

RF Theory and Design - Notes

Jeremiah Holzbauer Ph.D.

USPAS - Grand Rapids
June 2012

1 Introduction

This document is a short summary of the theory covered in the USPAS class Applied Electromagnetism for Accelerator Design. This is a living document, and will be expanded at a future date, please excuse small errors and sudden transitions. This document is Version 1.

2 Resonator Theory

This section will provide insights for use in further sections by producing a generic formalism to describe resonators with the explicit goal of treating resonating cavities with the same mathematical treatment. Starting from the simple harmonic oscillator, damping and driving terms will be added, and their effects derived. The treatment of electromagnetic resonators by this formalism will then be justified, and a special case of interest will be presented.

2.1 The Ideal Harmonic Oscillator

The ideal, linear oscillator is the most basic starting point for solving many problems in physics. This is especially true in accelerator physics because so many of the phenomena of accelerator physics can be treated as purely classical and having few confounding effects. The harmonic oscillator, without damping, has the form

$$\frac{d^2x}{dt^2} + \omega_0^2 x = 0 \quad (1)$$

where $x(t)$ is the oscillating position at time t , and ω_0 is the oscillation frequency. The general solution for this is characterized by an amplitude A and a phase ϕ

$$x(t) = A \cos(\omega_0 t + \phi). \quad (2)$$

2.2 The Driven, Damped Harmonic Oscillator

For this to be a useful model we must consider damping. The form of damping that is of interest to us (because it is the form of the losses in a resonator), is damping proportional to the change in “position”, which has a form of

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x = 0, \quad (3)$$

where the damping coefficient γ has the dimension of frequency. This type of equation has different forms of solution depending on the strength of the damping. If we choose the damping to be weak, the general solution to this equation has the form

$$x(t) = Ae^{-\gamma_1 t} \cos(\omega_1 t + \phi) \quad (4)$$

where

$$\omega_1 = \omega_0 \sqrt{1 - \frac{\gamma^2}{4\omega_0^2}} \quad (5)$$

and

$$\gamma_1 = \frac{1}{2}\gamma. \quad (6)$$

Note that adding damping shifts the resonator frequency based on how strong the damping is. From this we can define a “quality factor” Q which is related to the rate at which the resonator loses energy. Q is defined by the equation $Q = \omega_0/2\gamma$ with weak damping characterized by $Q \gg 1$. For a typical superconducting cavity, Q is generally $\sim 5 \times 10^9$, justifying the assumption of weak damping.

The effects of a driving term must also be considered. For an arbitrary driving term of the form $f(t)$, the differential equation becomes

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x = f(t), \quad (7)$$

with the solution (for $\gamma = 0$)

$$x(t) = x_0 \cos \omega_0 t + \frac{\dot{x}_0}{\omega_0} \sin \omega_0 t + \frac{1}{\omega_0} \int_0^t \sin(\omega_0(t-t')) f(t') dt' \quad (8)$$

where x_0 and \dot{x}_0 are the initial position and velocity. The third term of this solution gives the contribution from the driving term, and it is worthwhile to notice that a harmonic driving term with frequency equal to the resonant frequency of the oscillator will produce the largest oscillations, as expected.

2.3 Oscillator Behavior Near Resonance

It is most useful to examine the response of the weakly damped oscillator to driving near its resonant frequency because this is the desired case for resonator operation. Assuming a driving term that is purely harmonic ($f(t) = f_0 \cos \omega t$) and ignoring transient behavior yields a solution with a relatively simple form. These assumptions are very justified in almost every accelerator application; most changes made to the driving term of the cavity have a time scale much larger than the RF period (which is usually on the order of nanoseconds). This type of differential equation is solved quite simply by assuming a complex solution and writing

$$\frac{d^2 \Xi}{dt^2} + \gamma \frac{d\Xi}{dt} + \omega_0^2 \Xi = f_0 e^{-i\omega t} \quad (9)$$

with $x(t) = \Re \Xi(t)$. We are seeking a solution of the form $\Xi(t) = \Xi_0 e^{-i\omega t}$ where Ξ_0 is also a complex number of the form $\Xi_0 = |\Xi_0| e^{i\phi}$. Solving for the real variable of interest x gives $x(t) = \Re \Xi(t) = \Re(|\Xi_0| e^{-i\omega t + i\phi}) = |\Xi_0| \cos(\omega t - \phi)$. Plugging this form of the solution into the differential equation gives

$$(-\omega^2 - i\omega\gamma + \omega_0^2) \Xi_0 = f_0 \quad (10)$$

simplifying to

$$\Xi_0 = \frac{f_0}{\omega_0^2 - \omega^2 - i\omega\gamma} \quad (11)$$

with squared amplitude of

$$|\Xi_0|^2 = \frac{f_0^2}{(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2}. \quad (12)$$

The resulting behavior can be seen in Figure 1 for a variety of damping coefficients γ [1]. Again, it is easy to see that the maximum response will be shifted slightly depending on the strength of the damping. This shift can be neglected for $\gamma \ll \omega_0$, and as demonstrated earlier, this is a very good approximation for superconducting resonators. Another important feature of these curves is the characteristic width of each curve. This width ($\Delta\omega$), defined as the width at the level that is 3 dB below the maximum response, is equal to 2γ where γ is the damping parameter. An alternative and equivalent definition of the Quality Factor is $Q = \omega/\Delta\omega$.

The phase response of Ξ is also of interest. This phase ϕ can be interpreted as the difference in phase between the driving term and the response of the resonator, and is an important quantity for resonator control.

$$\tan(\phi) = \frac{\Im \Xi_0}{\Re \Xi_0} = \frac{\omega\gamma}{\omega_0^2 - \omega^2} = \left(\frac{\omega}{\omega_0}\right) \frac{1}{1 - \left(\frac{\omega}{\omega_0}\right)^2} \left(\frac{\gamma}{\omega_0}\right) \quad (13)$$

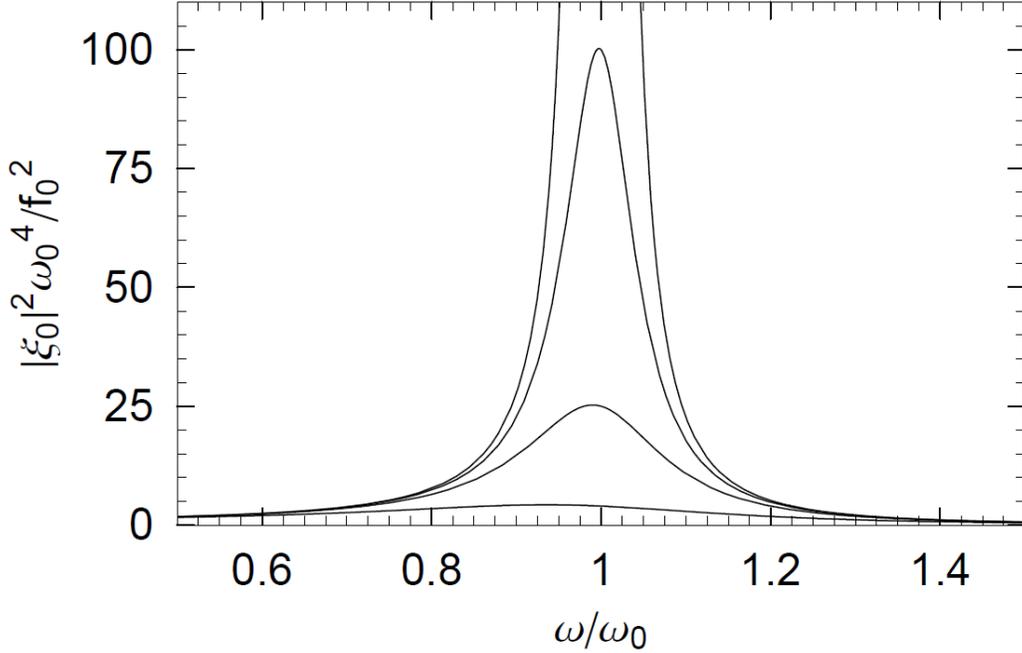


Figure 1: Resonant curves for damped, driven harmonic oscillator with $\gamma = 0.5, 0.2, 0.1, 0$

A plot of ϕ can be seen in Figure 2. The most important feature of this behavior is the nearly linear region near resonance. Most cavity control systems treat the cavity response as linear and must operate in this region to remain stable.

2.4 Special Case: The Duffing Equation

As we shall see, certain non-linear behavior in superconducting resonators in operation can be well modeled by the Duffing Equation. The Duffing Equation adds a weakly non-linear term to the restoring force such that

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + \omega_0^2 x + \alpha x^3 = f_0 \cos \omega t. \quad (14)$$

For the purposes of this application, it can be assumed both the damping, driving, and non-linear terms are small compared to the frequency ω_0 . Additionally, we will only look for solutions where $\omega \simeq \omega_0$. Using the standard van der Pol transformation [2], seen in Figure 3, we transform into a rotating coordinate frame. Using the transformations

$$u = x \cos \omega t - \frac{\dot{x}}{\omega} \sin \omega t \quad (15)$$

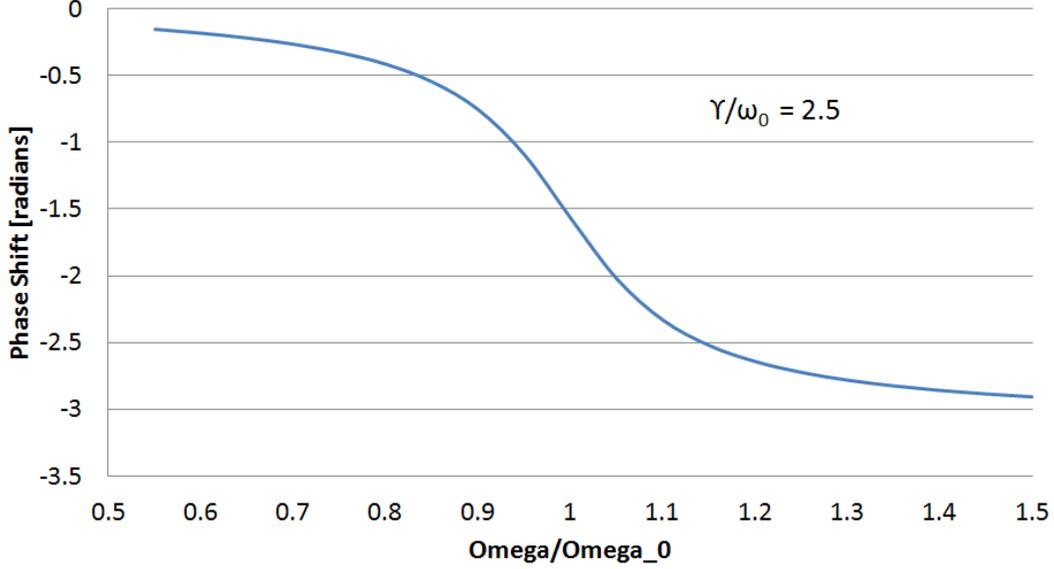


Figure 2: Oscillator phase shift compared to driving term versus detuning. Note the nearly linear region near resonance.

and

$$v = -x \sin \omega t - \frac{\dot{x}}{\omega} \cos \omega t, \quad (16)$$

we arrive at the following differential equations for u and v :

$$\begin{aligned} \dot{u} = \frac{1}{\omega} [& -(\omega^2 - \omega_0^2)(u \cos \omega t - v \sin \omega t) - \omega\gamma(u \sin \omega t + v \cos \omega t) \\ & + \alpha(u \cos \omega t - v \sin \omega t)^3 - f_0 \cos \omega t] \sin \omega t \end{aligned} \quad (17)$$

$$\begin{aligned} \dot{v} = \frac{1}{\omega} [& -(\omega^2 - \omega_0^2)(u \cos \omega t - v \sin \omega t) - \omega\gamma(u \sin \omega t + v \cos \omega t) \\ & + \alpha(u \cos \omega t - v \sin \omega t)^3 - f_0 \cos \omega t] \cos \omega t. \end{aligned} \quad (18)$$

Because we are assuming small non-linearities and constant frequency (ω), we are only interested in the average behavior of these functions. Averaging over a period of $2\pi/\omega$, we get

$$\dot{u} = \frac{1}{2\omega} \left(-\omega\gamma u + (\omega^2 - \omega_0^2)v - \frac{3}{4}\alpha(u^2 + v^2)v \right) \quad (19)$$

and

$$\dot{v} = \frac{1}{2\omega} \left(\omega\gamma v - (\omega^2 - \omega_0^2)u + \frac{3}{4}\alpha(u^2 + v^2)u - f_0 \right). \quad (20)$$

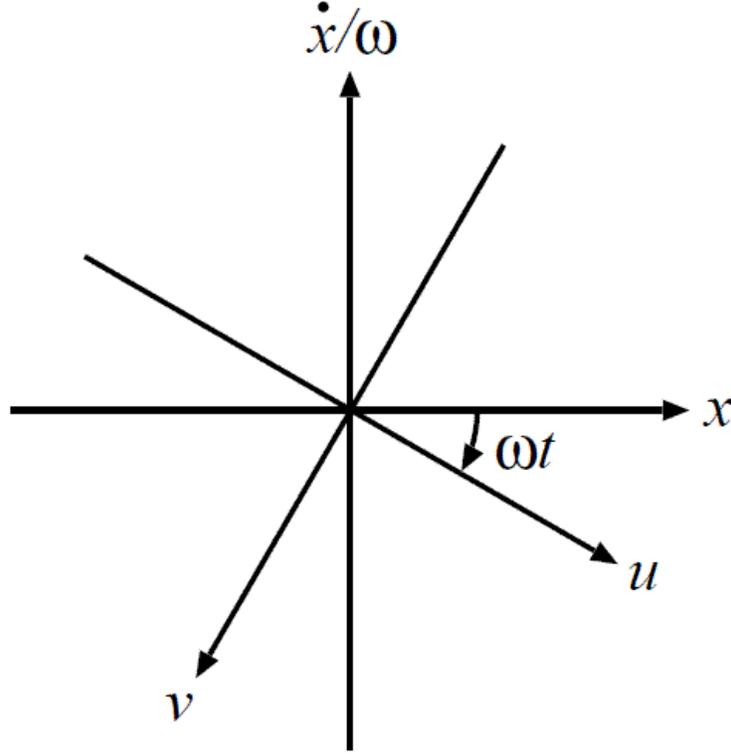


Figure 3: Van der Pol coordinate transformation into a frame rotating at ω .

In reality, these are not the values of interest for our resonator. For more meaningful results, we again transform these quantities into the magnitude ($r = \sqrt{u^2 + v^2}$) and phase ($\phi = \arctan(v/u)$) in the rotating frame. This results in the following differential equations:

$$\dot{r} = \frac{1}{2\omega} (-\omega\gamma r - f_0 \sin \phi) \quad (21)$$

and

$$\dot{\phi} = \frac{1}{2\omega} \left(-(\omega^2 - \omega_0^2)r + \frac{3}{4}\alpha r^3 - f_0 \cos \phi \right). \quad (22)$$

Again, we can simplify by assuming a steady state solution: $\dot{r} = \dot{\phi} = 0$. The resulting amplitude and phase plots can be seen in Figures 4 and 5, and the non-linear response is evident. Because of this detuning, both the amplitude and phase become multi-valued for certain ranges of detuning, and if the effect is severe enough, controlling the resonator behavior can become prohibitively challenging.

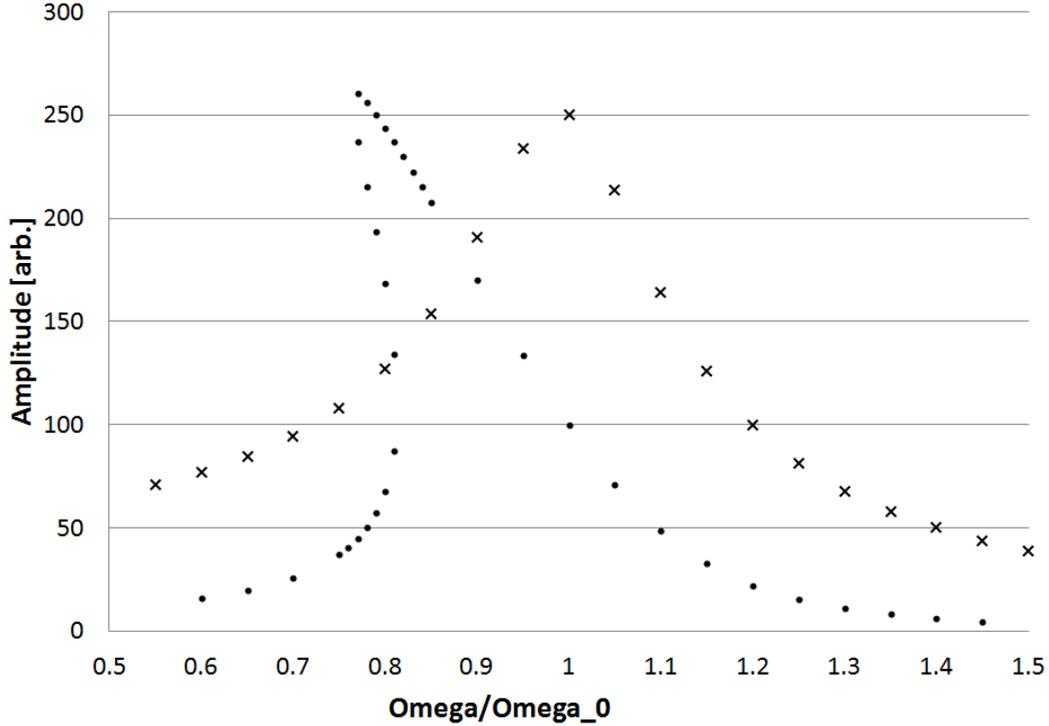


Figure 4: Oscillator amplitude versus detuning. Circles represent $\alpha = -0.002$, crosses represent $\alpha = 0$.

2.5 Electromagnetic Fields as a Harmonic Oscillator

To this point, I have focused on the formalism for a generic harmonic oscillator with damping and driving terms. Occasionally, I have motivated certain approximations and simplifications as relevant to accelerating resonators without proof, but now I wish to demonstrate the applicability of this formalism to electromagnetic resonators in preparation for the rest of the document. For a (much) more rigorous treatment of this material, the reader is directed to Chapter 7 of J.D. Jackson’s excellent book “Classical Electrodynamics” [3]. Maxwell’s Equations govern all electromagnetic behavior. In free space, these equations are

$$\nabla \cdot B = 0 \tag{23}$$

$$\nabla \times E + \frac{\partial B}{\partial t} = 0 \tag{24}$$

$$\nabla \cdot D = 0 \tag{25}$$

$$\nabla \times H - \frac{\partial D}{\partial t} = 0 \tag{26}$$

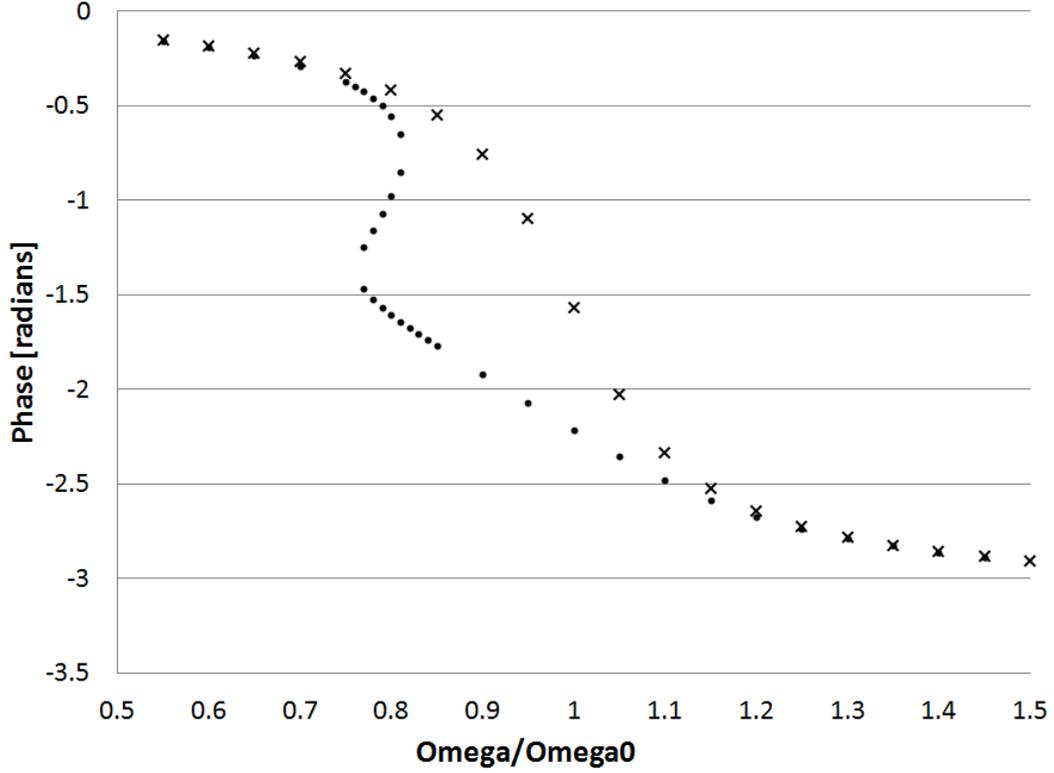


Figure 5: Oscillator phase versus detuning. Circles represent $\alpha = -0.002$, crosses represent $\alpha = 0$.

where E is the electric field, B is the magnetic field, and D (equal to $\epsilon_0 E$ in this case) is the electric displacement field. Assuming a harmonic time dependence of the form $e^{-i\omega t}$ gives

$$\nabla \cdot B = 0 \tag{27}$$

$$\nabla \times E - i\omega B = 0 \tag{28}$$

$$\nabla \cdot D = 0 \tag{29}$$

$$\nabla \times H + i\omega D = 0 \tag{30}$$

From this form, we can construct an arbitrary solution using Fourier superposition. Assuming that we are in a region where there are no losses (μ and ϵ are real and positive), Maxwell's Equations reduce to

$$\nabla \times E - i\omega B = 0 \tag{31}$$

$$\nabla \times B + i\omega\mu\epsilon E = 0 \tag{32}$$

where the divergence equations can be found from taking the divergences of these two equations. Combining these two equations we get the Helmholtz wave equations:

$$\nabla^2 E + \mu\epsilon\omega^2 E = 0 \quad (33)$$

$$\nabla^2 B + \mu\epsilon\omega^2 B = 0 \quad (34)$$

A solution to equations of this form is a plane wave of the form $e^{ikx-i\omega t}$, where $k = \sqrt{\mu\epsilon}\omega$ is the wave number. While solutions of this form are discussed many places at great length, the harmonic behavior allows us to use the same differential formulations to treat the resonator. For the remainder of this document, we will only consider harmonic behavior in vacuum where $\epsilon = \epsilon_0$ and $\mu = \mu_0$ with $c^2\epsilon_0\mu_0 = 1$.

3 A Simple Electromagnetic Resonator - The Pillbox Cavity

This section will provide a treatment of the most basic resonator shape, the pillbox cavity. The geometry and electromagnetic mode structure will be motivated and the performance of this cavity as an accelerating resonator will be quantified. Derivation and description of generic cavity figures of merit will be provided and derived for this geometry.

3.1 Cylindrical Waveguide Modes

Electromagnetic fields assume a certain form near a perfect conductor. Because a perfect conductor allows no field penetration into the surface and can sustain any currents required to expel the field lines, the fields at the surface of the perfect conductor must satisfy:

$$\hat{n} \times \vec{E} = 0, \quad \hat{n} \cdot \vec{H} = 0 \quad (35)$$

where \hat{n} is the conductor surface's normal vector. Let us consider a cylindrical waveguide of radius R that is infinite in length. Defining our cylindrical coordinate system in the usual way, we can use the generalized plane wave solution to the Helmholtz equation derived in the previous section, we can assume the fields take the form

$$\vec{E}(\vec{x}, t) = \vec{E}(\rho, \phi)e^{ikz-i\omega t} \quad (36)$$

$$\vec{H}(\vec{x}, t) = \vec{H}(\rho, \phi)e^{ikz-i\omega t}. \quad (37)$$

For the purposes of an accelerating cavity, it is desired to have longitudinal electric fields with no longitudinal magnetic fields, i.e., $B_z = 0$. Additionally, we know that the dependence of the fields on ϕ must be periodic, so we assume a dependence of the form $e^{im\phi}$ where m is an integer. This gives us a form for the electric field of

$$E_z(\rho, \phi, z, t) = f(\rho)e^{ikz-i\omega t-im\phi} \quad (38)$$

which we can plug into the wave equation:

$$\frac{1}{c^2} \frac{\partial^2 F}{\partial t^2} = \nabla^2 F \implies \frac{1}{\rho} \frac{d}{d\rho} \rho \frac{df}{d\rho} - \frac{m^2}{r^2} f + \left(\frac{\omega^2}{c^2} - k^2 \right) f = 0 \quad (39)$$

to get the form of $f(\rho)$. The solution of this equation is given by standard Bessel functions $f(\rho) = E_0 J_m(k_\perp \rho)$ where m is the order of the Bessel function and

$$k_\perp = \frac{1}{c} \sqrt{\omega^2 - c^2 k^2}. \quad (40)$$

To satisfy the boundary conditions at $\rho = R$, E_z must be zero. This requires that $k_\perp R$ is equal to one of the Bessel zeros, $j_{m,n}$ (where n is an integer indicating the zero), giving $k_\perp = j_{m,n}/R$. Eliminating k_\perp from our equations gives our definition of the wave number k for this geometry:

$$k_{m,n} = \pm \left(\frac{\omega^2}{c^2} - \frac{j_{m,n}^2}{R^2} \right)^{1/2}. \quad (41)$$

It is of interest to note that this gives a lower bound on the frequency for a traveling wave solution. For $\omega j_{m,n} c/R$, the wave number becomes imaginary, resulting in a solution known as an “evanescent” wave. This frequency is known as the “cut-off” frequency, with larger diameter waveguides able to support lower frequencies. This wave exponentially decays/increases as it propagates, and is of interest in the study of resonators only for study of mode propagation into and out of the resonator through beam pipes or other ports. For the remainder of this section, we will only consider traveling wave solutions. Now that the E_z has been defined and remembering that $B_z = 0$, we can use $\nabla \times E = i\omega B$ and $c^2 \nabla \times B = -i\omega E$ to find the remaining components of both fields: E_ϕ , E_ρ , B_ϕ , B_ρ . These component are

$$E_\rho = E_0 \frac{ik_{m,n} R}{j_{m,n}} J'_m \left(j_{m,n} \frac{\rho}{R} \right) e^{ik_{m,n} z - i\omega t - im\phi} \quad (42)$$

$$E_\phi = -E_0 \frac{mk_{m,n} R^2}{\rho j_{m,n}^2} J_m \left(j_{m,n} \frac{\rho}{R} \right) e^{ik_{m,n} z - i\omega t - im\phi} \quad (43)$$

$$B_\rho = E_0 \frac{m\omega R^2}{c^2 \rho j_{m,n}^2} J_m \left(j_{m,n} \frac{\rho}{R} \right) e^{ik_{m,n} z - i\omega t - im\phi} \quad (44)$$

$$B_\phi = E_0 \frac{i\omega R}{c^2 j_{m,n}} J'_m \left(j_{m,n} \frac{\rho}{R} \right) e^{ik_{m,n} z - i\omega t - im\phi} \quad (45)$$

where J'_m is the derivative of the m th order Bessel function. Standard convention designates these modes as TM_{mn} for Transverse Magnetic modes where the index m is the azimuthal order and index n is the radial order. It is worth noting that the boundary condition

$E_\phi(\rho = R) = 0$ is automatically satisfied because this component has the same radial dependence as E_z .

