cavity. A small fraction of the RF is coupled out and sent back to the LLRF system, where it is detected. This is called an open-loop operation. By implementing a feedback controller, it is possible to perform closed loop operation. In this case the detected pickup signal is compared against the given setpoint. Based on its difference the drive signal is computed and generated.
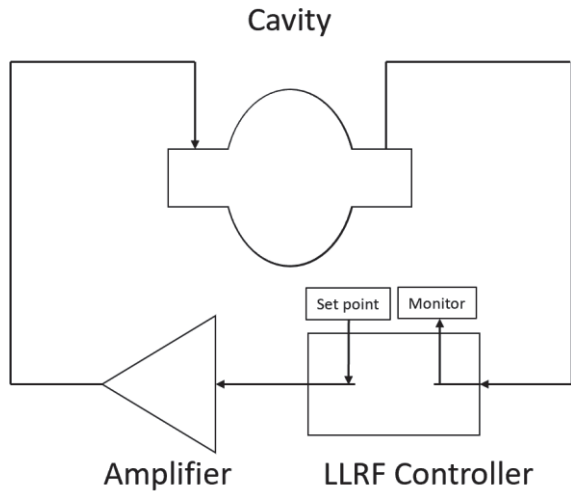
Figure 1: Most basic setup including an LLRF system.

3. LLRF System Overview

In order to be able to decide how to design and build an LLRF system, one has to understand first the nature of the object, which shall be controlled – in this case a SRF cavity. A cavity is a resonator, which can be modeled in a first approximation with an RCL electric circuit. It is shown in Figure 2.
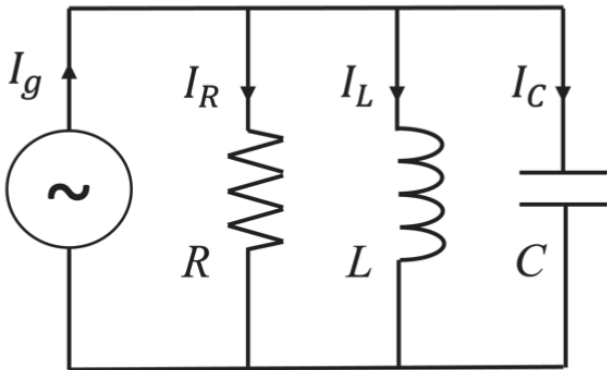

Figure 2: RCL circuit.

Before we explore the RCL model, let us define an important property of the cavity: the quality factor. In general, it is defined as the ratio between the energy stored in the cavity and the power dissipated from the cavity per RF cycle.

$$Q = \frac{2\pi f_0 W}{P_d}$$

There are different ways power can dissipate from a cavity, e.g. via losses due to the surface resistance. In this case the quality factor is called unloaded quality factor and is defined as

$$Q_0 = \frac{2\pi \frac{1}{2} C V_0^2}{T \frac{1}{2} \frac{V_0^2}{R}} \quad ,$$

where T is the time period of an RF cycle, C the capacitance, V_0 the amplitude of the oscillating voltage, and R the resistance. With $W = \frac{1}{2} C V_0^2$, $P_{diss} = \frac{V_0^2}{2R}$, $\omega_0 = \frac{1}{\sqrt{LC}}$, and $\omega_0 = 2\pi f_0$ we can rewrite the equation to:

$$Q_0 = \omega_0 RC = \frac{R}{L\omega_0} = \frac{\omega_0 W}{P_{diss}} \quad .$$

Beside the losses due to the surface resistance, there are external losses, e.g. via the power coupler. In this case the quality factor is called the external quality factor:

$$Q_{ext} = \frac{\omega_0 W}{P_{ext}}$$

When combining all losses, the quality factor is called the loaded quality factor:

$$Q_L = \frac{\omega_0 W}{P_{tot}}$$

In this case the total dissipated power is $P_{tot} = P_{diss} + P_{ext}$. One can compute the loaded quality factor as follows:

$$Q_L = \frac{1}{Q_0} + \frac{1}{Q_{ext}}$$

In the case of SRF cavities Q_0 is several orders of magnitude larger than Q_{ext}. Thus, Q_L is in the same order as Q_{ext}.

In order to reflect the real world better, we have to add a transition line to the RCL model, as shown in Figure 3.
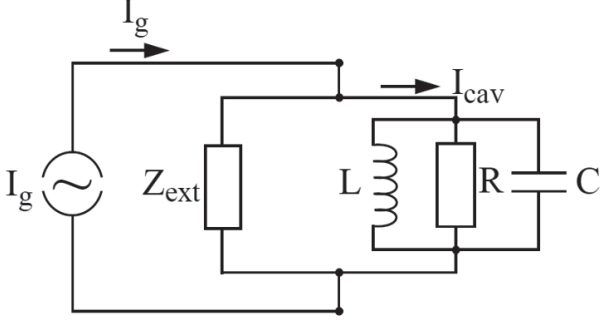

Figure 3: RCL circuit with transition line [1].

This models e.g. the cable connecting the power amplifier to the cavity. It is represented as the impedance Z_{ext} and is like a parallel resistor to R. The characteristic impedance of a coaxial cable is 50 Ω. Both can be replaced by the shunt impedance:

$$R_L = \frac{1}{R} + \frac{1}{Z_{ext}}$$

We can rewrite $Q_0 = \omega_0 R C = \frac{R}{L\omega_0} = \frac{\omega_0 W}{P_{diss}}$ to:

$$\frac{R}{Q_0} = \omega_0 = \frac{1}{\omega_0 C} = \sqrt{\frac{L}{C}}$$

Looking at this, we can see that $\frac{R}{Q_0}$ depends only on ω_0, C, and L. This means that it depends only on the cavity geometry and not on the surface resistance.
From

$$P_{diss} = \frac{1}{2}\frac{V_{cav}^2}{R} = \frac{V_{cav}^2}{R_{sh}}$$

we can see that the shunt impedance R_{sh} depends on the dissipated power. Comparing it to the RCL resistance R_{sh}, a factor of $\frac{1}{2}$ of the time average has to be included:

$$R = \frac{1}{2}R_{sh} = \frac{1}{2}\frac{r}{Q}Q_0$$

With this the normalized shunt impedance can be defined as:

$$\frac{r}{Q} := \frac{R_{sh}}{Q_0} = \frac{2R}{Q_0}$$

Furthermore, we can define the coupling between the cavity and the transmission line:

$$\beta = \frac{R}{Z_{ext}}$$

With $\frac{1}{R_L} = \frac{1}{R} + \frac{1}{Z_{ext}}$ we can rewrite this as:

$$R_L = \frac{R}{1+\beta}$$

Using the definition of the shunt impedance we get the two following equations:

$$Q_L = \frac{Q_0}{1+\beta}$$

and

$$\omega_{1/2} = \frac{\omega_0}{2Q_L}$$

From this we can learn that the loaded quality factor of a cavity Q_L can be manipulated by changing the coupling β. Furthermore, at the same time the half-bandwidth $\omega_{1/2}$ is manipulated. This fact is exploited at facilities. Two possible ways are changing the insertion depth of the input coupler antenna or the angle of a reflector plate inside an RF waveguide.

For the following we will focus on the pulsed operation mode with beam loading. During every pulse RF is fed into the cavity in order to charge it up. This time period is called filling or fill time. It is followed by a time period, which is called flattop. During this time the amplitude and phase of the RF inside the cavity are kept constant. The flattop is the time period designated for beam acceleration. After the flattop, the RF drive of the cavity is shut off. The RF field inside the cavity decays, thus it is called decay. A schetch of such an RF pulse is shown in Figure 4.
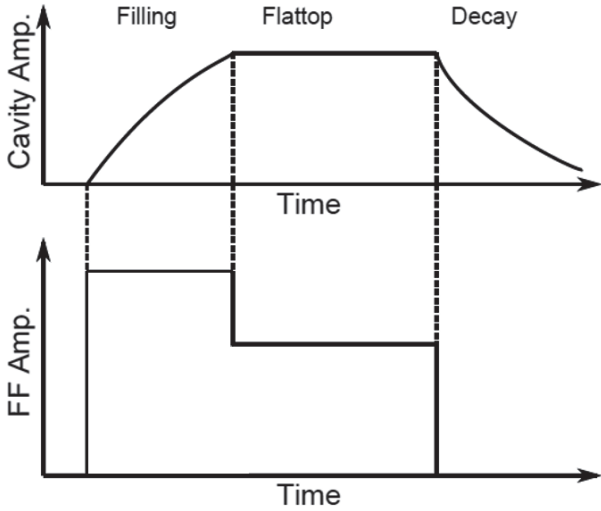
Figure 4: Cavity and drive amplitudes over time during an RF pulse [2].

Now let us assume that we would inject a beam (e.g. a train of electron bunches) into the cavity just at the beginning of the flattop. Furthermore, the beam would stop at the end of the flattop, just when the RF drive stops. The first electron bunch will draw energy from the RF field inside the cavity. This will lead to a drop of the cavity voltage. The second bunch will do the same, resulting in a further drop of the cavity voltage, and so on. One would see an effect, which looks similar like a decay after the flattop. This is called beam loading. In order to

compensate for this, one can increase the RF drive power during the beam transient time, which is in our case the whole flattop. If the compensation is just right, a flat flattop can be restored. This is desirable, since only in this situation every electron bunch would receive the same energy gain. A plot of this is shown in Figure 5.
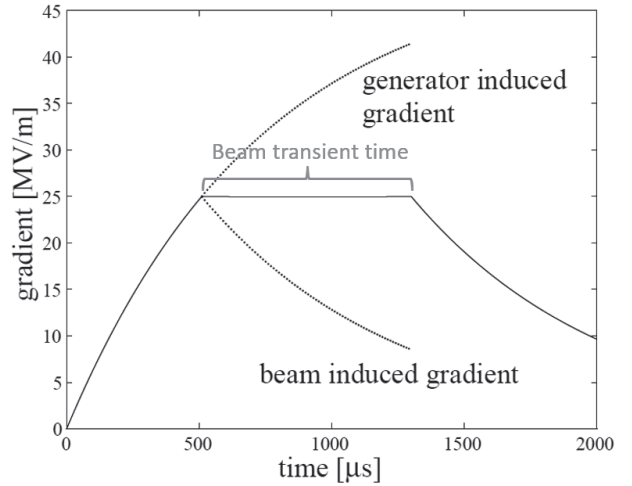

Figure 5: Cavity gradient over time for an RF pulse with compensated beam loading during the flattop [1].

