2017

A crack closure model and its application to vibrothermography nondestructive evaluation

Bryan Schiefelbein
Iowa State University

A crack closure model and its application to vibrothermography nondestructive evaluation

by

Bryan Edward Schiefelbein

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Engineering Mechanics

Program of Study Committee:
Stephen D. Holland, Major Professor
Ashraf Bastawros
Pranav Shrotriya

The student author and the program of study committee are solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Thank you to my wife. This degree would not have been possible without your unwavering love, support, and faith in my work. It is for our future that I am driven to succeed.

I would like to acknowledge my parents for their continuing love and support, as well as the work ethic, values, and drive they instilled in me at young age.

Thank you to my fellow grad students Tyler Lesthaeghe, Jyani Vaddi, and Elizabeth Gregory. Your willingness to answer questions and provide support when I was just starting did not go unnoticed.

I would like to acknowledge the role that Dr. Ashraf Bastawros and Dr. Stephen Holland had in my graduate school education. Their knowledge, expertise, and guidance has been invaluable in my learning and growth over the past several years.

Thank you to all of my friends, family, and colleagues who have played a role in me obtaining this degree. Without your support and encouragement, I would not be where I am today.

This material is based on work supported by the Air Force Research Laboratory Contract #FA8650-10D-5210, Task Order #023 and performed at Iowa State University.

ABSTRACT

Vibrothermography nondestructive evaluation (NDE) is in the early stages of research and development, and there exists uncertainty in the fundamental mechanisms and processes by which heat generation occurs. Holland et al. [30] have developed a set of tools which simulate and predict the outcome of a vibrothermography inspection by breaking the inspection into three distinct processes: vibrational excitation, heat generation, and thermal imaging. The stage of vibrothermography which is not well understood is the process by which vibrations are converted to heat at the crack surface. It has been shown that crack closure and closure state impact the resulting heat generation [39; 55; 56]. Despite this, research into the link between partial crack closure and vibrothermography is limited [40; 55; 56].

This work seeks to rectify this gap in knowledge by modeling the behavior of a partially closed crack in response to static external loading and a dynamic vibration. The residual strains left by the plastic wake during fatigue crack growth manifest themselves as contact stresses acting at the crack surface interface. In response to an applied load below the crack opening stress, the crack closure state will evolve, but the crack will remain partially closed. The crack closure model developed in this work is based in linear elastic fracture mechanics (LEFM) and describes the behavior of a partially closed crack in response to a tensile external load and non-uniform closure stress distribution. The model builds on work by Fleck [18] to describe the effective length, crack opening displacement, and crack tip stress field for a partially closed crack. These quantities are solved for by first establishing an equilibrium condition which governs the effective or apparent length of the partially closed crack. The equilibrium condition states that,

under any external or crack surface loading, the effective crack tip will be located where the effective stress intensity factor is zero. In LEFM, this is equivalent to saying that the effective crack tip is located where the stress singularity vanishes. If the closure stresses are unknown, the model provides an algorithm with which to solve for the distribution, given measurements of the effective crack length as a function of external load.

Within literature, a number of heating mechanisms have been proposed as being dominant in vibrothermography. These include strain hysteresis, adhesion hysteresis, plastic flow, thermoelasticity, and sliding friction. Based on experimental observation and theory, this work eliminates strain hysteresis, thermoelasticity, and plastic flow as plausible heating mechanisms. This leaves friction and adhesion hysteresis as the only plausible mechanisms. Frictional heating is based on the classical Coulomb friction model, while adhesion hysteresis heating comes from irreversibility in surface adhesion. Adhesion hysteresis only satisfies the experimental observation that heating vanishes for high compressive loading if surface roughness and the instability of surface adhesion is considered.

By understanding the fundamental behavior of a partially closed crack in response to non-uniform loading, and the link between crack surface motion and heat generation, we are one step closer to a fully predictive vibrothermography heat generation model. Future work is needed to extend the crack closure model to a two-dimensional semi-elliptical surface crack and better understand the distinction between frictional and adhesion heating.

CHAPTER 1. OVERVIEW

1.1 Introduction

Nondestructive evaluation (NDE) encompasses a number of measurement techniques aimed at detecting flaws or defects in a manner which maintains the integrity and functionality of the component being inspected. Within the field of NDE, there are a number of measurement modalities including ultrasound, eddy current, radiography, thermography, and visual inspection. Each modality uses either a passive or active form of excitation, and a sensing medium (sound, infrared radiation, x-rays) with which to detect any abnormalities. The detection of flaws requires some form of contrast when compared to the bulk material, or defect-free component. Cracks and voids block and scatter sound waves, alter x-ray absorption, and divert eddy currents. Thermography utilizes heat flow to provide this contrast, whether by active or passive excitation. Passive thermography uses ambient thermal gradients to gather information about defects, while active thermography requires a heat source to provide this thermal gradient. For example, in flash thermography a flash lamp is used to deposit heat onto a component's surface. As heat diffuses into the bulk, any defect that is encountered will alter the heat flow due to a change in thermal conductivity. The impeded heat flow will show up as a bright spot on the thermal image (or dark spot if the imaging system and heat source are on opposite sides of the component).

The characterization of fatigue cracking in structures is a topic that has been studied for decades. Fatigue cracks can propagate quickly due to loads below the yield strength

and result in mechanical failure of the component. This type of failure is especially important in safety critical components, such as those used in aerospace structures. The detection capabilities of most NDE techniques suffer due to the reduced contrast with the bulk material when crack closure is present [7; 66; 65; 44; 10]; sound waves are partially transmitted by contacting surfaces, eddy currents and heat are conducted through asperity contact points, and x-rays will not experience any change in absorption. Vibrothermography is one inspection technique which is particularly suited for the inspection of near-surface, partially closed fatigue cracks. Vibrothermography is a form of active thermography which utilizes a vibration excitation to induce a cyclic stress field at near surface defects. Cracks subjected to vibrations will generate energy in the form of heat due to interference between the opposing surfaces. The nature of this energy conversion is a topic that has been debated for some time [54; 46; 32], although many suppose that heat generation is the outcome of frictional rubbing (supported by the observation that heating occurs at locations along the crack where the surfaces are in tenuous contact [56]).

