Effects of surface damage on rf cavity operation

A. Hassanein, Z. Insepov, and J. Norem*
Argonne National Laboratory, Argonne, Illinois 60439, USA

A. Moretti, Z. Qian, and A. Bross
Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA

Y. Torun
Illinois Institute of Technology, Chicago, Illinois 60616, USA

R. Rimmer
Jefferson Laboratory, Newport News, Virginia 23606, USA

D. Li and M. Zisman
Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

D. N. Seidman and K. E. Yoon
Northwestern University, Evanston, Illinois 60208, USA

(Received 6 March 2006; published 8 June 2006)

We describe a model of damage in rf cavities and show how this damage can limit cavity operation. We first present a review of mechanisms that may or may not affect the ultimate fields that can be obtained in rf cavities, assuming that mechanical stress explains the triggers of rf breakdown events. We present a method of quantifying the surface damage caused by breakdown events in terms of the spectrum of field enhancement factors, \( \beta \), for asperities on the surface. We then model an equilibrium that can develop between damage and conditioning effects, and show how this equilibrium can determine cavity performance and show experimental evidence for this mechanism. We define three functions that quantify damage, and explain how the parameters that determine performance can be factored out and measured. We then show how this model can quantitatively explain the dependence of cavity performance on material, frequency, pulse length, gas, power supply, and other factors. The examples given in this paper are derived from a variety of incomplete data sets, so we outline an experimental program that should improve these predictions, provide mechanisms for comparing data from different facilities, and fill in many gaps in the existing data.

DOI: 10.1103/PhysRevSTAB.9.062001

PACS numbers: 29.17.+w, 52.80.Vp

I. INTRODUCTION

At least three different research and development efforts are independently studying the behavior of high gradient rf structures for accelerators. The Neutrino Factory and Muon Collider Collaboration (NFMCC) is looking at developing low frequency structures for muon cooling [1–5], the International Linear Collider is optimizing the performance of 1.3 GHz superconducting rf structures aimed at the design of a 1 TeV superconducting electron/positron collider [6], and the High Gradient RF Collaboration is studying high frequency \( (f > 10 \text{ GHz}) \) structures aimed at an electron-positron collider operating at energies higher than 1 TeV [7].

Muon ionization cooling in flight requires absorbers to reduce the muon momentum, accelerating fields to replace the lost momentum, and static solenoidal magnetic fields to focus the muon beams. The process is most efficient if both the magnetic fields and accelerating fields are high. Our experimental program studies high gradient rf in open and closed-cell cavities in a solenoidal field [1,8]. This program has primarily been at 805 MHz, but we are extending this work to 201 MHz, the frequency used in the muon ionization cooling experiment (MICE) [9]. This work has led to a small modeling effort to try to understand the mechanisms that limit accelerating gradients in rf structures in terms of mechanical stresses exerted by high local fields on the surface [2,3]. In addition, we have attempted to compare our data with the large volume of data in this very well studied field [10–13]. The field of breakdown in rf and DC structures has a very long history [10–16] and the models proposed have been diverse and contradictory. Reference [16] is an excellent summary of the field. While many have concluded that there may be more than one mechanism responsible, we argue that one mechanism can qualitatively account for much of the data, and experimental work can provide data that can make these predictions more precise.
We will argue that the peak surface field of high field rf structures can be described by the relation

$$E_{\text{surf}} = \frac{\sqrt{2\sigma/\varepsilon_0}}{\beta_{\text{eq}}},$$

where the maximum surface field $E_{\text{surf}}$ is determined by the tensile strength, $\sigma$, the permittivity constant, $\varepsilon_0$, and $\beta_{\text{eq}}$, described in detail below, is a field enhancement factor which depends on surface damage/condition [17]. In most cases, values of the enhancement factor are approximately $\beta \sim 100$. While it is not clear that only one mechanism is responsible for the wide variety of breakdown phenomena, we propose to study mechanical stress on surfaces because this mechanism undoubtedly contributes, and contributions from other effects can be considered after the effects of these stresses are understood. The effects of surface damage have been incorporated into the variable $\beta$ and will be discussed below. At very high frequencies ($f > 10–15$ GHz) the interaction of high surface current densities with defects and grain boundaries may also become significant [2]. While primarily relevant to normal conducting cavities, we believe this work also has some consequences for superconducting rf structures.

While there has long been a great deal of data, many laboratories and individuals have developed cleaning and handling methods without systematic comparisons. In this article, we first identify the parameters that are and are not involved in breakdown, discussing the dependence on frequency, cavity dimensions, state of conditioning, magnetic field, gas pressure, fatigue, temperature, pulse length, stored energy, available power, and other variables. Then we develop a formalism for parametrizing cavity damage due to surface changes during breakdown events, and show how three experimentally measurable spectra seem to determine completely the behavior of a given cavity, and can be used to estimate the operational limits of other cavities. We then use these spectra to estimate the maximum field, pulse length, geometry, and material dependence on cavity operation, as well as other useful parameters, and describe an experimental program that will permit more precise predictions from the model.

We have attempted to use existing data as efficiently as possible to explore the limits of this model. The arguments and data in this paper are somewhat preliminary and obtained from a variety of sources. Because there are no complete sets of data from a single cavity, however, we are forced to study a wide range of phenomena in different cavities and attempt to draw conclusions from data that are not entirely consistent. Since most of the data were taken for other purposes, well designed experiments could produce better data, a wider range of predictions, and higher precision. We outline an experimental program that would provide the required data.

II. BREAKDOWN MECHANISMS

Previous papers have outlined a model of breakdown based on electrostatic stresses producing fragmentation of highly stressed cavity materials [1,2,8]. This model is qualitatively consistent with a wide variety of data, but it is difficult to produce precise calculations without knowing more about the details of this fragmentation. For this reason we have begun to examine data from atom probe tomography (APT) which looks at ions extracted from solid surfaces at electric fields in the range $5 < E < 100$ GV/m [18,19]. Data from arcing and rf studies also provide useful information, however, and we can summarize some of this here. It is particularly interesting to look at the dependence of the maximum surface field on frequency, pressure, degree of conditioning, magnetic field, and temperature and we summarize recent data here.