Let us find out the required filling and flattop powers for this case.
In order to consider the beam-loading case, we have to extend the simple RCL circuit with an additional current source, (see Figure 6).
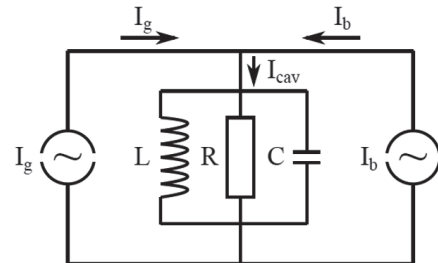

Figure 6: RCL circuit in the case of the beam transient time [1].

We can write the cavity current as the sum of currents of all three elements:

$$I_{cav} = I_C + I_R + I_L$$

Simply deriving it for time yields:

$$\dot{I}_C + \dot{I}_R + \dot{I}_L = \dot{I}_{cav}$$

With $\dot{I}_R = \frac{1}{R_L}\dot{V}_{cav}$, $\dot{I}_C = C\ddot{V}_{cav}$, and $\dot{I}_L = \frac{1}{L}V_{cav}$ this yields:

$$C\ddot{V}_{cav} + \frac{1}{R_L}\dot{V}_{cav} + \frac{1}{L}V_{cav} = \dot{I}_{cav}$$

With $\frac{1}{R_L C} = \frac{\omega_0}{Q_L}$ and $\frac{1}{LC} = \omega_0^2$ this equation can be rewritten as:

$$\ddot{V}_{cav} + \frac{\omega_0}{Q_L}\dot{V}_{cav} + \omega_0^2 V_{cav} = \frac{1}{C}I_{cav}$$

This is a second order differential equation. In order to solve it, we need to find a homogeneous solution and a particular solution. The homogeneous solution reads:

$$V_{hom} = e^{-\frac{\omega_0 t}{2Q_L}}(C_1 e^{i\alpha t} + C_2 e^{-i\alpha t})$$

with $\alpha = \omega_0\sqrt{1 - \frac{1}{4Q_L^2}}$.

A particular solution is:

$$V_{par} = \frac{R_L \hat{I} e^{i(\omega t + \Phi)}}{\sqrt{1 + \tan^2(\Phi)}}$$

with $\tan(\Phi) = R\left(\frac{1}{\omega L} - \omega C\right) = Q\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)$. The particular solution is also called the stationary solution. If the generator frequency ω is very close to the resonance frequency ω_0, the following approximation can be done:

$$\hat{V}_{par}(\Delta\omega) \approx \frac{R_L \hat{I}}{\sqrt{1 + \left(2Q_L \frac{\Delta\omega}{\omega}\right)}}$$

where $\Delta\omega = \omega_0 - \omega$. With this the general solution is:

$$V_{cav} = V_{hom} + V_{par}$$

$$V_{cav} = e^{-\frac{\omega_0 t}{2Q_L}}(C_1 e^{i\alpha t} + C_2 e^{-i\alpha t}) + \frac{R_L \hat{I} e^{i(\omega t + \Phi)}}{\sqrt{1 + \tan^2(\Phi)}}$$

Since $Q_L \gg 1$, we can approximate $\alpha \approx \omega_0$. For $C_1 = C_2 = -\frac{R_L \hat{I}}{2}$ we get:

$$V_{fill} = V_0\left(1 - e^{-\frac{t}{\tau}}\right)$$

with $V_0 \approx 2R_L I_g = \frac{r}{Q}Q_L I_g$ and $\tau = \frac{2Q_L}{\omega_0}$. Furthermore, for the beam transient time, where $I_{cav} = 2I_g - 2I_{b0}$, we get:

$$V_{flat} = \frac{r}{Q}Q_L\left(I_g\left(1 - e^{-\frac{t}{\tau}}\right) - I_{b0}\cos(\Phi_b)\left(1 - e^{-\frac{t-T_{inj}}{\tau}}\right)\right)$$

With this we now have two equations, describing the fill time and the flattop. Looking back at Figure 5, our requirement was to maintain a constant cavity voltage over the flattop. Thus, the following must hold:

$$\frac{dV_{flat}}{dt} = 0$$

Solving

$$\frac{d}{dt}\frac{r}{Q}Q_L\left(I_g\left(1 - e^{-\frac{t}{\tau}}\right) - I_{b0}\cos(\Phi_b)\left(1 - e^{-\frac{t-T_{inj}}{\tau}}\right)\right) = 0$$

yields:

$$I_g = I_{b0}e^{\frac{T_{inj}}{\tau}}$$

And finally:

$$V_{flat} = \frac{r}{Q}Q_L I_{b0}\left(e^{-\frac{T_{inj}\omega_0}{2Q_L}} - 1\right)$$

By inserting parameters similar to the ILC TDR (a fill time of 923 μs, a beam current of 5.8 mA, and a loaded quality factor of 5.44E6), we can produce the plot as shown in Figure 7, which is consistent with our expectations.
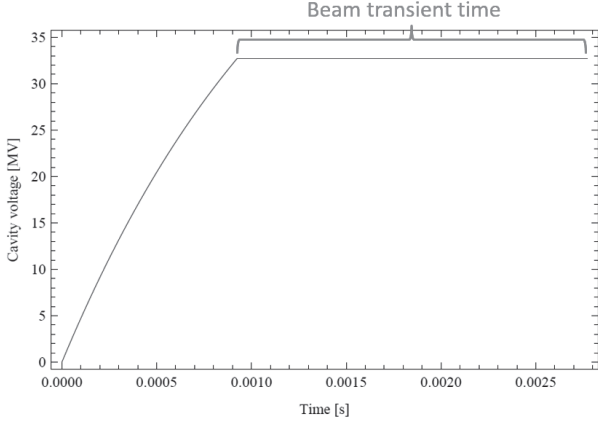

Figure 7: Cavity voltage during fill time and flattop with beam loading for parameters similar to the ILC TDR [2].

In order to compute the required power during the filling and the flattop, we can use $P = \frac{1}{4}\frac{r}{Q}Q_L I_g^2$, which yields:

$$P_{fill} = \frac{V_{cav}^2}{4\frac{r}{Q}Q_L\left(1 - e^{-\frac{T_{inj}\omega_0}{2Q_L}}\right)}$$

$$P_{flat} = \frac{V_{cav}^2}{4\frac{r}{Q}Q_L}\left(1 + \frac{\frac{r}{Q}Q_L I_{b0}}{V_{cav}}\right)^2$$

By inserting the following values $V_{cav} = 31.5\,\text{MV/m} \cdot 1.038m = 32.7\,\text{MV}$, $Q_L = 5.44 \cdot 10^6$, $T_{inj} = 923$ μs, $I_{b0} = 5.8$ mA, and $\Phi_b = 180°$ the powers read:

$$P_{fill} = 190\,\text{kW}$$

$$P_{flat} = 190\,\text{kW}$$

Of course, it was intentional that the filling and flattop powers are the same. The reason is that

for the ILC it is planned to use klystrons as high-power amplifiers, driving multiple cavities at once. These amplifiers show a non-linear saturation behavior in terms of output power versus input power. Thus, when staying at the same operation point over the entire RF pulse, the LLRF system does not have to cope with any non-linearities.

With the knowledge gained so far, it is possible to find a set of equations for optimal parameters:

The optimal coupling reads:

$$\beta_{opt} = 1 + \frac{\frac{r}{Q}Q_L I_{b0}}{V_{cav}}\cos(\Phi_b)$$

The minimum power for maintaining the cavity voltage reads:

$$P_{min} = \beta_{opt}\frac{V_{cav}^2}{\frac{r}{Q}Q_0}$$

The optimum tuning angle reads:

$$\tan(\Phi_{opt}) = -\frac{\frac{r}{Q}Q_{L,opt}I_{b0}}{V_{cav}}\sin(\Phi_b)$$

In case of SRF cavities we can simplify to:

$$Q_{L,opt} = \frac{V_{cav}}{\frac{r}{Q}I_{b0}\cos(\Phi_b)}$$

$$\Phi_{opt} = -\Phi_b$$

$$P_{flat,min} = \frac{V_{cav}^2}{4\frac{r}{Q}Q_{L,opt}} = V_{cav}I_{b0}\cos(\Phi_b)$$

In reality the cavities are detuned by the tuning angle Φ. One source for this is Lorentz force detuning. This is an effect which is caused by strong electro-magnetic fields within the cavity, causing the cavity walls to deform slightly.

Another source of detuning is microphonics. These are mechanicals vibrations caused by e.g. pumps, road traffic, etc., which are causing constantly changing deformations of the cavity. In Figure 8 the vector diagram of generator and beam-induced voltages in such a detuned cavity are depicted.
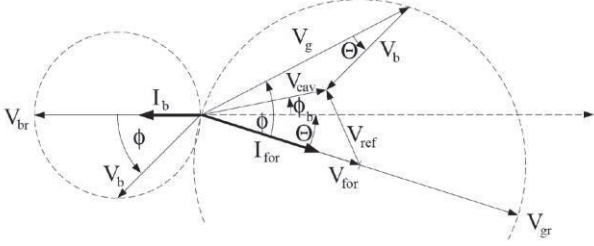

Figure 8: Vector diagram of generator and beam-induced voltages in a detuned cavity [1].

From this diagram we can see that $V_{cav} = V_g + V_b$, which we already have covered. But more importantly one can also see that:

$$V_{cav} = V_{for} + V_{ref}$$

This equation, as simple as it is, is very powerful. It is true for all times. Typically, at pulsed machines, it is convenient to plot the cavity, forward and reflected signals. It makes it very easy to spot, if something is wrong, e.g. with the calibration of a certain channel.