In contrast to other techniques, research in vibrothermography is at a relatively young stage. This means that there are opportunities to better understand, model, and predict the outcome of an NDE inspection and the physics of heat generation. This paper utilizes the foundation of fracture mechanics to describe the mechanics of a partially closed fatigue crack (like those seen in a vibrothermography test). This topic is of interest to both the NDE community, as well as the fracture mechanics community since the study of the mechanics of a partially closed crack below the crack opening stress is a topic which has seen little research.

Fracture mechanics and failure is a concept which underpins the entire field of NDE. The purpose of an NDE inspection is to detect critical flaws prior to the onset of catastrophic failure. Of course, in order to interpret the results of an NDE inspection and

make informed decisions regarding the continued use of a component, one must have an understanding of the behavior of the expected defect.

The field of fracture mechanics had its start with the pioneering work of Griffith and Irwin [26; 34] who characterized the growth of a crack by a balance between surface energy and elastic energy. The rate at which elastic energy is released by a growing crack, $\mathcal{G}$, must be balanced by the surface energy liberated by the newly created surfaces. The material resistance to fracture is the fracture toughness, $\mathcal{G}_c$, which is independent of the geometry or size of the component. Crack growth will occur when $\mathcal{G} = \mathcal{G}_c$. A related fracture criteria considers the stress fields induced at a sharp crack tip in response to external loading. With an atomically sharp crack, the stress concentration causes the strength of the material to be reduced significantly with respect to its theoretical strength. The stresses induced by this stress concentration can be represented by the stress intensity factor, K. There exists a critical stress intensity factor, K_{Ic}, above which fracture will occur. This is also a measure of fracture toughness.

This foundation, built on linear elasticity, was soon corrected to account for plasticity. In LEFM, the crack tip stress field exhibits a mathematical singularity. This is not physically meaningful, and the stresses near the tip must be bounded by the yield criterion of the material. There is a finite region around the crack tip where the material has yielded and plastically deformed in response to the high stresses. This plastic zone, or process zone, is a contribution to the large fracture toughness of ductile materials when compared with brittle materials. Some fraction of the energy associated with the propagation of a crack goes towards plastically deforming the material, instead of fracturing it.

These analyses were the first attempts to predict the onset of fracture in a quasi-static scenario. The addition of time-dependent loading leads to the concept of fatigue. Fatigue is the study of the weakening of a material in response to cyclic or repeated loading. When a crack is loaded in fatigue, its strength is lower than that for static

loading. The first stage of fatigue crack growth is nucleation and coalescence of voids and lattice defects. After a defect reaches a critical size, it will begin to propagate. The final stages of crack growth are catastrophic and unstable. Between nucleation and unstable fracture, there is a regime of stable growth. The rate of crack growth in this stage was first described by the Paris law [49], which defines an empirical relationship between crack growth rate and the range of applied stress intensity factors.

By the early 1970's, a behavior was observed in which fatigue cracks became partially closed post-fatigue. Elber [68; 14] detected this phenomenon by a "kink" in the load-displacement curve of a cracked body. He focused on the phenomenon termed PICC, in which a fatigue grown crack shuts back on itself due to the residual plastic strains surrounding the crack flanks. When a fatigue crack grows, the residual strains left behind by the stretched material in the plastic zone build up into what is known as the 'plastic wake'. This is a strip of strained material surrounding the crack flanks. When the crack undergoes tensile fatigue loading, the crack remains open over some fraction of the loading cycle, and the crack continues to grow. Once loading is removed, the residual strains cause the crack faces to close and induce a contact stress at the interface.

When partially closed, the tip of a fatigue crack experiences a reduced stress intensity factor. The load required to fully open the crack, σ_{op}, is a measure of the load shielded by the plastic wake. The driving force for fatigue utilized by Paris , $\Delta K = K_{max} - K_{min}$, must be corrected for this shielding effect. Elber defined the reduced stress intensity factor range, $\Delta K_{eff} = K_{max} - K_{op}$, which maintains that only loading above K_{op} contributes to fatigue crack growth. Fleck [18] utilized these concepts, and the use of weight functions, to understand how a partially closed fatigue crack opens in response to external loading. In this work, a closure mechanics model is developed for a one-dimensional crack based on Fleck's work. Closure mechanics, herein, refers to the response of a crack (in terms of forces and displacements) to an applied load.

1.2 Thesis organization

Chapter 2 investigates the application of fracture mechanics to a partially closed, one-dimensional fatigue crack. With the concept of fatigue crack closure, the weight function approach to fracture mechanics, and work on partially closed cracks by Fleck [18], a crack closure model is developed which successfully predicts the effective length of a partially closed crack (the length over which the crack is fully open), the apparent closure stress field acting on the contacting surfaces, and the opening displacement in response to an externally applied load. These predictions are validated with the finite element method (FEM).

In Ch. 3 a number of different heating mechanisms are investigated with respect to vibrothermography crack heating. This work is necessary in light of the range of different mechanisms which have been proposed in recent years. Strain hysteresis, thermoelasticity, plastic flow, adhesion hysteresis, and sliding friction are analyzed quantitatively and qualitatively through theory and compared to experimental observations. Through this analysis, it is clear that only friction and adhesion hysteresis are left as plausible mechanisms.

CHAPTER 2. A ONE-DIMENSIONAL FATIGUE CRACK CLOSURE STATE MODEL

Modified from a paper to be submitted to *International Journal of Fracture*.

Bryan E. Schiefelbein, Tianyu Yu, Ashraf Bastawros, Stephen D. Holland

2.1 Abstract

Crack closure is a phenomenon that arises from plastic deformation during fatigue. The phenomenon was first described by Elber [68; 14] as the partial closing of a crack after all external loading is removed. The following model describes the evolution of closure state in response to remote loading and local closure stresses, representing the influence of the plastic wake left after crack propagation. The fracture mechanics approach described by Fleck [18] is used as a foundation to develop a model in which the evolution of the crack closure state of a fatigue crack, subject to an non-uniform closure stress distribution, is predicted. The model is successfully validated numerically using the finite element method (FEM) for a one-dimensional through crack in a finite width plate. Using the closure state model as a foundation, an inversion algorithm is developed which evaluates the closure stresses given a series of remote loads and corresponding crack opening lengths. The model is also used to predict the redistribution of stresses within the closed region of the crack.