This derivation for *TM* modes is based on the condition that $B_z = 0$. One can follow the same approach starting with the condition of $E_z = 0$ to obtain another family of modes, Transverse Electric (*TE*) modes. As these are not useful for accelerator applications, they are beyond the scope of this work.

3.2 From Traveling Waves to Standing Waves

The goal of this section is to describe a cavity, not a waveguide. To this end, we will take a section of waveguide and terminate it with conducting faces at $z = 0$ and $z = L$. In order to satisfy these additional boundary conditions with *TM* modes, we take two modes with equal m and n but with opposite values of k . This is equivalent to combining identical waves traveling in opposite directions. The equations

$$\frac{1}{2}(e^{ik_{m,n}z} + e^{-ik_{m,n}z}) = \cos(k_{m,n}z) \quad (46)$$

$$\frac{1}{2}(k_{m,n}e^{ik_{m,n}z} - k_{m,n}e^{-ik_{m,n}z}) = ik_{m,n} \sin(k_{m,n}z) \quad (47)$$

can be plugged into the full field components, and both E_ρ and E_ϕ gain a factor of $\sin(k_{m,n}z)$. This means that the boundary condition requiring these components to be zero at the $z = 0$ plane is automatically satisfied, and the boundary condition at $z = L$ can be satisfied by requiring $k_{m,n}L = l\pi$ where l is an integer index that gives the longitudinal mode order. It is of interest to note that these new boundary conditions are now also satisfied for the magnetic fields. The resulting field components are

$$E_z = E_0 J_m \left(j_{m,n} \frac{\rho}{R} \right) \cos \left(\frac{l\pi z}{L} \right) e^{-i\omega t - im\phi} \quad (48)$$

$$E_\rho = -E_0 \frac{l\pi R}{j_{m,n} L} J'_m \left(j_{m,n} \frac{\rho}{R} \right) \sin \left(\frac{l\pi z}{L} \right) e^{-i\omega t - im\phi} \quad (49)$$

$$E_\phi = -E_0 \frac{iml\pi R^2}{\rho j_{m,n}^2 L} J_m \left(j_{m,n} \frac{\rho}{R} \right) \sin \left(\frac{l\pi z}{L} \right) e^{-i\omega t - im\phi} \quad (50)$$

$$B_\rho = E_0 \frac{m\omega R^2}{c^2 \rho j_{m,n}^2} J_m \left(j_{m,n} \frac{\rho}{R} \right) \cos \left(\frac{l\pi z}{L} \right) e^{-i\omega t - im\phi} \quad (51)$$

$$B_\phi = E_0 \frac{i\omega R}{c^2 j_{m,n}} J'_m \left(j_{m,n} \frac{\rho}{R} \right) \cos \left(\frac{l\pi z}{L} \right) e^{-i\omega t - im\phi}. \quad (52)$$

Note that this also changes the form of the wave number, resulting in a discrete dispersion relation instead of a continuous curve:

$$\frac{\omega_{m,n,l}^2}{c^2} = \pm \left[\left(\frac{l\pi}{L} \right)^2 + \frac{j_{m,n}^2}{R^2} \right]^{1/2}. \quad (53)$$

This shift to a discrete set of modes (conventionally referred to as TM_{mnl} modes) can be repeated for TE modes as well.

3.3 The “Fundamental Mode” and its Properties

For use as an accelerating cavity, the TM_{010} mode (called the “fundamental mode”) gives the best performance. The justification for this is simple when remembering that our desire was to have only the accelerating electric field component be non-zero near the beam axis. The field components for this mode are very simple:

$$E_z = E_0 J_0 \left(\frac{2.405\rho}{R} \right) e^{-i\omega t} \quad (54)$$

$$H_\phi = \frac{E_0}{\eta} J_1 \left(\frac{2.405\rho}{R} \right) e^{-i\omega t + i3\pi/2} \quad (55)$$

where

$$\eta = \sqrt{\frac{\mu_0}{\epsilon_0}} \approx 376.7 \, \Omega \quad (56)$$

is the impedance of free space. The frequency for this mode does not depend on the length of the cavity L , reducing to

$$\omega_{010} = \frac{2.405c}{R}. \quad (57)$$

For this mode to successfully accelerate beam, ports must be added for the beam to enter and exit the cavity. The rest of this section will not include these ports, as they are extremely challenging to model analytically. This is a perturbation best studied by computer simulations.

3.4 The Interaction of Conductors and Electromagnetic Fields

There is no material that will not dissipate power when an RF field is applied to its surface. The losses in a cavity are given by

$$\frac{dP}{dA} = \frac{1}{2} R_s |H|^2 \quad (58)$$

where dP/dA is the dissipated power per unit area and H is the local magnetic field. The remaining quantity is the surface resistivity. For a normal conducting material, this can be on the order of $m\Omega$, and for a high quality superconducting surface, this can be on the order of $n\Omega$. These losses are generated by the surface currents induced by the changing magnetic fields in the resonator. Remembering that, for an RF cavity,

$$E = E_0 e^{-i\omega t} \quad (59)$$

and defining the electrical conductivity σ such that

$$\vec{j} = \sigma \vec{E} \quad (60)$$

where \vec{j} is the surface current, we can show that, for a good conductor,

$$\nabla^2 \vec{E} = \tau_n^2 \vec{E} \quad (61)$$

where $\tau_n^2 = \sqrt{i\omega\sigma\mu_0}$. For a simple geometry, such as an infinite half-plane of conductor at $x > 0$ with an applied electric field that is only in the z direction, the solution to this equation is

$$E_z = E_0 e^{-\tau_n x} = E_0 e^{-\frac{x}{\delta}} e^{-\frac{ix}{\delta}} \quad (62)$$

where

$$\delta = \frac{1}{\sqrt{\pi f \mu_0 \sigma}}. \quad (63)$$

This quantity, δ , is called the “skin depth”, and gives a characteristic length for how deeply the electromagnetic fields penetrate the conductor. As an example of this, a copper surface with $\sigma = 6 \times 10^7$ S/m at a frequency of 1 GHz has a skin depth of approximately 2 μm . Also note that as the frequency increases, this skin depth decreases. Because this skin depth is so small compared to the wavelength of the resonator, it can be ignored in the treatment of the cavity properties. In the geometry above, the current density and losses can also be quantified. The current and field of this configuration are

$$j_z = j_0 e^{-\tau_n x}, H_y = H_0 e^{-\tau_n x}. \quad (64)$$

Using these, the surface impedance Z_0 can be calculated where

$$Z_0 = \frac{E_0}{I} \quad (65)$$

with I being the current. The losses come from the real part of the surface impedance, where the imaginary part of this is the reactance. Calculation yields

$$R_s = \sqrt{\frac{\pi f \mu_0}{\sigma}} = \frac{1}{\sigma \delta}. \quad (66)$$

Note that these calculations are done for normal conducting materials.

3.5 Cavity Figures of Merit - Accelerating Voltage

All cavities are judged by a small set of electromagnetic figures of merit. These figures of merit fall into two different categories: efficiency and scaling. Efficiency figures of merit are measures of how efficiently the cavity performs and do not scale with stored energy. The scaling figures of merit are all figures of merit that do scale with the stored energy

of the cavity. There are three major scaling figures of merit. First, and perhaps most important, is the Accelerating Voltage (V_{acc}). This is the measure of how much energy a charged particle passing through a cavity will gain, and it is defined as

$$V_{acc} = \int_{-\infty}^{+\infty} E_{acc}(z) \cos(\omega t + \phi) dz = \int_{-\infty}^{+\infty} E_{acc}(z) \cos\left(\frac{\omega z}{\beta c} + \phi\right) dz \quad (67)$$

where z axis is the beam axis, $\beta = v/c$, ϕ is a phase offset that maximizes V_{acc} and assuming that β doesn't change. This last approximation is equivalent to saying that the energy gained in the cavity is small compared to its totally energy. This approximation is not valid if a high degree of accuracy is required, say, for particle tracking. Practically, this can be calculated as

$$V_{acc} = \sqrt{\left(\int_{-\infty}^{+\infty} E_{acc} \sin(\omega t) dz\right)^2 + \left(\int_{-\infty}^{+\infty} E_{acc} \cos(\omega t) dz\right)^2}. \quad (68)$$

For the pillbox cavity, this integral has an analytic solution because the accelerating electric field is constant in z .

$$V_{acc} = \left| \int_{-\infty}^{+\infty} E_{acc}(z) e^{i\omega t + i\phi} dt \right| = E_0 \left| \int_0^L e^{i\omega_0 z / \beta c} dz \right| = LE_0 \frac{\sin\left(\frac{\omega_0 L}{2\beta c}\right)}{\left(\frac{\omega_0 L}{2\beta c}\right)} \quad (69)$$

The length of the cavity, at this stage, remains unspecified. An important consideration of the length is that this type of cavity is often operated in series as can be seen in Figure 6, with each cavity being π ahead of the previous one in phase. It is obvious to see that it is desirable to have the time it takes the particle to transit the cavity be equal to the time it takes the fields to reverse:

$$\frac{L}{\beta c} = \frac{\pi}{\omega_0}; L = \frac{\beta \lambda}{2}. \quad (70)$$

3.6 Cavity Figures of Merit - Transit Time Factor

Given a cavity length that is appropriate for the velocity of particle being accelerated, it is clear that there is still some loss in accelerating efficiency because the fields are varying in time. The maximum acceleration that could be expected from a cavity of this type would be

$$V_0 = \int_{-\infty}^{+\infty} |E_{acc}(z)| dz = LE_0. \quad (71)$$

The ratio of V_{acc} to V_0 gives the efficiency figure of merit called the Transit Time Factor (TTF). For the pillbox cavity with a length of $\beta\lambda/2$, the TTF is

$$TTF = \frac{V_{acc}}{V_0} = \frac{\sin\left(\frac{\omega_0 L}{2\beta c}\right)}{\left(\frac{\omega_0 L}{2\beta c}\right)} = \frac{\sin\left(\frac{\pi}{2}\right)}{\left(\frac{\pi}{2}\right)} = \frac{2}{\pi}. \quad (72)$$

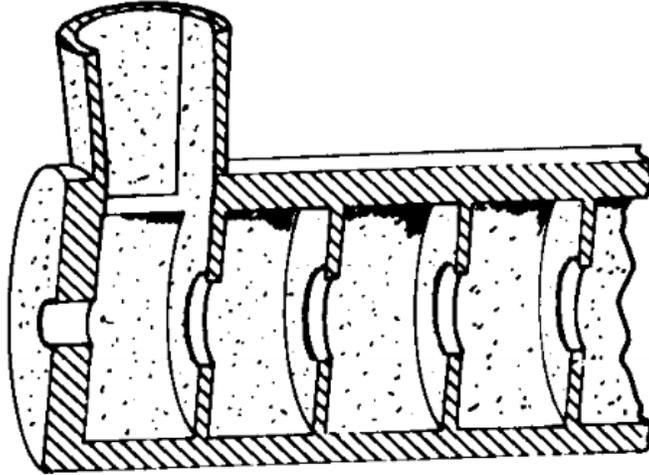


Figure 6: An example of pillbox type cavities coupled together.

It is important to note that, in reality, a given length of cavity L gives an *optimum* particle velocity β_{opt} . This means that the TTF varies with β for a given cavity geometry:

$$TTF(\beta) = \frac{\sin\left(\frac{\omega_0 L}{2\beta c}\right)}{\left(\frac{\omega_0 L}{2\beta c}\right)} = \frac{\sin\left(\frac{\pi\beta_{opt}}{2\beta}\right)}{\left(\frac{\pi\beta_{opt}}{2\beta}\right)}. \quad (73)$$

Figure 7 shows this curve for a cavity designed for $\beta_{opt} = 0.50$ for 2, 4, and 8 cavities in series. Notice that as the number of synchronized cavities (“gaps”) increases, the range of velocities that can be efficiently accelerated shrinks. This is of considerable importance for an accelerator that has to accelerate a beam with a velocity that changes significantly. While larger numbers of synchronized gaps mean a higher amount of acceleration per meter of constructed accelerator, this may require more cavities designed for a variety of different β_{opt} . Two extreme examples of this are heavy ion and electron accelerators. An electron is such a light particle that it very rapidly approaches the speed of light, thus only requiring one resonator geometry. To give an example of how fast this happens, remember that the International Linear Collider has a resonator designed to give roughly 30 MeV of acceleration with each resonator having 9 cells. This means that, from *rest*, an electron would be going 99%(!!!) the speed of light at the exit of the first resonator (even after including the reduction in acceleration efficiency when the electron’s speed is mismatched relative to the optimum beta). For a heavy ion accelerator, the situation is much different. For a beam of uranium (common in nuclear physics), over 500,000 times as much energy (almost 2 TeV!!!) is required to reach the same velocity, even in its highest charge state. Even using the most flexible accelerating resonators (2-gap structures), this would take at

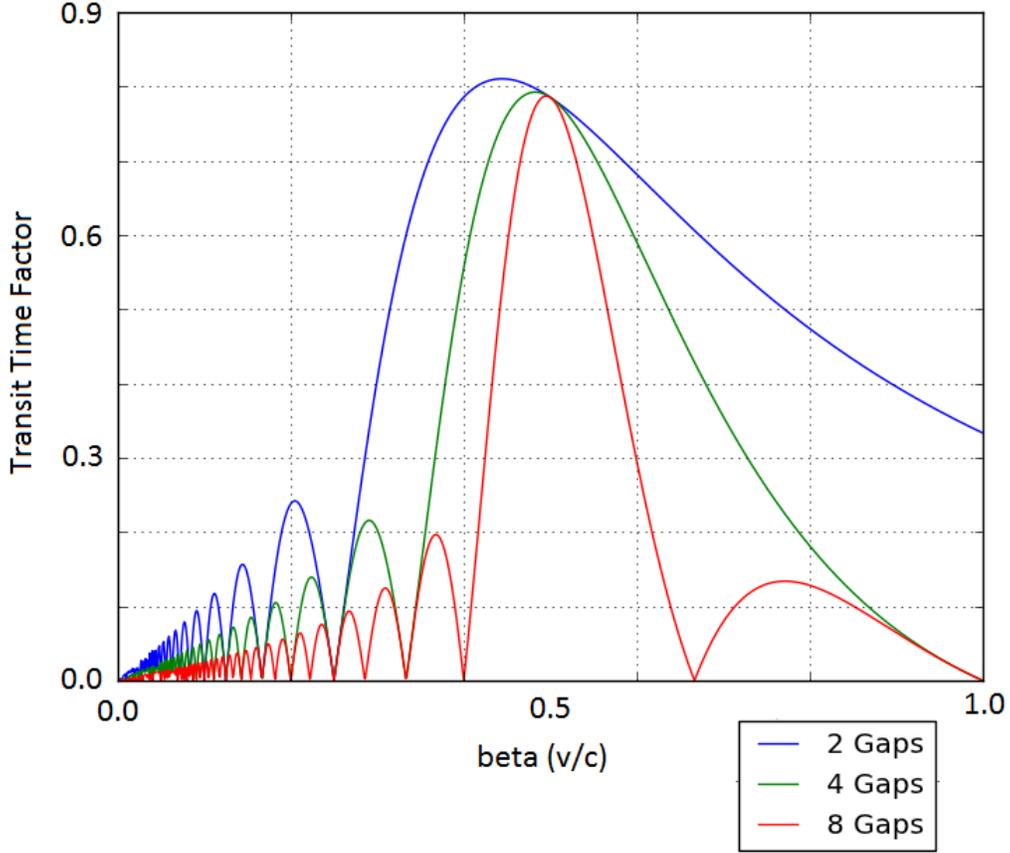


Figure 7: The Transit Time Factor (TTF) for 2, 4, and 8 cavities in series with $\beta_{opt} = 0.50$.

least 5 different cavity designs to achieve (and it would be ~ 4 km long!).

3.7 Cavity Figures of Merit - Peak Surface Electric and Magnetic Field

The remaining scaling figures of merit are the peak surface electric and magnetic fields. For the pillbox cavities, these values

$$E_{pk} = E_0 \quad (74)$$

$$H_{pk} = \frac{E_0}{\eta} J_1(1.84) = \frac{E_0}{647 \Omega} \quad (75)$$

are located on the end walls of the cavity.

3.8 Cavity Figures of Merit - Geometry and Quality Factor

The efficiency with which the resonator stores energy is of principal importance. For superconducting cavities, this is especially true because all dissipated power is exhausted into liquid helium. The figure of merit for this efficiency is the Quality Factor:

$$Q = \frac{\omega_0 U}{P_d}. \quad (76)$$

From our study of the harmonic oscillator, remember that we defined the quality factor as the ratio of the energy stored in the resonator to the energy lost per radian of oscillation. Inspection shows that this definition is consistent with the resonator definition. In addition to the Quality Factor, there is a figure of merit that does not depend on the surface resistance (R_s) called the Geometry Factor. This is defined as

$$G = R_s Q = \frac{\omega U}{\left(\frac{P_d}{R_s}\right)} \quad (77)$$

and is solely dependent on the geometry. This is a more useful figure of merit for cavity design because the R_s depends on many factors outside of the electromagnetic design such as material purity, surface quality, heat treatment of the cavity, and many more. This is also a useful figure to compare different geometries and cavities of different frequencies because it is independent of frequency [4, pp. 43-44]. To calculate the Geometry Factor for a given geometry, we must first calculate the stored energy (U) and dissipated power (P_d). Remembering that

$$\frac{dP_d}{dA} = \frac{1}{2} R_s \left| \vec{H} \right|^2 \quad (78)$$

where $\left| \vec{H} \right|$ is the magnetic field at dA , and that the Geometry Factor requires P_d/R_s the equation for the total dissipated power becomes

$$\frac{P_d}{R_s} = \frac{1}{2} \int_S \left| \vec{H} \right|^2 dA \quad (79)$$

where S includes all RF surfaces in the cavity. Note that we are assuming that R_s is constant over the surface, which is unlikely to be strictly true. While it is expected that its value will not vary much over the surface, accurate measurements of R_s and its variation over the surface of the cavity are extremely difficult. In any case, the variation will not likely be predictable, and should not affect cavity design. In addition to P_d , we need the stored energy in the cavity. In general, the time averaged stored energy density in the cavity is [5]

$$u = \frac{1}{2} \left(\epsilon |\vec{E}|^2 + \frac{1}{\mu} |\vec{B}|^2 \right). \quad (80)$$

At any given time, the energy is stored in the electric and magnetic fields of the cavity. However, the total energy is fixed, and in this time-averaged form, both electric and magnetic contributions are equal. This means we can simplify this density to

$$u = \frac{\epsilon}{2} |E|^2 = \frac{1}{2\mu} |B|^2 \quad (81)$$

which yields

$$U = \frac{\epsilon_0}{2} \int_V |E_0|^2 dV = \frac{1}{2\mu_0} \int_V |B_0|^2 dV. \quad (82)$$

This gives us

$$G = R_s Q = \frac{\omega_0 \mu_0 \int_V |\vec{H}|^2 dV}{\int_S |H|^2 dA}. \quad (83)$$

With some calculation, these integrals can be done for the pillbox cavity as described [4, p. 46]:

$$U = \frac{\pi \epsilon_0 E_0^2}{2} J_1^2(2.405) L R^2 \quad (84)$$

$$P_d = \frac{\pi R_s E_0^2}{\eta^2} J_1^2(2.405) R(R+L) \quad (85)$$

$$G = \frac{\omega_0 \mu_0 L R^2}{2(R^2 + RL)} = \eta \frac{2.405 L}{2(R+L)} = \frac{453 \frac{L}{R}}{1 + \frac{L}{R}} \Omega. \quad (86)$$

By remembering that we desire

$$L = \frac{\beta \lambda}{2} \quad (87)$$

we find that, for a pillbox cavity,

$$\frac{L}{R} = \frac{\beta \pi}{2.405} \quad (88)$$

which gives a Geometry Factor for ALL optimized pillbox cavities of

$$G = 257\beta \Omega. \quad (89)$$

As a comparison, the Facility for Rare Isotope Beams (FRIB) calls for a cavity (of a totally different geometry) designed for $\beta = 0.29$. An equivalent pillbox cavity would have $G \approx 74 \Omega$ with the current FRIB cavity design having $G = 76 \Omega$. While a pillbox cavity is impractical at these velocities (it would require a cavity that has a diameter over 5 times its length), this demonstrates the usefulness of this figure of merit for comparison between cavities.

3.9 Cavity Figures of Merit - R/Q

The third major efficiency figure of merit is the R/Q . This is a measure of how efficiently the cavity transfers its stored energy to the beam passing through it. This is the ratio of the *effective* shunt impedance to the cavity's quality factor, and this is a point that needs to be stressed. There are many definitions of the shunt impedance depending on context or even convention, and I will be using the common accelerator conventional definition [4, p. 47]:

$$R_a = \frac{V_{acc}^2}{P_d} \quad (90)$$

in units of Ω per cell. This is further complicated by the fact the the majority of this document will be spent discussing cavities that, while one "cell", have two accelerating gaps. From all this, the definition of the R/Q is

$$R/Q = \frac{V_{acc}^2}{\omega_0 U}. \quad (91)$$

One final note is that this is the effective shunt impedance. While I will omit the *eff* subscript from R , I will be using V_{acc} instead of V_0 , which is implied by "shunt impedance" in some circles. For a pillbox cavity, this can be readily calculated

$$R/Q = 150 \Omega \frac{L}{R} = 196\beta \Omega. \quad (92)$$

3.10 Pillbox Cavities for Low Velocity Particles

The relationship between R/Q and β indicates that pillbox type cavities would perform best for high velocity beams, and would perform quite poorly for very low velocity beams. This can be seen by comparing the performance of a FRIB cavity design for $\beta = 0.29$, a pillbox $R/Q \approx 57 \Omega$ compares poorly to the $R/Q = 110 \Omega/\text{gap}$ of the current design. In addition to efficiency, there are mechanical drawbacks to using a pillbox cavity at low velocity. As has been shown, the optimum length of the accelerating gap for a pillbox cavity is $\beta\lambda/2$, and I will now posit that what this actually means is that the distance between *adjacent, synchronized gaps* should be

$$L_{gap-to-gap} = \left(n + \frac{1}{2}\right) \frac{\beta\lambda}{2} \quad (93)$$

where n is an integer. Pillbox cavities require very little space between synchronized gaps, thus for them (and a vast majority of resonators) n will be 1. This gap-to-gap distance is a more general principle of accelerating resonators, and thus can be used to compare different geometries. This scaling gives interesting results for pillbox cavities designed for very low velocity beams. FRIB calls for the first resonator to be optimized for $\beta = 0.041$. With a conservative gap-to-gap distance of 0.1 m, this gives a cavity with $f = 61.5$ MHz,

and a cavity radius of $R = 0.534$ m. This cavity has very poor mechanical qualities from such a large ratio of radius to gap size, in addition to the loss in accelerating efficiency due to the small gap size.

4 Coaxial Resonators

In the previous section, we derived the mode structures and properties of resonators constructed from a terminated section of cylindrical waveguide. The scaling of these structures means they become undesirable for acceleration of very low velocity particles. Coaxial resonators originate from the mode structure of coaxial transmission lines. These resonators offer a practical option for lower velocity particle acceleration because they can achieve the lower fundamental frequency required for a lower optimum velocity of beam while maintaining efficiency of acceleration. In payment for this, these cavities have more complex geometries that require more material, construction, and processing cost and time. Their mode structures, scaling parameters, and figures of merit will be derived and compared with the pillbox cavity type, and their use in heavy ion accelerators will be justified.

4.1 Coaxial Transmission Line Geometry

The addition of a cylindrical inner conductor to a cylindrical waveguide significantly alters the mode structure. First, modes with longitudinal electric fields are no longer supported, and the modes of interest have no longitudinal magnetic fields, giving these modes their designation as *TEM*, or *Transverse Electric and Magnetic*. There are two types of termination that are of interest for this type of geometry, short and open. A short termination is the same type that was used in the previous section for pillbox cavities, a conducting plane in the $\rho - \phi$ plane. This gives a magnetic field maximum and a zero in the electric fields at the boundary. The other type, an open, is the addition of a perfect *magnetic* conducting plane in $\rho - \phi$. This provides a zero in the magnetic field and a maximum in the electric field at the boundary. While this is, technically, non-physical, it is an excellent starting place for analytical calculations. An easy justification for the contained nature of modes using an effective open boundary condition can be seen from considering a length of coaxial line, shorted at one end, and connected to a cylindrical waveguide with the same radius as the outer conductor at the other end. As long as the frequency of the coaxial mode was below the cutoff frequency of the cylindrical waveguide, it would be reflected off the transition and bound in the coaxial section. Clearly, there would be some non-ideal field profile on the length scale of the evanescent decay length, but this is only a perturbation of the mode shape. It should be noted that while you can solve for pillbox modes with open terminations, these structures will give, in reality, unbound modes that are completely impractical.