For the following let us consider one more time the differential equation for a driven LCR circuit and let us write is as follows:

$$\ddot{V}(t) + \frac{\omega_0}{Q_L}\dot{V}(t) + \omega_0^2 V(t) = \frac{\omega_0 R_L}{Q_L}\dot{I}(t)$$

The driving current I_g and Fourier component I_b of a pulsed beam are harmonic with the time dependence $e^{i\omega t}$. Therefore, we separate the fast RF oscillation from the real and imaginary parts of the field vector (also see Figure 9):

$$V(t) = \big(V_r(t) + iV_i(t)\big) \cdot e^{i\omega t}$$

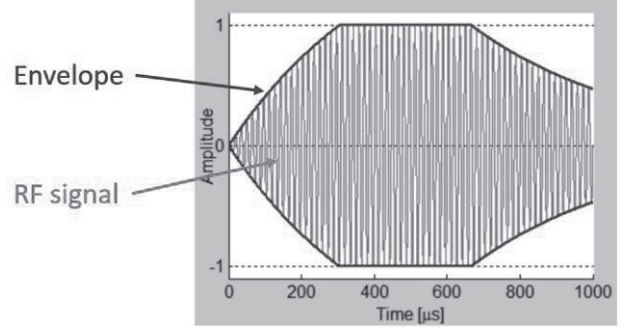$$I(t) = \big(I_r(t) + iI_i(t)\big) \cdot e^{i\omega t}$$

Figure 9: RF signal (red) and its envelope (blue) [3].

If we now insert this into the equation above and omit the second-order time derivatives $\ddot{V}(t)$, we get the first-order differential equation for the envelope:

$$\dot{V}_r + \omega_{1/2}V_r + \Delta\omega V_i = R_L\omega_{1/2}I_r$$

$$\dot{V}_i + \omega_{1/2}V_i + \Delta\omega V_r = R_L\omega_{1/2}I_i$$

with the cavity bandwidth $\omega_{1/2} = \frac{\omega_0}{2Q_L}$ and the cavity detuning $\Delta\omega = \omega_0 - \omega$. We can rewrite these two equations in state space formalism as:

$$\frac{d}{dt}\begin{pmatrix}V_r\\V_i\end{pmatrix} = \begin{pmatrix}-\omega_{1/2} & -\Delta\omega\\\Delta\omega & \omega_{1/2}\end{pmatrix}\cdot\begin{pmatrix}V_r\\V_i\end{pmatrix} + \begin{pmatrix}R_L\omega_{1/2} & 0\\0 & R_L\omega_{1/2}\end{pmatrix}\cdot\begin{pmatrix}I_r\\I_i\end{pmatrix}$$

This is called the cavity differential equation. With $A = \begin{pmatrix}-\omega_{1/2} & -\Delta\omega\\\Delta\omega & \omega_{1/2}\end{pmatrix}$, $B = \begin{pmatrix}R_L\omega_{1/2} & 0\\0 & R_L\omega_{1/2}\end{pmatrix}$, $x = \begin{pmatrix}V_r\\V_i\end{pmatrix}$, and $u = \begin{pmatrix}I_r\\I_i\end{pmatrix}$, we can write this equations in short as:

$$\dot{x}(t) = A \cdot x(t) + B \cdot u(t)$$

Until now we have considered the cavity differential equation continuous in time, which is true for reality. For simulations with e.g. a computer, the cavity differential equation discreate in time is more useful. It reads:

$$\begin{bmatrix} V_{i,n} \\ V_{q,n} \end{bmatrix} = \begin{bmatrix} 1 - T\omega_{1/2} & -T\Delta\omega \\ T\Delta\omega & 1 - T\omega_{1/2} \end{bmatrix} \begin{bmatrix} V_{i,n-1} \\ V_{q,n-1} \end{bmatrix} + T\omega_{1/2}R_L \begin{bmatrix} I_{i,n-1} \\ I_{q,n-1} \end{bmatrix}$$

This differential equation was used to program a single cavity simulator, which is presented in the lecture to demonstrate pulsed operation with ILC TDR-like parameters. Furthermore, the use cases for a low and a high Q_L are discussed. Figure 10 shows the GUI of the simulator.
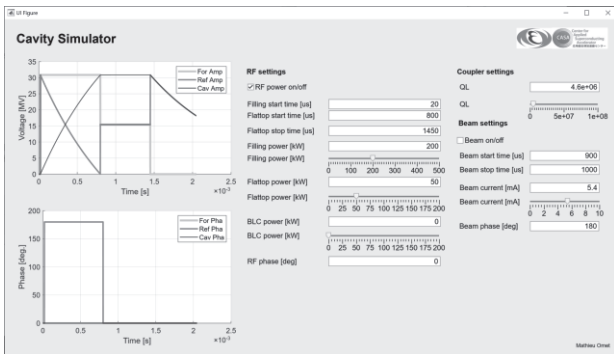
Figure 10: GUI of single cell cavity simulator.

4. LLRF Systems

LLRF systems can be separated into two categories: analog LLRF systems and digital LLRF systems.

In the case of the analog LLRF systems, the signal detection, signal processing, feedback control, and drive signal generation are all realized using analog components. These systems are typically designed, optimized, and built for a specific purpose. When requirements change, they are hard to modify. Furthermore, one needs additional hardware for data monitoring and recording.

In the case of digital LLRF systems, signals are converted from analog to digital using analog to digital converters (ADCs). The signal processing, including the feedback controller as well as the drive signal generation is realized in a digital way. The digital drive signal is then converted to an analog signal, using digital to analog converters (DACs). The advantage of digital LLRF systems is that these are very flexible. If a requirement changes, the digital algorithm can be reprogrammed. There are also some modern algorithms, which would be very difficult and costly to implement in an analog way. Since the signals are already processed digitally, it is easy to make them available for monitoring and recoding. Furthermore, digital systems can be maintained remotely to a great extent.

Generally, digital LLRF systems can be subdivided into three different flavors: 1. 19-inch module-based systems, 2. Crate-based system, and 3. Mixed system.

An example for a 19-inch module-based system is the LCLS-II LLRF system. Its prototype version is shown in Figure 11.
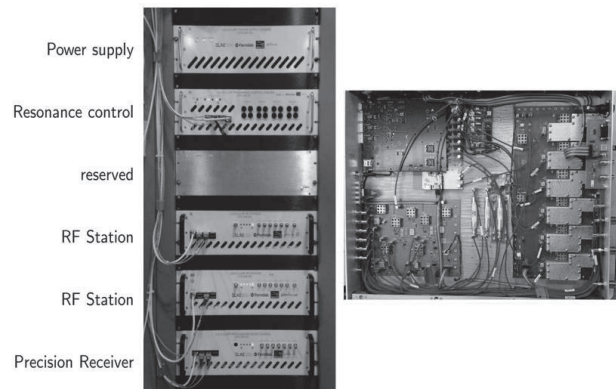

Figure 11: LCLS-II prototype LLRF system at FNAL CMTS [4].

On the right side of the picture, on can see the inside of a 19-inch module. This hardware was developed and built specifically for its purpose and is very well optimized.

An example for a crate-based system, is the MicroTCA.4-based LLRF system at the European XFEL at DESY. Figure 12 shows on the left side a system inside the tunnel underneath a cryomodule and on the right side the crate and four cards as an example. The
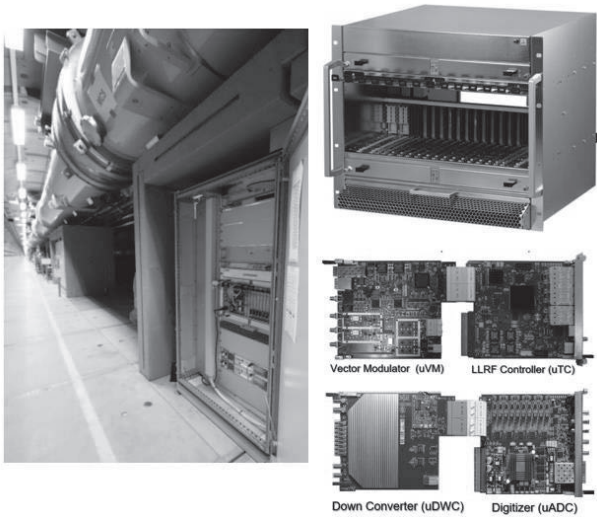
Figure 12: MicroTCA.4-based LLRF system at the European XFEL at DESY [5].

Figure 13: MicroTCA.0-based LLRF system at cERL at KEK [6].

crate offers slots at the front and at the back. In the back cards with analog hardware are inserted. In the front, cards with digital hardware are inserted. The cards are interconnected via a connector. This way e.g. the down converter card is combined with a digitizer card and the LLRF controller card is combined, with the vector modulator card. The modular nature of such systems allows an easy replacement of a broken card. Furthermore, all cards are of-the-shelf components. A wide range of cards optimized for different purposes are commercially available. Another advantage of such kind of system is, that it is very compact.

An example for a mixed system is the MicroTCA.0-based LLRF system at the cERL at KEK. Figure 13 shows on the left side the racks of several system within the temperature-controlled hut and on the right side the components of one LLRF system. In this case the down conversion is done with a 19-inch module, whereas the controller is realized on a card within a MicroTCA.0 crate. Such kind of system allows the combination of the best of both other system flavors.

Figure 14 shows an example system architecture of a digital LLRF system. As described above, the analog input signals are digitized using ADCs. The digital signals are processed by e.g. filters and a feedback controller on a Field Programmable Gate Array (FPGA). It is possible to implement also further algorithms and calculations in order to allow special functionalities. The generated drive signal is converted to analog using a DAC. The FPGA is typically interconnected with a local CPU. This is basically a small computer, connected to the control system of the facility. It allows data communication with remotely connected computers, which are used either for monitoring and operation or for data acquisition.
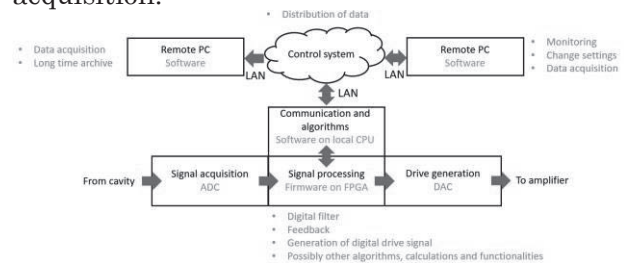

Figure 14: Example system architecture of a digital LLRF system.