2.2 Introduction

Since the pioneering work of Griffith [26], the scientific community has made great strides in understanding the mechanisms of fracture and fatigue. This great body of work is what allows for the design, manufacture, and safe use of materials and structures in fields with safety critical components such as aerospace, infrastructure, medical, and energy. Fatigue effects plague many of the critical structures and components in these fields; there is no shortage of catastrophic failures whose origin lies in the growth of fatigue cracks.

It was Paris who proposed a law which represents the steady state growth of a fatigue crack [49], allowing for a quantitative prediction of fatigue life, but it was not until the work of Elber [68; 14] that the concept of fatigue crack closure was pursued. Fatigue crack closure is a term which describes the effect in fatigue crack growth where the crack shuts back on itself after remote loading is removed. This closure can be due to plasticity, roughness, or crack wedging due to particles or oxide formation [36; 1; 5].

In this work, we focus on the mechanism known as plasticity induced crack closure (PICC). When a crack is subjected to tension, large stresses are induced near the tip. These stresses cause local yielding in a small zone around the tip, known as the plastic zone or process zone [1]. When a crack is grown under fatigue, there is a cyclic loading and unloading, during which plasticity builds up near the tip. As the crack propagates, it must do so through the previously stretched material. This results in a buildup of stretched material along the crack flanks. When external loading is removed, this residual strain works to shut the crack and can lead to contact between the opposing faces. With closure present, any applied external load will first have to overcome the closure effects to continue to grow the crack [68; 14]. Figure 2.1 shows a fatigue crack exhibiting PICC. When fully open, there is a process zone around the crack tip. When the loading is removed, the residual plastic wake causes the crack to close back on itself.

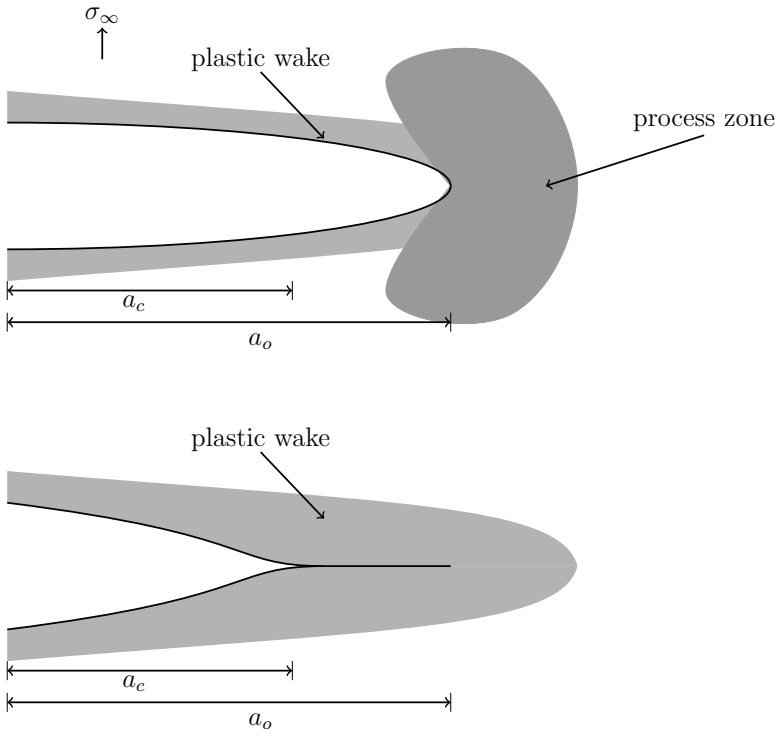

Figure 2.1 Fatigue crack exhibiting PICC under external loading (top) and without external loading (bottom)

Understanding of fatigue crack closure has been used to correct the crack growth rates described by Paris. Below a level, designated S_{op} or K_{op}, the closure effects shield the crack tip from external loading, reducing the ΔK felt by the crack tip and retarding growth [5; 59; 61]. The effective stress intensity range $\Delta K_{eff} = K_{max} - K_{op}$ is the range of loading which acts as the driving force for crack growth.

Despite this vast body of knowledge in fracture and fatigue, little work has been done to describe the mechanics of partially closed, non-propagating fatigue cracks. In a paper by Budiansky and Hutchinson [8] the authors state, "... we skip over the details of the behavior between $K = 0$ and K_{open}, during which the edge of the contact region sweeps in toward the crack tip," stating that it would be "absurd" to try and corroborate the model with experimental results in light of the approximations made by the Dugdale-Barenblatt strip-yield model [13]. While not experimental in nature, this corroboration

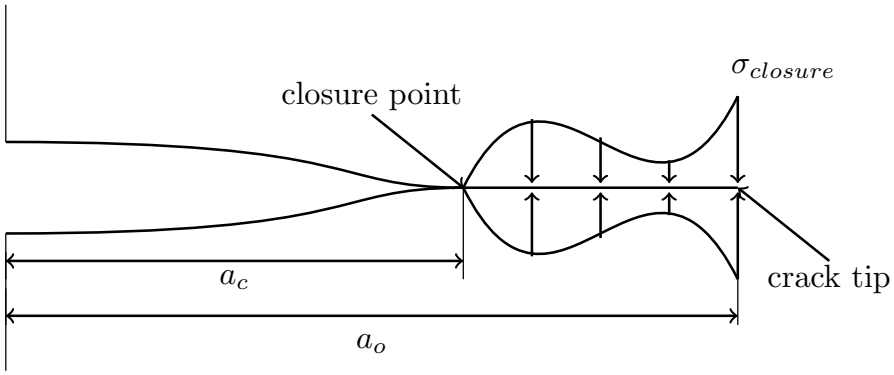

Figure 2.2 Geometry of a partially closed fatigue crack

is precisely what this work seeks to achieve. We develop an analytical model describing the mechanics of a crack between $K = 0$ and $K = K_{op}$ and corroborate it numerically through FEM.