A. Existing data

A recent collection of data from a variety of sources shows the frequency dependence seen in modern rf systems. This is shown in Fig. 1. This plot shows the average surface field seen in the cavity and the local field present on the surface of field emitters, as measured by the properties of field emitted electrons. The surface field and the local field at asperities is related by the enhancement factor, $\beta = E_{\text{local}}/E_{\text{surf}}$. We see that the average surface fields rise with frequency, but the local surface fields remain constant at about $6–7$ GV/m independent of frequency, which seems to be due to changes in $\beta$ between different systems that we attempt to explain in the remainder of this paper. DC systems have been shown to break down at an average surface field of about $160$ MV/m [10,11], and the local
field at asperities of 6–7 GV/m is consistent with the field where tensile stress becomes equal to tensile strength for copper. This picture is consistent with there being no significant dependence of the maximum gradient on the frequency, if it is assumed that it is the local fields of emitters that are responsible for breakdown, and not the average surface fields. Note that some of the scatter in the points can be due to differing definitions of an acceptable breakdown rate.

The breakdown levels discussed in this paper are the operating fields reached in accelerating structures after being fully conditioned, with operating conditions and pulse lengths appropriate to their (frequency dependent) filling times. Since the breakdown rate is a function of the accelerating field and pulse lengths, and the tolerable breakdown rates vary from one facility to another depending on the ultimate use of the cavity, there is some ambiguity or systematic error in comparing data from different cavities built for different uses. While more precise methods of comparing cavities could and should be developed, the maximum operating field and local fields are the most relevant parameters in this model.

The primary picture of the frequency dependence of rf breakdown is from Kilpatrick [20] who published in the 1950s. These studies were done with early rf cavities that seem to have been comparatively roughly constructed by modern standards, and it is seen that recent cavities generally exceed the “Kilpatrick Limit/Criterion” by about a factor of two. Nevertheless, the scaling law seems to produce roughly the frequency dependence seen in modern data.

Figure 1 also shows the gradients at which field evaporation of copper would occur and the range where field emission of electrons occurs. It is seen that the local fields required for field emission of electrons and fields capable of damaging surfaces are only different by a factor of probably less than two. This paper attempts to explain the surface field limits shown in Fig. 1 for copper structures. Some aspects of the behavior of superconducting cavities are also described using this model described here, as shown in Sec. IV, and these cavities can operate in modes similar to copper systems.

**B. The breakdown model**

The model of breakdown triggered by tensile stresses in the material has been discussed elsewhere [1,2,8]. In this model, fracture of the surface due to electrostatic forces triggers the event. The fragment produced is then heated and ionized by field emitted electron beams to produce a small local plasma [21]. The lossy plasma produced then couples the electromagnetic energy of the cavity to the wall, triggering a breakdown event, and ultimately converting most of the stored energy to heat. Experimental evidence for this is obtained from field emitted beams, which show a maximum local surface field at the tips of asperities of $E_{\text{local}} = 7 \, \text{GV/m}$ in a wide variety of applications. These values of $E_{\text{local}}$ are obtained from fitting the Fowler-Nordheim emission curve through data on radiation or dark current levels as a function of electric field for structures like those shown in Fig. 1. At these fields the electrostatic stress is equal to the tensile strength of copper. The maximum surface field that can be obtained in any structure seems to be equal to

$$E_{\text{surf}} = \sqrt{\frac{2\sigma}{\varepsilon_0}} \sim 7 \, \text{(GV/m)},$$

where $\beta_{\text{eq}}$ is determined by the damage left by the breakdown event. A method for evaluating $\beta_{\text{eq}}$ will be presented below.

Data from materials science supports this model. In APT, small samples of materials are subjected to surface fields from 2–150 GV/m, and the ions produced are identified, permitting computer reconstruction of the material. Long experience with this technology [22] has shown that samples frequently fracture at comparatively low fields (2–10 GV/m). when first exposed to high fields. We are actively pursuing this problem.

There are a number of questions that require more experimental data, however, such as possible contributions from other mechanisms, field emission, heating from field-emission currents, fatigue, plasma spots, and the behavior of metals under high fields in general. In addition, how mechanical forces apply in complex systems, at the nanoscale, is not well understood. It is possible to understand the interactions of materials under high fields, but serious study in this area is just beginning.

Adsorbed gas or loosely bound oxides have often been proposed as the trigger for breakdown, essentially assuming that this gas is ionized and produces a lossy plasma. The oxide is, in fact, generally harder than the pure metal, and the weakest point in an oxide coating is possibly where the oxide is coupled to the pure metal. We regard fragments of oxide from a metal surface to be equivalent to the metal for the purposes of our model. Data from atom probe tomography on copper and niobium surfaces at high field show failure of this type.

Other mechanisms that have been proposed include plasma spots, field emission, and multipactor. Plasma spots, which have been seen on the surface of a number of cavities, are the basis of a model proposed by Wilson [23], and these predictions have been found useful [24]. Field emission is the most visible result of the operation of high gradient surfaces [21,25]. Multipactoring, the resonant amplification of electrons produced by sequential acceleration and secondary emission of electrons in time with the rf fields in the cavity, has often been associated with breakdown events [23,26]. Although this mechanism appears in some models, the evidence that this process has a significant effect on breakdown thresholds is not strong.
III. PARAMETRIZING SURFACE DAMAGE

When a breakdown event occurs, we expect that the asperity that triggered the event (usually also a field emitter) is destroyed, as was shown in Ref. [8]. It is common, however, for molten metal from one breakdown site to be transported some distance across the cavity to make other secondary emitters [8]. The production of secondary emitters ultimately limits the operating gradient. There should be a threshold in available energy (stored energy plus input power within some time interval) below which breakdown sites are destroyed, and above which more breakdown sites are created, therefore limiting the maximum operational field. We would expect this threshold to be larger than the energy required to melt and eject the volume of copper comparable to those seen in microphotographs of active surfaces.

We assume that a small plasma is produced at the point of the trigger and this lossy plasma produces an avalanche effect by transmitting the electromagnetic energy of the cavity into the walls. We are not aware of any systematic measurements of breakdown where the radiated energy, wall heating, etc. were measured.