5. Signal Sampling

Converting analog signals into digital ones is crucial for realizing digital LLRF systems. As mentioned in the previous sections, ADCs are used to this end. As shown in Figure 15, the ADC measures the amplitude of an analog signal at a certain point in time and outputs the according number. This is called a sample point. The time between two consecutive sample points is the sample period Ts. ADC typically have a clock input, which allows to trigger the sampling.
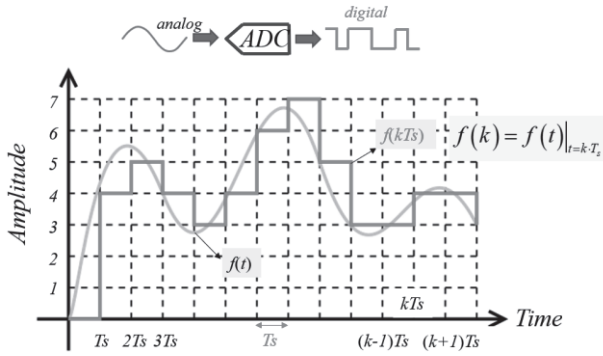

Figure 15: General idea of digitally sampling an analog signal [7].

As prerequisite for the following, we need to define the coordinates we will be using. When dealing with sinusoidal signals, amplitude and phase are naturally the quantities used for description. It turns out that these are not optimal for performing computations on an FPGA. For this purpose, the best choice is the representation in In-phase (or real part) I and Quadrature (or imaginary part) Q. Figure 16 shows the representation of a coordinate in terms of amplitude and phase as well as I and Q. The relations are:

$$I = A \cos(\Phi)$$

$$Q = A \sin(\Phi)$$

$$A = \sqrt{I^2 + Q^2}$$

$$\Phi = \text{atan}(\frac{Q}{I})$$
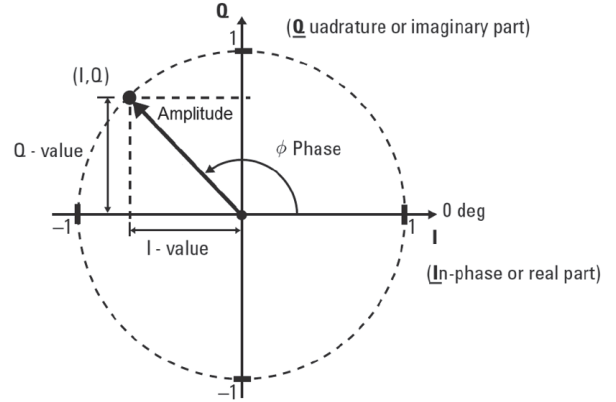
Figure 16: Relation between amplitude and phase and in-phase and quadrature representations [1].

Furthermore, we have to consider the frequency of the RF signal f_{RF} and the sampling frequency f_s. If the frequency of the RF signal is too high, it cannot be reconstructed (aliasing). We have to obey the Nyquist-Shannon theorem. It states that, if

$$f_s > 2f_{RF}$$

a perfect reconstruction of f_{RF} is quarantined. In the case of ILC, the f_{RF} is 1.3 GHz. Typically, ADCs can sample with tenths to a few hundreds of MHz. In order to overcome this issue, the RF signal can be down converted in frequency. In general, the RF signal can be described as:

$$S_{RF}(t) = A_{RF} \cdot \sin(2\pi \cdot f_{RF} \cdot t + \Phi_{RF})$$

For the following let us assume $\Phi_{RF} = 0$. Furthermore, let us assume we create a second signal with a local oscillator with a slightly higher frequency (e.g. f_{LO} = 1.31 GHz). This can described as:

$$S_{LO}(t) = A_{LO} \cdot \sin(2\pi \cdot f_{LO} \cdot t + \Phi_{LO})$$

In order to keep it simple, let us assume $A_{LO} = 1$ and $\Phi_{LO} = 0$. Now we mix both signals, yielding:

$$S_{LO \cdot RF}(t) = \sin(2\pi \cdot f_{RF} \cdot t) \cdot \sin(2\pi \cdot f_{RF} \cdot t)$$

This we can rewrite as:

$$S_{LO \cdot RF}(t) = \frac{1}{2}(\cos(2\pi \cdot (f_{LO} - f_{RF}) \cdot t) - \cos(2\pi \cdot (f_{LO} + f_{RF}) \cdot t))$$

At this point let us assume we apply a low-pass filter, which cuts off the frequency component $f_{LO} + f_{RF}$. The resulting signal is:

$$S_{IF}(t) = \frac{1}{2}(\cos(2\pi \cdot (f_{IF}) \cdot t))$$

where $f_{IF} = f_{LO} - f_{RF}$. IF stands for intermediate frequency. In our example $f_{IF} = 10$ MHz. This frequency can easily be sampled using an ADC. It is very important to note is that this method preserves the original amplitude and phase information. Figure 17 shows a schematic of the down conversion steps.
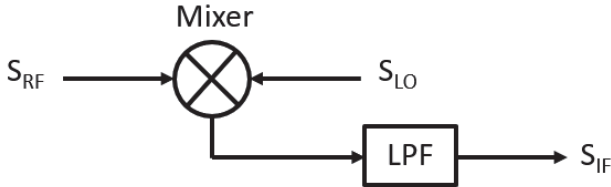
Figure 17: Diagram of the down conversion process.

There are different sample methods. In the following we will discuss IQ sampling as well as under sampling and over sampling.

IQ sampling is the most straight forward. In this case $f_s > 4 \cdot f_{IF}$. This means in terms of I and Q:

$$f_{IF}(0) = Q$$

$$f_{IF}\left(\frac{\pi}{2}\right) = I$$

$$f_{IF}(\pi) = -Q$$

$$f_{IF}\left(\frac{3\pi}{2}\right) = -I$$

We can also identify:

$$\begin{pmatrix} I \\ Q \end{pmatrix}_n = \begin{pmatrix} \cos(\Delta\Phi_n) & -\sin(\Delta\Phi_n) \\ \sin(\Delta\Phi_n) & \cos(\Delta\Phi_n) \end{pmatrix} \cdot \begin{pmatrix} f_{IF,n+1} \\ f_{IF,n} \end{pmatrix}$$

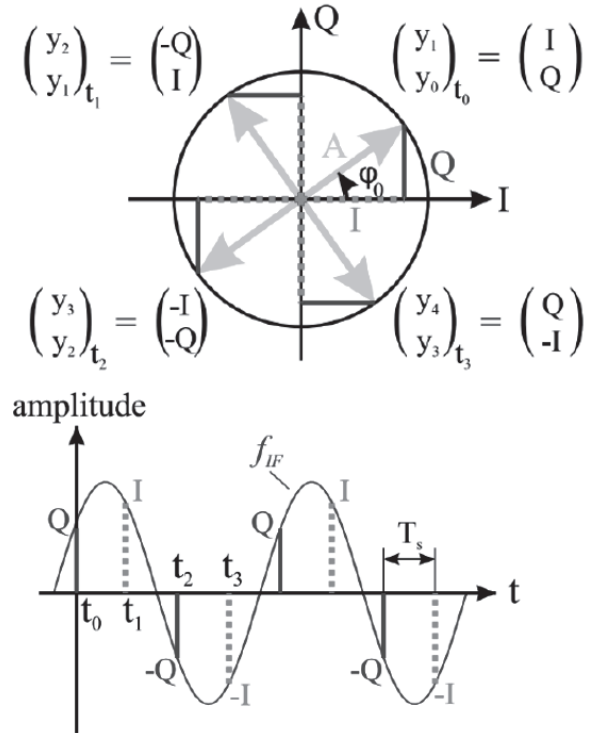This method is also illustrated in Figure 18.

Figure 18: IQ sampling in the complex plane (top) and in time domain (bottom) [8].

Let us write down the general ratio between the sampling frequency and the intermediate frequency as:

$$\frac{f_s}{f_{IF}} = m$$

with $\Delta\Phi = \frac{2\pi}{m}$ (see Figure 19). As we have discussed above, the case of $m = 4$ is called IQ sampling. Furthermore, it is defined that $m < 2$ corresponds to under sampling and $m > 2$ corresponds to over sampling. Thus, IQ
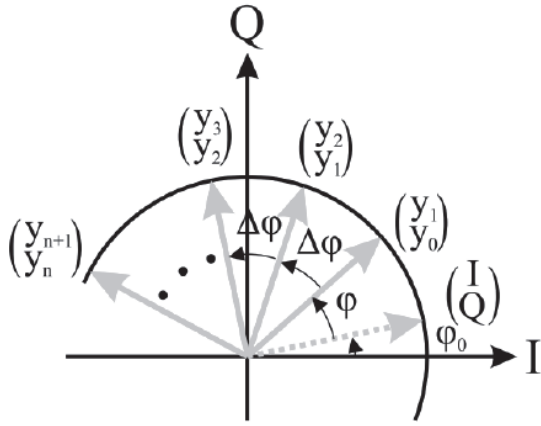
Figure 19: Example of general IQ sampling, here over sampling [8].

sampling is a special case of over sampling. In general, we can compute I and Q with:

$$\begin{pmatrix} I \\ Q \end{pmatrix}_n = \frac{1}{\sin(\Delta\Phi + \Phi)} \begin{pmatrix} \cos(n\Delta\Phi + \Phi) & -\cos((n+1)\Delta\Phi + \Phi) \\ -\sin(n\Delta\Phi + \Phi) & \sin((n+1)\Delta\Phi + \Phi) \end{pmatrix} \cdot \begin{pmatrix} y_{IF,n+1} \\ y_{IF,n} \end{pmatrix}$$

I and Q can also be computed via the following sums:

$$I = \frac{2}{m} \sum_{n=0}^{m-1} y_n \cos\left(\frac{2\pi n}{m}\right)$$

$$Q = \frac{2}{m} \sum_{n=0}^{m-1} y_n \sin\left(\frac{2\pi n}{m}\right)$$

The advantage of under sampling are relaxed requirements for the ADC due to the lower sampling rate. This translates also to relaxed requirements for the FPGA due to the lower data rate. Both may result in cost reductions. Furthermore, in under sampling it is possible to detect IF signals with a higher frequency than the ADC sampling rate.