While the closure state of a crack has been historically represented by the term K_{op} (a parameter which is useful in fatigue life predictions), this description is inadequate in understanding the full closure state of the crack. The closure state refers to the set of parameters needed to describe the state of the partially closed crack, namely the point of closure (or effective length), the crack tip displacement, and the stress field in the vicinity of the effective or apparent crack tip. Figure 2.2 shows a simplified schematic of a partially closed crack. The residual plastic zone shown in Fig. 2.1 manifests itself as the closure stresses in Fig. 2.2. For this reference state, with no external tension, the closure point or effective length, designated a_c, is located where the closure stresses begin. This represents the location along the crack where the faces first come into contact.

When subjected to an external tensile load the closure point will shift, and the stress field ahead of the crack tip will no longer be the original closure stress distribution; the apparent closure stress field will be a combination of the original closure stresses and the tensile stress field induced by the remote load. The application of a remote load will

cause the crack faces to open some amount, referred to as the crack tip or crack flank opening displacement.

The effective crack length, or closure point, represents the point (or contour, in the case of a 2D crack) on the crack at which the stress intensity factor due to the remote load and the stress intensity factor due to the closure stresses are in equilibrium [18]. If we imagine this point acts as a virtual crack tip, this can also be thought of as the crack length for which the crack tip singularity vanishes (since the fractured surfaces can support a limited amount of tension). Following the work of Fleck [18], we utilize weight functions to compute the stress intensity factor and crack tip displacements due to the local closure stresses and the remote loading. The closure point, crack displacement, and apparent closure stress field are described in terms of integrals of the product of these weight functions and the remote and closure stress fields.

2.3 Closure mechanics

2.3.1 Background

In LEFM, the stress field ahead of a crack tip is dictated by the stress intensity factor, K. The stress intensity factor is a fracture parameter that depends on the crack geometry and the loading. The general form of K for uniform loading is given by, [1],

$$K = Y\sigma_\infty\sqrt{\pi a}, \tag{2.1}$$

where Y is a correction factor which depends on the crack geometry, sample dimensions, and loading configuration. Expressions for Y are typically evaluated empirically and tabulated for common crack geometries [19]. The remote load is given as σ_∞ and the half crack length as a. Once K is known, the local stresses, strains and the crack opening displacement can be evaluated as a function of the radial position r and angular position θ from the crack tip.

The stress distribution σ_{yy} around the crack tip is given by,

$$\sigma_{yy} = \frac{K_I}{\sqrt{2\pi r}} f(\theta), \tag{2.2}$$

where $f(\theta)$ is an function that describes the angular distribution of the stress around the crack tip. Since we will be primarily interested in the stress field in the plane of the crack, this term will go to unity.

To analyze the role of the presumably non-uniform closure stress field on the crack surfaces, the stress intensity factor is derived from a point load solution. The solution takes the form of a weight function, as proposed by Rice [58] and Bueckner [9]. The weight function is unique in that it is specific to a crack geometry, and is independent of the loading conditions. This means that, if a weight function is known for a crack under the influence of one load, then the weight function is known for the crack under any arbitrary load. The general form of the one-dimensional weight function [20] is given by Eq. 2.3 for a crack of length a,

$$m(x,a) = \frac{2}{\sqrt{2\pi(a-x)}} \left[1 + M_1 \left(1 - \frac{x}{a}\right)^{1/2} + M_2 \left(1 - \frac{x}{a}\right) + M_3 \left(1 - \frac{x}{a}\right)^{3/2} \right], \tag{2.3}$$

where M_1, M_2, and M_3 are geometric correction factors and x is the coordinate position, measured from the center of the crack. If the function $m(x,a)$ is known for a crack, the stress intensity factor is given by,

$$K = \int_0^a \sigma(x) m(x,a) dx, \tag{2.4}$$

where $\sigma(x)$ is the stress distribution acting on the crack faces and a is the half crack length. The weight function represents the stress intensity factor at the crack tip due to a point load, and is directly related to the crack tip stresses induced by a point load (by Eq. 2.2). The coefficients M_1, M_2, and M_3 of the weight function equation are tabulated for a variety of loading configurations and crack geometries [16; 17].

Within the framework of LEFM is the assumption of small-scale yielding (SSY), which implies that plastic yielding at the crack tip is small relative to the elastic stress field and the specimen dimensions. It is advantageous to validate that SSY is indeed valid for the range of parameters we are concerned with. According to Zehnder [69], the SSY condition is valid as long as the plastic zone is less than about 20 % of the crack length.

Consider a semi-circular fatigue crack in a specimen of Nickel super alloy Inconel 718, with half length $a = 2\ mm$ and yield stress $\sigma_Y = 1182\ MPa$. With $K_I = \frac{2}{\pi}\sigma_\infty\sqrt{\pi a}$ and subject to a remote tensile stress of $\sigma_\infty = 240\ MPa$, the plastic zone size is [1],

$$r_p = \frac{1}{\pi}\left(\frac{K_I}{\sigma_Y}\right)^2 \approx 33.42\ \mu m. \tag{2.5}$$

Near the surface, the stress field will behave as plane stress. In other words, $r_p/a \approx 1.7\%$ which is well within the limits of small-scale yielding.

2.3.2 Effective length

Consider a one-dimensional crack of length a_o (Fig. 2.3). There exists a closure stress acting over the region a_c to a_o. With certain regions of the crack open (not in contact) and other regions closed (faces in contact), we can imagine a point at which the crack transitions from open to closed. We will refer to this point as the effective crack length, apparent crack length, or closure point. When an external load is applied, the partially closed crack opens by some increment δa. The stress field ahead of the new closure point (at $a_c + \delta a$) will be a combination of the original closure stresses and the stresses induced by the remote load.