The approximate magnitudes of these energies can be estimated from data taken in the 805 MHz open-cell and pillbox cavities described in Refs. [1,8]. The pillbox cavity stored an energy of $U_s = \int e_0 E^2/2dV$, which is about 5 J at full field ($\sim 25$ MV/m), and the 12 MW power supply could contribute some additional energy in a discharge that occurred in a few hundred ns. Since we see many shallow craters and copper droplets with radii of 100 µm in our cavity, one can estimate the amount of energy, $U$, to melt and expel this volume of copper as $U = V \rho (c\Delta T + L)$, where $V$, $\rho$, $c$, $L$ and $\Delta T$ are the volume, density, specific heat (385 J/kg°C), heat of fusion (2.05 $\times$ 10$^5$ J/kg), and temperature increase required to melt the copper, on the order of 1000°C. Craters or droplets of this size both represent about 10 mJ of energy in the copper, thus it seems that only a few percent of the available electromagnetic energy goes into melting copper and creating potential secondary emitters. We have also shown that the spectrum of secondary emitters is peaked at very low enhancement factors [8], so the production of likely secondary breakdown sites requires significant energy. Although the mechanism is unclear, we assume that damage to the surface is determined by the energy in the discharge.

Since the stored energy of cavities of equal field increases with their volume, and the dimensions of cavities are inversely proportional to their operating frequency, we expect that the operating field will be related to the operating frequency.

In order to understand the conditioning process and the limiting operating field it is useful to look at the density of field emitters in the form of a function of the spectrum of enhancement factors. We define three functions $s_1(\beta)$, $s_2(\beta)$, and $s_3(\beta)$, which describe this damage and are illustrated in Fig. 2. These functions have the dimensions of [(number of emitters)/(unit area)(unit interval in $\beta$)]. They describe, respectively, the surface before any power is applied to the structure, the damage caused during a breakdown event, and the surface that develops when a long conditioning process leads to a cavity operating at the highest surface fields.

(i) When first produced, the cavity has an initial spectrum, $s_1(\beta)$, of field enhancements, $\beta$, at emission sites, the hottest of which are burned off during conditioning. These emitters have been studied in DC systems as part of superconducting rf development [27,28].

(ii) Every breakdown event produces surface damage, and the spectrum of enhancements produced in breakdown events is described by, $s_2(\beta, U)$, with $U$ the energy available during the breakdown event. We assume that this function should be directly proportional to the breakdown energy coupled to the wall. Discharges with more energy will produce more emitters, and/or emitters with larger enhancements. At low operating fields (breakdown energies), the hottest emitters are destroyed and not replaced. Increasing the cavity field (breakdown energy) will produce breakdown events with increasing field enhancements until the hottest emitters are effectively replaced during breakdown. We have measured $s_2(\beta)$, as shown in Fig. 3 [8]. We assume that $s_3$ is proportional to the energy available in the discharge.

\[
\begin{align*}
s_1(\beta, \text{material}) & \quad \text{The Initial distribution of field enhancement factors.} \\
\beta & \quad \text{After the hottest emitters have been burned off.} \\
\beta & \quad \text{But each burnt emitter produces a spectrum of secondary emitters.} \\
\beta & \quad \text{Secondary emitter spectrum not the same as the initial distribution} \\
\beta & \quad \text{The spectrum of secondaries determines the operating field.}
\end{align*}
\]
(iii) Ultimately a fully conditioned cavity should develop a spectrum of enhancement factors, $s_3(\beta, U)$, which at low $\beta$ values is dominated by the spectrum $s_2(\beta, U)$, but with all the asperities with high $\beta$ burned off. At some value, $\beta_{\max}$, there should be a sharp discontinuity in the spectrum. Since these functions determine the behavior of structures during conditioning and operation, they, and the effects they cause, should be measurable using a variety of techniques.

A. Damage from breakdown events

If we define $\beta_{eq}$ as the highest enhancement factor that is tolerable on the surface, the cavity performance will improve or degrade depending on the integral

$$\int_{\beta_{eq}}^{\infty} A_2(\beta, A, U, \tau, B, \text{material}) d\beta = p,$$

where $A$ is the active area, and $p$ is number of breakdown sites hotter than the one that was destroyed. This point determines the maximum operating field of the cavity and we assume that, when $p$ is greater than or equal to 1, the operating conditions of the cavity cannot be stable. One measurement of the spectrum of secondary emitters has been made for $\sim5$ J discharges, in Ref. [8], but it would be useful if systematic measurements were done as a function of the energy $U$. The nature of these functions is shown in Fig. 2.

This model argues that extensively conditioned cavities should show very large numbers of emitters with small enhancement factors and a sharp cutoff of the enhancement factor spectrum at

$$\beta_{eq} \sim \gamma_{(GV/m)} / E_{\text{surf, max}},$$

which should be experimentally detectable. Using the 805 MHz cavity extensively described in Ref. [1], we were able to measure the intensity of emitters using a solenoidal field to confine the dark current electrons. The narrow range of intensity distribution of emitters detected, shown in Fig. 3, shows that only a small fraction of emitters on a surface actively produce measurable dark currents.

If the spectrum $s_3(\beta)$ were a complicated function of many variables, it would not be particularly useful. We assume, however, that in many cases, the dependence on external parameters are simply factorable. For example, if breakdown event $K$ has twice the deposited energy of breakdown event $L$, we would expect that $s_3(\beta, U_K) = 2s_3(\beta, U_L)$, since twice as much deposited energy should produce twice as much damage. More generally, we assume that the dependence on energy deposited to the walls, $U$, can be written as $s_3(\beta, U) \propto U_s(\beta)$. It is possible in this way to determine, from a wide variety of data, how the contributions from different parameters can be factored. We will show that the active area and deposited power in an event seem to factor in this way. This is shown in Fig. 4.

If, guided by the data in Fig. 3, we parametrize $s_2(\beta) = a e^{-b \beta}$, it is possible to solve the integral

$$\int_{\beta_{eq}}^{\infty} A_2 e^{-b \beta} d\beta = 1$$

to obtain

$$A_2 e^{-b \beta_{eq}} / b = 1,$$

which can be solved for $\beta_{eq}(a, b)$. From this simple
parametrization we find that the enhancement factor is of the form

$$\beta_{eq}(a, b) = -\ln(b/Aa)/b,$$

where all the experimental variables (stored energy, pulse length, etc.) enter through the natural log. The constant $b$ has been measured in Ref. [8], and the value was found to be 0.03, which we use in this article. Using this expression then becomes a question of inserting realistic expressions into the variable $a$, which should be proportional to the energy in the discharge, and should primarily determine the spectrum of damage. We can calculate the maximum field for a given rf structure from the maximum local field, which we assume is determined from the tensile strength by the expression

$$\varepsilon_0 E_{\text{surf}}^2/2 = \sigma,$$

where $\sigma$ is the tensile strength, in Pa, of the material, and $E_{\text{max}}$, in V/m, is the maximum surface field. We will use the relation,

$$E_{\text{surf}} = \sqrt{2\sigma/E_0/\beta_{eq}} \sim \sqrt{2\sigma/E_0/\ln(b/Aa)/b},$$

in a number of examples that are presented below. The dependence of the maximum surface field on the surface damage parameter, $a$, which is proportional the magnitude of the damage, is shown in Fig. 5. As damage increases, the maximum surface field would be expected to decrease approximately following this curve. The actual values of the field that are predicted depend on how the constant $a$ is factored.