The advantages of over sampling are e.g. to have more sampling points per IF period. This results in noise reduction due to the averaging in the calculation of the I and Q values. Beside this, the choice of the IF location in the first Nyquist zone is more flexible.

6. Digital Signal Processing and Implementation

At ILC it is planned to drive the cavities in groups of 39 with individual single klystrons. One group of cavities with the corresponding klystron and LLRF system is called an RF station. Since the klystron can receive only one input signal, and its output is distributed to the multiple cavities, a so-called vector-sum control has to implemented within the LLRF system. To this end the probe signal of every cavity of the RF station is send to the LLRF system and is digitized. Then, the I and Q values of the individual signals are summed up, resulting in the vector sum. Figure 20 shows its representation in the complex plane for a case of eight cavities. This calculation is realized on the FPGA.
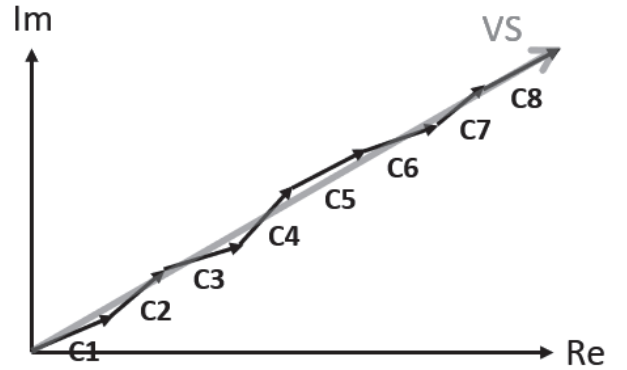

Figure 20: Representation of the vector sum in the complex plane.

An FPGA is chip, in which arbitrary logic circuits can be realized. The major advantage of FPGAs is that they can be reprogrammed. Modern FPGAs feature in addition processor cores, so that demanding computations can be performed without occupying resources in terms of e.g. logical cells. Since the FPGA acts more like an integrated circuit, the time required for algorithms or computations is defined by its design and is always constant. In

other words, FPGAs allow actual real-time algorithms. ("Real-time" does not mean instantaneously. It means the time between the input and the output is always the same.)

Regardless, if working alone or in a team, a structured way of work is essential for the success of a project. For the implementation of the firmware of the FPGA the following workflow is recommended.

First, the requirements of the firmware to be implemented have to be defined and documented. It is useful to make flow charts of algorithms and subcomponents for a better understanding and for checking signal widths. By this signal overflow can be avoided, which otherwise can lead to dangerous situations during operation.

In a next step the code of the firmware is created. There are two major possibilities to approach this. The first is to write the code (e.g. VHDL) by hand. Depending on the project, this may take time. But it allows to control and optimize every aspect of the algorithm. The second way is to use software like e.g. MathWorks Simulink. It allows to build algorithms via a flow chart. From the flowchart, VHDL code is created automatically. This makes quick prototyping possible. The disadvantage is, that one does not have full control over predefined subcomponents. Thus, optimization is impossible in some cases. Furthermore, the automatically generated VHDL code is very hard or sometimes even impossible to be understood by humans. Of course, it is also possible to combine both approaches of code generation.

After the code was generated, a so-called test bench should be programmed. This is additional code, which allows to test the VHDL code in a simulator. During such a test, all possible inputs should be fed to the algorithm and the output should be consistent with the requirements. This way malfunctions of the algorithm in some edge cases can be found and fixed.

After testing within the test bench was successful, the code should be tested and debugged on the target hardware. Target hardware means ideally test hardware, which is identical to the production hardware. This step is very important, since the simulation is only a rough approximation to reality. When the testing and debugging on the target hardware was successful, the code can be deployed to the production hardware.

If requirements for the firmware have changed, the entire implementation and testing workflow has to be passed through again.

7. Controller Theory

The controller is the essential part of the firmware. Its design and setup have a great impact on what RF stability can be achieved in the end. In order to talk about controller theory, we first have to define, what a transfer function is. It is the ratio of the Laplace-transforms of the output signal $Y(s)$ and the input signal $X(s)$ of a linear time-invariant system (see Figure 21). It reads:
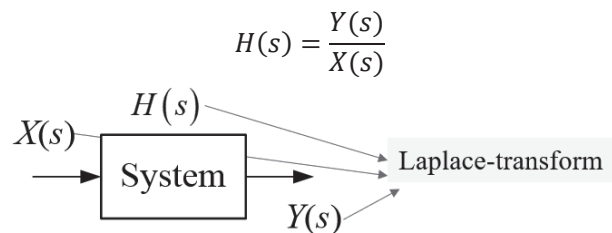
$$H(s) = \frac{Y(s)}{X(s)}$$

Figure 21: Definition of the transfer function [9].

$H(s)$ is a representation of the system, since we can compute the output for a given input.

The Laplace transform reads:

$$H(s) = \int_{t=0}^{+\infty} h(t)e^{-st}\,\mathrm{d}t$$

Note that $h(t)$ is in the time domain and $H(s)$ in the complex frequency domain. The inverse Laplace transform reads:

$$f(t) = L^{-1}\{F(s)\} = \frac{1}{2\pi j}\int_{s=a-j\omega}^{a+j\omega} F(s)\cdot e^{st}\,\mathrm{d}s$$

Figure 22 shows the properties of the Laplace transform as well as Laplace transform pairs.

Figure 22: Properties of Laplace transform and Laplace transform pairs [9].

In order to learn how to obtain a transfer function, let us consider an RC circuit as an example as shown in Figure 23.

Let us write down the current:

$$I(t) = C\cdot\frac{\mathrm{d}u_c(t)}{\mathrm{d}t} = \frac{u(t) - u_c(t)}{R}$$
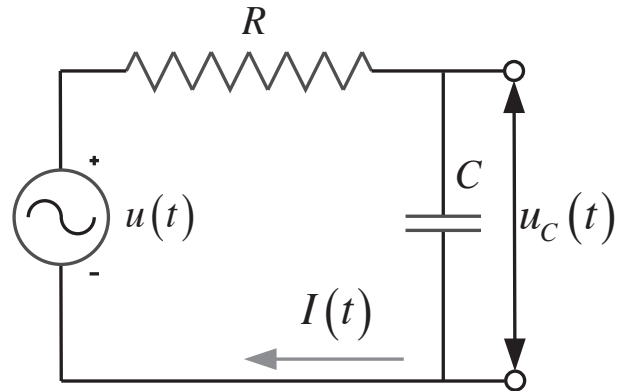

Figure 23: RC circuit [9].

It can be rearranged to:

$$\frac{\mathrm{d}u_c(t)}{\mathrm{d}t} + \frac{u_c(t)}{RC} = \frac{u(t)}{RC}$$

With the Laplace transform $f'(t)\leftrightarrow sF(s)$ we get:

$$sU_c(s) + \frac{U_c(s)}{RC} = \frac{U(s)}{RC}$$

And with this:

$$\frac{U_c(s)}{U(s)} = \frac{\frac{1}{RC}}{s + \frac{1}{RC}} = H(s)$$

which is the transfer function of the RC circuit. From this we can deduct the system output from a given input. Let us assume a unit step $U(s) = \frac{1}{s}$:

$$U_c(s) = \frac{\frac{1}{RC}}{s + \frac{1}{RC}}\frac{1}{s}$$

And from this the system output in time domain, with $RC = \tau$ being the time constant:

$$u_c(t) = \int_0^t \frac{1}{\tau}e^{\frac{-x}{\tau}}\,\mathrm{d}x = \left(e^{\frac{-x}{\tau}}\right)\Big|\begin{matrix}x = t\\ x = 0\end{matrix} = 1 - e^{\frac{-t}{\tau}}$$

If we go back to the differential equation

$$\frac{\mathrm{d}u_c(t)}{\mathrm{d}t} + \frac{u_c(t)}{RC} = \frac{u(t)}{RC}$$

and solve it, we will find that $u_c(t) = 1 - e^{\frac{-t}{\tau}}$.

In the next step we want to transform the transfer function $H(s)$ from the complex frequency domain to the transfer function in frequency domain $H(j\omega)$.

$$H(s)|_{s=j\omega} = H(j\omega) = |H(j\omega)|e^{j\,H(j\omega)} = |H(j2\pi f)|e^{j\,H(j2\pi f)}$$

In this equation we can nicely see the separation of the amplitude in dependence of the frequency $|H(j2\pi f)|$ and the phase in dependence of the frequency $e^{j\,H(j2\pi f)}$. When plotting these two functions, we get something like what is shown in Figure 24. These plots are called Bode plots.
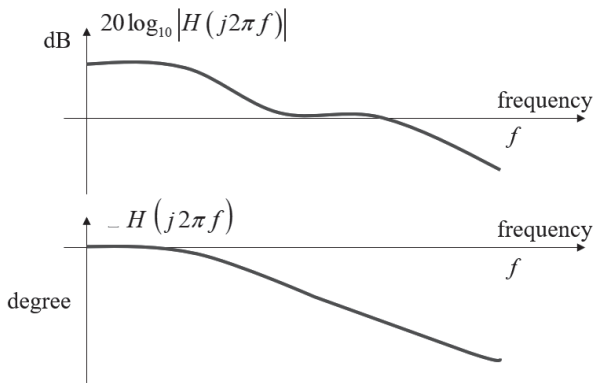

Figure 24: Example of Bode plots [9].