The effective crack length is the point at which the external stresses and closure stresses are in equilibrium. To find this point, we follow a similar procedure to that of Fleck [18]. When modeling the growth of a fatigue crack with crack closure, the closure is represented by the parameter K_{op}. This parameter represents the loading that must

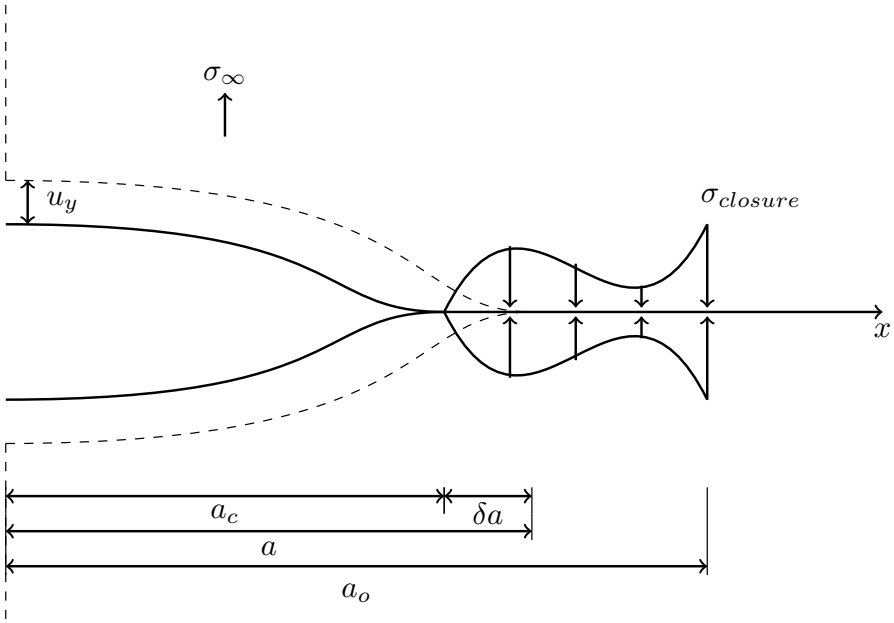

Figure 2.3 Partially closed crack geometry

be applied to overcome the closure effects and fully open the crack. The crack is thought to be fully open when the externally applied K equals K_{op}. In other words, the effective tip reaches the physical crack tip when the quantities K and K_{op} are in equilibrium. Therefore, when we apply an external load and a closure load to the crack, the effective crack tip will be the point at which K due to the remote loading and K due to closure are in equilibrium,

$$K_{op}(a) + K_{cl}(a) = 0. \tag{2.6}$$

This simple reasoning is extremely powerful and allows for a description of the evolution of the closure state of the crack [1]. The stress intensity factors are found using Eq. 2.4,

[1]The reasoning behind this equilibrium condition is rooted in the crack tip singularity. Imagine that a crack of length a_o is partially closed to length a. We can imagine that this really behaves as a "virtual" crack of length a. Since we know that in reality the surfaces cannot support tension, any tension at the tip will cause the crack to open further. Referring to the definition of the crack tip stress field (Eq. 2.2), the crack tip stresses will vanish when the effective stress intensity factor $K_{eff} = K_{op} + K_{cl}$ at the tip is zero.

$$\sigma_\infty \int_0^a m(x,a)dx + \int_{a_c}^a \sigma_{closure}(x)m(x,a)dx = 0. \tag{2.7}$$

Here, σ_∞ is the external loading applied to the crack geometry (in this case we assume uniform tension) and $\sigma_{closure}$ is the closure stress distribution induced by the residual strains surrounding the crack flanks. The right hand term that includes $\sigma_{closure}$ represents the shielding effect that $\sigma_{closure}$ has on a crack tip located at a, while the left hand term represents the driving force on a crack tip located at a due to the external load.

If the crack is loaded by the crack opening stress σ_{op},

$$\sigma_{op} \int_0^{a_o} m(x,a_o)dx + \int_{a_c}^{a_o} \sigma_{closure}(x)m(x,a_o)dx = 0, \tag{2.8}$$

such that the full crack length a_o satisfies the equilibrium equation 2.7.

2.3.3 Opening displacement

The crack flank displacement due to uniform loading for a through crack in an infinite medium is given by [62],

$$u_y = \frac{\kappa+1}{4\mu}\sigma_\infty\sqrt{a^2-x^2}, \tag{2.9}$$

with,

$$\kappa = \begin{cases} 3-4\nu & \text{(plane strain)} \\ \frac{3-\nu}{1+\nu} & \text{(plane stress)} \end{cases}, \tag{2.10}$$

where μ is the shear modulus and ν is the Poisson ratio. Considering displacements very near the crack tip ($x/a \approx 1$), the equation becomes,

$$u_y = \frac{K_I(\kappa+1)}{2\mu}\sqrt{\frac{r}{2\pi}}, \tag{2.11}$$

where the stress intensity factor is due to mode I loading.

The crack tip stress field (Eq. 2.2) induces a displacement field behind the crack tip, given by Eq. 2.11. This displacement represents the opening of the crack faces due to the uniform external load. The residual strain due to PICC will work against this remote load and reduce the displacement of the crack faces. The influence of the closure stress (whose origin is the residual strains) is found through the point load solution. We can imagine that the closure stress will work against the remote load to shut the crack faces. The full displacement field is the superposition of the displacement due to the external loading and the displacement due to the closure stresses.

The definition of the weight function is given in [58],

$$m(x,a) = \frac{E'}{2K_I(a)}\frac{\delta u_y(x,a)}{\delta a}, \tag{2.12}$$

where E' is the effective modulus. For plane stress $E' = E$ and for plane strain $E' = \frac{E}{1-\nu^2}$.

Solving for u_y [37],

$$u_y(x,a) = \frac{1}{E'}\int_x^a K_{eff}(\alpha)m(x,\alpha)d\alpha, \tag{2.13}$$

where $K_{eff} = K_{op} + K_{cl}$ and we have neglected the factor of 2 to obtain the half opening displacement. This equation is only valid near the tip.

2.4 Closure model applications

2.4.1 Closure stress inversion

Imagine a fatigue crack with unknown closure state. If we are able to probe the crack in some way to evaluate the effective length as a function of external load, we

can uniquely determine the closure stress distribution required to produce this opening behavior. First, we use the equilibrium condition from Eq. 2.7, forward substitution, and the assumption that the closure stress field is a series of uniform bins.