4.2 The Ideal Quarter Wave Resonator

First, we will discuss a coaxial resonator with one short termination (at $z = L$) and one ideal open termination (at $z = 0$), seen in Figure 8. Coaxial lines are typified by a radial electric field and an azimuthal magnetic field, with a $1/\rho$ dependence in amplitude. The traveling wave formulation of this gives

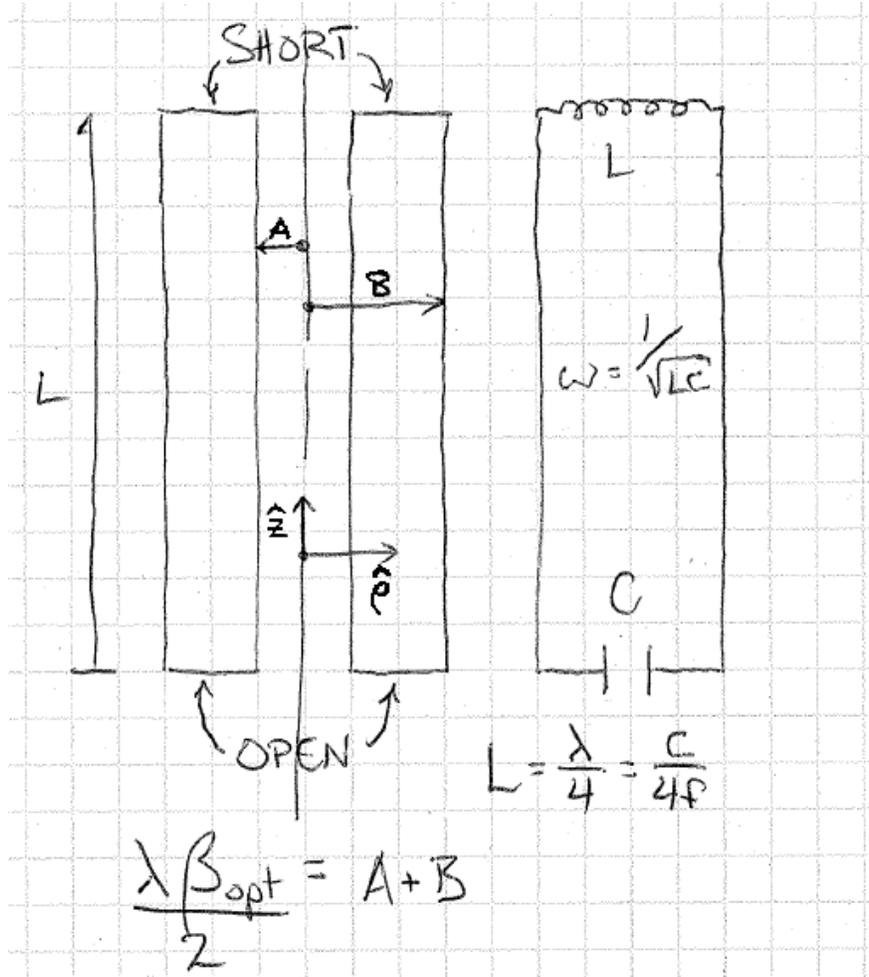


Figure 8: Schematic of the ideal QWR.

$$E_{\rho} = -\frac{I_0 \mu_0 c}{2\pi \rho} e^{i\omega t + ikz} \quad (94)$$

$$B_{\phi} = \frac{I_0 \mu_0}{2\pi r} e^{i\omega t + ikz}. \quad (95)$$

Again, summing two identical traveling waves with opposite directions of propagation

$$e^{ikz} + e^{-ikz} = 2 \cos(kz) \quad (96)$$

$$e^{ikz} - e^{-ikz} = 2i \sin(kz) \quad (97)$$

we can satisfy the boundary conditions to get

$$E_\rho = \frac{E_0 a}{\rho} \cos\left(\frac{p\pi z}{2L}\right) \sin(\omega t) \quad (98)$$

$$B_\phi = -\frac{E_0 a}{\rho c} \sin\left(\frac{p\pi z}{2L}\right) \cos(\omega t) \quad (99)$$

where p is an integer giving the order of the mode in z , E_0 is the peak surface electric field (located at $z = 0$ and $\rho = a$), a and b are the inner and outer conductor radii, and $\omega = pc\pi/2L$. From this, it is trivial to see that the peak surface magnetic fields is E_0/c and is located at $z = L$ and $\rho = a$. Rearranging this last equation gives us $L = \lambda/4$ for the lowest mode, $p = 1$, giving the cavity its name: Quarter Wave Resonator (QWR). The accelerating axis for this cavity is at $z = \phi = 0$, giving an ideal QWR accelerating field seen in Figure 9. Using these field distributions, it is straight-forward to find an ideal QWRs figures of merit. The stored energy, U , is

$$U = \frac{\epsilon_0}{2} \int_V |E|^2 dV \quad (100)$$

$$= \frac{\epsilon_0 E_0^2 a^2}{2} \int_a^b \frac{1}{\rho} d\rho \int_0^{2\pi} d\phi \int_0^L \cos^2\left(\frac{\pi z}{2L}\right) dz \quad (101)$$

$$= \frac{\epsilon\pi E_0^2 \lambda}{8} \ln\left(\frac{b}{a}\right) \quad (102)$$

remembering that $L = \lambda/4$. The dissipated power, P_d , is also straight-forward to calculate:

$$\frac{P_d}{R_s} = \frac{1}{2\mu_0^2} \int_S |B|^2 dA = \frac{\pi E_0^2 a^2 \epsilon_0}{2\mu_0} \left[\frac{\lambda}{4} \left(\frac{1}{a} + \frac{1}{b} \right) + 2 \ln\left(\frac{b}{a}\right) \right] \quad (103)$$

remembering to sum the contributions from the inner conductor, outer conductor, and the short plate. The dissipated power on the open termination is zero because the magnetic field there is zero. These numbers allow us to calculate the geometry factor G :

$$G = \frac{c\pi\mu_0 \ln\left(\frac{b}{a}\right)}{\left[\frac{\lambda}{2} \left(\frac{1}{a} + \frac{1}{b} \right) + 4 \ln\left(\frac{b}{a}\right) \right]}. \quad (104)$$

The calculation of the optimum beta is less straight-forward for this geometry compared to the pillbox cavity. The integral is made complicated by the form of the accelerating field which gives an equation for the accelerating voltage

$$V_{acc} = 2 \int_a^b \frac{E_0 a}{x} \sin\left(\frac{\omega x}{\beta c}\right) dx \quad (105)$$

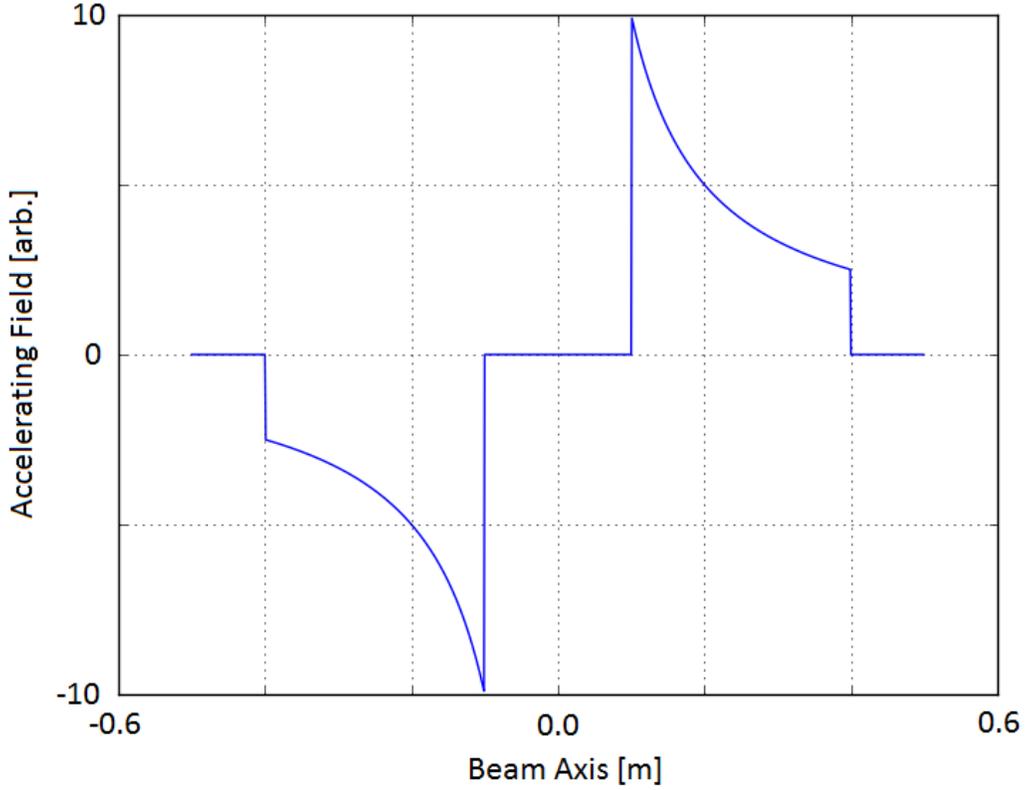


Figure 9: An example of the accelerating electric field for an ideal coaxial resonator. The beam axis passes radially through the cavity at $z = 0$.

that does not have an analytic solution. We can sidestep this problem by remembering that we can find the optimum velocity by solving for when the derivative of V_{acc} is zero.

$$\frac{\partial V_{acc}}{\partial \beta} = 2E_0a \int_a^b -\frac{\cos\left(\frac{\omega x}{\beta c}\right) \frac{\omega x}{\beta^2 c}}{x} dx = \frac{2E_0a\omega}{\beta^2 c} \left[\int_b^a \cos\left(\frac{\omega x}{\beta c}\right) dx \right] \quad (106)$$

which gives

$$\frac{\partial V_{acc}}{\partial \beta} = \frac{2E_0a}{\beta} \left[\sin\left(\frac{\omega a}{\beta c}\right) - \sin\left(\frac{\omega b}{\beta c}\right) \right] = 0. \quad (107)$$

Using the trigonometric identity

$$\sin u - \sin v = 2 \cos\left(\frac{u+v}{2}\right) \sin\left(\frac{u-v}{2}\right) \quad (108)$$

we obtain

$$\cos\left(\frac{\omega(a+b)}{2\beta c}\right) \sin\left(\frac{\omega(b-a)}{2\beta c}\right) = 0. \quad (109)$$

This gives conditions of

$$n\beta\lambda = b - a; \quad \frac{n\beta\lambda}{2} = b + a \quad (110)$$

where n is an integer. The first equation give solutions where the gap size $(b - a)$ is extremely large, large enough for the particle to see at least one full RF period. Thus, these solutions are likely to be minima, and will be disregarded. The second equation gives the relationship that we expected for $n = 1$, that

$$\frac{\beta\lambda}{2} = b + a \quad (111)$$

which is identical to the relationship assumed for the pillbox cavity (remembering that $b + a$ is the gap-to-gap distance). For an ideal QWR, knowing the frequency and optimum particle velocity desired specifies two of the three variables for the cavity, with the gap size $b - a$ being the only unspecified value. A plot of R/Q and G versus gap size $(b - a)$ can be seen in Figure 10. In this plot, it is easy to see that a large gap size results in a more efficient cavity. This is tempered, however, with the knowledge that the larger the gap for a given β_{opt} , the smaller the inner conductor diameter. The remaining figures of merit for the ideal QWR are the peak surface electric and magnetic fields. These are both on the inner conductor, and their magnitudes scale like $1/a$. The limit to how much we can reduce a as we increase R/Q and G is driven by the requirement of minimizing these peak surface fields.

From Figure 11, we can see that the ideal open termination is not truly representative of the electric field region of a practical QWR. While a detailed treatment of QWR design is outside the scope of this document, an issue of some importance is the asymmetry of this geometry not present in the ideal QWR fields. This asymmetry causes steering electric and magnetic fields on the beam axis, especially when the frequency of the cavity is increased above ~ 200 MHz. This steering can be corrected to a degree [6], but QWR usage is generally limited to applications at lower frequencies, and therefore lower velocities.

4.3 The Ideal Half Wave Resonator

If, instead of a short termination at $(z = L)$ and an ideal open termination at $(z = 0)$, you have a short termination at $(z = L/2)$ and another short termination at $(z = -L/2)$ (seen in Figure 12), the mode structure changes, giving you a higher frequency cavity for a given length. The field profiles change to

$$E_\rho = \frac{E_0 a}{\rho} \cos\left(\frac{p\pi z}{L}\right) \sin(\omega t) \quad (112)$$

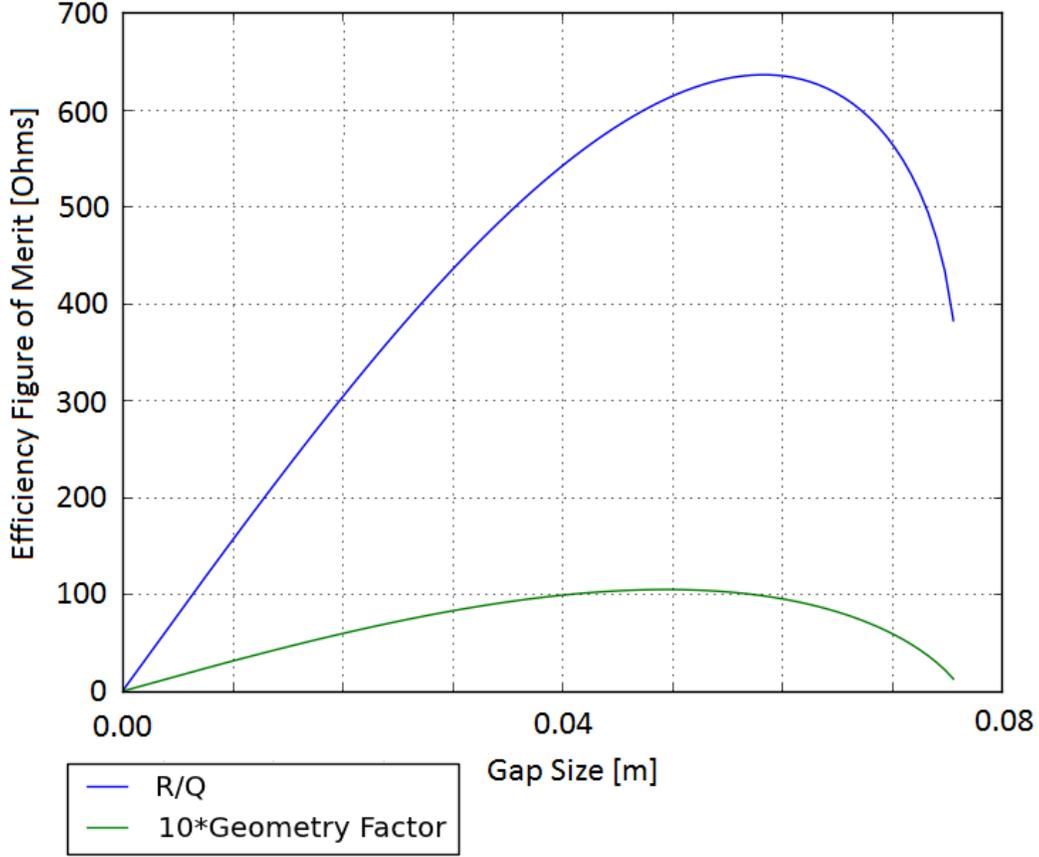


Figure 10: R/Q and G versus gap size for an ideal 80.5 MHz, $\beta = 0.041$ QWR.

$$B_\phi = -\frac{E_0 a}{\rho c} \sin\left(\frac{p\pi z}{L}\right) \cos(\omega t) \quad (113)$$

where p is an integer giving the order of the mode in z , E_0 is the peak surface electric field, a is the inner conductor radius, and $\omega = pc\pi/L$. Rearranging this last equation gives us $L = \lambda/2$ for the lowest mode, $p = 1$, giving the cavity its name: Half Wave Resonator (HWR). These equations are all very similar to a QWR, and in fact are only different because of the change in definition of the length, L . An ideal HWR is effectively an ideal QWR, but reflected about the QWR's ideal open boundary condition. This means that almost all of the derived quantities for the QWR can be used here with minor modification.



Figure 11: FRIB Quarter Wave Resonator designs.

These modifications give

$$U = \frac{\epsilon_0}{2} \int_V |E|^2 dV \quad (114)$$

$$= \frac{\epsilon_0 E_0^2 a^2}{2} \int_a^b \frac{1}{\rho} d\rho \int_0^{2\pi} d\phi \int_{-L/2}^{L/2} \cos^2\left(\frac{\pi z}{L}\right) dz \quad (115)$$

$$= \frac{\epsilon\pi E_0^2 a^2 \lambda}{4} \ln\left(\frac{b}{a}\right). \quad (116)$$

$$\frac{P_d}{R_s} = \frac{1}{2\mu_0^2} \int_S |B|^2 dA = \frac{\pi E_0^2 a^2 \epsilon_0}{\mu_0} \left[\frac{\lambda}{4} \left(\frac{1}{a} + \frac{1}{b} \right) + 2 \ln\left(\frac{b}{a}\right) \right] \quad (117)$$

$$G = \frac{c\pi\mu_0 \ln\left(\frac{b}{a}\right)}{\left[\frac{\lambda}{2} \left(\frac{1}{a} + \frac{1}{b} \right) + 4 \ln\left(\frac{b}{a}\right) \right]} \quad (118)$$

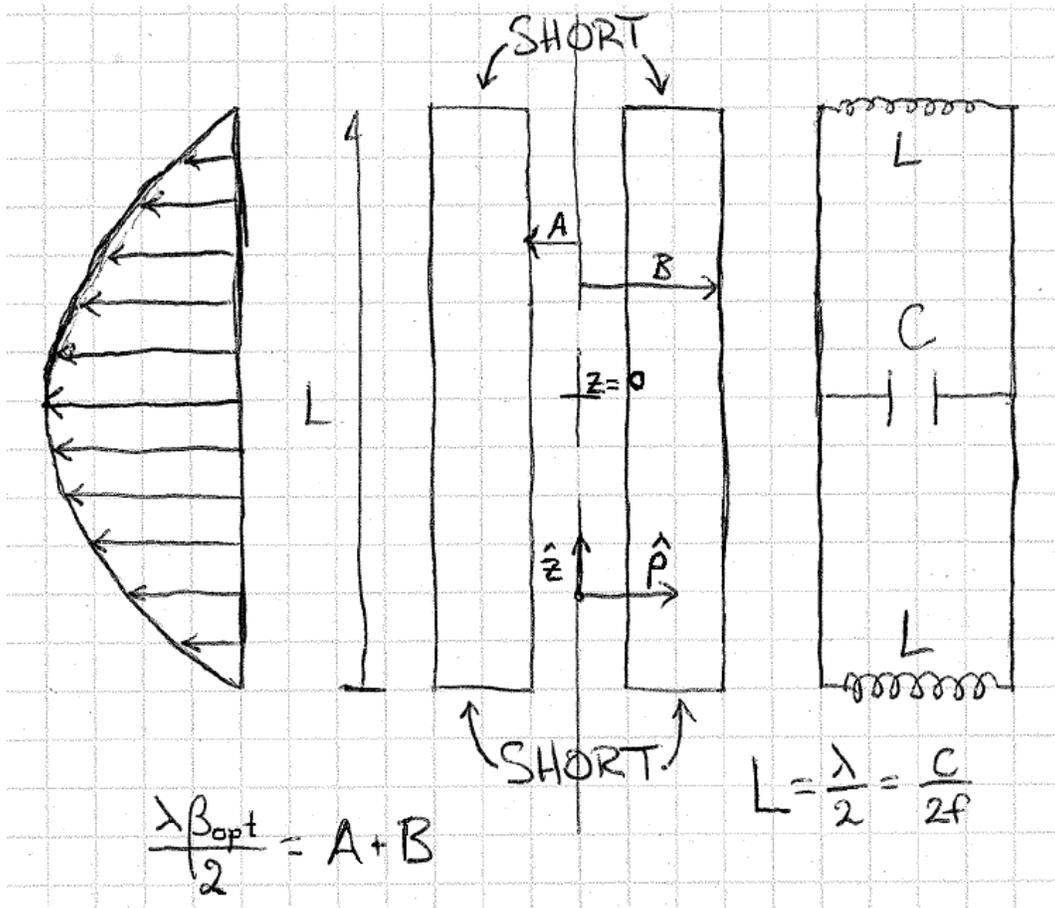


Figure 12: Schematic of an ideal HWR.

remembering that $2L = \lambda$. Additionally,

$$V_{acc} = 2 \int_a^b \frac{E_0 a}{x} \sin\left(\frac{\omega x}{\beta c}\right) dx \quad (119)$$

$$\frac{\beta \lambda}{2} = b + a. \quad (120)$$

remain identically valid for the HWR. It should be noted that, for a given length, the frequency of the structure has doubled. In addition, because the losses are driven by the magnetic fields, the losses for a given acceleration will double going from a QWR to a HWR. This will double the HWR's R/Q , but because the stored energy also doubles, G remains the same. In addition, the location of the peak surface electric and magnetic fields remain the same, with the caveat that there are not *two* peak surface magnetic fields, at

$z = L/2$ and $z = -L/2$. The benefit of the HWR geometry is the symmetry along the long axis, meaning that steering fields are only generated from manufacturing errors as opposed to the inherent geometry itself. This means that the frequency limitations of the QWR do not apply to HWR, allowing acceleration of higher velocity particles.

4.4 Coaxial Cavity Efficiency at Low Velocity

While V_{acc} for coaxial cavities is not analytic, the possible acceleration (V_0) is. For a QWR, this leads to

$$V_{acc} = V_0 \times TTF = 2E_0 a \ln\left(\frac{b}{a}\right) \times TTF \quad (121)$$

which leads to

$$\frac{R}{Q} = TTF^2 \times \frac{8 \ln\left(\frac{b}{a}\right)}{\pi^2 \epsilon_0 c} \approx TTF^2 \times \ln\left(\frac{b}{a}\right) \times 300 \Omega \quad (122)$$

which does not scale with β . The TTF at optimum velocity for most designs will be $0.85 < TTF < 0.90$, and does not scale strongly with β . For most coaxial designs, b/a is between 2.5 and 3.5, although the exact definitions of a and b break down for realistic cavity designs. This means that coaxial cavities maintain their acceleration efficiency for very low velocity particles while pillbox type cavities become increasingly less efficient. To directly compare an ideal HWR to an ideal pillbox cavity, the HWR's R/Q must be halved to account for the HWR having two gaps compared to a single pillbox. This is an interesting comparison because it can be used to indicate the β where the R/Q of the pillbox cavity is equal to the R/Q of a coaxial cavity, giving some idea of the velocity where one geometry should replace the other:

$$\beta_{transition} = \frac{0.5 \times R/Q_{HWR}}{R/Q_{pillbox}} \approx \frac{150 \Omega \times \mathcal{F}}{150 \frac{\pi}{2.405} \Omega} \approx 0.75 \times \mathcal{F} \quad (123)$$

where

$$\mathcal{F} = TTF^2 \times \ln\left(\frac{b}{a}\right). \quad (124)$$

\mathcal{F} is dependent on the design of the cavity, but for most HWR designs will be between 0.8 and 1, giving a transition β of ~ 0.7 . This scaling is born out in the parameters of realistic cavity designs for applications around these velocities. While the R/Q scales like β for a pillbox and is roughly constant for coaxial cavities, the Geometry Factor is more complex. For an ideal pillbox, $G = 257\beta \Omega$. For a HWR, assuming a reasonable b/a of 3, the dependence is more complex:

$$G_{HWR} \approx \frac{1300\beta}{5.3 + 4.4\beta} \Omega \quad (125)$$

which goes to zero for small β , and is only $\sim 120 \Omega$ for $\beta = 1$. A plot of this behavior can be seen in Figure 13. For comparison, an ideal HWR with $\beta = 0.5$ would have a Geometry

Factor of $\sim 80 \Omega$, where the realistic FRIB design for a similar cavity has a Geometry Factor of $\sim 100 \Omega$. This is because the ideal HWR geometry is not very efficient; a more complex inner and outer conductor geometry provides this increase in efficiency.

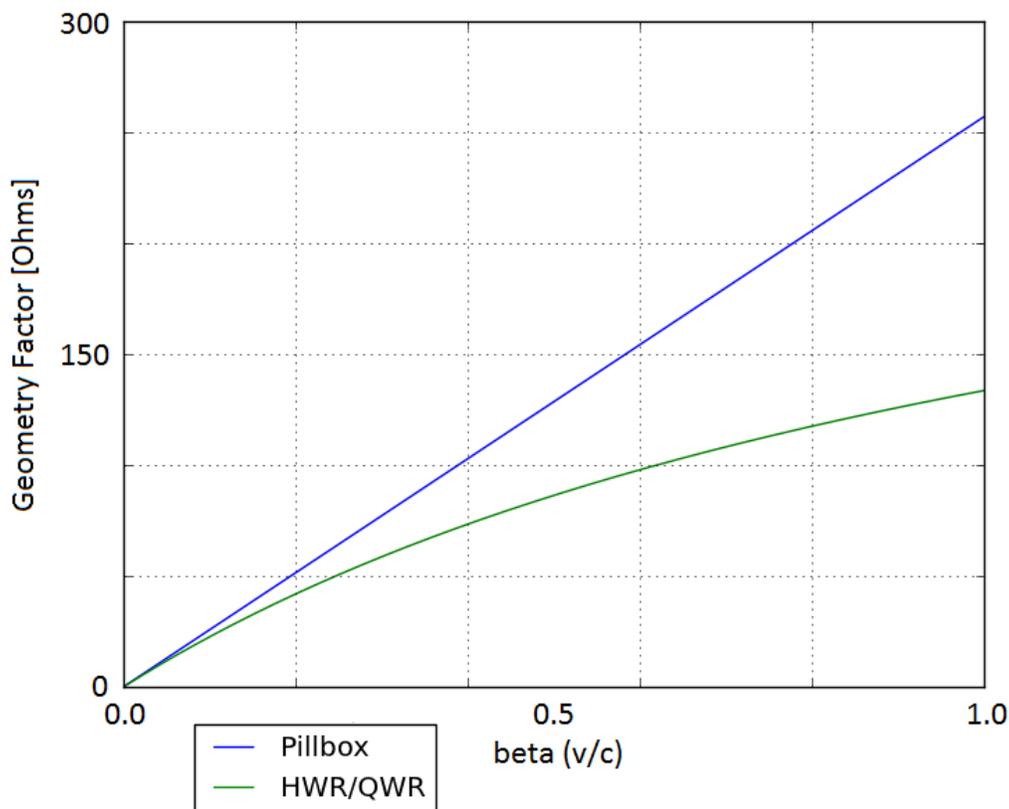


Figure 13: Geometry Factor plotted versus β for the ideal cavity geometries.

4.5 An Aside: The Three-Quarter Wave Resonator

An extension of the logic used to get the parameters for the HWR can be used for a more novel type of cavity geometry. It has been proposed [7] to operate a QWR-like geometry (open and short boundary conditions) in the $p = 2$ mode, giving $L = 3\lambda/4$. The particles would be accelerated through the electric field maximum that is closer to the short plate, away from the open boundary. This would combine the benefits of the QWR geometry (open end for processing and inspection after fabrication) while avoiding the steering fields present in a traditional QWR. This cavity type is called a Three-Quarter Wave Resonator, or 3QWR. Comparing a HWR and 3QWR at the same frequency and beta, it is clear that

the 3QWR will have the same G while the R/Q will be 1/3 lower.