If we come back to the RC example and assume a time constant of 1 ms, we get Bode plots as shown in Figure 25.
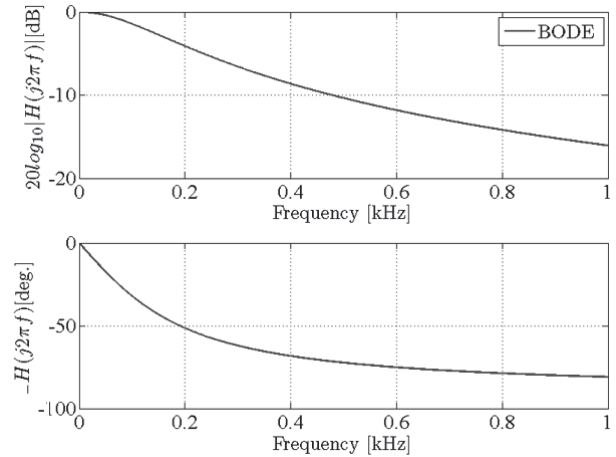

Figure 25: Actual Bode plots for the RC example with a time constant of 1 ms [9].

Let us consider the case of a sinusoidal signal as an input signal. We will find, that e.g. for an input of $\cos(2\pi \cdot 50t)$, we will get an output of $|H(j\omega)|\cos(2\pi \cdot 50t + H(j\omega))$, where $|H(j\omega)|$ is the amplitude of the output signal and $H(j\omega)$ its phase shift. Software like MathWorks Simulink allows us to simulate such a system. In the case shown in Figure 26, the input signal has a frequency of 160 Hz.
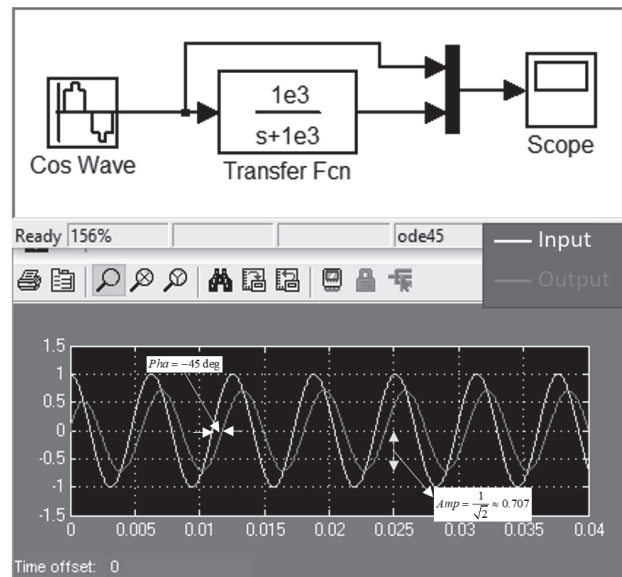

Figure 26: Signal simulation of a 160 Hz sine wave through a RC circuit with a time constant of 1000 ms [9].

From the figure we can clearly see the phase shift, but we can also recognize a change in amplitude. When calculating the output signal, we find that it reads:

$$u_c(t)_{f=160\text{ Hz}} = \frac{1}{\sqrt{2}}\cos\left(2\pi \cdot 160t - \frac{\pi}{4}\right)$$

If we increase the frequency of the input signal to 600 Hz, we will find that:

$$u_c(t)_{f=600\text{ Hz}} = \frac{1}{4}\cos\left(2\pi \cdot 600t - \frac{5\pi}{12}\right)$$

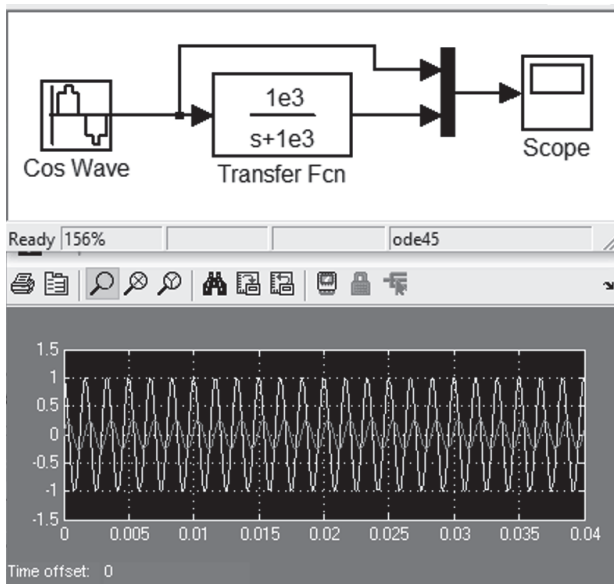with the corresponding simulation shown in Figure 27.

Figure 27: Signal simulation of a 600 Hz sine wave through a RC circuit with a time constant of 1000 ms [9].

This means that the amplitude and the phase shift of the output signal depends on the frequency of the input signal.
We can model a cavity using a RCL circuit, as already seen above. It its case we can write:

$$Z(s) = \frac{\frac{s}{C}}{s^2 + \frac{1}{RC}s + \omega_0^2} = \frac{\frac{R\omega_0}{Q}s}{s^2 + \frac{1}{Q}s + \omega_0^2}$$

If we assume a resonance frequency of 1.3 GHz and a quality factor of $Q = 1.3 \cdot 10^6$, we get the Bode plots as shown in Figure 28. This case corresponds to a SRF cavity.
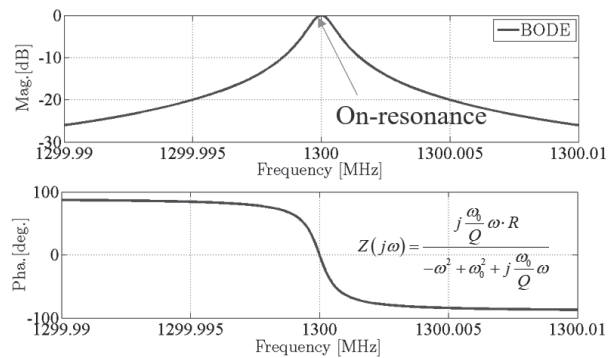

Figure 28: Bode plot for a 1.3 GHz cavity with Q=1.3E6 [9].

In case of a normal conducting cavity, the quality factor is much lower. Let us assume $Q = 7000$ and compare both Bode plots in Figure 29. We can clearly see that the bandwidth of the cavity is related with the quality factor.
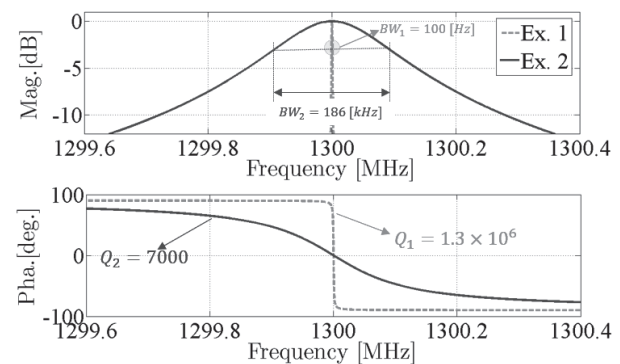

Figure 29: Bode plot for a 1.3 GHz cavity with Q1=1.3E6 and Q2=7E3 [9].

Until now we have considered the transfer function for a single element or device. In the following we want to describe a basic control

system in a transfer function representation, as shown in Figure 30. There $K(s)$ represents the controller, $P(s)$ the plant one wants to control, and $F(s)$ the detector or sensor, which is used to measure the response of the plant.
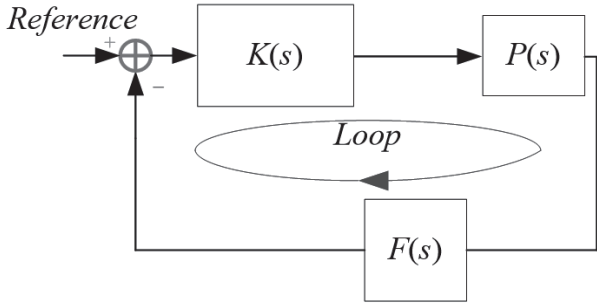

Figure 30: Schematic of a simple control system in a transfer function representation [9].

In order to be able to calculate the transfer function of a whole system, we need to know how to combine the transfer functions of single elements. Figure 31 shows an overview of the rules for the serial, the parallel and the feedback case.
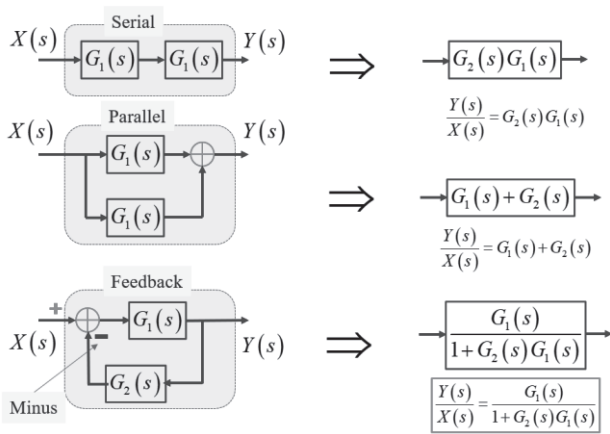

Figure 31: Rules on how to combine single transfer functions [9].

Following these rules, let us calculate the transfer function of the feedback loop shown in Figure 30. To this end, we first combine $K(s)$ and $P(s)$. Then we apply the rule for the feedback system. Figure 32 shows the process step by step.
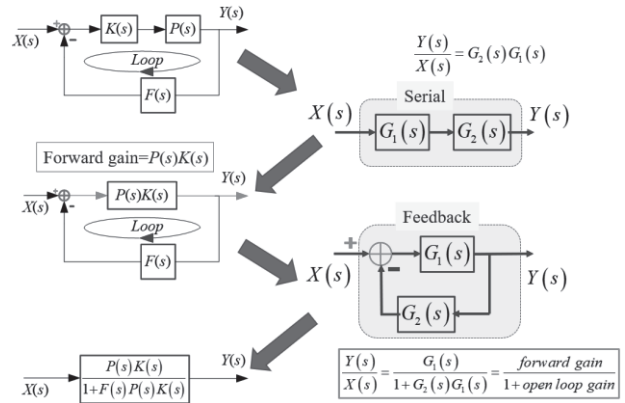

Figure 32: Calculating the transfer function of a simple feedback system [9].