Imagine a closure stress distribution which begins at a_c and ends at a_o. It should be obvious from Fig. 2.2 (along with Eq. 2.7) that the effective crack length is a_c for $\sigma_\infty = 0$. Now, a tensile remote load $\sigma_1 = \delta\sigma$ is applied. Under the influence of this load, the crack opens such that the effective crack length will increase from $a_0 = a_c$ to $a_1 = a_c + \delta a$. For simplicity, assume that the closure stress in the region between a_c and $a_c + \delta a$ is constant. To satisfy Eq. 2.7 we must have,

$$\sigma_1 \int_0^{a_1} m(x, a_1)dx + \sigma_{c1} \int_{a_0}^{a_1} m(x, a_1)dx = 0, \tag{2.14}$$

or,

$$\sigma_{c1} = -\sigma_1 \frac{\int_0^{a_1} m(x, a_1)dx}{\int_{a_0}^{a_1} m(x, a_1)dx}. \tag{2.15}$$

An additional tensile remote load $\sigma_2 = \sigma_1 + \delta\sigma$ is applied, such that the effective crack length is measured to be $a_2 = a_1 + \delta a$ ($\delta\sigma$ and δa are not necessarily the same as above). If we also assume the closure stress between a_1 and a_2 is a constant (let us call it σ_{c2}), then equilibrium will be a combination of: i) the remote load $\sigma_2 = \sigma_1 + \delta\sigma$, ii) the closure stress σ_{c1} between a_0 and a_1 (solved for in Eq. 2.15), and iii) the unknown closure stress σ_{c2} between a_1 and a_2. From Eq. 2.7,

$$\sigma_2 \int_0^{a_2} m(x, a_2)dx + \sigma_{c1} \int_{a_0}^{a_1} m(x, a_2)dx + \sigma_{c2} \int_{a_1}^{a_2} m(x, a_2)dx = 0, \tag{2.16}$$

or,

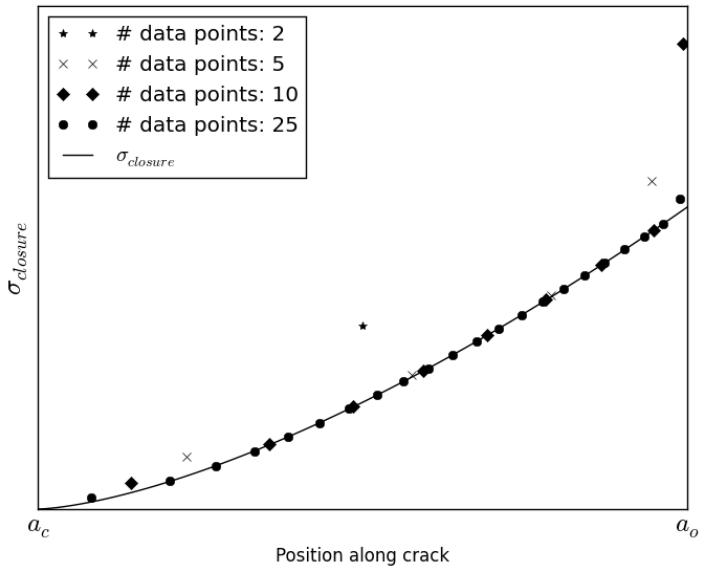

Figure 2.4 Comparison of the closure stress inversion to the assumed closure stress distribution. Predictions are made for 2, 5, 10, and 25 data points (the number of loads at which the effective length is measured).

$$\sigma_{c2} = \frac{-\sigma_2 \int_0^{a_2} m(x, a_2)dx - \sigma_{c1} \int_{a_0}^{a_1} m(x, a_2)dx}{\int_{a_1}^{a_2} m(x, a_2)dx}. \tag{2.17}$$

We can generalize this result for a series of applied external loads σ_i and for a series of constant closure stresses σ_{cj} between a_j and a_{j-1},

$$\sigma_{ci} = \frac{-\sigma_i \int_0^{a_i} m(x, a_i)dx - \sum_{j=1}^{i-1} \sigma_{cj} \int_{a_{j-1}}^{a_j} m(x, a_i)dx}{\int_{a_{i-1}}^{a_i} m(x, a_i)dx}. \tag{2.18}$$

Solving Eq. 2.18 iteratively represents the inversion algorithm to determine an unknown closure stress. It is trivial to show that the inversion algorithm converges to the correct solution as the step size is decreased. Figure 2.4 shows this convergence for a number of different step sizes.

To validate this prediction, the forward model described in Sect. 2.3.2 is used (with an assumed closure stress) to find effective length as a function of remote load. The

remote load and effective length results from the forward model are used with Eq. 2.18 to obtain the original closure stress distribution.

Another method for inverting the closure stress distribution is by noticing that Eq. 2.7 takes the form of a Volterra integral equation of the first kind,

$$-K_{op}(a) = \int_{a_c}^{a} \sigma_{closure}(x) m(x,a) dx, \tag{2.19}$$

where the weight function $m(x,a)$ is the kernel and $\sigma_{closure}$ is the unknown function. It is important to note that this integral equation exhibits a singular kernel, specifically one with a diagonal singularity ($m(x,a) \to \infty$ for $x = a$). To overcome this singularity, we can use a subtraction technique outlined in [2; 51]. If we subtract $m(x,a)\sigma_{closure}(a)$ from the right hand side of Eq. 2.19,

$$-K_{op}(a) = \int_{a_c}^{a} (\sigma_{closure}(x) - \sigma_{closure}(a)) m(x,a) dx + \sigma_{closure}(a) \int_{a_c}^{a} m(x,a) dx, \tag{2.20}$$

where the rightmost term is evaluated numerically with an adaptive Gaussian quadrature which can handle endpoint singularities (such as QAGS) [50]. When $x = a$, the term $\int_{a_c}^{a} (\sigma_{closure}(x) - \sigma_{closure}(a)) m(x,a) dx$ will vanish, effectively eliminating the singularity. This makes it possible to use the Nyström or quadrature method to solve for $\sigma_{closure}(x)$ [12] (see Appendix A.1 for details on the solution to Volterra integral equations).

Figure 2.5 shows the convergence of both methods. The first method (uniform bin method) is recommended as it converges more quickly than the quadrature method.