4.6 Another Aside: Power Transmission Choices Based on Frequency

The frequency scalings we’ve derived for the cylindrical and coaxial waveguides give some idea for the type of transmission line that would be required to transmit energy to these cavities. For the cylindrical waveguide, we derived a cut-off frequency of

$$\omega_{\text{cutoff}} = \frac{2.405c}{R} \quad (126)$$

where R is the radius of the pipe. For the FRIB cavities, operating at 80.5 and 322 MHz, this would require a pipe radius of 1.43 and 0.36 m, far too large to be practical. For the coaxial waveguide, there is no cutoff frequency, making this geometry the desired transmission line for this application.

5 Superconducting Material Properties

Oxygen-Free, High Conductivity (OFHC) copper and niobium have been the materials of choice for the vast majority of accelerators because of the efficiency offered. Copper is an excellent conductor of both electricity and heat, allowing for low losses and easy evacuation of the joule heat generated. Niobium is a poor conductor at room temperature, but becomes superconducting below 9.2 K. The operational implications of using these materials as well as the theory behind their losses will be discussed below.

5.1 Superconducting Material Theory

As has been shown, the skin depth δ and the surface resistance R_s for a normal conducting material are

$$\delta = \frac{1}{\sqrt{\pi f \mu_0 \sigma}} \quad (127)$$

$$R_s = \frac{1}{\sigma \delta}. \quad (128)$$

Superconducting materials behave very differently from normal conducting materials, and require a very different treatment to derive equivalent quantities. A superconducting material is typified by a sudden transition between standard normal conducting physics to a DC resistivity of zero at a certain temperature (below 9.2 K for bulk niobium). Below this critical temperature (T_c), it becomes energetically favorable for electrons to be combined into “Cooper pairs” which act as a superfluid, allowing current to flow through the material with no losses [8]. The current theory of this behavior is BCS theory [9, 10], and its predictions agree well with measured results [11, 12]:

$$R_{BCS} = 2 \times 10^{-4} \frac{C_{RRR}}{T} \left(\frac{f}{1.5} \right)^2 e^{\left(-\frac{17.67}{T} \right)} \quad (129)$$

as a good fit for the BCS component of the surface resistance of niobium where R_{BCS} is in ohms, T is the temperature in Kelvin, and f is the frequency in GHz. C_{RRR} is a factor based on the material purity, 1 for reactor grade niobium, and approximately 1.5 for $RRR = 300$ niobium. From this, it is easy to see that it is desirable to operate superconducting cavities at the lowest temperature possible. In practice, R_{BCS} isn't the only contribution to the surface resistance. The remaining resistance is a material/cavity processing dependent resistance called the "residual resistance" R_{res} that does not depend on temperature or frequency. While residual resistances of below 1 n Ω have been achieved, 10 n Ω is more reliably reproducible for large scale cavity fabrication [13]. If the cavity is operated at or below 2 K, the residual losses tend to dominate performance. An approach for estimating the skin depth of a superconducting material can be derived [4, pp. 85-88] by using a two-fluid model for the electrons in the material, one fluid composed of the Cooper pairs, and the other the unbound electrons. This gives a penetration depth (called the London Length [14]) of

$$\lambda_L^2 = \frac{m}{n_s e^2 \mu_0} \quad (130)$$

where m and e are the mass and charge of the electron, and n_s is the density of the superconducting electron pairs. For niobium, this penetration depth is on the order of tens of nanometers, meaning that the RF performance is dominated by the quality of a very thin layer on the surface of the material. The temperature dependence for λ_L has been empirically found as [15]

$$\lambda_L(T) \approx \lambda_L(0) \left[1 - \left(\frac{T}{T_c} \right)^4 \right]^{-1/2}. \quad (131)$$

While it becomes energetically favorable for electrons to form Cooper pairs below T_c , this is offset by the application of a magnetic field. Magnetic flux is expelled in the transition from normal to superconducting (called the Meissner Effect [16]), but the fields still penetrate the material on the order of the London Length. BCS theory approximates the temperature dependence of this critical magnetic field as

$$H_c(T) = H_c(0) \left[1 - \left(\frac{T}{T_c} \right)^2 \right]. \quad (132)$$

Finally, niobium is a Type-2 superconductor. This means that the transition from superconducting to normal conducting in the presence of a magnetic field actually has two critical magnetic field levels, H_{c1} and H_{c2} , in contrast to a Type-1 superconductor in which the superconductivity breaks completely at one field level. This behavior can be seen in Figure 14. One of the reasons niobium is used for superconducting accelerators is the relatively high T_{c1} , one of the highest for any Type-2 superconductor.

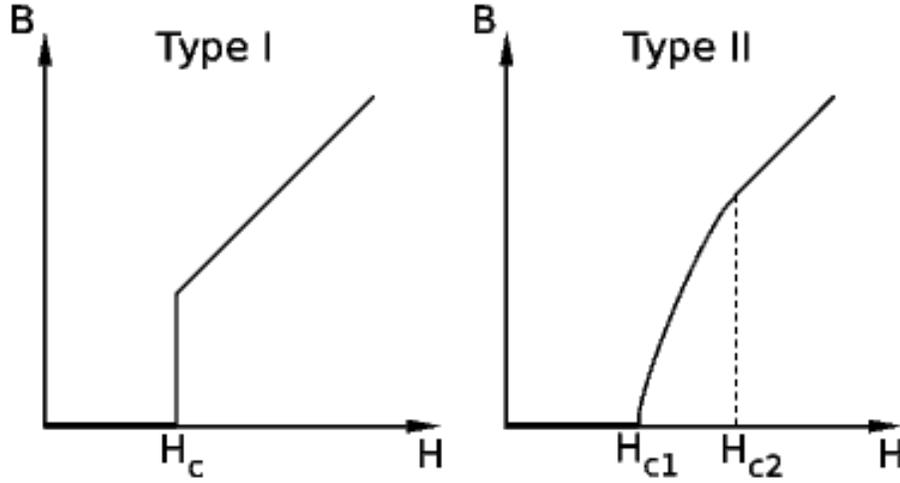


Figure 14: Magnetic field just inside superconducting material around transition for Type 1 and 2 materials. Note the sudden penetration for Type 1 and the gradual break down for Type 2.

5.2 The Advantages of Superconductivity

Assuming that it is operated at a temperature well below 9.2 K (not challenging with liquid helium's boiling point of 4.2 K at 1 bar), superconducting cavities can be operated with very low losses. The surface resistance for a well prepared bulk niobium can be as low as single nano-ohms, as compared to micro-ohms for a well prepared copper cavity. This means that for the same cavity and field level, the losses are lowered by 6 (!) orders of magnitude. Some of this is lost because of the requirement to operate at very low temperature. The maximum achievable efficiency for heat transfer depends on the input and output temperature and is given by the Carnot Efficiency:

$$\eta_c = \frac{T_C}{T_H - T_C} \quad (133)$$

where T_H and T_C are the temperature at which the heat is exhausted and extracted in Kelvin. For 4 K operation, this gives an *optimum* operation efficiency of .013. This should be combined with the technical efficiency of cryogenics systems ($\nu \approx 0.2$ [4, p. 48]), resulting in an increase in required wall power by two to three orders of magnitude. This means that the six orders of magnitude decrease in material losses is, in practical terms, three orders of magnitude gain in wall power efficiency. For applications which call for the cavities to be operated in continuous wave (CW) mode, the cost savings of superconducting operation can be substantial, despite the complex technology involved.

5.3 The Challenges of Superconducting Materials

While the decrease in required power is substantial, there are drawbacks to using superconducting cavities. The cryogenic engineering talent and facilities required is a dramatic investment, and can be a significant portion of the overall cost of an accelerator. Additionally, the ultimate performance and achievable accelerating gradient is limited by the peak surface magnetic field. Even for an extremely well prepared cavity, the peak surface magnetic field is ~ 220 mT, and for low velocity cavities, even with state of the art processing, this number is closer to 120 mT. Because the losses in a superconductor also increase strongly with frequency, this means that superconducting cavities are limited to operation under 2 GHz for practical reasons. Even the International Linear Collider which is designed to operate at 1.3 GHz and 2 K has moved to a pulsed design to limit the load on the cryogenic systems. Additionally, there are many processing and fabrication considerations for these cavities to achieve a low surface resistance with little to no radiation. These considerations will be discussed later in this document.

5.4 Practical Concerns: RRR and Impurities

Procurement of the appropriate material for an accelerating cavity is an important step in obtaining good performance. While the RF losses in a cavity are small, the specific heat of niobium at those temperatures is also quite low. If the thermal conductivity is small, local heating can occur at high fields, breaking the superconductivity in what is known as a thermal quench. The thermal conductivity of the niobium is dominated by impurities and crystal structure, which is measured by the RRR:

$$RRR = \frac{\sigma(4.2K)}{\sigma(300K)} \propto \kappa \quad (134)$$

where σ is the electrical conductivity and κ is the thermal conductivity. Material can be readily obtained from suppliers with RRR of over 250, considered good enough to give good cavity performance. This material is produced by repeated electron beam melting in vacuum to drive off impurities, and can also be improved by baking over 1200°C in vacuum with either titanium or Yttrium as a getter to remove impurities.

6 Coupler Theory and Design

Cavities are ultimately powered by a power plant, like all other devices that require electricity. This wall power is used to power both the low power RF signal generators and the amplifiers (klystrons, solid state, or other types) that generate the large amounts of power needed to fill the cavity. This power must be effectively transformed from traveling wave power in the transmission lines to the standing wave mode of the cavity. In this section, the basic design and some considerations of these couplers will be discussed in the context

of different applications. These topics will be expanded upon with more specific design and theory for couplers used for low beta superconducting resonators like those being designed for the FRIB driver linac.

6.1 Coupler Figures of Merit

Types of transmission lines have already been discussed, and coaxial waveguide has been demonstrated to be the only practical solution at low frequency. From this, we can design two types of couplers, probe (capacitive) and loop (inductive). Probe couplers use the transmission line's inner conductor to excite the cavity electric fields in the local region. There are different ways to quantify the coupling strength, but the overlap between the excited fields from the inner conductor of the coupler and the fields in the cavity drive the coupling strength. Thus, the probe type coupler is more effectively used in the high electric field region of the cavity. The desired coupling strength is then achieved by changing the length of the coupler inner conductor. Magnetic field coupling is achieved by attaching the coupler inner conductor to its outer conductor, making a loop. Magnetic fields are excited normal to the loop, and the strength of the coupler is determined by the confluence of these excited fields and the cavity mode's magnetic field distribution. The desired coupling strength is achieved by varying the area of the loop exposed to the cavity (moving the coupler in or out of the cavity) or changing the angle between the cavity field and the normal of the loop (rotating the coupler). The coupler strength can be quantified by the rate at which energy enters or leaves the cavity through the coupler. This emitted power, P_e , can be normalized to the stored energy in the cavity and used to form an effective quality factor for the coupler called the "external Q":

$$Q_{ext} = \frac{\omega U}{P_e}. \quad (135)$$

The ratio of the external Q to the cavity Q is called the coupling factor, β (not to be confused with the relativistic $\beta = v/c$)

$$\beta = \frac{Q_0}{Q_{ext}} = \frac{P_e}{P_d}. \quad (136)$$

In reality, what will be measured if a cavity is connected to a Vector Network Analyzer (VNA) will be a quality factor given by the *total* losses in the cavity. The total losses in the cavity are given by

$$P_T = P_d + P_e + P_t \quad (137)$$

where P_e is the power emitted from the input coupler (port 1), P_t is the power emitted from the transmitted power coupler (port 2), and P_d is the power dissipated in the cavity itself. This gives us a "loaded" quality factor of

$$\frac{1}{Q_L} = \frac{1}{Q_0} + \frac{1}{Q_{ext,1}} + \frac{1}{Q_{ext,2}} \quad (138)$$

which gives

$$Q_0 = Q_L(1 + \beta_1 + \beta_2). \quad (139)$$

It is desired to measure Q_0 directly, but in steady state, only the scattering parameters (S_{11} and S_{21}) can be measured. The scattering parameters are defined as

$$|S_{11}|^2 = \frac{P_r}{P_f} \quad (140)$$

$$|S_{21}|^2 = \frac{P_t}{P_f} \quad (141)$$

with P_f being the forward power and P_r being the *total* reverse power from the cavity, the vector sum of the reflected and emitted traveling waves. β_1 and β_2 can be expressed in terms of these scattering parameters [4, p. 48]

$$\beta_2 = \frac{|S_{21}|^2}{1 - |S_{11}|^2 - |S_{21}|^2} \quad (142)$$

$$\beta_1 = \frac{1 - |S_{11}|}{1 + |S_{11}|} \quad (143)$$

or

$$\beta_1 = \frac{1 + |S_{11}|}{1 - |S_{11}|} \quad (144)$$

depending on coupling strength. For $\beta_1 < 1$ (Equation 143), the coupling is considered “undercoupled,” for $\beta_1 > 1$ (Equation 144), the coupling is considered “overcoupled.” The remaining case, $\beta = 1$, is the “matched” condition. In this configuration, there is no reverse power because the emitted and reflected power from port 1 are equal in magnitude and opposite in phase. These equations are derived assuming the measurement is being done in a steady-state and that $\beta_1 \gg \beta_2$ so that P_t can be dropped while considering the behavior of Port 1. It can be seen from these equations that some additional knowledge is required for accurate interpretation of the measured values, any measurement of $|S_{11}|$ can result in two different β_1 results. This information can be found by changing the parameters of the measurement. Observing the change in $|S_{11}|$ because of adjusting the coupling is a known way to determine if the cavity is over or under coupled. For example, increasing the length of a probe coupler should increase the coupling strength β . If $|S_{11}|$ increases because of this change, the coupler is now overcoupled. If $|S_{11}|$ drops, the coupler started as undercoupled. This type of change must be done with care to avoid confusion from crossing $\beta = 1$. Additionally, the P_f can be modulated, and the behavior of P_r and P_t can be used to determine the coupling. This method is only realistic when the time constant of the cavity

$$\tau = \frac{Q_L}{\omega} \quad (145)$$

is large enough to be measured easily. Because Q_L is so small when the cavity is normal conducting (~ 1000), it is only possible to measure the coupling by modulation of the forward power when the cavity is superconducting, with $Q_L \approx 1 \times 10^7$ or higher. For a realistic cavity, two couplers are used. The “input” coupler is a high power coupler designed to be either close to matched or overcoupled depending on the situation. The second coupler, called the “pickup” coupler, is used as a field probe, allowing simple measurement of the cavity stored energy. This coupler is significantly undercoupled so that the power radiating out of this coupler (P_t) is very low, so it can be treated as a perturbation, greatly simplifying the mathematical analysis of the measured results. A useful rule of thumb when choosing the pickup coupling is to ensure that P_t is always below 1 mW, even at the extreme limit of cavity performance.

6.2 Bandwidth Considerations

The bandwidth of the cavity is determined by the total energy loss in operation, giving the equation

$$\Delta f_{BW} = \frac{f_0}{Q_L}. \quad (146)$$

Because Q_0 is so large for superconducting cavities, this means that the cavity bandwidth is often dominated by the choice of $Q_{ext,1}$. When testing a cavity in controlled environment, the cavity frequency can be allowed to drift as needed. This means that the bandwidth of the cavity can be kept quite small, often as close to matched as possible, because the RF circuitry is flexible enough in the configuration to not require a large bandwidth. This keeps the power requirements as small (and inexpensive) as possible, simplifying the testing hardware required. For operation in a real machine, the cavity frequency must be tightly controlled for stable operation. This means that the RF control systems must react to the changing cavity conditions while keeping the amplitude, phase, and frequency of the cavity within tight, preset parameters. This is a much more difficult task, and the cavity bandwidth must be large enough to allow this stable operation of the cavity. This means heavily overcoupling the cavity, meaning that P_e dominates cavity losses and $P_r \approx P_f$. How overcoupled the cavity must be is determined by size and time scale of the frequency perturbations the cavity is likely to experience as well as the control systems used.

6.3 High Power Concerns

When the cavity is heavily overcoupled, P_f is much larger than P_d . This means that the input couplers must be able to handle full power in the forward and reverse direction simultaneously, and this leads to stringent cleaning standards for couplers, much like the cavities. Couplers are often cleaned and conditioned much like normal conducting cavities are, and their cooling can be a serious issue. Couplers designed for such high power and so heavily overcoupled will often be a significant perturbation to the cavity fields, and their

effect on cavity frequency and performance must be taken into account during standard cavity design. Additionally, the coupling is chosen based on an assumption of Q_0 , but in reality, this can vary from cavity to cavity. The ability to vary the input coupling in situ adds mechanical complexity to coupler design, but allows fine tuning of each cavity to either minimize power requirements on an overperforming cavity or allow operation of a poorly performing cavity. This benefit extends to Dewar testing, where Q_0 is varying more widely than in real operation. A variable coupler allows for testing or conditioning at a higher field by keeping the cavity close to matched for a much wider range of cavity performance.

7 Electromagnetic Cavity Simulation

The ideal HWR geometry is an instructive tool for estimating cavity performance because it is analytically solvable. This geometry is, however, not practical for use in a real accelerator. There are many features that must be added and optimized that cannot be treated analytically, and must be simulated. Modern electromagnetic simulation software can simulate the parameters of almost any geometry and evaluate its figures of merit, often only limited by the amount of computing power available. In this section, the basic methods that these solvers use to simulate these complex geometries will be discussed. In addition, more practical considerations and methods used to perform the design simulations will be presented for use in the rest of this document.

7.1 Finite Element Solvers

In order to solve a complex geometry like that of a realistic cavity, the volume of the cavity is broken into a discrete set of interlocking simple geometries, called elements. Tetrahedral elements are the most commonly used for cavity design because they often give the most accurate representation of the complex surfaces involved in cavity design. An example of this can be seen in Figure 15, generated in ANSYS-APDL. The fields in the cavity are then assumed to have a simple representation in the volume of the element, expressed as a low-order (often 1st or 2nd) function of space. An example of the structure of these elements can be seen in Figure 16. Because the fields in the cavity are assumed to be monochromatic, the electric fields in the cavity can be described by the Helmholtz Equation:

$$(\nabla^2 + k^2)\vec{E} = 0. \tag{147}$$

This equation is then discretized and applied to each element. Elements “interact” at fixed places, generally along the outside edges, called nodes. The parameters of each element are iterated with the goal of minimizing the global stored energy. The study and development of these solvers is an active and growing field, but for the basic methods and mathematical foundations of these solvers, “Electro-magnetics and Calculations of Fields” by Ida and

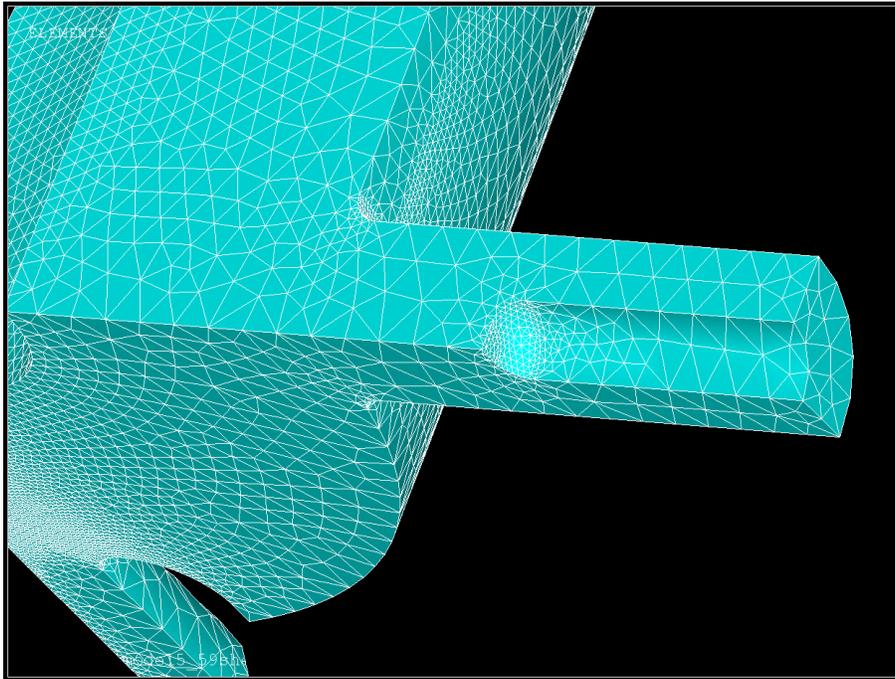


Figure 15: A cavity vacuum space meshed with tetrahedral elements, a coupler probe is in the foreground with a beam port cup visible in the background.

Bastos [26] or “Finite-element Methods for Electromagnetics” by Humphries [27] are recommended as starting places. What this procedure allows is the accurate representation of the fields in a complex geometry with evaluation of the electromagnetic figures of merit.

7.2 Geometry Creation

The first step of any simulation is defining the problem. Because the geometries involved are very complex, it was chosen to use a Computer-Aided Design (CAD) program, SolidWorks [28], to create the needed geometry. This is because the native geometry creation in the electromagnetic solver used for this project (and most other solvers) works by the creation and Boolean combination of primitive objects (cone, spheres, rectangular boxes) or swept two dimensional shapes. After the geometry is completed, it must be exported in a form that can be read by the simulation software. The format of choice for this was SAT, for ease of both export and import. Depending on the complexity of the geometry, it may be necessary to use different SAT [29] file settings or versions (as of publication, there were 18 versions) for the simulation software to properly interpret the geometry.

Representation of field with linear elements in 3d

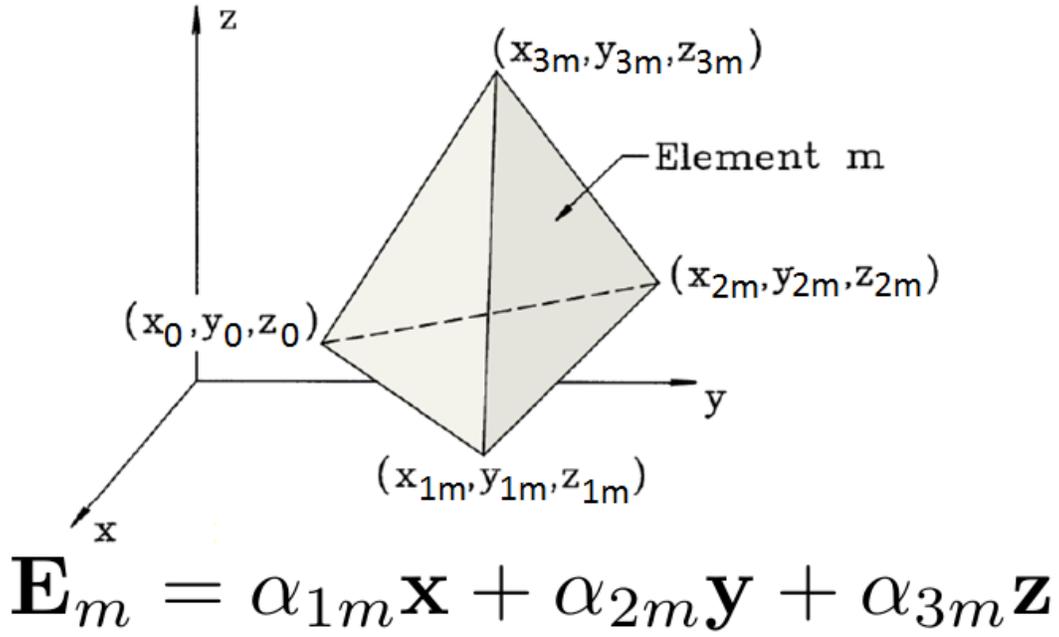


Figure 16: Structure of a first order (linear) tetrahedral element. Image credit to F. Krawczyk.

7.3 Simulation Procedure

Assuming that the geometry has been successfully created and imported into simulation software, the basic electromagnetic simulation occurs in four major stages: Setup, Meshing, Solving, and Post Processing. Each of these stages carries its own importance and subtleties, but in the following subsections, the work flow for each stage in the context of the simulation software used for the work presented later will be outlined. The simulation software used for the vast majority of the work presented is ANSYS-APDL [30]. Complete scripts detailing the execution of the described procedures can be found in the Appendices. It should be noted that, while the general process can be applied, the details of repeating this work in different simulation software (CST-Microwave Studio [31], ANALYST [32], HFSS [33], ACE3P [34]) will very likely be quite different, i.e., *Caveat Lector*.

7.4 An Aside - Computer Resource Management

Computer resource management is an important setup for large scale simulations. This includes storing solutions files on the fastest available disk drive available. Using solid state drives in place of traditional palette drives can decrease simulation time by 30 - 40%, depending on the type of solution being done, assuming a large enough solid state drive can be acquired. Eigenmode simulations tend to be memory limited, as opposed to processor speed limited, meaning that investments in upgrading available memory speed and size can speed simulations considerably. Additionally, the temporary files generated during simulation and results files can be extremely large. If the disk drive you are using is 32-bit formatted, it will be unable to store these files. Some simulation software have a feature that ensures no files are larger than the 4 Gigabyte limit for FAT32 formatted drives. A 32-bit architecture also limits the amount of memory that can be allocated to a single program to 2 Gigabytes as well, so moving to 64-bit architecture removes both of these limitations. Depending on the software in use, multiple processor cores can be used in parallel to increase solving speed. When starting a simulation, ensure that the maximum number of processors, largest amount of memory, and fastest hard drive space available are allocated for the job.

7.5 Simulation - Setup

Given that the skin depth for high conductivity materials (and especially superconducting materials) is so small, the fields *in* the conductor and the losses generated can be treated as a perturbation. That leads to two standard types of boundary conditions, Perfect Electric Conductors (PEC), and Perfect Magnetic Conductors (PMC). The PEC boundary condition indicates that $E_{\parallel} = 0$, and $H_{\perp} = 0$, and is appropriate for conductor surfaces and some boundary conditions depending on the mode structure you are trying to simulate. The PMC boundary condition indicates that $E_{\perp} = 0$ and $H_{\parallel} = 0$, and is used for a majority of symmetry planes. All exterior surfaces of the model must be defined as one of these. Starting from a basic imported vacuum space, it is very useful to define groups of important surfaces for later reference. These groups can include symmetry planes, conducting planes of the same conductivity, or areas to be used for special types of analysis. For the HWR model seen in Figure 17, generated in SolidWorks, there are two groups of critical surfaces, the symmetry planes seen in blue and the cavity RF surface in gray. The symmetry planes are defined as PMC, and the RF surfaces are defined as PEC. Additionally, any surface can be assigned a conductivity. This will be used to perturbatively calculate losses based on the magnetic fields on that surface. While the cavity RF surfaces are generally assigned the same conductivity, surfaces of interest can be assigned different conductivities based on their material type. This can be very useful for specialized investigations. Numerical constants can also be set for convenient use later. Grouping their declarations early in a script eases later modifications. The type of element must be defined as a high-frequency

eigenmode type element, and the order of this element must be set. Material properties are also set here, but for a simple high-frequency eigenmode solution, $\epsilon = \epsilon_0$ and $\mu = \mu_0$ are the only required definitions.