In the real world, feedback systems are much more complicated. It is not uncommon to have systems with several loops as shown in Figure 33.
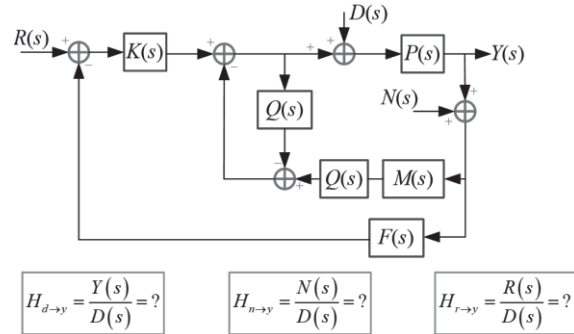

Figure 33: Example of a more complicated feedback system with multiple loops [9].

The transfer function of such a system can be computed by using the Mason's Gain Rule. In it M is the transfer function of the system. M_j is the gain of one forward path. j is an integer representing the forward paths in the system. $\Delta_j = 1-$ the loops remaining after removing path j. If none remain, then $\Delta_j = 1$. $\Delta = 1 + \Sigma$ non-touching loop gains taken two at a time $-\Sigma$ non-touching loop gains taken three at a time $+\Sigma$ non-touching loop gains taken four at a time $-$ etc. The resulting rule reads:

$$M = \frac{\sum_j M_j \Delta_j}{\Delta}$$

In the following let us evaluate the impact of a feedback loop in comparison to the open loop case. To this end we will compare the transfer functions. Until now we have considered ideal systems. In the real world, we have to deal with disturbances. Thus, let us add a disturbance to the simple feedback system and write down the transfer function. Figure 34 shows this for the closed loop case on the left and the open loop case on the right.
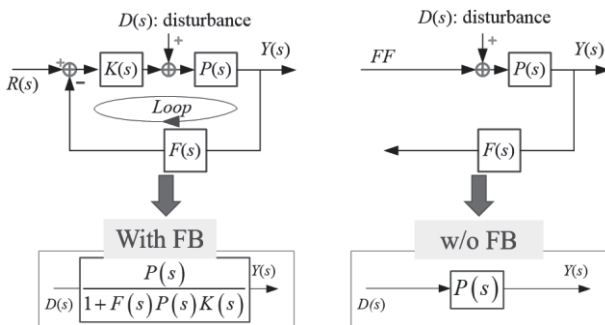

Figure 34: Closed and open loop systems with disturbances [9].

Let us give the plant in both cases some meaning and write down the transfer functions in the complex frequency domain as well as in the frequency domain (see Figure 35).
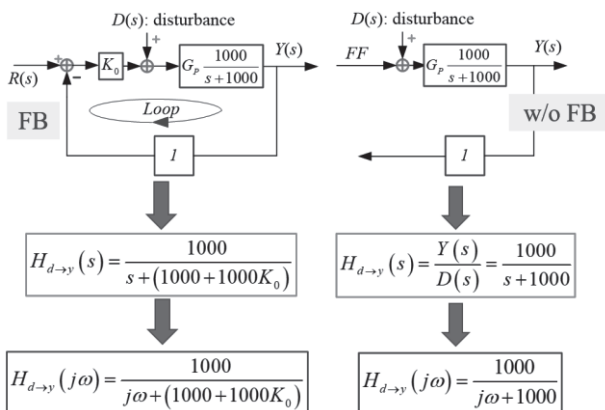

Figure 35: Transfer functions of closed and open loop systems with disturbances [9].

A good way of comparing these results is to look at the Bode plots. Figure 36 shows the closed and open loop cases combined. In the closed loop
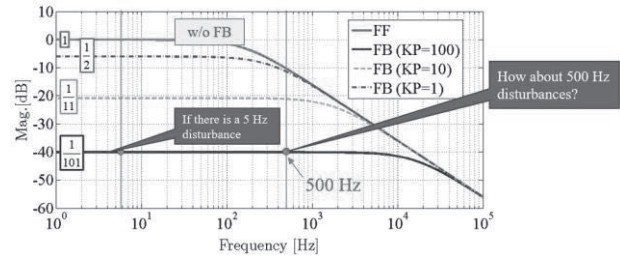

Figure 36: Combined Bode plots of the closed and open loop cases [9].

case, the proportional gain KP is varied. We can clearly recognize an offset of the graphs for $f \rightarrow 0$ Hz. In case of the open loop, there is no disturbance rejection for low frequencies. For higher frequencies we do see a rejection, but this is only due to the chosen low-pass behavior of the plant. In the feedback case, the higher the gain is chosen, the more disturbances are rejected, which holds for lower and also higher frequencies.

Beside the ability to suppress disturbances, a feedback loop has to be stable. There are several stability criteria: e.g. root locus, solving the characteristic equation, the Routh-Hurwitz stability criterion, etc. The most simple and intuitive definition of stability is: A stable system is a dynamic system with a bounded response to a bounded input (see Figure 37).
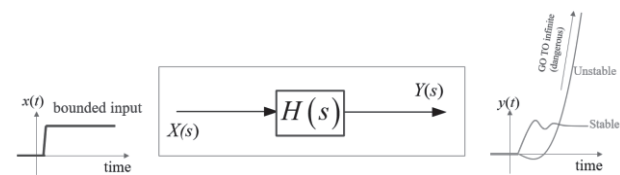

Figure 37: Schematic of a stable and a unstable dynamic system [9].

Since we cannot try all bounded input signals on our system, let us take a look at the characteristic equation. It is the denominator of the transfer function.

$$H(s) = \frac{K(s - z_1)(s - z_2) \dots (s - z_{n-1})(s - z_n)}{(s - p_1)(s - p_2) \dots (s - p_{n-1})(s - p_n)}$$

A necessary and sufficient condition for a feedback system to be stable is that all the poles $(p_1, p_2, p_3, \ldots, p_n)$ of the system transfer function have negative real parts. $(z_1, z_2, z_3, \ldots, z_n)$ are called the zeros of the transfer function. Thus, the following system would be stable:

$$H(s) = \frac{s + 1.5}{s^3 + 4s^2 + 6s + 4}$$
$$= \frac{s - (-1.5)}{[s - (-1 + i)][s - (-1 - i)][s - (-2)]}$$

Three poles: $-1 \pm i$ and -2.
And the following system would be unstable:

$$H(s) = \frac{s + 1.5}{s^3 - 2s^2 + 4}$$
$$= \frac{s - (-1.5)}{[s - (1 + i)][s - (1 - i)][s - (-2)]}$$

Three poles: $1 \pm i$ and -2.
Figure 38 shows the poles and zeros in the pole-zero map. The right half plane poles are unstable.
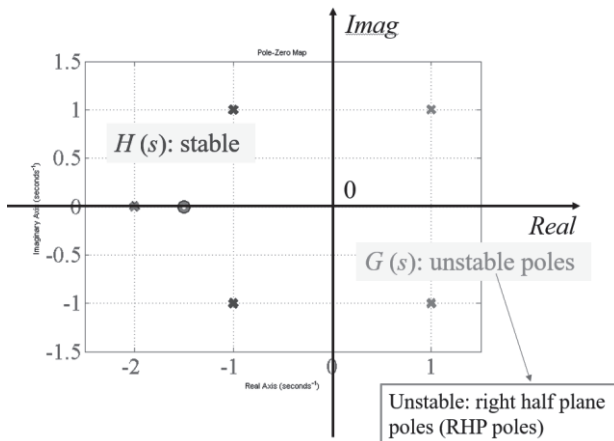

Figure 38: Example of a pole-zero map [9].

In some cases, the characteristic equation can be very complicated and it is not easy to evaluate the poles. In addition, one also wants to know, if a given feedback system is stable at a certain feedback gain. To answer this question, it is popular to evaluate the Bode plot.

The closed loop is stable, if the open loop gain at -180 degree (or -540 degree, etc.) is less than 1 or 0 dB (see Figure 39). Furthermore, by this method it is also possible to evaluate the gain margin, since a larger margin means a better robustness.
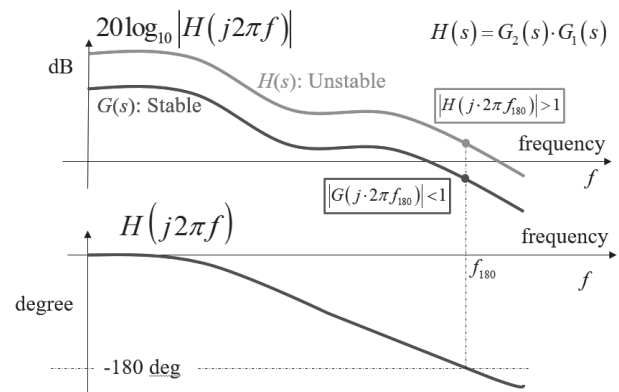

Figure 39: Example bode plot for a stable and unstable system [9].

Another way to evaluate the stability is to plot the Nynquist diagram. It is the plot of the imaginary part versus the real part of the transfer function in the frequency domain. An example is shown in Figure 40.

Bode $\quad H_{OL}(j\omega) = |H_{OL}(j\omega)| e^{j\angle H_{OL}(j\omega)}$
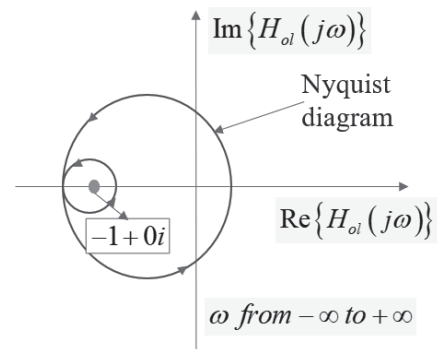Nyquist $\quad H_{OL}(j\omega) = \text{Re}(H_{OL}(j\omega)) + j \cdot \text{Im}(H_{OL}(j\omega))$

Figure 40: Example of a Nyquist diagram [9].