2.4.2 Apparent closure stress field

In fracture mechanics, the LEFM crack tip stress field solution exhibits a singularity at the crack tip (Eq. 2.2). Since this is not physically possible, corrections are made to bound the stresses. Within this plastic zone or process zone, the crack tip stresses are

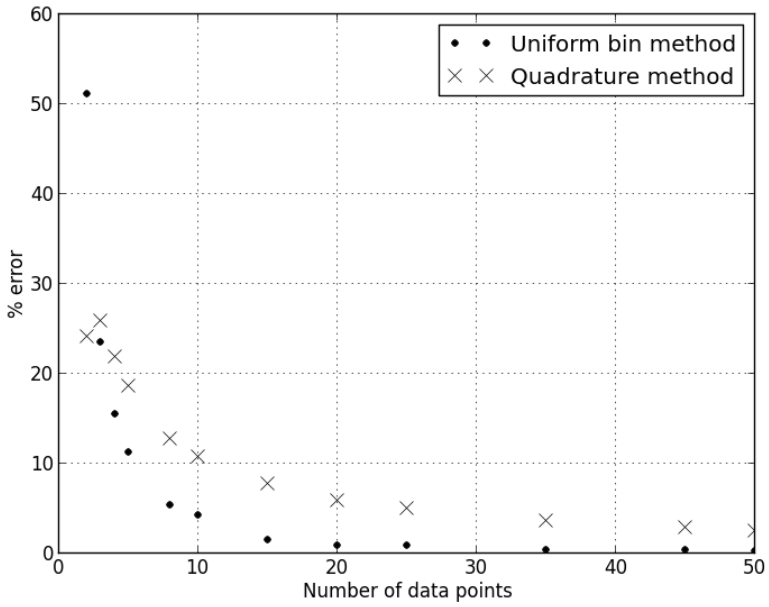

Figure 2.5 Convergence of the closure inversion algorithm for the uniform bin method and the quadrature method.

limited by the material's yield strength. Limiting stresses within this region also leads to a redistribution of stresses outside of this zone [1].

One approach is given by Irwin [1]. Irwin states that the size of the plastic zone r_p (inside of which the stress field is given by the yield strength σ_Y) is found by the equilibrium,

$$r_p \sigma_Y = \int_0^{r_p} \sigma_{yy} dx, \tag{2.21}$$

where σ_{yy} is the crack tip stress field given in Eq. 2.2. The original crack tip stress field is in equilibrium with the redistributed stress field inside of the region bounded by r_p.

2.4.2.1 Approximate redistribution

We propose a similar approach to correcting the apparent closure stress field for the partially closed fatigue crack. A correction is needed due to the fact that, while the crack faces are held in compression, the amount of tensile stress they can support is bounded by the closure stress field.

Consider a partially closed crack of initial length a_c, with a closure stress field present in the region a_c to a_o (as in Fig. 2.3). When the crack is loaded, it opens from a_c to a, relieving the stresses between a_c and a. If we consider the singular stress field of Eq. 2.2, the combined stress field ahead of the effective crack tip at a is,

$$\sigma_{eff} = \sigma_{closure} + \sigma_{yy}, \tag{2.22}$$

where the crack tip stress field approaches infinity at a. Our definition of the effective crack tip (Section 2.3.2), however, states that the singularity vanishes at this point. We should have,

$$\sigma_{eff}(a) = 0. \tag{2.23}$$

Let us define a redistributed crack tip stress field σ'_{yy} which satisfies Eq. 2.23,

$$\sigma'_{yy}(a) = -\sigma_{closure}(a). \tag{2.24}$$

Similar to the approach taken by Irwin, there will be some redistribution ahead of the crack tip (see Fig. 2.7). We require that the stress field be redistributed such that the crack opening stress always remains constant. The crack opening stress is the remote tensile stress required for the crack to open fully, or equivalently for the effective length or closure point to reach the physical crack tip a_o. The crack opening stress is defined in the context of this model (per Eq. 2.7) as,

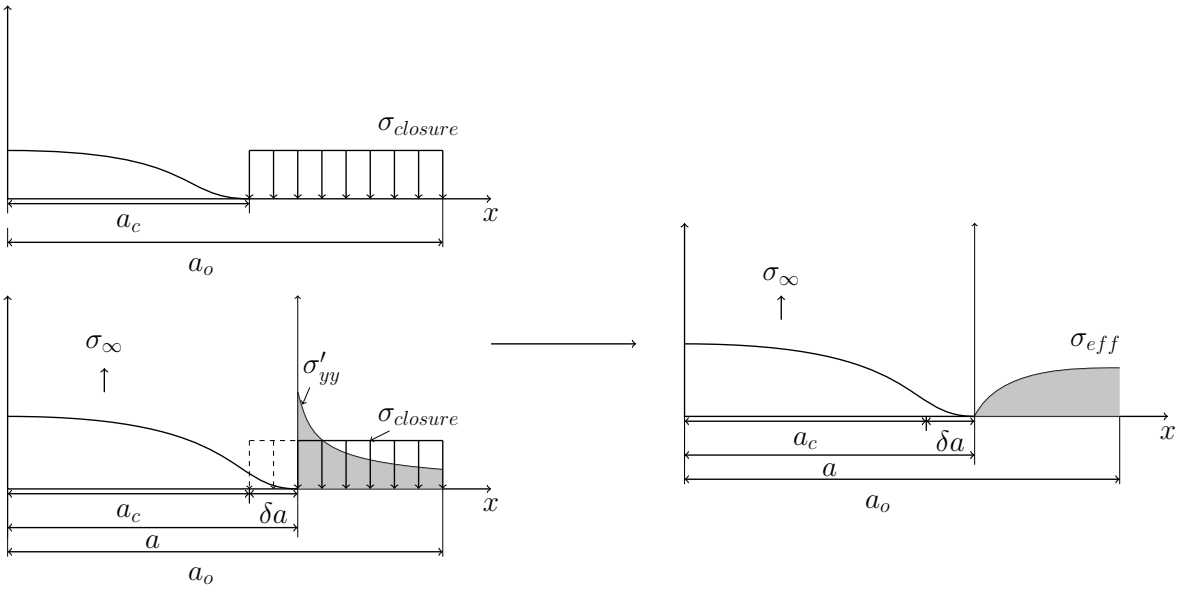

Figure 2.6 Partially closed crack configuration, with apparent crack tip stress field.

$$\sigma_{op}\int_0^{a_o} m(x,a_o)dx + \int_{a_c}^{a_o} \sigma_{closure}(x)m(x,a_o)dx = 0. \tag{2.25}$$

If the crack is loaded by σ_∞, opening it to a, there will be some redistributed field ahead of the effective tip a. If we apply an additional load $\sigma_{op} - \sigma_\infty$, the crack will fully open and we will find that the effective length is a_o,

$$(\sigma_{op} - \sigma_\infty)\int_0^{a_o} m(x,a_o)dx + \int_a^{a_o} \sigma_{closure}(x)m(x,a_o)dx + \int_a^{a_o} \sigma'_{yy}(x)m(x,a_o)dx = 0. \tag{2.26}$$

This equation represents the equilibrium from Eq. 2.7, applied to a crack which has already been opened by a load σ_∞ (taken as the new reference state). Instead of the original closure stress $\sigma_{closure}$ acting on the crack surfaces, there will be some apparent closure stress field $\sigma'_{yy} + \sigma_{closure}$. Here, σ'_{yy} is the tensile stress field induced at the apparent crack tip due to the remote load σ_∞.