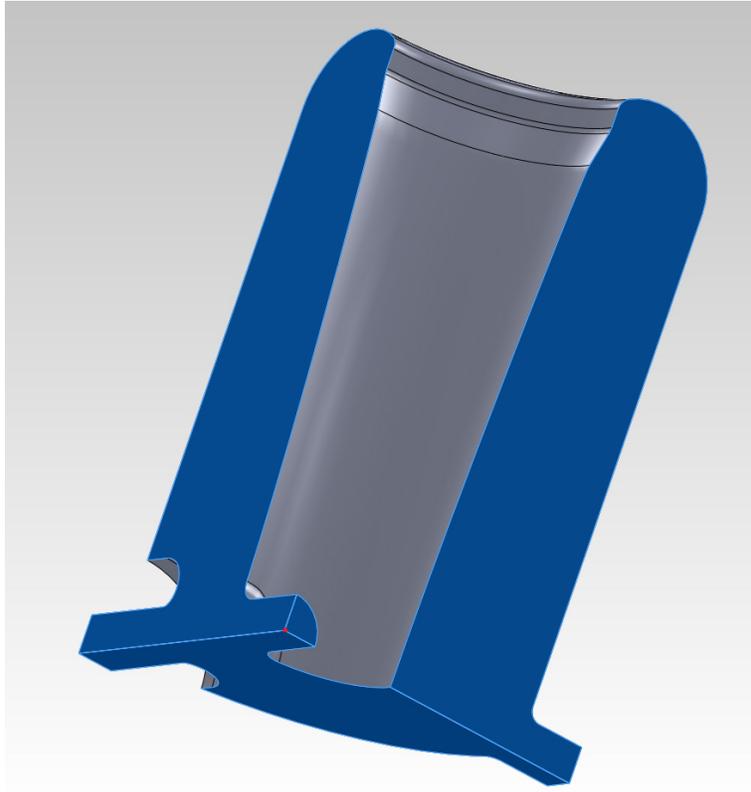


Figure 17: HWR vacuum model. Symmetry planes can be seen in blue, Perfect Electric Boundaries can be seen in gray, the cavity origin is marked in red.

7.6 Simulation - Meshing

Finite element simulations require the cavity volume to be represented by a collection of smaller volumes called elements. The process of defining the size and shape of these elements is called meshing. These elements should be, at largest, one tenth the wavelength you expect to simulate to improve accuracy [26]. The accuracy of the remaining figures of merit are strongly influenced by the quality of the mesh. Curved surfaces are especially sensitive to meshing as the elements on the surface must be small enough to accurately represent the curvature. A rule of thumb for small curved features is the element size should be smaller than one third the radius of curvature. An example of inadequate meshing can

be seen in Figure 18.

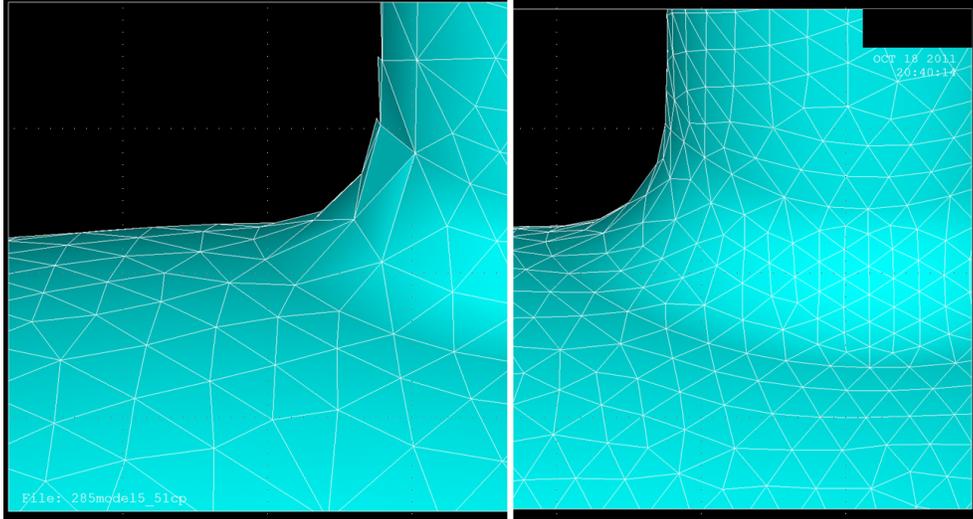


Figure 18: Comparison of different meshing on a curved surface. Left: Poor meshing on a curved surface. Right: Same surface with improved mesh density. Images generated using ANSYS.

While most software has advanced automatic meshing algorithms, additional manual refinement may be required based on the geometry being simulated. The first, most common, mesh optimization step is the setting of a maximum element edge length. This can be done for both whole volumes and on areas of known high surface fields. For the volume sizing, it is important to set a global maximum value for an element edge length to prevent drastic gradients in element size. The standard rule of thumb for electromagnetic simulation calls for a maximum element edge length of $\lambda/10 \approx 9$ cm for 322 MHz. For complex geometries like the ones used for HWRs, additional surface refinement may be required to accurately calculate surface losses and peak surface fields. For accurate surface fields on the small curved surfaces, elements as small as 1 mm were found to be needed in some locations of the cavity. If the transition between these element sizes is not carefully controlled, poor mesh quality can result. Setting a global maximum on the element size is a powerful tool for regulating this transition. A major factor in mesh quality tolerance is the order of element chosen (i.e., what *order* function is used to represent the fields in an element). As can be seen in Figure 19, increasing the mesh density reduces frequency error down to a baseline inaccuracy that is solver dependent. For ANSYS-APDL, this error was found to be roughly 1 part in 10^5 , or $\sim \pm 50$ kHz out of 322 MHz. This level of accuracy was achieved with second order elements at a much lower density than for first order elements. While the solve times for the same accuracy is roughly equivalent, the use of second order elements greatly reduces the meshing tolerances while giving similar or

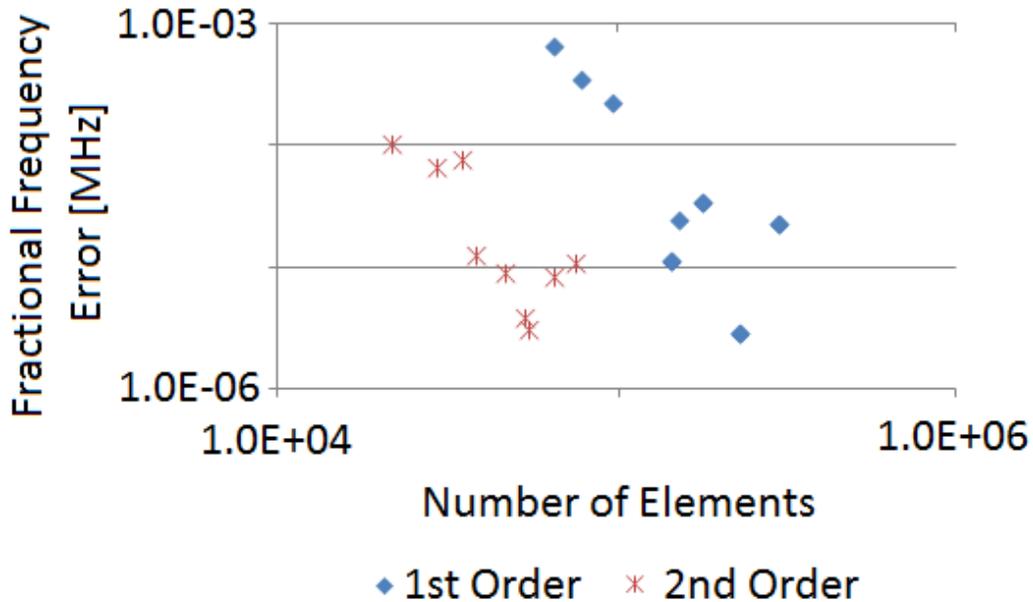


Figure 19: Fractional frequency error for different mesh densities in identical geometries for first and second order elements.

improved surface field accuracy. There are sometimes limitations from the meshing that can help drive the decision to strive for first or second order meshing. Some geometries are so complex that their mesh densities must be high for geometric fidelity. Depending on the available computing hardware, this type of geometry can often require far more mesh elements than could be solved with second order elements, but can be solved with first order elements at a loss of accuracy. In this situation, all results, especially surface fields, must be interpreted with this in mind. When the computing hardware allows, this type of complex geometry should be solved using second order elements if possible, especially if surface fields are of particular interest. Meshing must also be approached carefully with geometries that are very simple. Meshing software will often generate very large elements in these geometries if allowed, and care must be taken to ensure they do not grow large enough to sacrifice accuracy. This is an ideal situation to use a global element edge length limit.

Additional care must be taken when generating geometry to avoid causing meshing problems. While this is a more serious issue for mechanical simulations, and will be discussed in greater detail in the relevant section, there are certain features that should be avoided if possible, and these are places of optimization to generally reduce the size of simulations without great sacrifice for accuracy. CAD geometries are represented in simulation

software as a continuous and closed collection of surface, each bounded by lines. Mesh structure must respect the shape and boundary of these surfaces, and care should be taken to not create very small areas unnecessarily, or otherwise collect many boundaries together without purpose. An example of a mesh that is overly dense at surface intersections can be seen in Figure 20. This type of feature can be tolerated if the solution time isn't excessive, but when working on a very large simulation, it may be beneficial to find instances like this and correct the geometry to allow for more optimal meshing.

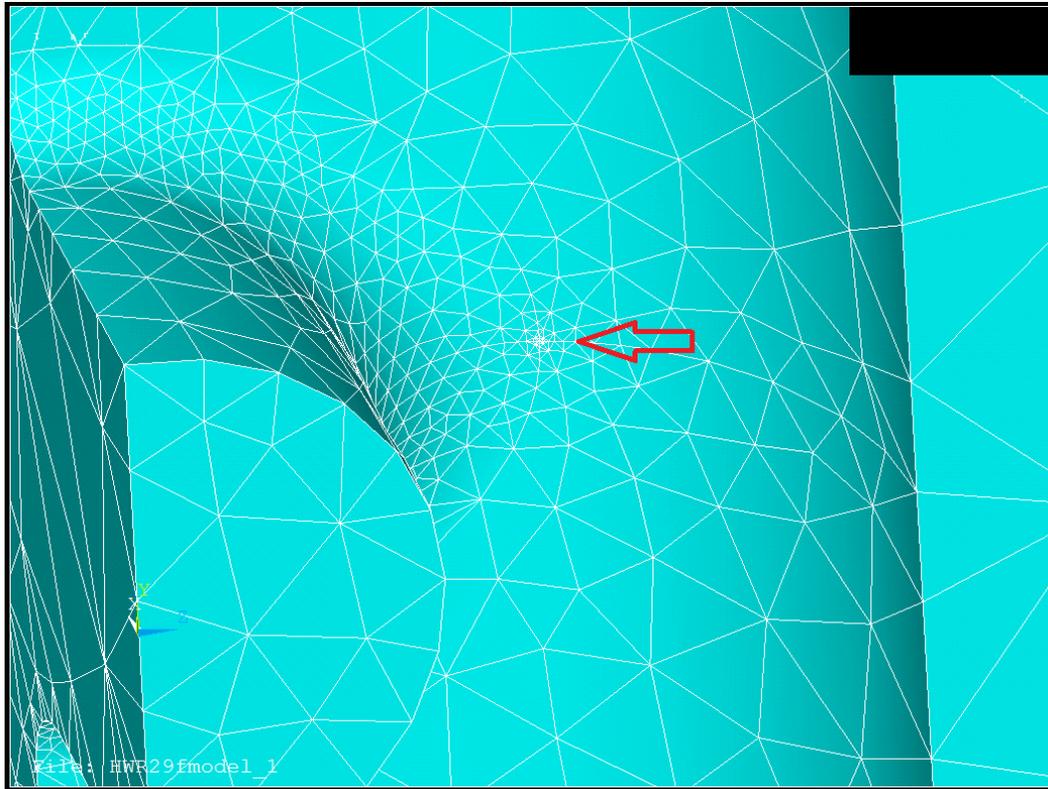


Figure 20: The intersection of four surfaces can force excess mesh density, indicated by the red arrow. Images generated using ANSYS.

Automatic meshing algorithms often make meshing decisions based strictly on the geometry, not on the structure of the mode being simulated. This is understandable because the information is not available to the program before the simulation is complete. Some programs like ANALYST have automatic iteration algorithms that use solution data to adjust and optimize the meshing until the solution has converged to the accuracy desired. For most simulations software, foresight can dramatically improve the quality of the simulation results. We've touched on this for surface fields and especially external surfaces

with a large curvature, but this is also quite important for the accelerating field profile. Large elements on the beam axis result in a non-physical accelerating field because the fields cannot be properly interpolated between elements to give a smooth shape. Improving the density of the meshing around the beam axis gives a more accurate accelerating field profile. Both of these behaviors can be seen in Figure 21.

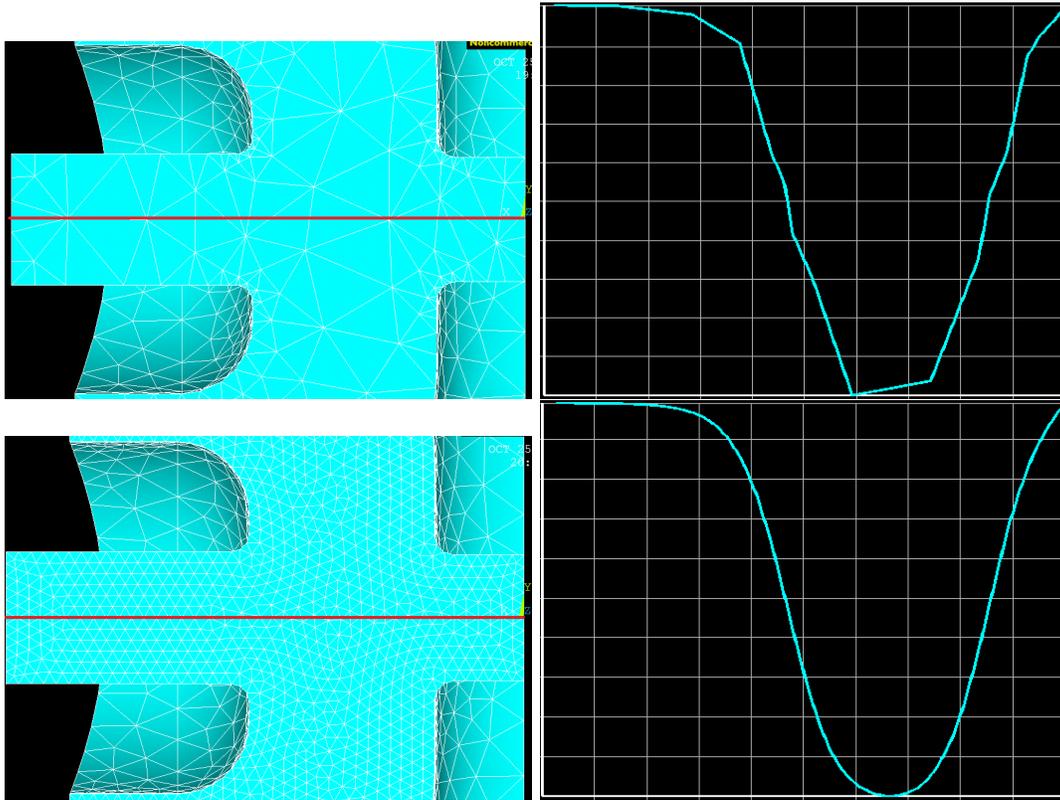


Figure 21: Top Left: Poorly meshed cavity, beam axis marked in red. Top Right: Resulting accelerating electric field. Note the non-physical roughness that results from the large element size. Bottom Left: Cavity with improved meshing on beam axis which is marked in red. Bottom Right: Resulting accelerating electric field. Note the improved field profile that results. Images generated in ANSYS.

7.7 Simulation - Solving

Once the cavity has been meshed and appropriate boundary conditions (PEC, PMC, conductivities, etc.) have been specified, most solvers require some options set to calculate an eigenmode. For ANSYS-APDL, the solver needs to be set to modal, and a frequency range given for the mode(s) desired. The solver will search for eigenvalues (resonant frequencies)

for the cavity element matrix from the low end to the high end of the range until the requested number of modes are found. For most purposes, this range should only be as large as needed to encompass all the variation in the single frequency anticipated. Searching for a range that is $\sim \pm 10\%$ of the desired resonant frequency is often large enough to still find the resonance even after relatively large geometry changes. For complex structures, this range may also include unintended trapped modes in small features like couplers or tuners, so it is often advisable to scan a larger frequency range for modes before focusing exclusively on the mode of interest. After launching the solution process, the solver will often list the amount of memory needed for the solution. “In core” solutions are done with the entire simulation resident in memory. This greatly speeds simulation time and is the desired mode of operation. “Out of core” solutions require the solver to transfer data between memory and the hard disk, a much slower process. This is one of the prime reasons simulation programs should be allowed access to as much memory as possible.

7.8 Simulation - Post Processing

The first step after a new simulation finishes is verifying that the solve has completed properly. While there are various intuitive methods for this, the most effective is often plotting the fields in the cavity. For most geometries, the mode structure is known, at least in concept. Qualitative comparison of the solved fields with this general form (comparing a HWR simulations to the general structure of the ideal HWR mode, for instance) can give a good indication of the quality of the solve. This will also identify surfaces requiring additional mesh refinement, an example of which can be seen in Figure 22. In this example, the coarse meshing resulted in a non-physical roughness to the surface fields, with one element having artificially high fields. Improved meshing removes a great deal of this roughness, and gives a more accurate peak surface field. Once the solution has been evaluated and found to be sufficient, extraction of the basic cavity figures of merit can begin. This starts with extraction of bulk properties of the mode and cavity, such as total dissipated power, total stored energy, and cavity volume. These values must be scaled depending on the amount of symmetry used in the cavity. Some solvers automatically do this, but in ANSYS-APDL, it must be done by hand. It should be noted that all stored energy dependent quantities like dissipated power and peak surface fields must be normalized to the stored energy. ANSYS-APDL solves the cavity mode at some nominally small amount of stored energy, and this value must be used to scale the cavity figures of merit. Once the dissipated power, frequency, and stored energy are known, the cavity quality factor can be calculated:

$$Q = \frac{2\pi fU}{P_d}. \quad (148)$$

This can then be multiplied by the surface resistance given in the setup phase of the simulation to calculate the geometry factor:

$$G = R_s \times Q. \quad (149)$$

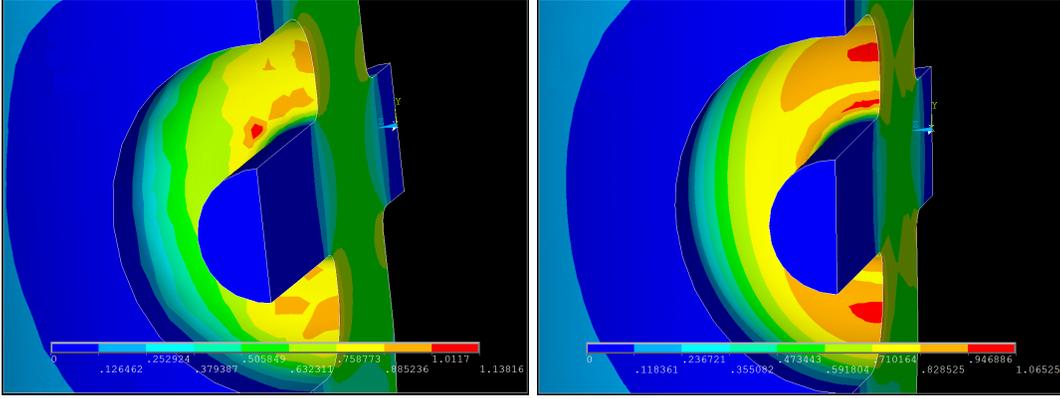


Figure 22: Left: Magnitude of the surface electric field on a beam port cup with inadequate meshing for surface field accuracy. Right: Identical geometry with improved meshing and more realistic surface fields.

The peak surface fields can be found in ANSYS-APDL by making an element table of the field in question. This table can be sorted by field magnitude and the maximum value can be extracted. Again, this is the field value for the nominally low stored energy of the simulation, so E_{pk}/\sqrt{U} and B_{pk}/\sqrt{U} should be calculated for later scaling when the proper stored energy is known. The remaining figures of merit require V_{acc} , and the bulk of the remaining post processing will revolve around its calculation. The first step is to extract the accelerating electric field component along the beam axis. In ANSYS-APDL, this is done via the PATH command, specifying the beginning, end, and field component desired. This gives the value of the accelerating field at a number of discrete points along the line specified, and this number must be large enough for this to be a good approximation of the actual (continuous) accelerating field profile. The first calculation is

$$V_0 = \int_{-\infty}^{+\infty} |E_{acc}(z)| dz = \sum_i^N |E_{acc,i}| \Delta z \quad (150)$$

which gives the maximal acceleration possible, V_0 . The next step is calculating the accelerating voltage, which includes the field's variation in time. This calculation uses the general form

$$V_{acc} = \sqrt{\left(\int_{-\infty}^{+\infty} E_{acc} \sin(\omega t) dz\right)^2 + \left(\int_{-\infty}^{+\infty} E_{acc} \cos(\omega t) dz\right)^2} \quad (151)$$

which does not depend on the choice of ($z = 0$). This becomes

$$V_{acc}(\beta) = \sqrt{\left(\sum_i^N E_{acc,i} \sin\left(\frac{\omega z_i}{\beta c}\right) \Delta z\right)^2 + \left(\sum_i^N E_{acc,i} \cos\left(\frac{\omega z_i}{\beta c}\right) \Delta z\right)^2} \quad (152)$$

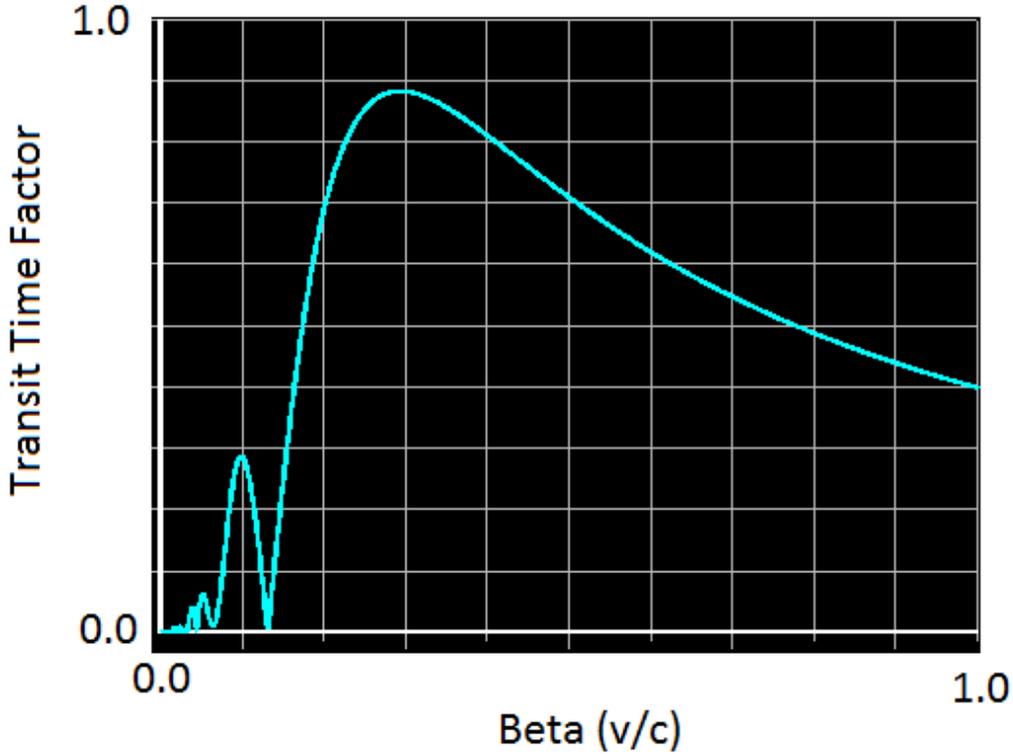


Figure 23: Plot of transit time factor versus particle beta.

in practical terms, although this can be simplified sometimes by the choice of ($z = 0$) and cavity symmetry. For instance, a HWR with the origin placed in the geometric center of the cavity can neglect the \cos term because of symmetry. This formula depends on β , although does not directly give the optimum β . To find this, $V_{acc}(\beta)$ is calculated for β ranging from 0.001 to 1 by steps of 0.001. Taking the maximum of this set gives the optimum β with enough accuracy for essentially all applications. This data can be normalized to V_0 to give the transit time factor curve, an example of which can be seen in Figure 23. Once the optimum β has been found, V_{acc}/\sqrt{U} is calculated, giving the last of the desired field quantities. This is used to calculate R/Q :

$$R/Q = \frac{1}{\omega} \times \left(\frac{V_{acc}}{\sqrt{U}} \right)^2. \quad (153)$$

To find the figures of merit for the cavity at full field, some normalization must be chosen. For the design of the FRIB cavities, it was convention to choose a V_{acc} desired for a cavity and normalize to that. This determined the stored energy (U) from V_{acc}/\sqrt{U} , and therefore the peak surface fields from the equivalent ratios already calculated. A detailed ANSYS-

APDL script used for all of this work can be seen in Appendix A. To this point, the basic electromagnetic figures of merit have been calculated. Sometimes, more details or complex problems needed to be solved, and these often can be solved in post processing. This type of problem often involves field levels or losses on certain parts of the cavity (RF joint, plunger, flange). These problems are solved by selecting the areas, elements, and nodes related to the area in question, and repeating the same process used to extract the fields or losses for the full cavity. If necessary, the losses can be scaled to a different resistivity to simulate a small normal conducting part of the cavity and the quality factor recalculated for the new losses.