While it is likely that the spectra $s_2(\beta)$ have a shape like a Maxwell-Boltzman distribution, we can only measure these functions over a comparatively narrow range, and have little knowledge of their shape where they have not been measured. Thus, some caution is required when using predictions of this model, particularly in the overall normalization of the data. When the spectra $s_1$, $s_2$, and $s_3$ are better measured, improved predictions will be possible. In this paper we try to fit the largest variety of data possible to determine dependencies on one parameter at a time. If we believe all cavity performance is determined by these functions, the effects of the functional dependence should be quite clear.

The measured spectrum $s_2(\beta)$ may or may not give an accurate picture of what damage may look like for individual breakdown events. In these events, the damage would be expected to be highly position dependent, with the most active secondary breakdown sites very close to, or part of, the initial breakdown site. The data in Fig. 3 shows damage rather far removed from the breakdown sites, and likely showing a somewhat different spectrum than would be obtained by integrating over the entire cavity area. Nevertheless, since this data is the best guide available, we use it in the examples below. We find that the predictions obtained from this spectrum agree well with a wide variety of experimental data.

**B. The fully conditioned state**

As shown in Fig. 2, after all active asperities are burned off, there should be a discontinuity in the equilibrium enhancement factor spectrum, $s_3(\beta)$, at $\beta_{eq}$, due to the instability of asperities with larger values of $\beta$. This can be expressed by parametrizing this function in the form

$$s_3(\beta, a, b, d) \sim \frac{ae^{-b\beta}}{e^{(\beta-\beta_{eq})/d} + 1},$$

where $a$ and $d$ are constants to be fitted from data. In this case, the numerator is essentially the measured distribution from Ref. [8] and described above, and the denominator produces the Fermi-Dirac distribution function, which is equal to 1 below $\beta_{eq}$ and 0 above $\beta_{eq}$, with the width of the transition region described by the constant, $d$ [29]. The Woods-Saxon potential function, used in nuclear physics, has similar properties [30].

The parameters of $s_3(\beta)$, particularly in the region of the discontinuity at $\beta_{eq}$, are very important to the behavior of the cavity, and are experimentally accessible in a number of ways. The most direct is a measurement of the intensity of emitters in a thoroughly conditioned cavity. We would expect to observe the distribution of emitter strengths multiplied by the Fowler-Nordheim emission curve, which we approximate with $E^n$, in this case $E^{16}s_3(\beta)$ as derived in Ref. [1]. This product produces a fairly narrow range of emitter strengths, constrained above by the absence of hot sources and below by the Fowler-Nordheim emission law.

We can measure the distribution of emitter strengths from the optical density of dark current beams in Fig. 6, and compare this with the estimate obtained from the function $s_3(\beta)$. The results of this are shown in Fig. 7.
This figure plots the function $s_3(\beta)$ in a way that shows the discontinuity at $\beta_{eq}$ as a change of slope. The optical density measured for the photograph in Fig. 6, shown in an inset, is fitted with a curve, and points from this curve are compared with the expected spectrum of dark current beam intensities $E^{16}s_3(\beta)$. It is important to note that the optical density, measured from glass slides, has an uncertain horizontal calibration due to the effects described in Ref. [1]. Nevertheless there is a close match between the measured and predicted spectrum.

Examining the shape of $s_3(\beta)$ in the region around $\beta_{eq}$, we see that the emitter density is roughly proportional to $\beta$ with a high negative exponent. Since the breakdown rate is a function of the local electric field, $\beta E_{surf}$, this model would predict that the breakdown rate for fully conditioned cavities show a dependence like $E_{surf}^n$, with $n$ large. This is, in fact, what is measured, as shown below in Sec. IV D.

IV. USING THE MODEL

It is possible to compare this model with the huge volume of data on rf cavity operation and vacuum breakdown that has been produced in the past 100 years. As mentioned above, however, the predictive power is limited by uncertainties in the experimental measurements of $s_1$, $s_2$, and $s_3$. It is useful to look at one parameter at a time to learn how to factor the functional dependence of the variables (stored energy, area, pulse length) that are involved. While it should be possible to use this model to explain all rf structure behavior with good precision, in this paper we will only outline these procedures and present preliminary results.

The aim is to show that the model is at least compatible with all good data from rf structures, and able to quantitatively predict much of it. Our primary interest is in high gradient, low frequency rf systems operating in strong (2.5 T) magnetic fields, and a detailed knowledge of a wide range of parameters should help us do the necessary extrapolations.

A. Degree of conditioning

RF structures must cope with imperfections and cumulative damage throughout their operational lives. This damage may come from normal or fault conditions and should not require maintenance.

When a cavity is first fabricated, the surface is usually dominated by a few potentially hot breakdown sites. In order to operate at the maximum gradient it is necessary to burn these off. As they burn off, the enhancement factors of active emitters will decrease. The equilibrium condition will be reached when the emitters produced during breakdown events are, on average, as active as the ones being destroyed.

When cavities are first turned on it is necessary to condition them, by slowly increasing the operating field as breakdown events occur at some tolerable level. Data from earlier studies have shown that during this process the enhancement factors $\beta = E_{local}/E_{surf}$ and $E_{surf}$ are both changing, however the product $E_{local} = \beta E_{surf}$ is constant. This is shown in Ref. [1]. Better data on conditioning have been collected at KEK that show that, as the average field

FIG. 6. (Color) The pattern of dark current radiation damage on a glass slide at the exit window of an open-cell rf cavity. These data are described in Ref. [1]. The area used for the analysis of optical density is shown in the rectangle.