In real-world systems the proportional and integral (PI) controller is very popular. It can be modeled as shown Figure 41. In some cases, a differential part is added, extending it to a PID controller.
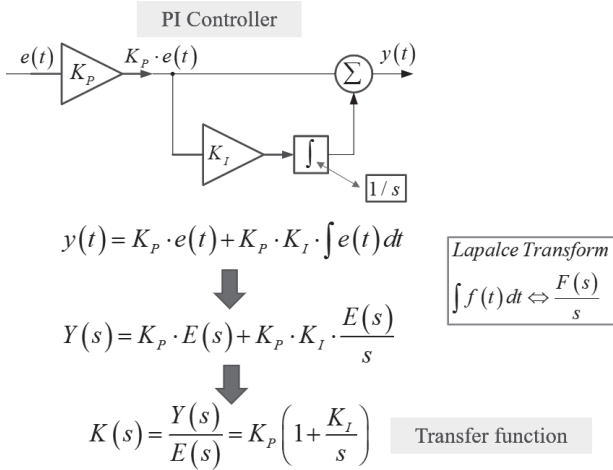

Figure 41: Modelling of the PI controller [9].

A more modern and very useful type of controller is the multiple input multiple output (MIMO) controller (see Figure 42). Beside the PID controller functionality, it can also be used e.g. for the cancellation of cross coupling between inputs as well as for the cancellation of a passband mode.
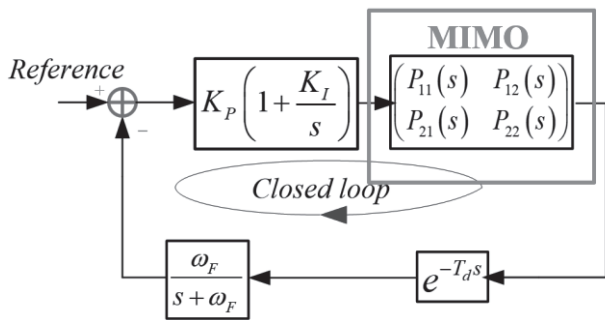

Figure 42: Feedback loop with MIMO controller [9].

8. Example Features of an LLRF System

Typically, an LLRF system is a subsystem of a facility as e.g. a test stand or a large scale particle accelerator. At such facilities a personal protection system (PPS) is required. The goal is to prevent any harm, injury to or death of humans working at the facility of being around it. E.g., if a particle accelerator is being operated, it is forbidden to enter the accelerator tunnel. If someone opens a door to the tunnel during operation, the operation is stopped automatically. In addition to a PPS, especially in larger scale facilities, a machine protection system (MPS) is implemented. Its goal is to prevent the accelerator or in most cases subcomponents from damaging or destroying itself. If e.g. arcing in a klystron is detected, the LLRF system receives a signal to stop the generation of the drive signal. Since in such a case no accelerating field can be generated anymore in the cavities, the energy of transmitted beam would be wrong, resulting in downstream beam loss. In order to prevent this, the beam is also turned off automatically in such a case. The signal triggering the stop of operation of subsystems is called an interlock signal. All LLRF systems must have interlock capability. Since it is imperative to work reliably, such interlock functionality should be hardwired in hardware or firmware. Figure 43 shows a possible solution. Note, that an implementation in software is not reliably enough.
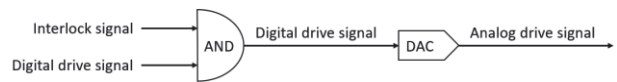

Figure 43: Possible implementation of interlock functionality in firmware.

Beside this, the LLRF system should have a certain degree of exception prevention and handling. A simple example is to implement a limiter on the drive signal. This way e.g. overdriving the high-power amplifier or quenching a cavity can be prevented. In

addition, it is advisable to implement a limiter on the set point amplitude. This way an unexperienced operator of the LLRF system cannot enter values, which could be potentially dangerous. An example for active exception handling would be quench detection. The loaded quality factor Q_L of a cavity can be calculated from the pickup signal during the decay time. In case of a quench, the Q_L value drops significantly. The LLRF system can compute the Q_L value for every RF pulse and when it falls below a predefined threshold, the drive signal is cut, and an interlock signal is generated.

At ILC it is planned to utilize 9-cell SRF cavities for beam acceleration. Since these cavities have nine cells, they have nine different eigenmodes. For beam acceleration only the so-call π-mode is suitable. Thus, the goal is to feed all power only to this mode. All other modes should be suppressed. This can be achieved by e.g. Notch filters at the ADCs for the $\frac{8\pi}{9}$-mode. It is also possible to add an additional filter for the $\frac{7\pi}{9}$-mode in a MIMO controller, as mentioned above.

Another feature typically implemented within the LLRF system is the fast cavity tuning. SRF cavities are built and assembled at room temperature, but operated at 2 K. During the cooldown the cavity is typically relaxed in order to prevent any possibilities of unwanted deformations. The cavity was pre-tuned at room temperature to roughly reach the design frequency at 2 K. In order to actually reach the design frequency, the cavity can be elastically deformed in a controlled manner, using motor tuners. These motors are slow stepper motors and are left in position once the cavity has been tuned. Beside the motor tuners, the cavity is also equipped with a piezo tuner. This is used for the compensation of fast frequency changes, induced by e.g. Lorentz force detuning or

microphonics. Figure 44 shows pictures of three different piezo configurations.
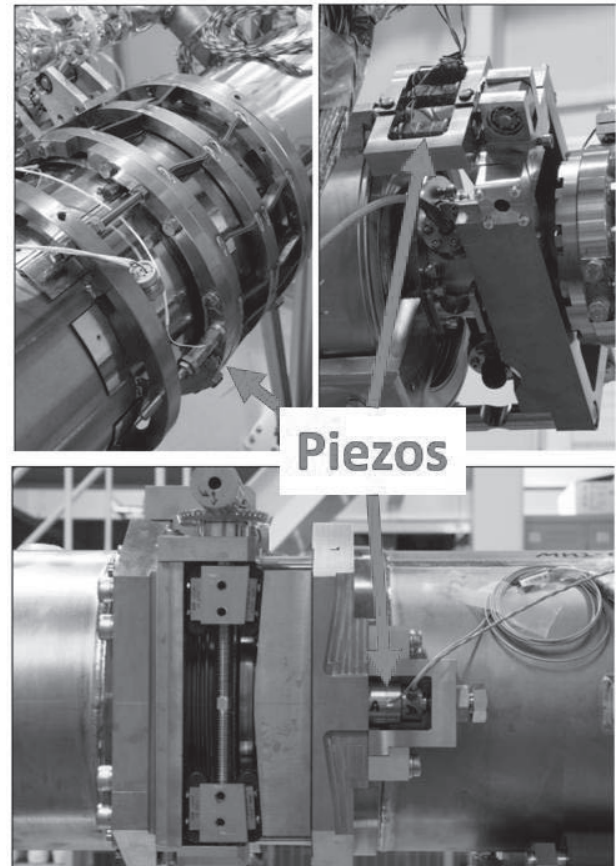

Figure 44: Three different of piezo configurations on cavities during the S1-global project [10].

As stated at the beginning, the goal of the LLRF system is to control the amplitude and phase within cavities to be better than the required stability. Thus, a LLRF system should have the feature of computing the RF stability. Depending on the facility, different kinds of stability have to be computed. In a single cavity control scheme, of course the stability of the single cavity is important. For the amplitude the stability is typically defined as $\Delta A/A$, where ΔA is the standard deviation of the amplitude and A the average of the amplitude. For the phase $\Delta\Phi$ is used. These values are computed only for a certain time period, as e.g. for the

flattop region or the beam transient time. In this case the stability is called intra train stability. In case of vector sum control, the stability of the vector sum amplitude and phase is computed in the same way. It is also possible to define a measure for the stability from pulse to pulse. This is called the inter train stability.

9. Summary

Due to the limited length of this lecture, it is impossible to show and explain all important aspects of LLRF in detail. It is a subject one has to study for an extended time. Nevertheless, with the overview given, everyone who is planning to get involved with LLRF should now understand what he or she should learn. First, learn about the target facility. What are the requirements for e.g., stability, etc.? Plan carefully, how the LLRF system can be integrated. Second, learn the required theoretical background as e.g. cavity theory, RF theory, signal processing theory, controller theory, etc. Third, learn about analog hardware. Fourth, learn about digital hardware. Fifth, learn about firmware. Sixth, learn about software.

Every of these topics is an individual field of expertise. But as an LLRF expert, one has to know the most important aspects of every topic in order to design, implement, commission, and operate an LLRF system.

References

[1] T. Schilcher, *Vector Sum Control of Pulsed Accelerating Fields in Lorentz Force Detuned Superconducting Cavities*, PhD Thesis, 1998

[2] M. Omet, *Digital Low Level RF Control Techniques and Procedures Towards the International Linear Collider*, PhD Thesis, 2014

[3] S. Pfeiffer, *LLRF Controls and Feedback*, The CERN Accelerator School, 2016

[4] L. Doolittle et al., *High Precision RF Control for SRF Cavities in LCLS-II*, SRF2021, 2017

[5] J. Branlard et al., *Installation and First Commissioning of the LLRF System for the European XFEL*, IPAC17, 2017

[6] T. Miura et al., *Digital LLRF Control System for cERL*, 13th Annual Meeting of Particle Accelerator Society of Japan, 2016

[7] https://pt.wikipedia.org/wiki/Sinal_digital#/media/Ficheiro:Digital.signal.svg

[8] T. Schilcher, *Applications in Digital Signal Processing,* CAS – CERN Accelerator School: Course on Digital Signal Processing, Sigtuna, 2007

[9] F. Qiu, *Tutorial on control theory*, 10th International Accelerator School for Linear Colliders, Japan, 2016

[10] C. Pagani et al., *Test results of the international S1-global cryomodule*, Proceedings of the 8th Annual Meeting of Particle Accelerator Society of Japan, 2011