7.9 Coupling Power to the Cavity

An obvious but non-trivial feature of any cavity is a port for coupling power from an external generator into the cavity. Coupler design drives the size and positioning of the RF port. For FRIB, it was decided to use electric probe coupling for both input and pickup couplers, ninety degrees rotated from the beam axis, as can be seen in Figure ???. While each of these couplers will perturb the cavity fields, couplers designed for high power are often much larger than testing couplers, and will perturb the cavity more to achieve the desired stronger coupling. These disturbances will shift the resonant frequency of the cavity and must be taken into account during design, although the shift itself is often measured experimentally. Good coupler and cavity interface design also ensures that regions of high surface fields are not created around the coupler port, although the location of the RF port on the outer conductor means this is not generally challenging. Additionally, the coupling strength as a function of probe length can be simulated with a standing wave simulation. A typical geometry for this type of simulation can be seen in Figure 24. The coupling strength of the port can be simulated by solving the model with the termination of the coupler port (indicated by a red arrow in Figure 24) as both PEC and PMC, and extracting the fields on the boundary. This is called the Current-Voltage method [35, 36], and assumes that the fields in the coupler port are purely TEM at the termination. At the PEC termination, the electric field is zero because the TEM only allows fields transverse to the surface, but these are forced to zero by the PEC condition. The distribution of the magnetic fields can be seen in Figure 25, and it has the characteristic $1/\rho$ dependence of a coaxial TEM mode. An identical distribution for the electric field is seen when the cavity is simulated with a PMC termination. The Current-Voltage method uses a lumped circuit for the coaxial transmission line, and calls for the calculation of the “voltage” between the inner and outer conductor on a PMC boundary and the “current” across the termination for a PEC boundary. Using this model, we can calculate Q_{ext} as

$$Q_{ext} = Q_i + Q_v \quad (154)$$

where

$$P_i = \frac{\omega U}{Q_i} = \frac{1}{2\eta} \int_S |\vec{E}|^2 ds \quad (155)$$

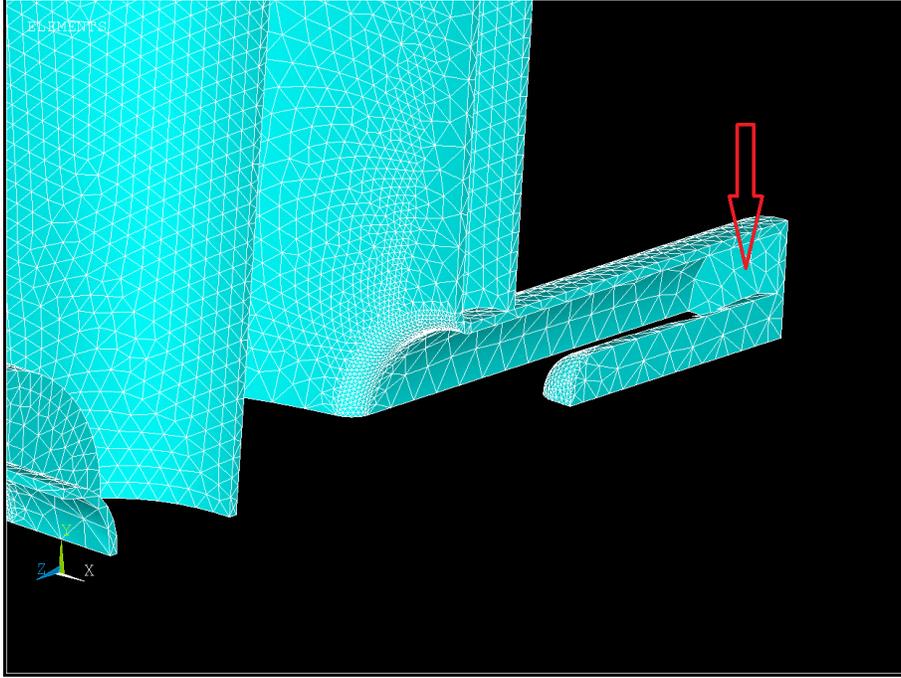


Figure 24: The coupler probe can be seen in the foreground connected to the RF port by the surface indicated by the red arrow. This is a quarter of the geometry cut on the symmetry planes.

and

$$P_v = \frac{\omega U}{Q_v} = \frac{\eta}{2} \int_S |\vec{H}|^2 ds \quad (156)$$

following the standard definition of a quality factor. The form of a TEM wave in a coaxial transmission line gives a relatively simple form for Q_{ext} :

$$Q_{ext} = \frac{\omega U}{\pi \ln\left(\frac{b}{a}\right)} \left(\frac{\eta}{(aE_{pk})^2} + \frac{1}{\eta(aH_{pk})^2} \right) \quad (157)$$

where a and b are the inner and outer conductor radii for the coupler port, $\eta = 377 \Omega$ is the impedance of free space, and E_{pk} and H_{pk} are the peak fields on the termination for the PMC and PEC boundary conditions, respectively.

Comparison of simulation and measured data for a prototype $\beta = 0.53$ HWR can be seen in Figure 26. The details of measuring the Q_{ext} on a real cavity will be discussed later. The most error-prone part of this measurement is the transformation of a given coupler tip length to penetration into the cavity volume. While this is easy to do in a CAD model, errors in manufacturing and positioning add significant error to this part of the

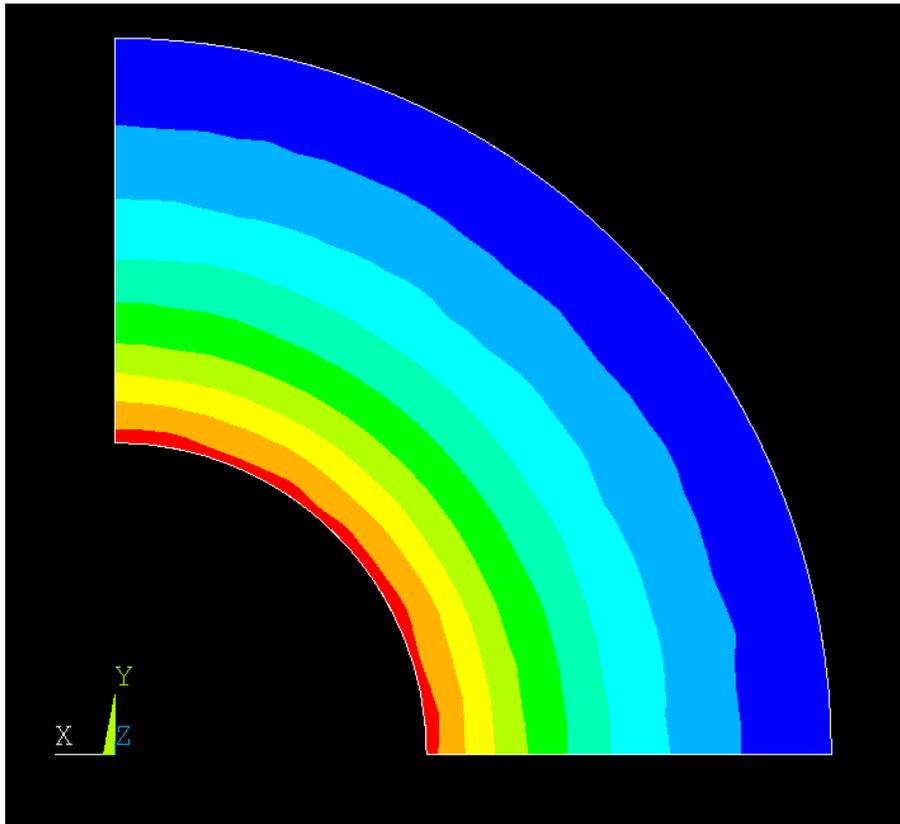


Figure 25: Plot of the magnitude of the magnetic field on the PEC termination of a coupler port. Simulated with planes of symmetry in x and y.

measurement. This can be seen in the data plotted in Figure 26, where the measured values (thin black trend lines) and the simulated data (large black squares) are offset from each other by as much as 5 mm. This can even be seen in the shift between different measurement runs which are a similar distance apart. The slope of the measured and simulated data agrees well, however. This is because the errors in fabricating the port diameter and measuring the coupler probe diameter are much smaller, and can be simulated accurately. It is important to notice that at very negative penetration into the cavity (very small probe lengths), the trend line changes significantly from both the measured and simulated data. This is because the approximation used in the simulation breaks down significantly for very small probe lengths, the modes no longer have enough space to transform into a pure TEM shape before reaching the termination plane of the RF flange. To accurately model this behavior, a more advanced theoretical method or a traveling wave simulation is required. These additionally require a much more accurate description of the geometry involved, including the coupler feed through and flange shape.

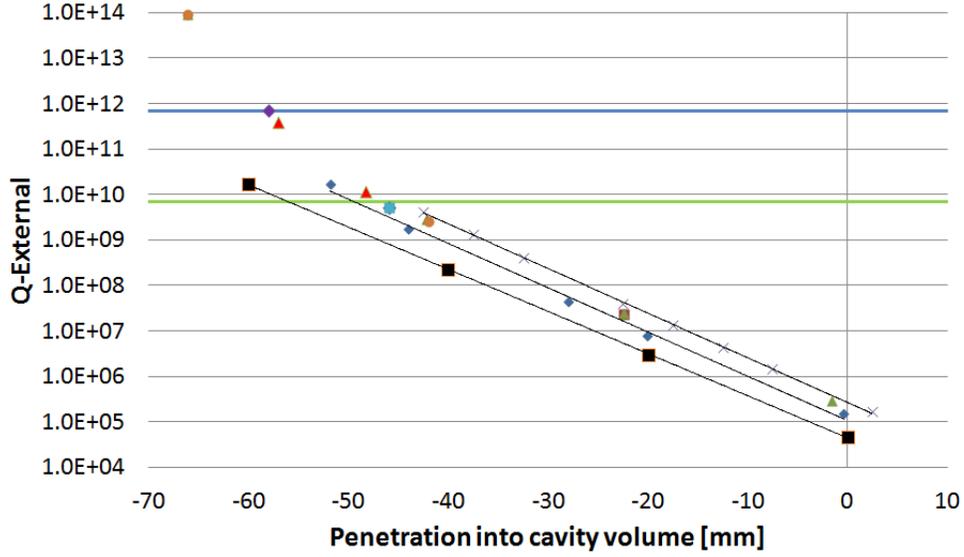


Figure 26: Plot of measured and simulated Q-externals for a 0.53 HWR prototype. The simulated data is represented by the large black squares. The rest are repeated rounds of measured data.

7.10 Etching Coefficient Estimation

A major step in cavity processing is a bulk etching of the RF surfaces. The goal of this etching is to remove around $150 \mu\text{m}$ of material from the RF surface that has been damaged by cavity fabrication, improving cavity performance. This etch generally changes the cavity frequency enough so as to require that this shift be well understood so that the correct final frequency can be repeatably achieved. Using Slater's Theorem (see Section 8.4)

$$\frac{\Delta f}{f_0} = \frac{1}{4U} \int_{\Delta V} (\epsilon_0 E^2 - \mu_0 H^2) dV \quad (158)$$

this shift can be perturbatively estimated. Because the simulation procedure requires the volume be broken into discrete elements, this formula requires modification to:

$$\frac{\Delta f}{\Delta T} = \frac{f_0}{4U} \sum_i (\epsilon_0 E_i^2 - \mu_0 H_i^2) dA_i \quad (159)$$

where ΔT is the thickness of the etched layer, A_i is the area of the external face of element i , and the fields E_i and H_i are the fields from surface element i . For an ideal HWR, this can be calculated directly from Slater's theorem: $-1383 \text{ Hz}/\mu\text{m}$. This large negative shift is intuitive because etching of an ideal HWR effectively makes the cavity longer, directly

lowering the cavity frequency. In the process of cavity optimization, the changes to the electric field region increase its contribution to this shift greatly, producing a less negative coefficient. For the HWR design being developed here, the simulated etching coefficient is $-206 \text{ Hz}/\mu\text{m}$. In reality, this calculation makes assumptions that are unlikely to be true. The most important one is that the etching is uniform over the cavity surface, but etching is a quite complex and dynamic process, involving both fluid dynamics and thermodynamical issues. What can be said is that, in an ideal etching setup with very good temperature and flow control, the achieved frequency shift should approach the simulated value. Of additional interest is using etching as a way to fine-tune the frequency of a cavity. This process is called differential etching, etching shift versus the acid fill level can be seen in Figure 27 for an example HWR. This process has been demonstrated and used for cavity production for QWRS at TRIUMF [37] and elsewhere. From this plot, it can be seen that

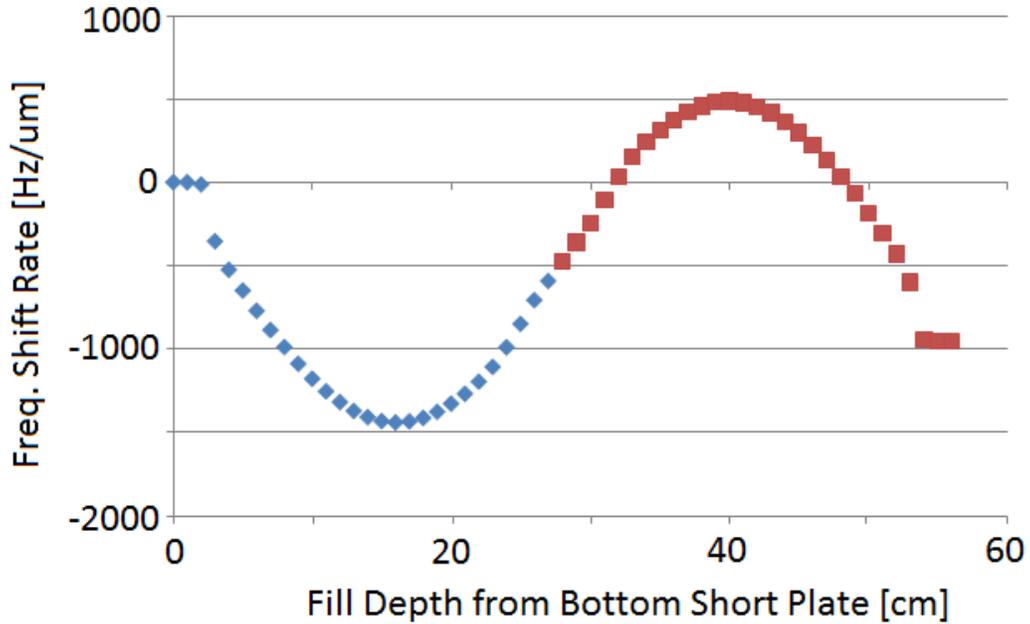


Figure 27: Plot of simulated etching coefficient for a prototype $\beta = 0.53$ HWR filled with acid to different levels. The blue diamonds represent simulation data, and the red squares are this data mirrored assuming symmetry.

a much more positive frequency shift can be achieved if the cavity is filled such that the full electric field region is etched while only one of the magnetic field regions are etched. Experience gained from fabricating a cavity design is required to determine if such a cavity tuning process would need to be developed.

8 Coupled Electromagnetic and Mechanical Simulations

A mature cavity design starts with an electromagnetic model, but must account for various practical concerns. During normal operation, the cavity walls may be subjected to small forces, which will result in small perturbations to the shape of the cavity. This deformation will change the cavity vacuum space, and therefore the resonant frequency. This can have significant controls implications and must be evaluated and optimized before cavity production. Because matching the simulation setup to the exact conditions the cavity will see in operation is difficult, several rounds of simulation and prototyping can sometimes be required to achieve the accuracy and performance desired. In this section, the procedure for this type of simulation, the types of pressures simulated, and the results of these simulations for the HWR developed in the last section will be discussed in detail.

8.1 Generalized Procedure

The mechanical simulation process has many parallels to the electromagnetic simulation procedure. The major steps remain Geometry Creation/Setup, Meshing, Solving, and Post-Processing. From the cavity vacuum geometry, the SHELL command in SolidWorks will create the niobium space that represents the cavity walls. This command has the option to shell with multiple thicknesses in different parts of the cavity, and this can be used to model cavities made from materials of various thicknesses. From this, the helium vessel can be constructed around the cavity, using different bodies for different materials. Once the material space has been generated, it is exported as a SAT (or other appropriate file type) file for importation into ANSYS. Once in ANSYS, the different bodies must be “glued” together. This command (“VGLUE”) takes different volumes that are next to each other and joins them for simulations purposes. Before this command is given, two volumes may be touching, but where they touch they both have their individual surfaces with no reference between them. After the command, the redundant surfaces are removed so that the different volumes *share* a common surface, meaning that the mesh and mechanical forces will be shared across the boundary, effectively binding them together. From the material space of the cavity, seen in Figure 28, the vacuum space must be recreated. This is done by using the existing geometry as the base for the remaining areas needed to create the continuous set of areas that define the vacuum space: symmetry faces and ends of ports. Once these are created and the vacuum space has been created, all volumes must be meshed. Because there are volumes that share the same areas, most notably the vacuum space and the cavity space, special care must be taken to mesh in the correct order with the correct refinement to give a satisfactory mechanical and electromagnetic solution. For this purpose, it is highly recommended to use second order elements for the electromagnetic simulation. Using the very fine mesh required for a satisfactory electromagnetic solution will almost certainly force there to be an extremely high number of elements in the mechanical simulations, unnecessarily stretching the limits of even a very high powered computer.

Other than this caveat, the procedure for performing the electromagnetic simulation is the same as described earlier in this document. Once the frequency is simulated, the next step

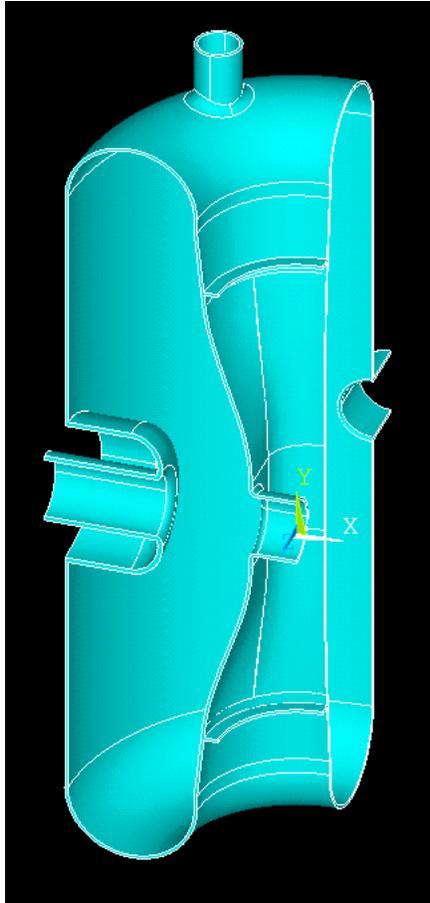


Figure 28: Material space after importation into ANSYS-APDL.

is the setup of the mechanical simulation. This includes giving all volumes the appropriate material properties, including the vacuum space. This may seem counter-intuitive, but treating the vacuum space like a material is essential for accuracy. This is because this process preserves the meshing in the vacuum space, allowing it to be perturbed by the mechanical deformation, providing a far more accurate result than remeshing after the mechanical simulation. The material properties used for the various mechanical simulations can be seen in Table 1, all of which are for cryogenic temperatures.

The remaining steps before starting the simulation are applying the pressure to be studied and the boundary conditions. The forms of the pressures will be discussed later in this section. The boundary conditions are of critical importance for this type of simulation.

Table 1: Material Properties used for Simulation.

| | Elastic Moduli Pa | Major Poisson's Ratio - | Mass Density kg/m ³ |
|---------------------|----------------------|----------------------------|-----------------------------------|
| “Vacuum-ium” | 1 | 0.38 | 0.01 |
| titanium | 1.23E11 | 0.338 | 4510 |
| niobium | 1.25E11 | 0.301 | 8570 |
| 304 Stainless Steel | 2.00E11 | 0.33 | 7860 |
| niobium titanium | 1.24E11 | 0.32 | 5700 |

First, any faces on symmetry planes must be constrained to stay in that plane, i.e. fixing the variable defining the plane while letting the two other degrees of freedom float. Then the simulation must be fixed in space. If you are using three symmetry planes, then this is automatically done. If two or less planes of symmetry are being used, at least one part of the cavity (a point that is on *all* symmetry planes used) must be fixed. While this will not change the simulation result, it will prevent the solution from diverging if the applied force isn't symmetric. Additionally, other parts of the geometry can be fixed, based on the situation you are trying to simulate. An example of this is the relationship between a HWR and its tuner. While simulating how much force the tuning will require, the beam ports cannot be fixed, but when doing some other simulation that assumes the tuner is in place, fixing the beam port flanges in their normal direction is a good approximation. If more detailed studies are required an effective spring boundary condition can be applied to simulate, for instance, a weakly compliant tuner. After solving for the deformation, the meshing in the material space can be discarded, and the vacuum space material should be reset to true vacuum. The deformation should be kept by updating the coordinates of the mesh nodes. Once the simulation is switched back to modal from structural, the standard electromagnetic simulation procedure should be used to get the new frequency. The difference between the first and second frequency is the frequency shift for the applied pressure.

8.2 Geometry Modification for Simulation

Without careful consideration of the geometry, the number of elements in a mechanical simulation can become needlessly large. Therefore, serious thought must be put into a cavity solid model before simulations can proceed. A mechanically designed model will include many small surfaces/features that are required for fabrication, but will make accurate simulation effectively impossible. It is necessary to make certain geometrical changes to the model before it is simulated to ensure both accuracy as well as efficiency. Some of these geometry features will not exist in the finished cavity, but are required for fabrication. An

example of this would be a small cut in preparation for welding of two subassemblies that will be obliterated by the welding process, leaving a much smoother surface. This type of feature must be modified to accurately represent the final geometry as well as possible. In non-sensitive areas, this type of feature can be safely removed entirely, which will give great benefit in simulation performance with almost no sacrifice in the result's accuracy. The second type of geometry feature that can be changed are small details that are totally unrelated to electromagnetic performance and have little to no impact on mechanical performance. Excellent examples of this type of feature are bolt holes and chamfers on parts of the cavity vessel. These features will exist in the finished cavity, but do not affect the cavity vacuum and do not change the mechanical properties of the cavity assembly. These features are often quite small as well, leading to dense local mesh that is not required. Removing these two types of features can drastically improve simulation efficiency by reducing meshing requirements, an example of this process can be seen in Figure 29. It is, indeed, even common for these types of features to forestall any simulation because of the complexity of meshing them. Many of these features can include overlapping volumes or other small/non-physical features that will need to be fixed before any work can be done. Additionally, even if they can be meshed, the density required may be so high as to exceed the limits of the available simulation hardware. Additionally, simplifications can be made to allow estimates of less critical figures of merit. An excellent example of this is the removal of bellows from a mechanical simulation. Their removal greatly simplifies the model by removing all of the associated small/delicate features while allowing relatively accurate simulations. This is achieved by replacing the bellows with a contact body with an effectively equivalent spring constant, or depending on the level of deformation under study, they can be safely removed all together. This modification must include more careful deliberation of the consequences involved, and the increased error of this modification. Additionally, small gaps between volumes must be removed to allow "gluing" during meshing.

8.3 Accuracy Level

Because the mesh is kept during the mechanical simulation, the simulated frequency shift is extremely accurate. This is crucial because some of the effects being simulated only shift the frequency by tens to hundreds of hertz. As can be seen in Figure 30, these types of shifts can be accurately simulated to and below this level.

8.4 An Aside: Slater's Theorem

Cavity perturbation analysis is an extremely powerful technique for cavity design and simulation. An analytical description of perturbations in resonant frequency due to perturbations in cavity geometry was first developed by Slater [40]. His description is built on classical resonator invariants which are constant under adiabatic changes to the resonator.

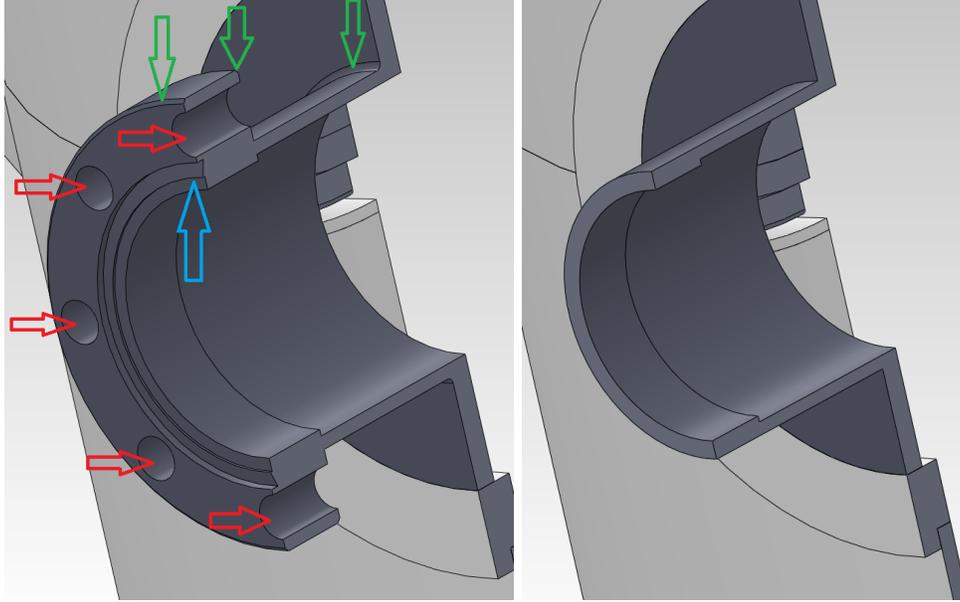


Figure 29: Left: Flange before simplification for simulation. Bolt holes are indicated by red arrows, fillets and chamfers by green arrows, and a complex feature (vacuum seal knife edge) is indicated by a blue arrow. Right: Simplified flange ready for inclusion in a simulation.

This assumption is valid under almost all cavity design considerations where the time scale of the changes to be studied are at least five orders of magnitude longer than the RF period. For the classical harmonic oscillator, the stored energy at any time is given by

$$U = \frac{\omega^2 m x^2}{2} + \frac{p^2}{2m} \quad (160)$$

where p and x are canonical variables. This means that the area of the ellipse described by this equation is an adiabatic invariant. Modifying this equation into the standard form of the ellipse

$$1 = \frac{x^2}{a^2} + \frac{p^2}{b^2} = \frac{x^2}{\frac{2U}{\omega^2 m}} + \frac{p^2}{2mU} \quad (161)$$

we find the area of the ellipse is

$$A = \pi ab = \frac{U}{f} \quad (162)$$

where f is the resonant frequency. Given that A is constant, we find that

$$dU = (A)df = \frac{Udf}{f}. \quad (163)$$

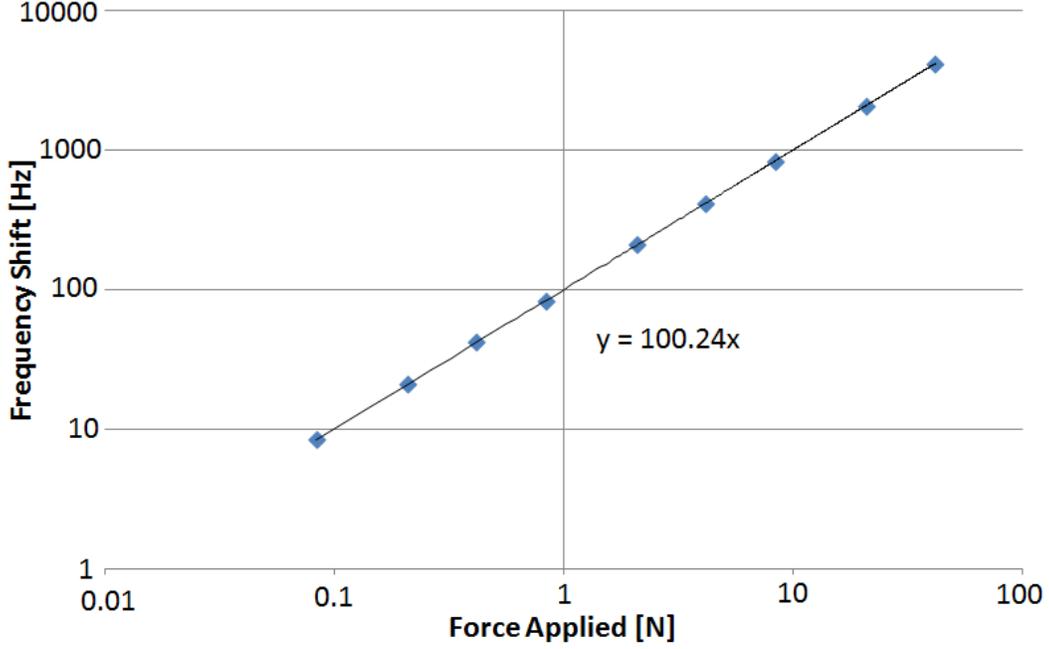


Figure 30: Frequency shift versus applied pressure for a mechanical simulation. Note the linearity down to very small frequency shifts.