This figure plots $s_3(\beta)$ in a way that shows the discontinuity at $\beta_{eq}$ as a change of slope. The optical density measured for the photograph in Fig. 6, shown in an inset, is fitted with a curve, and points from this curve are compared with the expected spectrum of dark current beam intensities $E^{16}s_3(\beta)$. It is important to note that the optical density, measured from glass slides, has an uncertain horizontal calibration due to the effects described in Ref. [1]. Nevertheless there is a close match between the measured and predicted spectrum.

Examining $s_3(\beta)$ in the region around $\beta_{eq}$, we see that the emitter density is roughly proportional to $\beta$ with a high negative exponent. Since the breakdown rate is a function of the local electric field, $\beta E_{surf}$, this model would predict that the breakdown rate for fully conditioned cavities show a dependence like $E_{surf}^n$, with $n$ large. This is, in fact, what is measured, as shown below in Sec. IV D.

IV. USING THE MODEL

It is possible to compare this model with the huge volume of data on rf cavity operation and vacuum breakdown that has been produced in the past 100 years. As mentioned above, however, the predictive power is limited by uncertainties in the experimental measurements of $s_1$, $s_2$, and $s_3$. It is useful to look at one parameter at a time to learn how to factor the functional dependence of the variables (stored energy, area, pulse length) that are involved. While it should be possible to use this model to explain all rf structure behavior with good precision, in this paper we will only outline these procedures and present preliminary results.

The aim is to show that the model is at least compatible with all good data from rf structures, and able to quantitatively predict much of it. Our primary interest is in high gradient, low frequency rf systems operating in strong (2.5 T) magnetic fields, and a detailed knowledge of a wide range of parameters should help us do the necessary extrapolations.

A. Degree of conditioning

RF structures must cope with imperfections and cumulative damage throughout their operational lives. This damage may come from normal or fault conditions and should not require maintenance.

When a cavity is first fabricated, the surface is usually dominated by a few potentially hot breakdown sites. In order to operate at the maximum gradient it is necessary to burn these off. As they burn off, the enhancement factors of active emitters will decrease. The equilibrium condition will be reached when the emitters produced during breakdown events are, on average, as active as the ones being destroyed.

When cavities are first turned on it is necessary to condition them, by slowly increasing the operating field as breakdown events occur at some tolerable level. Data from earlier studies have shown that during this process the enhancement factors $\beta = E_{local}/E_{surf}$ and $E_{surf}$ are both changing, however the product $E_{local} = \beta E_{surf}$ is constant. This is shown in Ref. [1]. Better data on conditioning have been collected at KEK that show that, as the average field
rises by a factor of two, the enhancement factor decreases by the same amount and the product, $\beta E_{\text{surf}}^2$, is constant during conditioning. This agrees with the argument that the local field is the dominant variable. The KEK data are shown in Fig. 8 [31].

**B. Materials**

Constraints due to materials have never been systematically explored in rf structures, although carefully measured data are beginning to become available [32,33]. There are many relevant electrical and mechanical parameters and there are not yet sufficient data to completely predict or optimize the performance of a given material.

A significant problem with much of the experimental data is that breakdown may be determined in many cases by surface contamination, which can consist of micron sized dielectric and metallic particles. This problem is less significant in rf cavities than in DC breakdown tests, particularly if the rf structures have been conditioned at high power for long periods, under $10^{-8}$ Torr vacuum. Variations in the work function, used to estimate the local field, due to the adsorption of gases on a surface are not considered in these calculations.

The primary motivation to study breakdown is to see if it is possible to increase the breakdown limits seen in a variety of experimental situations. Since the surface field should be primarily determined by the mechanical properties of materials, it should be possible to find materials that permit higher gradients. The maximum surface field (for a copper structure) is determined by the relation

$$E_{\text{surf}} = \left(\frac{2T}{\epsilon_0} \sim 7\,\text{GV/m}\right) / \beta_{\text{eq}}$$

where the maximum local field of 7 GV/m is determined by the tensile strength, $T$, of copper. There seem to be two ways to improve the performance of high gradient structures: (i) using higher tensile strength materials that would survive higher local gradients would increase the local electric field, $E_{\text{local}}$, and (ii) finding materials that produce more optimized spectra during cavity damage, $s_2(\beta)$, lowering the values of $\beta$ produced in breakdown events. Both of these options seem possible.

Existing data on the dependence of breakdown on surface materials show a general trend that softer materials (gold, silver) break down easily, and harder materials (stainless steel, tungsten, molybdenum) seem to have higher breakdown thresholds (and conditioning times) [32,33]; this is shown in Fig. 9. Data on DC breakdown as a function of material are also available; however, systematic studies are old, and not done in a high vacuum environment or with care to ensure that the surface was not contaminated with particulates. Figure 10 shows the sort of damage that is found in cavities. Damage in cavities is hard to measure and difficult to parametrize; however, atom probe tomography is designed to study systematically the behavior of pure materials, alloys, and coatings at high surface fields in a clean environment [19].

It should also be possible to achieve higher surface gradients by finding materials, which, when melted and
splashed around the cavity, would cool in a way that did not produce sharp asperities. Since, according to Ref. [34], the enhancement factor is inversely proportional to the local radius of these asperities, it should be possible to decrease the enhancement factors produced when the metal droplets cool, since the surface morphology of the splashes is due to surface tension, cooling rate, viscosity and other parameters.

C. Pulse length

The model proposed above becomes particularly simple in the case of pulse length. Since the cavity damage should be directly proportional to the energy deposited in a cavity, which is, in turn, proportional to the pulse length in a breakdown event, one can directly compare two different equilibrium configurations with different discharge energy. Following the derivation in Sec. III A, since the absorbed energy is proportional to the pulse length, \( \tau \), the result should take the form

\[
E_{\text{max}} \sim -0.03 \times \frac{7 [\text{GV/m}]}{\ln[0.03/(3AP\tau)]},
\]

where \( P \) is the available power and \( A \) the area. The number 0.03 comes from measurements of \( s_2(\beta) \) [8]. Assuming appropriate values for the variables gives values for the pulse length dependence that fall in the range \( \tau^{-1/6} \) to \( \tau^{-1/4} \), which generally agrees with SLAC experiments [33], see Fig. 11.

Since we assume the trigger of a breakdown event is mechanical stress, there should be little dependence on where in the pulse the event occurs. This is, in fact, what is seen.

D. Breakdown probability and rate

The operating limits of a structure will depend very strongly on many parameters and it is very useful to be able to predict and control this behavior, which is closely related to the ultimate limits of the structure.