Using the known form of dU from Maxwell's Equations and assuming a finite (if small) change in volume (ΔV), we get Slater's Theorem:

$$\frac{\Delta f}{f_0} = \frac{\Delta U}{U} = \frac{1}{4U} \int_{\Delta V} (\epsilon_0 E^2 - \mu_0 H^2) dV. \quad (164)$$

8.5 Pressure Sensitivity

During operation, the cavity will have a pressure differential across its walls. This is because, while the vacuum space is at extremely low pressure, the helium space is at ~ 28 torr (the pressure of 2 K liquid helium). This pressure differential deforms the cavity, shifting the cavity frequency. The deformation can be seen in Figure 31. As expected, the largest deformation occurs in the weakest places, the outer conductor between the RF and beam ports and the inner conductor just above the beam port. Immediately applying Slater's Theorem we can see that deformation in the electric field region will give an opposite contribution to the frequency shift to the shift from deformation in the magnetic field region. The LC circuit model of the cavity gives an intuitive reason for this scaling.

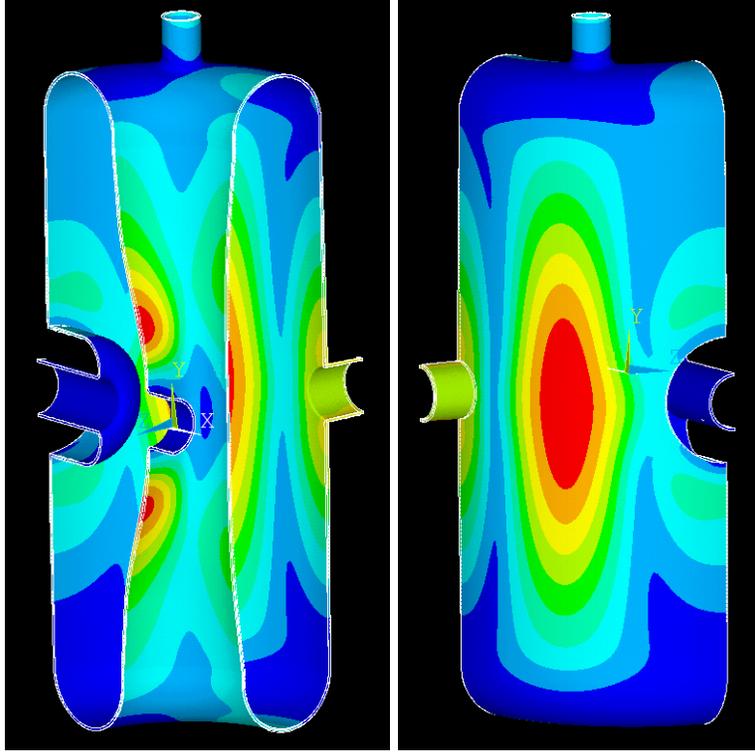


Figure 31: Left: Deformation of the inner conductor due to pressure differential. Right: Deformation of the outer conductor due to a pressure differential (5000 Pa).

In this model, the frequency of the cavity is given by

$$\omega^2 = \frac{1}{LC} \quad (165)$$

where L is the inductance and C is the capacitance. Removing a small volume in the high magnetic field region will decrease the inductance, increasing the resonant frequency. Removing a small volume in the high electric field region will *increase* the capacitance, reducing the cavity frequency.

This behavior means that, if modification is necessary, stiffening can be used to balance these two contributions, resulting in a near-zero frequency shift coefficient. While a fixed frequency shift can be accounted for, the helium pressure can only be maintained to $\sim \pm 1$ torr at a frequency of 1 Hz. This fast varying pressure must not vary the cavity frequency too much, or controlling the cavity in operation becomes quite difficult. This is why the FRIB specifications require that the final coefficient $|df/dP|$ be below 2 Hz/torr. The bare cavity gives a shift of -4.07 Hz/torr. This agrees with what would be predicted by Slater's Theorem because most of the loss in volume ($-dV$) is in the electric field region.

8.6 Lorentz Force Detuning

The fields in the cavity will interact with the induced charges and currents on the RF surfaces. The detuning from these forces is therefore called Lorentz Force Detuning (LFD) [41]. Again using Slater's Theorem, we find that

$$\frac{\Delta f}{f_0} = \frac{1}{4U} \int_{\Delta V} (\epsilon_0 E^2 - \mu_0 H^2) dV = -\frac{1}{U} \int_{\Delta V} (P) dV \quad (166)$$

where P is the LFD pressure. It is worth noting that the term $(P)dV$ is always a positive quantity, a change in sign of the pressure P changes the direction of deformation, changing the sign of dV . This means that any contribution to this frequency shift will be negative. This means that the compensation techniques that could be used for correcting the pressure sensitivity do not apply here, so we must rely on straightforward stiffening of the cavity. The LFD frequency shift scales with the stored energy in the cavity, and therefore the field level squared. Thus, the LFD coefficient, K_L , is defined as

$$K_L = \frac{\Delta f}{(\Delta E_{acc})^2}; E_{acc} = \frac{V_{acc, \beta_{opt}}}{\beta_{opt} \lambda}. \quad (167)$$

This type of detuning is well modeled by the Duffing Equation, discussed earlier along with the basic controls implications of this effect. Because the FRIB HWRs are to be run with a duty factor of 1, this effect is not the dominant concern it can be for pulsed machines [42, 43], so the use of advanced compensation methods and control systems is not anticipated for FRIB. FRIB specifications do, however, call for $|K_L|$ to be below 3 Hz/(MV/m)² for the HWRs. The bare cavity, as designed earlier, gives an initial $K_L = -3.06$ Hz/(MV/m)². The deformation resulting from this force can be seen in Figure 32.

8.7 Cavity Tuning

For the $\beta = 0.29$ and 0.53 HWRs, it was decided to tune the cavity in operation by deforming the beam port cups by application of pressure on the beam ports themselves. The design of this tuner requires an estimate of the physical range and force required to give the FRIB specified tuning range of 150 kHz. For tuning measurements, all distances and forces given are for movement of *one* port while the frequency shift is for *both* ports moves symmetrically. Simulation of this effect is done by applying force to the beam port flange, and gives coefficients of $df/dF = -126$ kHz/kN and $df/dx = 599$ kHz/mm for full tuning range parameters of 150 kHz, 0.25 mm, and 1.185 kN. While this does provide a potentially large tuning range, the resolution of the tuner is also of great importance. For the controls, the desired frequency resolution of the tuner is 1 Hz, meaning that the tuner must have a resolution of ~ 2 nm or ~ 8 mN. It should be noted that this frequency sensitivity could have been estimated from the earlier beam port cup sensitivity study. This predicted a frequency sensitivity of 820 kHz/mm, and it is expected that a deformed,

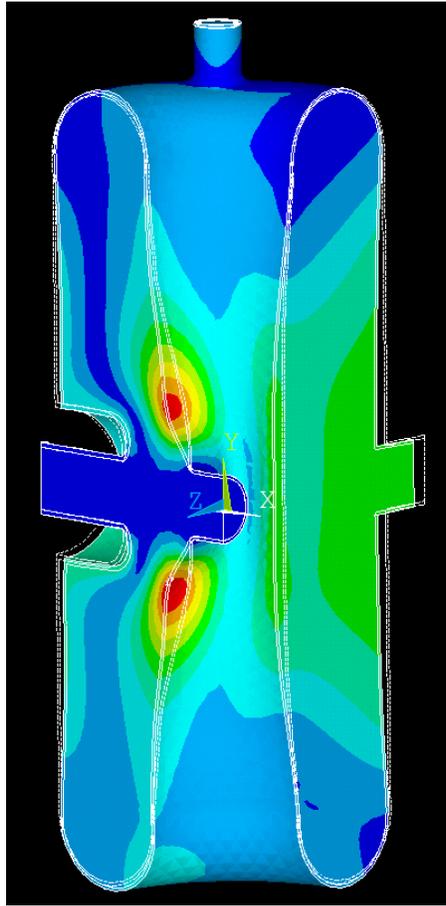


Figure 32: Deformation of the cavity with the Lorentz Force applied at full accelerating voltage.

as opposed to fully moved, beam port cup would be slightly less sensitive. The deformation from the applied tuning pressure can be seen in Figure 33.

9 Cavity Testing Theory and Setup

A critical part of cavity design is feedback from prototyping and testing of realistic cavities. This is because cavity performance is a combination of RF design, material selection and handling, fabrication technique, cavity processing, and other, more subtle effects. There are a many steps in the fabrication process that must be taken into account in the design process that can only be quantified on a real cavity, such as weld shrinkage, coupling strengths, superconducting surface resistance (and therefore quality of cavity processing),

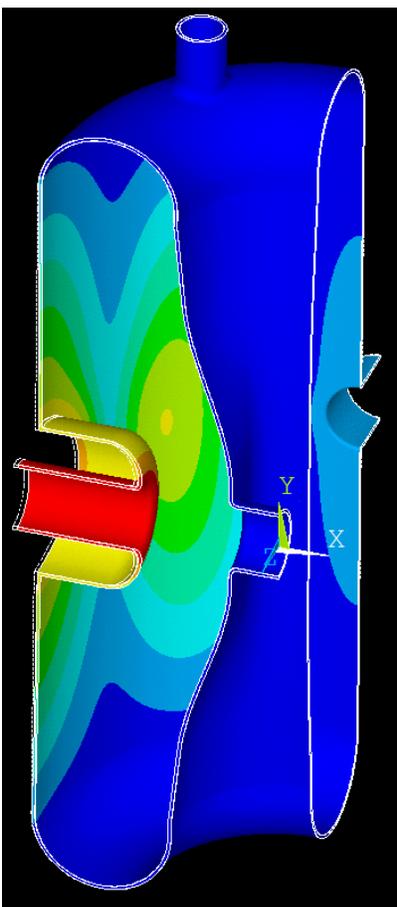


Figure 33: Simulation of cavity deformation for an applied tuning force.

and many others. Cavity testing measures the cumulative effect of all of these factors. In this section, the theory and setup of a cavity test will be discussed. Additionally the data measured during a specific HWR test will be examined along with the details of translating this data into relevant cavity parameters.

9.1 Testing Theory

The basic theory of a cavity test centers around measuring the quality factor of the cavity at a variety of stored energies. While the cavity is designed to optimize many figures of merit, most of them cannot be directly measured during a cavity test. These include peak surface fields, R/Q , β_{opt} , and the Geometry Factor, which must be inferred using simulation data. A cavity test is performed with two power couplers, one close to matched to minimize the required forward power (“input” coupler) and one very weakly coupled

(“pickup” coupler). Power is coupled to the cavity with the input coupler, while the pickup coupler is used strictly as a diagnostic tool. Cavity tests are done with a phase lock loop matching the driving frequency to the cavity resonance. The basic testing phase lock loop arrangement can be seen in Figures 37 and 38.

This RF circuit allows a signal generator to track the cavity resonant frequency by mixing the transmitted power signal with the forward power signal, giving a DC signal whose amplitude is based on the difference between the driving frequency and the cavity resonant frequency. With the cavity driven in steady-state on resonance, only the cavity frequency (f), forward power (P_f), reverse power (P_r , vector sum of the reflected and emitted traveling waves from the input coupler), and the transmitted power (P_t , from the pickup coupler) can be measured. Using equations from earlier

$$|S_{11}|^2 = \frac{P_r}{P_f} \quad (168)$$

$$|S_{21}|^2 = \frac{P_t}{P_f} \quad (169)$$

$$\beta_2 = \frac{|S_{21}|^2}{1 - |S_{11}|^2 - |S_{21}|^2} = \frac{Q_0}{Q_{ext,2}} = \frac{P_t}{P_d} \quad (170)$$

$$\beta_1 = \frac{1 \pm |S_{11}|}{1 \mp |S_{11}|} = \frac{Q_0}{Q_{ext,1}} = \frac{P_r}{P_d} \quad (171)$$

and

$$Q_0 = Q_L(1 + \beta_1 + \beta_2) \quad (172)$$

we see that we don’t have enough information to calculate Q_0 and U . To do this, we must determine the Q_L and the coupling (whether β_1 is greater or less than 1). Both of these can be obtained from modulating the forward power. The most straightforward method is to suddenly shut off the forward power, so that the equation for the stored energy in the cavity becomes

$$\frac{dU}{dt} = -P_L = -\frac{\omega U}{Q_L} \quad (173)$$

which gives

$$U(t) = U_0 \exp\left(-\frac{\omega t}{Q_L}\right). \quad (174)$$

For a sufficiently large Q_L (true for most Dewar tests), a decay time τ_L can be measured for a given U/U_0 (usually about a -6 dB fall), giving Q_L . Additionally, the transitory response of P_r determines whether the cavity is over or under coupled. This effect can be seen in Figure 34. In this way, we have enough information to calculate not only β_1 and β_2 , but also Q_0 . Given this, we can calculate $Q_{ext,2}$, allowing us to transform our measurement of P_t into a measurement of U . While it is often assumed that $Q_{ext,1}$ and

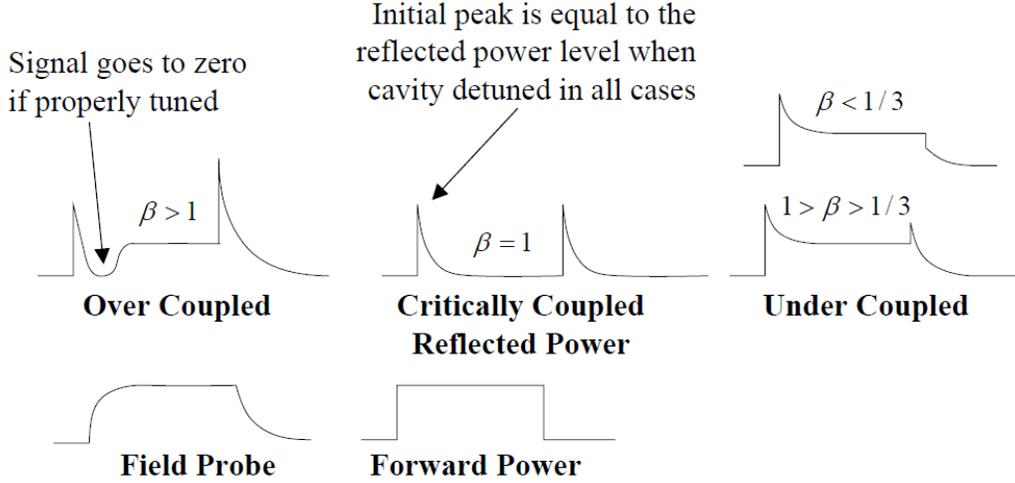


Figure 34: Transient behavior of the reverse power for different coupling strengths. All curves are in response to the shown forward power behavior. Image credit to Tom Powers, TJNAF.

$Q_{ext,2}$ do not change dramatically over the course of a Dewar test, it is recommended to periodically repeat the modulated measurements to ensure these values have not drifted during testing. With knowledge of $Q_{ext,2}$, every set of power measurements can be used to directly calculate Q_0 and U :

$$P_d = P_f - P_r - P_t; \quad U = \frac{Q_{ext,2}}{\omega} P_t \quad (175)$$

where the first equation is the cavity energy balance, and

$$Q_0 = \frac{\omega U}{P_d}. \quad (176)$$

This method, known as the “direct” method, has the disadvantage of using P_r extensively. Especially far from unity coupling, $P_f - P_r$ is likely to be quite small, and error prone. Fortunately, there is another method for calculating Q_0 and U . While the first method does not require the use of $Q_{ext,1}$, this method uses this previously measured number in place of P_r . This method is known as the “indirect” method, because, while measuring $Q_{ext,1}$ involves P_r , this usage makes it less sensitive to the uncertainty of its measurement. Plugging the definition of $Q_{ext,1}$ into our equations, and making the usual assumption that $\beta_2 \ll 1$, we find that

$$Q_0 \approx \frac{Q_{ext,1}}{2\sqrt{\frac{Q_{ext,1}}{Q_{ext,2}} \cdot \frac{P_f}{P_t} - 1}}. \quad (177)$$

It should be noted that once $Q_{ext,2}$ and $Q_{ext,1}$ are calculated with modulated measurements, they are assumed to be constant for following measurements. This, however, makes each CW measurement over-specified, and this can be used to check the measurement's accuracy. In principle, both the direct and indirect methods should give the same answer. Starting from

$$\frac{Q_{ext,2}}{Q_{ext,1}} = \frac{\beta_1}{\beta_2}. \quad (178)$$

it is straight-forward to derive (assuming, as usual, that P_t is small compared to P_f)

$$\sqrt{\frac{Q_{ext,2}}{Q_{ext,1}}} |S_{21}| = 1 \pm |S_{11}|. \quad (179)$$

Plotting the left side of this equation versus the right side of the equation using the data ($|S_{21}|$ and $|S_{11}|$) and the assumed values for $Q_{ext,2}$ and $Q_{ext,1}$ gives an easy graphical method of checking the consistency of the measured values. Depending on the coupling, $|S_{11}|$ will vary from 1 to -1 , so the plotted data should, ideally, lay on a simple triangle. Thus, the plot of these quantities is an excellent way to test whether the values of $Q_{ext,1}$, $Q_{ext,2}$, P_r , P_f , and P_t are consistent. Any systematic deviation between the plotted data and this ‘‘Duality Triangle’’ indicates a problem with the data. This generally happens when the assumption of constant coupling is invalid, which is the case for the test presented in Section 4.4, a dramatic shift in coupling between test day 1 and test day 2 had to be accounted for.

9.2 Measurement Procedure

For a given cavity configuration, there are generally two phases to measuring the cavity's parameters. The first consists of the modulated measurements required to calibrate the CW measurement, and this is generally done at low field levels to avoid dynamic behavior that can occur at higher fields. After the field level is set, generally around $E_{pk} \sim 1$ MV/m, the phase of the transmitted power signal going to the mixer is adjusted to maximize P_t . This is important because, while there is a phase shift from driving the cavity off resonance, there is also a phase shift from the cable and devices used to do the measurement. Having a mechanism for shifting this phase (physical or electronic) is crucial not only for compensating the overall phase difference between P_t and the signal sampled from P_f (modulo 2π), but also for fine-tuning before every measurement. This is because the phase of the circuit can (and will) drift during the course of the cavity test because of thermal or mechanical changes in the components and transmission lines. For a perfectly adjusted circuit, the cavity is driven directly on resonance, maximizing U and therefore P_t for a given P_f . Having properly adjusted the phase, the forward power is shut off, and the dissipated power is plotted versus time, generally on a spectrum analyser. The decay time is recorded for a given power difference (generally -6 dB), and the initial P_t is recorded.

It is important to accurately correlate this measurement with a set of CW measurements taken immediately after the modulated measurement. Combining these modulated and CW measurements allows the calculation of $Q_{ext,1}$ and $Q_{ext,2}$ as detailed in the previous subsection, and these are used to interpret the following CW measurements.

The second phase of measurement is to perform CW measurements at different P_f . The forward power is increased in steps of 1 – 3 dB, and several CW data points are recorded at every step, after adjusting the phase shift. The higher the cavity fields, the smaller the steps in P_f should be, to get the best resolution in this region. This is because the most dynamic behavior happens here, and sudden shifts in Q_0 can occur. This is generally continued until limited either by cavity quench, maximum available P_f , or high radiation levels from field emission. It should be noted that many values are recorded with each data point, not just power levels. This includes data such as date and time, readings from temperature and pressure sensors, liquid helium level, and more. Any interesting phenomena are then investigated, time permitting, such as the detailed behavior of sudden drops in Q_0 , field emission onset levels, quench field level, etc. If desired, time can be spent conditioning the cavity at high fields in an attempt to improve performance. This is most generally done when there is high radiation levels in an attempt to destroy or otherwise shut off the serious field emission sites causing the radiation. In cases of high radiation, staying at high power or pulsing to even higher power can be successful in improving performance, allowing access to higher field levels. Once all of the desired data are collected for a given cavity configuration, then the configuration is changed and the process is repeated. The nature of the cavity configurations depends strongly on the exact phenomenon being studied, but a common example of this would be measurement at different temperatures as the cavity is pumped down to 2 K.

9.3 Calibration of the RF Network

It is important to remember that the RF circuit has losses at almost every point. This means that the powers measured by the power meters (Cal Points E, F, and G, seen in Figure 35) have additional losses in them from the RF circuit compared to the values we need to calculate the cavity parameters. These power values would, ideally, be recorded at the cavity (Cal Points B and C). Calibrating the RF circuit is the process of measuring the effective losses between these locations. Calibration measurements are all done using a Vector Network Analyzer (VNA), carefully calibrating the measurement equipment at each step. All measurements are done at four frequencies: the design frequency, the design frequency ± 0.5 MHz, and the cavity frequency as measured during the cavity test. It should also be noted that all losses should be measured in decibels, this allows direct addition of losses in series. This allows interpolation of the calibration for the test data as required. This calibration takes place in two phases for logistical reasons. While all parts of the RF circuit outside the Dewar are room temperature and can be considered relatively stable if left in place, the transmission lines and connections inside the Dewar

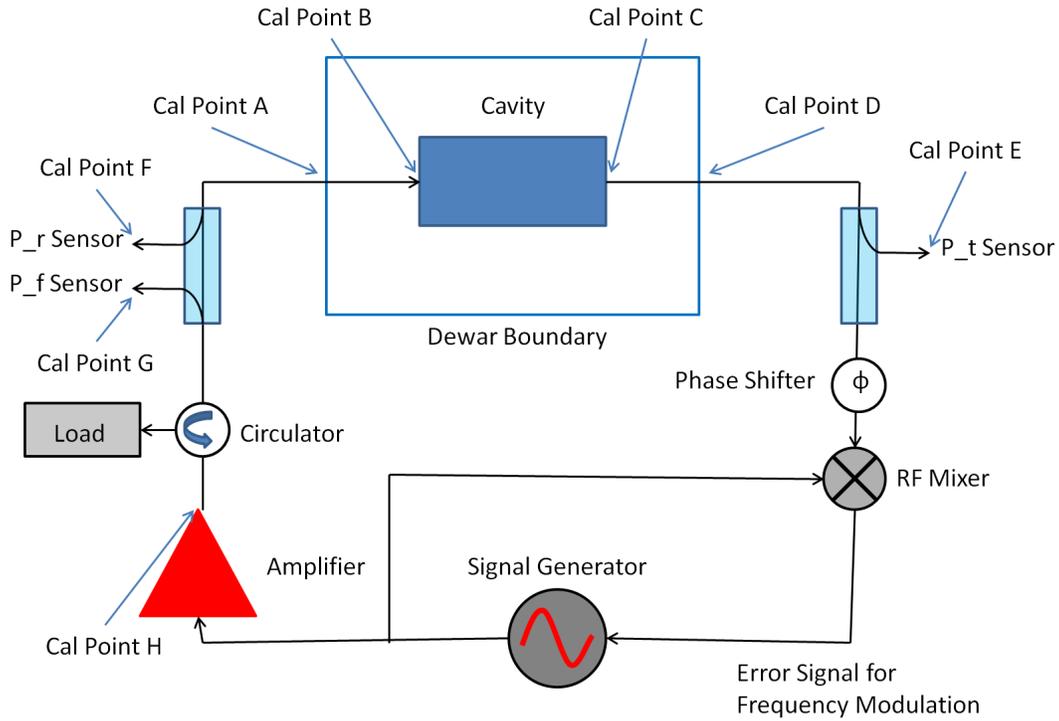


Figure 35: Basic RF circuit for a phase lock loop for cavity testing with the calibration points marked.

must be calibrated while the cavity is cold because of the significant thermal shifts. This is generally measured directly after the completion of a cavity test, while the circuit is as thermally stable as possible. This is done by connecting Cal Point A to Port 1 of a VNA through a calibrated cable, then measuring the round-trip losses (S_{11}), repeating for Cal Point D. These values are called the “Cold Calibration”. The rest of the process is called the “Warm Calibration” because it involves measurement of the rest of the RF circuit, which is at room temperature. The first step for this is calibration of the P_t and P_r power sensors. Because we have already measured the losses inside the Dewar, all that remains is measuring the losses from Cal Point D to Cal Point E for P_t and from Cal Point A to Cal Point F for P_r . These losses will include the cable losses as well as the attenuation from the directional couplers used to sample these signals. This measurement is done by connecting the Dewar end to Port 1 of the VNA and the power sensor end to Port 2 of the VNA, and measuring the losses (S_{21}). The final calibration for these power sensors is the warm losses plus *half* of the appropriate round-trip cold losses. The most complex calibration is for P_f . This is because what we desire is the forward power at Cal Point B (again, what the cavity actually sees) as measured at Cal Point G for a given power at Cal Point H (right

out of the amplifier). The first step accounts for the losses from the amplifier to the cavity by connecting Cal Point H to Port 1 of the VNA and Cal Port A to Port 2 of the VNA, then *calibrating out* the through losses. Port 2 of the VNA is then moved *without changing the calibration* to Cal Point G while terminating Cal Point A with a $50\ \Omega$ load, where the power loss (S21) is measured. The power measured, though, is the forward power at Cal Point A seen at Cal Point G. This means that, for the final calibration of P_f , the warm calibration number must have half the appropriate cold calibration *subtracted* to properly reflect the forward power at the cavity.