In this model, both breakdown probability and rate are determined by the density of asperities with enhancement...
An asperity on the surface of an rf cavity will see an alternating electric field of magnitude $\beta E_{surf}$, which will produce a train of mechanical pulses at frequency $2f$, where $f$ is the frequency of the rf excitation. Since we assume that the magnitude of the electrostatic force can be close to the ultimate tensile strength of the material, it is reasonable to assume that fatigue effects could be detectable in cavity breakdown data.

The relation that governs fatigue life for examples where there is a high degree of strain (plastic deformation) is the Manson-Coffin relation [38],

$$\Delta \epsilon_p/2 = \epsilon'_f (2N_f)^{-c},$$

where $\Delta \epsilon_p/2$ is the plastic strain amplitude, $\epsilon'_f$ is the fatigue ductility coefficient, $2N_f$ is the number of strain reversals to failure, and $c$ is a material property in the range of 0.5 to 0.7.

For lower (elastic) strain amplitudes, the appropriate relation is

$$\Delta \epsilon_e/2 = (\sigma'_f / E)(2N_f)^{-c},$$

where $\sigma'_f$ is the fatigue strength coefficient, and $E$ is the modulus of elasticity. For elastic strain, the exponent $c$ is much smaller, perhaps 0.07–0.14 [38].

These relations show that the fatigue lifetime of a given sample depends on the strain amplitude, with a large range of exponents (−0.07 to −0.7). As a result of this, variations in strain amplitude can produce very large changes in fatigue lifetime.

**F. Correlated breakdown events**

The breakdown model described herein predicts that breakdown events at high fields will produce damage, and this damage will produce high $\beta$ asperities under high electrostatic stress, which can cause other breakdown events, which are correlated. This behavior has been seen in both the Fermilab linac [37], and the NLC cavity prototypes [35]. If the timing of a subsequent breakdown event was a completely random process, one would expect that the breakdown rate, $R$, would go as $R \sim e^{-ut}$, where $u$ is the time constant for whatever mechanism was involved, and $t$ is the time. On the other hand, if fatigue is the dominant factor, the lifetime, $u$, of a given breakdown site will depend on the local strain amplitude at the location of the defect that causes failure, and the process could be more complex.

If the defects that caused failure were distributed randomly through the asperity, the failure would be sensitive to a range of strain amplitudes, and would have a range of fatigue lifetimes. The lifetime of these samples would be the sum of a distribution of lifetimes that could, in principle, be calculated from the distribution of defects and strain throughout the sample. Thus, defects distributed heterogeneously throughout the structure would experience different

---

**E. Fatigue**

At high powers it is reasonable to expect cyclic stress and strain to contribute, through fatigue, to the behavior of rf systems. Because of the high surface fields used in atom probe tomography, samples are sensitive to similar effects. We can outline a simple application of fatigue models to nanoscale samples.

---

**FIG. 13.** (Color) Comparing breakdown rates as a function of pulse length for fixed electric field in the Fermilab linac (open squares) [37], and SLAC/NLC prototype (solid circles and triangles) [33], with estimates. The time scale for the Fermilab data is ms, and the SLAC data is ns. Both have a $\tau^5$ dependence over widely different time scales.

---

Factors large enough to fracture with a given surface field. As shown in Sec. III, the density of these asperities can be measured by a number of methods. Using the density obtained from measurements of the dark current beamlets in the glass plate, the width of the high $\beta$ cutoff was fitted, and the slope was found to be on the order of $\sim \beta^{-25}$. Thus, small changes in the electric field of the cavity would increase the density of active emitters by $\sim E^{25}$, therefore both the breakdown rate and probability for breakdown should have a very sharp threshold.

Breakdown rates have been measured in NLC and CLIC prototype cavities and found to have a $\sim E^{25}$ [35] or $E^{26}$ dependence [36], which is consistent with this model, see Fig. 12.

Assuming that the breakdown rate goes as $R \sim E^{25}$, and the pulse length dependence of the maximum electric field goes as $E \sim \tau^{-1/5}$, one would expect that the breakdown rate for constant field should look like $R \sim \tau^{25/5} = \tau^5$. Data taken during the Fermilab linac conditioning exhibit this behavior, as shown in Fig. 13 [37]. Breakdown rates in SLAC/NLC prototypes also follow a rough $\tau^5$ dependence [33] over widely different time scales.
strain amplitudes and a distribution of these defects could produce a complex lifetime curve.

In Fig. 14, we show how a distribution of lifetimes could sum to produce the breakdown number vs interval between successive sparks taken during the initial operation of the Fermilab linac [37]. Any single time constant does not fit the data, but the sum of a distribution of time constants can fit the data.

In continuing operation of the Fermilab linac over a period of seven years, this \( R \sim 1/t \) trend of breakdown rate with time has continued, indicating that the model may be relevant over longer time scales [39].

Failures in atom probe tomography samples frequently occur during initial operation, and these failures are not confined to the tips of the samples, where the stresses are presumably largest. Birdseye and Smith [22] have shown an example of a known defect far away from the tip, experiencing much less than the maximum strain, which seems to be the trigger for failure.

G. Predicting the Maximum Field—Scaling

Each rf structure, driven by a power supply, is a unique system, with many parameters that affect the breakdown process; frequency, geometry, pulse length, stored energy, material, electromagnetic field, and coupling method. We argue that the damage is primarily a function of only a few of these parameters and this permits an attempt at developing simple scaling laws.

If we factor out the contribution of energy, neglect the variation due to magnetic field, pulse length and material, and extrapolate from the measurements of damage in the 805 MHz pillbox cavity described in Sec. III, we can assume, following [8], that the overall damage spectrum \( A_S(\beta) \) for the structure takes the form

\[
A_S(\beta, U_c) \sim 1.5AU_c e^{-0.03\beta},
\]

where \( U_c \) is the stored energy of the cavity (\( \sim 5 \) J), and \( A \) is the active area of one end of the cavity. Then we assume that the maximum surface field the cavity can support, over an area \( 2A \), is determined by

\[
\int_{\beta_{eq}}^{\infty} 3Ae^{-0.03\beta}d\beta = 1,
\]

which can be solved for

\[
\beta_{eq} = -\ln\left(\frac{0.03}{3AU_c}\right)/0.03 = \frac{7\text{[GV/m]}}{E_{\text{max}}},
\]

or

\[
E_{\text{surf}} = \frac{-0.03 \times 7\text{[GV/m]}}{\ln(0.03/(3AU_c))}.
\]

This model should be able to explain the frequency dependence seen in comparisons of data, Fig. 1. These comparisons, however, inevitably incorporate data from a wide variety of cavity geometries (single or multicell, standing or traveling wave), power supplies, cavity Q, tolerable breakdown rate, and state of conditioning. These relations make it possible to understand how the...
interior environment of cavities changes with differences in design and operation, and may produce more accurate estimates of the maximum field that can be produced under given conditions.

Because the cavity parameters enter in the logarithm, the operating field is somewhat insensitive to the cavity parameters. We show in Fig. 15, the maximum surface field as a function of frequency using this model, which qualitatively agrees with the Kilpatrick limit. There should be some minimum energy required to produce damage, which should produce a high frequency limit to damage. At high frequencies, however, the current density for skin currents in the wall becomes a constraint, and the limiting mechanism seems to be fatigue. This effect has been derived by Wilson, who found that the maximum gradient rises with frequency as $f^{1/8}$, and may be on the order of 300 MV/m [40].

**H. Comparison with DC breakdown**

An enormous volume of data has been collected over the past century on the subject of DC breakdown. Unfortunately, it is not clear how much of this data is relevant to this problem. Although studied for 105 years, with an extensive international literature, there has been no agreement on the cause of DC breakdown. Although a number of models have historically been proposed [14–16], these have not convincingly explained the existing data, while we argue that the models described in Refs. [1–3] do explain the data.

Breakdown in both rf and DC structures seems to be due to mechanical stress becoming equal to the tensile strength. The DC analog of Fig. 1, which compares breakdown fields with system dimensions, is Fig. 16, where the local and average breakdown fields are plotted as a function of the gap length [14]. The data show that DC breakdown occurs at the same local field as in rf structures. The average surface field varies in a complex way with the system dimensions, which is discussed at length, and found to be due to geometrical factors in Ref. [14]. They interpret the change in enhancement factors in terms of the combined effect of an enhancement due to local microscopic projections on the cathode and the associated macroscopic changes in the electric field distribution at larger gap spacings. Larger dimensions would also involve larger transverse areas, including a larger sample of field emitters, which would inevitably include some with higher surface enhancement factors.

The model presented herein should be able to predict enhancement factors and average surface fields for DC systems, based on stored energy and geometry, however the geometries used to obtain the data in Fig. 16 are not known.

**I. Atom probe tomography sample failure**

As discussed above, the sharpened needles used as samples in field-emission microscopy and atom probe tomography fail frequently when first exposed to high pulsed electric fields. We assume that these samples have some microroughness, which is removed by preferential field evaporation. The process is described in a recent paper [41]. Experimental data obtained from room temperature copper samples has shown that there are large fluctuations in the evaporation rate, microflashes, and very high surface fields present in new samples; however, when smooth field evaporation is obtained from a sample, it can be exposed to air and reexposed to high fields without a significant conditioning period. This may imply that microroughness is the cause of the conditioning process.

The atom probe tomography environment provides a unique way of studying the interactions of high electric fields with surfaces. While generally used for analysis of bulk samples, surface analysis using this technique is also interesting and relevant. This work is continuing.

**J. Light and power switching**

An enormous volume of electrical power is ultimately controlled by 100–250 V switching systems, and, like many aspects of breakdown, the exact mechanism involved in triggering the initial current flow is not well understood.

The first measurements of vacuum breakdown, made in 1900–1905, showed that, for gaps small enough that gas breakdown was impossible, breakdown still occurred at high fields, due to mechanisms the authors assumed were operating on the surface of the material [10,11]. Since the
surface phenomena for both processes occur at surface fields of 100–150 MV/m for clean, but not perfect, surfaces, we believe that the mechanism discovered by these authors is, in fact, the vacuum breakdown mechanism responsible for rf cavity behavior, and our explanation of breakdown applies to their data. It is interesting to note that the same physical mechanism seems to be responsible for the initial flow of current in rf cavity breakdown, DC vacuum breakdown, and any time an electrical switch is closed (below about 300 V in air). Thus the mechanism is very common.

K. Discussion and future work

We believe the model described above can be useful in explaining a wide range of cavity operational limits, and are studying a number of extensions of the model. We describe a number of other applications where consistent experimental data exists and useful comparisons with the model are possible. The ultimate goal of our effort is to show that the model is consistent with all good data; however, here we only indicate that the model is generally consistent other work.

1. Gas pressure

The vacuum pressure of a cavity is probably the most easily controlled variable; however, it has not been shown that rf breakdown is particularly sensitive to pressure. Gas breakdown is a well understood phenomenon, and if this is allowed to occur, the cavity will also break down. On the other hand, even a fairly rough vacuum will prevent sufficient atomic densities to allow an electron avalanche, while high pressures can prevent gas avalanches by producing so much electron scattering that electrons never reach ionization energies. Small gaps can prevent gas avalanches by not allowing electrons sufficient space to accelerate. Between the high and low pressure data is a region where breakdown of the gas masks the breakdown at the surface.

Gas pressure enters this model because the heating of a fragment by intense field emitted electron beams can, in principle, be reduced if the field emitted electron beams that ionize the fragment are attenuated by high pressure gas. We show in Fig. 17 how data at very low pressures demonstrate that there is negligible pressure dependence to average surface breakdown fields using low pressure N₂ gas [1] which would be consistent with our model. The high pressure breakdown has been studied in air [10,11], and He or H₂ [43]. Much denser gases, such as SF₆ have long been known to suppress vacuum breakdown, and we can explain this due to the electron drag of the very dense gas preventing field emitted electrons from depositing their power into emitted surface fragments.

2. Geometrical dependence of damage

Since damage is produced by a complex physical process, the shape of the cavity, where the damage occurs, and how the damage products are ultimately distributed around the structure can affect structure performance. There are a variety of geometrical effects that must be considered.

From measurements in our pillbox cavity, it is possible to estimate the angular distribution of material expelled from a breakdown site by looking at the deposited copper on the Be windows. Preliminary data, with some uncertainty in the location of the breakdown sites, is shown in Fig. 18. Material is ejected with enough kinetic energy to uniformly cover a titanium window tens of cm away [1].