10 Challenges for Cavities in Testing and Operation

There are many practical concerns to take into account when designing a cavity besides the standard figures of merit. Based on the geometry, it is possible (almost impossible to suppress entirely, in practice) to have stable resonant trajectories for stray electrons in the cavity (themselves almost impossible to suppress). This behavior can severely complicate machine operation and even limit achievable field and quality factor. Beyond this, processing and fabrication are imperfect processes, and can introduce their own operational complications. This section will discuss these more common challenges that are often encountered while testing or operating a cavity and their mitigation.

10.1 Multipacting

The name multipacting is a compound word used to shorten “multiple impacting”, and this name is a good description of this phenomenon. Stray electrons in the cavity vacuum space interact with the cavity fields and strike the cavity walls. The impact can generate more electrons emitted from the surface, which will also have some trajectory based on their energy, angle of emission, and RF phase. The number of electrons emitted is quantified by the Secondary Emission Yield (SEY), which depends on many factors including electron energy, surface preparation, and impact angle. The electron’s trajectory also extracts energy from the cavity, depositing it into the cavity walls. Based on the surface it strikes, the energy of impact, and the angle of impact, the average number of ejected electrons varies, and can be over 1.4 [46], even in a well processed cavity. The energy this cascade extracts is trivial, in principle, until it is sustained for many generations. If the trajectories of the electrons becomes closed, i.e. return to near the original emission point (and RF phase), the cascade continues, eventually consuming all additional RF power injected into the cavity.

10.2 Avoiding Multipacting Before Testing/Operation

There are multipacting considerations both in design and operation. In the design, care must be taken not to introduce the conditions for strong multipacting, mostly characterized

by regions of parallel surfaces, especially items like couplers, plungers, and stub tuners. The strongest multipacting occurs when there is a wide range of trajectories in location, energy, and RF phase that satisfy the resonant condition. These barriers are called “hard” barriers because they are extremely difficult to get beyond in testing/operation. Particle tracking simulations [46, 47] for predicting multipacting have been increasing in accuracy in recent years, and have become reliable for predicting and avoiding strong multipacting barriers in cavity designs. However, because of the perfection of the models used, these simulations often find far more barriers than will present themselves in practice. Fabrication errors are often beneficial, in this case. The random errors added to the surface often break the resonant condition for the weaker multipacting bands, although this is not a strong enough benefit to loosen manufacturing tolerances. Additionally, careful surface preparation during cavity fabrication and processing can reduce the SEY, automatically narrowing multipacting bands. For HWRs of this geometrical type, multipacting studies have been done [46] and found no significant barriers (besides the rinse port plungers, which were later removed from the design). This has been born out in testing where no significant multipacting has been encountered.

10.3 Dealing with Multipacting During Testing/Operation

Once encountered in operation, there are two common strategies to overcome this limitation. The first involves rapidly adjusting the RF phase, frequency, or forward power to try to “jump” past the barrier. Because multipacting is a resonant behavior, perturbations to the conditions in the cavity can disturb the electron cascade long enough for the cavity to be brought to a higher power, ending the resonant condition. This is easier in operation than in testing because the much higher coupling available means that the cavity field can be ramped much more rapidly. It is uncommon to have significant multipacting barriers at high fields, so this method will often allow testing of the cavity high field behavior. Care must be taken to not reduce the cavity fields below the barrier, or risk becoming trapped in the barrier again. The second strategy for multipacting is reducing the local SEY. This is achieved by allowing the multipacting barrier to condition itself. The electron bombardment will cause gasses to desorb from the surface, generally decreasing the strength of the multipacting band. While this process has obvious limits, strong barriers encountered at the beginning of testing can be significantly weakened to allow testing to proceed. Any venting of the cavity or further processing includes the danger of reintroducing these barriers.

10.4 Field Emission

Field emission is caused by a stray electron in the cavity gaining energy from the cavity fields and impacting the cavity wall with enough energy to generate an X-Ray instead of more electrons. Because this is not a resonant behavior, it requires a different source

of electrons, usually particulate (“field emitters”) on the cavity surface not removed by processing. These field emitters locally enhance the electric field, causing a stream of electrons to flow into the cavity at high fields. Because of this behavior, field emission is a problem at high fields, unlike multipacting. At low fields, the fields on the emitters aren’t enough to generate large numbers of electrons, and they don’t gain enough energy from their trajectory to generate significant numbers of X-Rays. As the fields in the cavity increase, both the number of electrons and the energy they gain increase until cavity performance begins to suffer. The energy extracted by the electrons during their transit can significantly reduce the cavity quality factor, and can limit achievable fields by requiring more RF power, both from extracted energy and increased coupler mismatch to the cavity. If the field emission is strong enough, the energy deposited by the electrons on the cavity wall can be enough to cause a local thermal quench. Once the superconductivity is broken locally, the cavity stored energy is rapidly dissipated there, turning the field emission off. Field in the cavity can be restored once the wall has cooled and become superconducting again. There are several common ways of mitigating this problem. During processing, high-pressure rinsing with ultra-pure water forceably removes larger particulate from the surface, significantly reducing the sources for electrons at high fields, while clean room preparation and assembly reduces reintroduction of these particulate after cleaning. During operation, the field emitters can be used to process themselves. By operating close to the limit of operation (limited by available power or radiation limits), often pulsing to higher fields as well, enough power can sometimes be introduced to strong emitters to cause their destruction. This process works best for small numbers of strong emitters, and has obvious limits. While strong emitters can be suppressed by careful fabrication, processing, and testing, stray electrons can never be eliminated from the cavity. With a non-zero population of electrons in the cavity from things such as residual gas, cosmic rays, and weak field emitters, it is common to see minor field emission at very high fields. Care must be taken during design of testing facilities to include adequate radiation shielding to allow testing of even the worst cavities.

10.5 Q-Disease

Hydrogen is introduced into the cavity material at many points along the cavity processing. The strongest source is etching, especially if the temperature of the acid blend used isn’t well controlled. The hydrogen itself is not thought to degrade cavity performance. However, between 40 K and 120 K [48], niobium reacts with hydrogen to form several kinds of hydrides on the surface of the cavity. These hydrides are dielectric, and extremely lossy. In operation, the quality factor of the cavity is severely reduced, and degrades further as the field is increased. The formation process of these hydrides is relatively slow; if the dangerous temperature region is passed through in under an hour [48], little to no degradation is seen. This results in interesting requirements for the design of cryomodule cryogenic systems, a balance between efficient use of cryogenics and rapid cooling of the

cavities through the dangerous temperature region. Additionally, a degassing bake can be used to drive off the hydrogen from the material, significantly reducing the danger of Q-Disease. This bake is generally done at 600°C for roughly 10 hours in high vacuum. Once observed in testing or operation, there is no known remedy for Q-Disease besides a warming to 200 K and more rapid cooling cycle.

11 Electromagnetic and Mechanical Simulations

Three dimensional computational simulation provides a powerful tool during the design of accelerating resonators. Properly done, both electromagnetic and mechanical simulations can be performed to great accuracy, allowing exploration of different designs and their implications before prototyping. Most problems can even be solved on a modern, powerful desktop computer, although more powerful codes and computer infrastructure exists when needed. However, the interconnection of simulation and prototyping in the design of a resonator cannot be overlooked, nor overstated. As the interface between powerful three dimensional geometry creation programs and flexible simulation software is improved, the design of these complex structures will become more rapid and accurate. The strength of simulations lies in the rapid design optimization, the cost-effective analysis of many options. This relies, however, on the relative simplicity and exactness of the computerized model. Many real life considerations cannot be simulated, and prototyping must inform these decisions. This interaction becomes even more important during the design of complex geometries like HWRs. This is because the analytical formulas and simple two-dimensional simulations traditionally used for cavity design are not powerful enough to guide this type of design. The new generation of powerful simulation software must be matched with the human experience of prototyping and testing these new, complex geometries for future projects that will rely on them to succeed.

12 On Simulation Scripting and Optimization

Cavity design requires a solid understanding of both the electromagnetic behavior of a design and the mechanical complexities of fabrication. This knowledge can be used to design the cavity geometry with the proper variables and features required to efficiently optimize the design while allowing practical and reliable fabrication. While prototyping almost invariably uncover unexpected factors and problems that require further cavity design iterations, proper parameterization greatly speeds cavity optimization. The optimization time itself can be greatly reduced with simulation scripting and automation. The scripts seen in the Appendices required significant time and effort to create, but once completed they greatly facilitated rapid simulation and optimization. This scripting is possible because ANSYS-APDL is designed to make geometric references consistent and allows very flexible geometric reference operations. An excellent example of this flexibility can be seen

in this section from Appendix A:

```
!Calculates Volume parameters
vsum
!Select volume 1
vsel,s,,1
!Define cavity vacuum as "cavity"
cm,cavity,volu
!Select "cavity"
cmsel,s,cavity
!Select areas associated with cavity vacuum
aslv,s
!Unselect symmetry planes
asel,u,loc,x,0
!asel,u,loc,y,0
asel,u,loc,z,0
!asel,u,,4
!Define PEC walls
cm,cavwall,area
aplot
```

This section of code defines two objects for use later in the script: the cavity vacuum volume *cavity* and the cavity RF surfaces *cavwall*. These references are used extensively through the rest of the scripts to set boundary conditions and for post-processing. It is important to note the extreme flexibility of the code used to do this. For essentially any cavity type, this code will work properly if the correct volume number is used in line 4 and the appropriate symmetry planes are selected in lines 12-14, an almost trivial amount of work. This is just one example of this scripting philosophy that can be seen in the Appendices, and it allows rapid generation of the appropriate scripts for a wide variety of cavity designs, greatly speeding simulation. This means that the time required for parameter scans used to optimize cavity performance can be reduced to the time it takes for the computation. Proper scripting of the meshing can also reliably reduce this time as well. Some code systems like ACE3P are being designed to be fully scriptable, from geometry creation to post-processing, using high-power computing clusters to speed simulation time. This suite of software shows great promise for allowing rapid simulation and optimization of even the most complex geometries. Simulation software that allows scripting and optimization like this is understandably preferable for serious optimization projects.

References

- [1] G. Stupakov, "Lecture notes on Classical Mechanics and Electromagnetism in Accelerator Physics", United States Particle Accelerator School, Revision 20, 2011

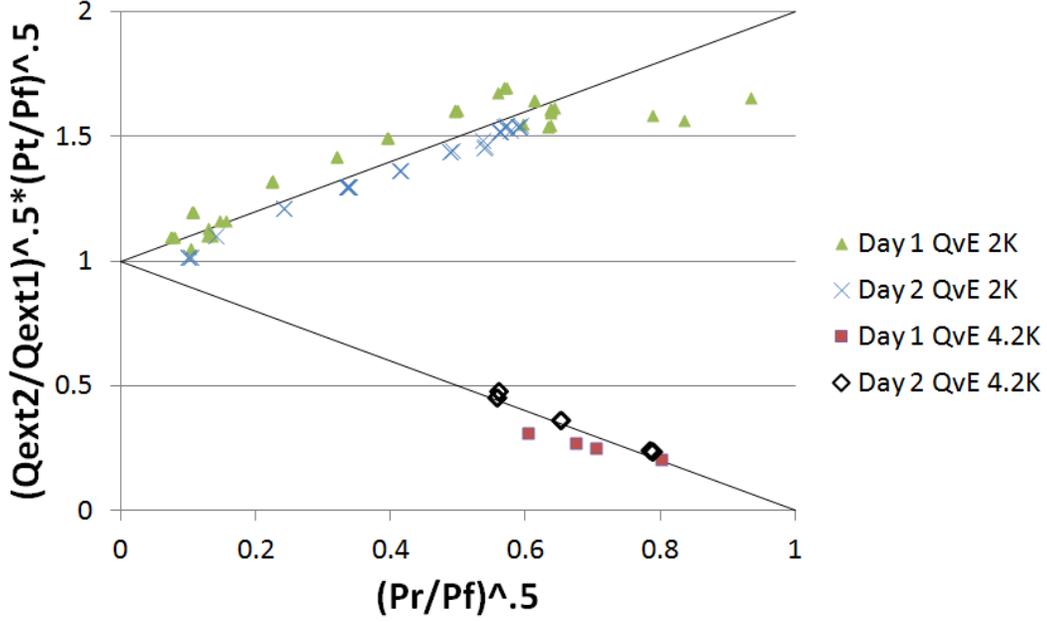


Figure 36: Both days of testing data plotted on the Duality Triangle to verify self consistency.

- [2] B. Van der Pol, J. Van der Mark, “Frequency demultiplication”, Nature, 120, 363-364, (1927)
- [3] J.D. Jackson (1999), “Classical Electrodynamics, 3rd Edition”, John Wiley and Sons Inc.
- [4] H. Padamsee, J. Knobloch, T. Hayes (2008), “RF Superconductivity for Accelerators, 2nd Edition”, John Wiley and Sons Inc.
- [5] J.D. Jackson (1999), “Classical Electrodynamics, 3rd Edition”, John Wiley and Sons Inc., p. 298
- [6] A. Facco, V. Zvyagintsev, “Beam Steering in Superconducting Quarter-Wave Resonators: An Analytical Approach”, Phys. Rev. ST Accel. Beams 14, 070101 (2011)
- [7] E. Zaplatin et al., “Low-Beta Superconducting RF Three-Quarter Wavelength Cavities”, ESS 02-125-L, April 2002, Juelich
- [8] M. Tinkham, “Introduction to Superconductivity, 2nd Edition”, Dover Publications, Inc., p. 44

- [9] M. Tinkham, “Introduction to Superconductivity, 2nd Edition”, Dover Publications, Inc., Chapter 3
- [10] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957)
- [11] H. Padamsee, J. Knobloch, T. Hayes (2008), “RF Superconductivity for Accelerators, 2nd Edition”, John Wiley and Sons Inc., p. 88
- [12] W. Hartung, “Introduction to Superconducting Cavities: Lecture Notes for USPAS Course on Beam Measurements and Experiments at NSCL”, Lecture notes, US Particle Accelerator School, Michigan State University, June 2007
- [13] A. Facco, “SRF Low-beta Accelerating Cavities for FRIB”, Lecture notes, Phys 905 - The Accelerator Physics of FRIB, Michigan State University, October 2011
- [14] F. and H. London, Proc. Roy. Soc. (London) A149, 71 (1935)
- [15] M. Tinkham, “Introduction to Superconductivity, 2nd Edition”, Dover Publications, Inc., Chapter 1
- [16] W. Meissner and R. Ochsenfeld, Naturwissenschaften 21, 787 (1933)
- [17] The Argonne Accelerator History Document Collection, www.aai.anl.gov/history/project_pages/atlas.html
- [18] ATLAS Home Page, www.phy.anl.gov/atlas/
- [19] Laboratori Nazional de Legnaro Home Page, www.lnl.infn.it
- [20] TRIUMF National Laboratory Accelerator Physics, www.triumf.ca/research/research-topics/accelerator-physics
- [21] SPIRAL-2 at GANIL homepage, <http://www.ganil-spiral2.eu/spiral2>
- [22] Stopped and reaccelerated beams at NSCL, www.nscl.msu.edu/exp/sr
- [23] Soreq Applied Research Accelerator Facility, http://www.soreq.gov.il/default_EN.asp
- [24] C. Piel et. al., “Phase 1 Commissioning Status of the 40 MeV Proton/Deuteron Accelerator SARAF”, Proceedings of EPAC08, Genoa, Italy, THPP038
- [25] H. Padamsee, J. Knobloch, T. Hayes (2008), “RF Superconductivity for Accelerators, 2nd Edition”, John Wiley and Sons Inc., p. 48
- [26] N. Ida, J.P.A. Bastos, “Electromagnetics and calculation of fields”, Second Edition, Springer Verlag, N.Y., 1996

- [27] S. Humphries, “Finite-Element Methods for Electromagnetics”, www.fieldp.com/femethods.html
- [28] SolidWorks Corp., Dassault Systems, www.solidworks.com/default.htm
- [29] Developed by Spatial Corporation, www.spatial.com
- [30] ANSYS Inc., www.ansys.com
- [31] CST Computer Simulation Technology AG, www.cst.com
- [32] AWR Corp., www.awrcorp.com
- [33] ANSYS HFSS, www.ansoft.com/products/hf/hfss
- [34] Advanced Computations Department, SLAC National Accelerator Laboratory, https://slacportal.slac.stanford.edu/sites/ard_public/acd/Pages/Default.aspx
- [35] W. Hartung, “LEP 1500 Plans Collapse and Other Stories or What I Learned on my CERN Vacation,” Report NS/RF-92-1701, National Superconductor, 160000 Lincoln Avenue, Brentwood, Colorado 80523 (1992).
- [36] W. Hartung and E. Haebel, “In Search of Trapped Modes in the Single-Cell Cavity Prototype for CESR-B,” Proceedings of the 1993 Particle Accelerator Conference, Washington, D.C., May 1993, Vol. 2, pp. 898-900.
- [37] V. Zvyagintsev, et. al., “SCRF Development at TRIUMF”, RUPAC 2010, Protvino, Russia
- [38] H. Padamsee (2009), “RF Superconductivity: Volume II: Science, Technology, and Applications”, John Wiley and Sons Inc.
- [39] SRF2011, conferences.fnal.gov/srf2011, 15th International Conference of RF Superconductivity, July 25-29, 2011, Chicago IL, USA
- [40] J.C. Slater, “Microwave Electronics”, D. Van Nostrand Company, New York, 1950, pp. 80-81.
- [41] “Handbook of Accelerator Physics and Engineering”, edited by A. Chao and M. Tigner, 3rd Printing (2006), World Scientific Publishing Co. Pte. Ltd., p. 580
- [42] M. Grecki et. al., “Compensation of Lorentz Force Detuning for SC Linacs (with Piezo Tuners)”, Proceedings of EPAC08, Genoa, Italy, MOPP129
- [43] R. Carcagno et. al., “First Fermilab Results of SRF Cavity Lorentz Force Detuning Compensation Using a Piezo Tuner”, Proceedings of SRF2007, Peking University, Beijing, China, TUP57

- [44] H. Padamsee, J. Knobloch, T. Hayes (2008), “RF Superconductivity for Accelerators, 2nd Edition”, John Wiley and Sons Inc., p. 120
- [45] T. Powers, “Theory and Practice of Cavity RF Test Systems”, Technical Document, Thomas Jefferson National Accelerator Facility
- [46] L. Ge et. al., “Multipacting Simulation and Analysis for the FRIB Superconducting Resonators Using TRACK-3P”, Proceeding of LINAC10, Tsukuba, Japan, September 12 - 17, 2010, THP092
- [47] R.L. Geng, “Multipacting Simulations for Superconducting Cavities and RF Coupler Waveguides”, Proceedings of the 2003 Particle Accelerator Conference, Portland, OR, May 12-16, 2003
- [48] K. Saito, “Cavity Preparation”, Lecture notes from ILC School 2009
- [49] J. Popielarski, FRIB, Private Communications, November, 2011

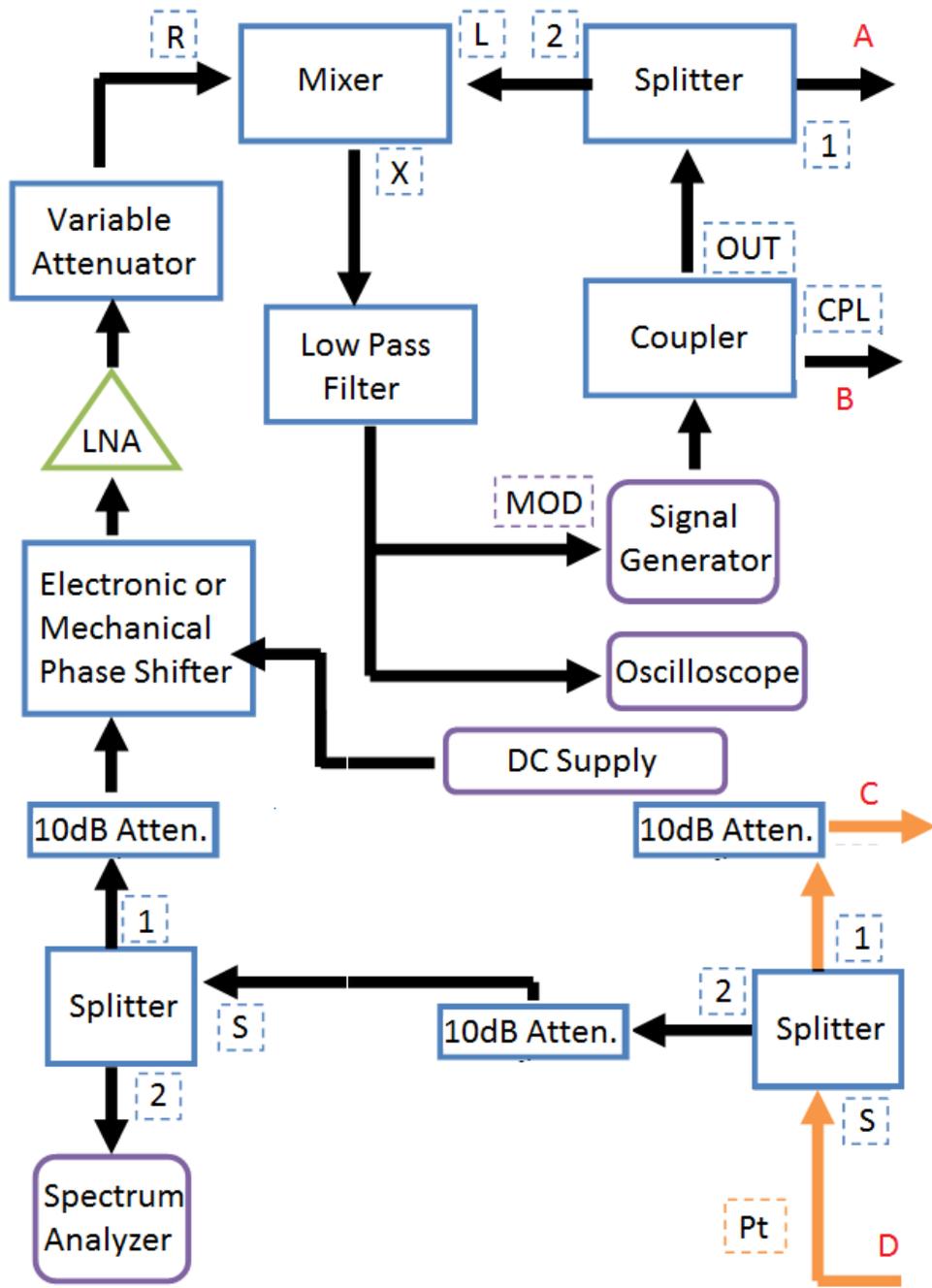


Figure 37: One half of the RF circuit for a phase lock loop for cavity testing (circulator, isolators, and RF diodes excluded). Letters in Red connect to the other half of the RF circuit, seen on the next page.

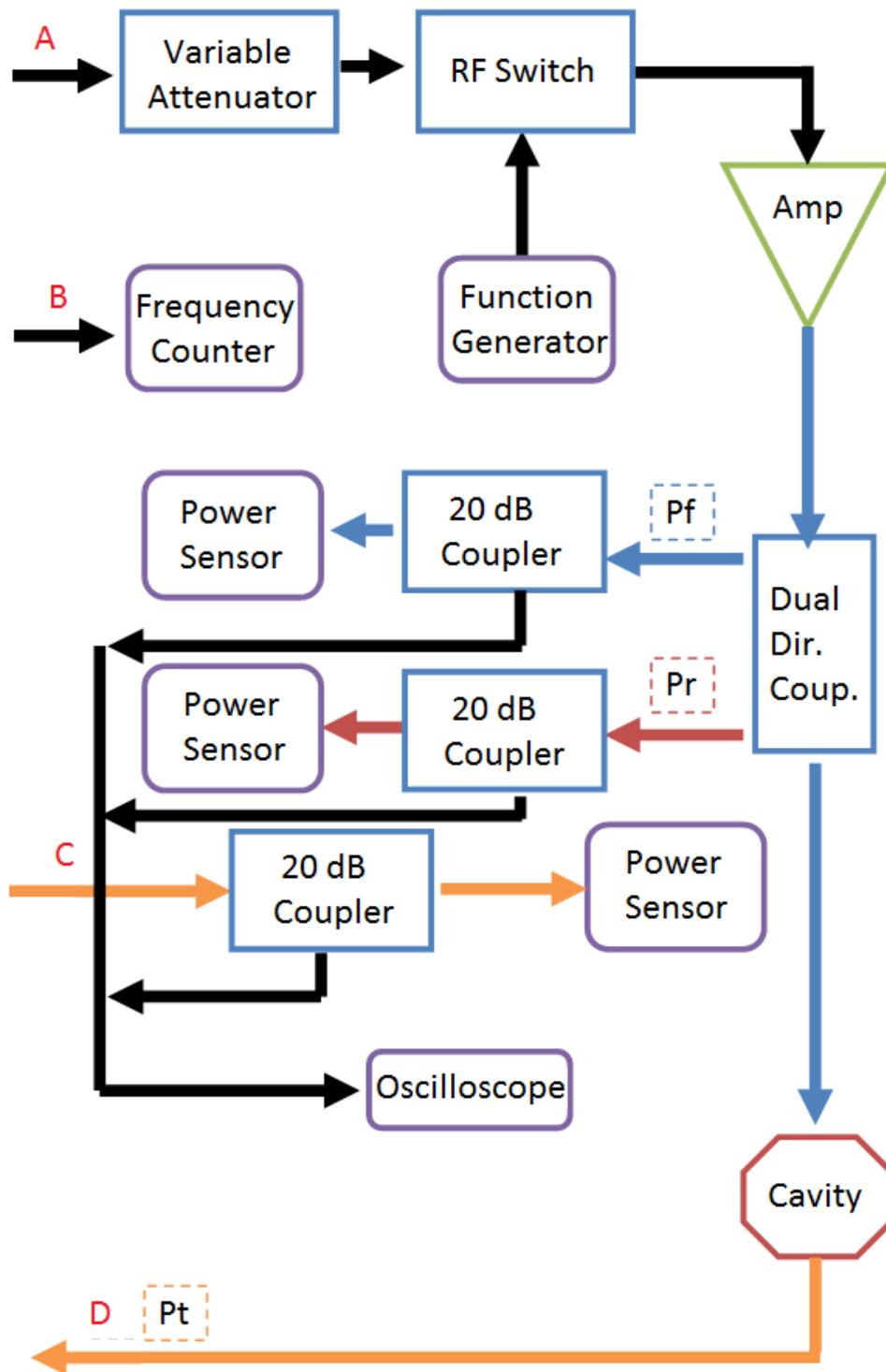


Figure 38: One half of the RF circuit for a phase lock loop for cavity testing (circulator, isolators, and RF diodes excluded). Letters in Red connect to the other half of the RF circuit, seen on the previous page.