FIG. 17. (Color) Maximum surface fields as a function of pressure. Structures with smaller surface area seem to operate at slightly higher fields. Data are from Refs. [1,10,11,43].

FIG. 18. (Color) The figure shows (a) the cavity from Ref. [8] with the location of the breakdown events, (b) the window covered with spray from breakdown events, and (c) the approximate angular distribution of the spray.
This behavior shows that the damage functions $s_2(\beta)$ are highly position dependent, and integrals of this function over the cavity surface might be difficult to perform with precision.

2. Temperature and magnetic field dependence

The dependence of breakdown and temperature has been discussed in detail in recent experimental and theoretical papers [3,32], with the conclusions that the temperature of the material affects the breakdown process weakly, and this weak dependence is consistent with the mechanical stress being the primary cause. Likewise, mechanical stress also seems to explain, qualitatively, effects of magnetic field [8]. This paper outlines how mechanical effects from $\mathbf{j} \times \mathbf{B}$ forces can “unscrew” emitters, and a simple model based on this mechanism roughly explains the data.

3. High surface currents

There is evidence that at very high frequencies, the primary constraint on the maximum field is not breakdown damage. This is reasonable since there must be some threshold in discharge energy below which there is no significant damage, and this should ultimately change the behavior of small structures.

At high frequencies, there is an additional constraint to the maximum fields that can be generated in rf cavities. High surface current densities in cavities generate a limit on the maximum magnetic field that can be in contact with the wall, and both thermal and fatigue limits have been evaluated by Wilson [40]. These limits primarily apply at high frequencies ($f > 10 \text{ GHz}$) and seem to imply a maximum surface heating of 100 °C. The exact mechanism for this limit is not known. We note that the interactions of high current densities with grain boundaries and defects seem to be important but are not well understood [2,44].

4. Surface preparation

There is an extensive lore devoted to surface preparation in rf structures, comparing different machining, cleaning, and polishing techniques. The arguments, presented in Sec. III, lead to the conclusion that, for a well conditioned cavity at least, the surface and the ultimate performance of the structure is dominated by the damage produced by breakdown events experienced in the conditioning process. Since the spectrum of emitters that exist before the cavity breakdown events experienced in the conditioning process, the cavity at least, the surface and the ultimate performance of Sec. III, lead to the conclusion that, for a well conditioned rf structures, comparing different machining, cleaning, and polishing sources over the cavity surface might be difficult to perform with precision. That can produce potential breakdown sites. The enhancement factors of these sites have not been measured. This source of damage implies that simply providing a clean cavity surface may not be sufficient to avoid conditioning and breakdown damage.

5. Superconducting rf

The superconducting option has been selected for the International Linear Collider [6]. In addition to the critical magnetic field limit that applies to all superconducting rf structures, there is also a limiting electric field limit that is a result of surface imperfections and contamination. Since the performance of these structures cannot be reliably predicted, it seems important to understand how the processes seen in normal structures apply to superconducting ones.

While superconducting rf systems can, in principle, be made such that electric field limits do not apply, many recent, carefully made structures are limited by electric fields [27,45]. It is common that a period of “high power processing,” is required to neutralize active emitters. This processing seems to have exactly the same purpose as the conditioning process of normal cavities and follows the model shown in Fig. 2. The maximum tolerable dark current in superconducting rf structures is seen at surface fields of roughly 4 GV/m (evaluated from the slope of the radiation vs electric field curves as described in Refs. [1,45,46]), which is about half of the local field seen in copper cavities. This limits the maximum surface field to $E_{\text{surf}} = (4 \text{ GV/m})/\beta$, where $\beta$ is determined by the most active field emitter in the distribution $s_1(\beta)$. Superconducting rf performance has been compared with copper cavities in Fig. 1.

V. AN EXPERIMENTAL PROGRAM

Since the model proposed here provides an explanation of almost all aspects of the operation of rf structures, it should be easy to verify if these predictions are accurate and the guidance provided by the model is useful. If the relation

$$E_{\text{surf}} \leq \sqrt{\frac{2\sigma/\varepsilon_0}{\beta_{\text{eq}}}}$$

describes high gradient structures, one can improve the performance of rf structures by studying materials with high tensile strengths and somehow produce damage that conspires to have low values for $\beta_{\text{eq}}$. Because each structure and power source are somewhat unique, the nature of these experiments has produced many relatively uncoordinated measurements that are difficult to compare with each other. Thus it seems useful to describe an experimental program that can check and extend this model in an opti-
mum way. The elements of this program should include the following:

(i) Overall modeling of all aspects of the breakdown process: including triggers, energy balance, and material effects. Because each facility is unique, a database of experimental results would be very useful. The trigger mechanism may be complicated and should be better understood.

(ii) Measurements of rf structures over a wide variety of cavity breakdown configurations, material, coatings (including submonolayer), with parametric studies, paying particular attention to tensile strength and melting point. Measurement of surface damage spectra \( s_1(\beta), s_2(\beta), s_3(\beta) \) in situ, and comparison with measurements made in field-emission microscopes.

(iii) Atom probe studies of the behavior of metals, and surface failure at high electric fields. These should include studies of control of the surface, including oxides and metallic coatings.

(iv) Study of high current densities in materials. This would include modeling and experiments with scanning tunneling potentiometers and exploding wires.

It seems likely that there is considerable improvement in linac performance possible with optimized materials and design.

VI. SUMMARY

After reviewing existing data and models of rf breakdown, this paper proposes a comparatively simple method of predicting cavity performance based primarily on tensile strength and surface damage and we show how this model seems to be compatible with existing data and may be useful in predicting future results. This model should apply to all rf structures, independent of frequency, material, and geometrical design. The damage produced in a cavity by breakdown events seems to be one of the factors that determines its ultimate performance and we produce a method of quantifying this damage and experimentally observing the effects of cavity damage in a variety of applications. We then show how the model can be refined and used to explain a variety of old and new results. We also outline an experimental program that would refine and improve the model.

ACKNOWLEDGMENTS

We have had many useful discussions with P. Wilson, V. Dolgashev, and S. Tantawi of SLAC. We would like to thank Steve Geer along with many members of the Fermilab Accelerator and Technical Divisions for help in this work. J. Sebastian, now at Questek LLC, has also been very helpful. This effort was supported by the Office of High Energy Physics of the U.S. Department of Energy, under Argonne Contract No. W-31-109-ENG-38.

[46] M. Kelly (private communication).