Solving for σ_{op} in Eq. 2.25 and substituting it in Eq. 2.26,

$$\int_{a_c}^{a_o} \sigma_{closure}(x) m(x, a_o) dx + \sigma_\infty \int_0^{a_o} m(x, a_o) dx = \int_a^{a_o} \sigma_{eff}(x) m(x, a_o) dx, \tag{2.27}$$

where $\sigma_{eff} = \sigma'_{yy} + \sigma_{closure}$. We can rewrite Eq. 2.27 as,

$$\int_a^{a_o} \sigma'_{yy}(x) m(x, a_o) dx = \int_{a_c}^{a} \sigma_{closure}(x) m(x, a_o) dx + \sigma_\infty \int_0^{a_o} m(x, a_o) dx. \tag{2.28}$$

This equation simply states that, if we apply a load of σ_∞, the crack tip stresses σ'_{yy} must satisfy this relation if σ_{op} is to be held constant irrespective of the order in which loading is applied.

The choice of constraint leading to Eq. 2.28, namely the requirement that σ_{op} is constant, is not the only possible constraint we can choose to enforce. The first requirement, that $\sigma_{eff}(a) = 0$, is a product of the equilibrium statement that the closure point will be located where the stress fields vanish, but the second constraint on σ_{op} stems from the fact that, whether loading is applied in many steps or in a single step, we must end up at the same place. Applying a load in incremental steps requires us to solve Eq. 2.7 sequentially, using the result from the previous step. Since Eq. 2.7 is a function of the stresses applied to the crack surfaces, this leads to an intermediate scenario where $\sigma_{closure}$ is no longer the stress field present at the crack surface. It is $\sigma_{closure}$ combined with some stress field induced by the loading at that particular step.

We propose the following form for this redistributed crack tip stress field,

$$\sigma'_{yy} = \frac{K'_I}{\sqrt{2\pi(r + \Delta)}}, \tag{2.29}$$

where $r = x - a$ is the distance from the effective crack tip. This stress field must satisfy,

1. $\sigma_{eff}(a) = \sigma_{closure}(a) + \sigma'_{yy}(a) = 0$ (Eqs. 2.22 and 2.23)

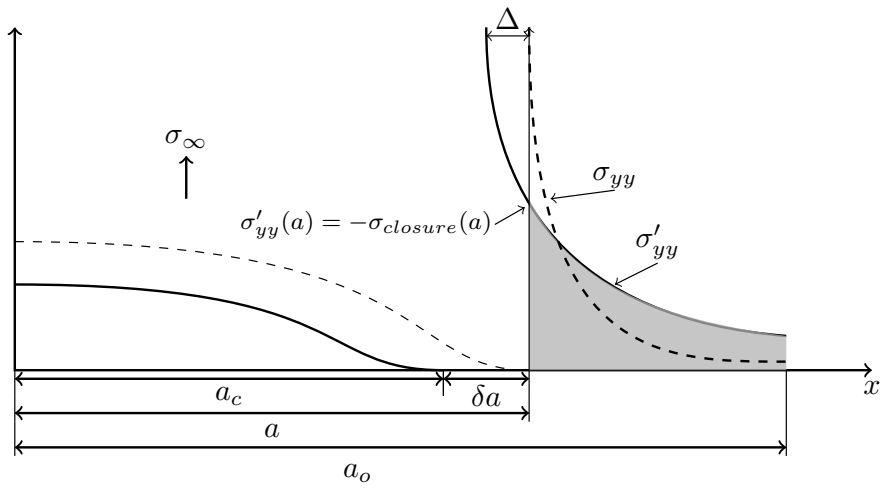

Figure 2.7 Partially closed crack configuration, with effective crack tip stress field.

2. $\int_a^{a_o} \sigma'_{yy}(x) m(x, a_o) dx = \int_{a_c}^{a} \sigma_{closure}(x) m(x, a_o) dx + \sigma_\infty \int_0^{a_o} m(x, a_o) dx$ (Eq. 2.28)

Using the proposed redistributed stress field in Eq. 2.29, condition (1) yields,

$$\Delta = \frac{1}{2\pi} \left(\frac{K'_I}{\sigma_{closure}(a)} \right)^2 . \tag{2.30}$$

Using the result from Eq. 2.30 and the proposed redistributed stress field (Eq. 2.29) along with condition (2),

$$\int_a^{a_o} \frac{K'_I}{\sqrt{2\pi \left(x - a + \frac{1}{2\pi} \left(\frac{K'_I}{\sigma_{closure}(a)} \right)^2 \right)}} m(x, a_o) dx = \int_{a_c}^{a} \sigma_{closure}(x) m(x, a_o) dx + \sigma_\infty \int_0^{a_o} m(x, a_o) dx, \tag{2.31}$$

where Eq. 2.31 is solved numerically. The redistributed stress field is shown qualitatively in Figs. 2.6 and 2.7.

2.4.2.2 Full solution

When considering the redistribution of crack tip stresses for a partially closed crack, we can be more rigorous by noticing that the equilibrium introduced in Sect. 2.3.2 should hold whether the external load is applied all at once or in increments.

Consider the crack which, when subject to an external load of σ_∞, opens from a_c to a. We superimpose an additional load $\delta\sigma$ which opens the crack from a to $a+\delta a$ (as in Fig. 2.7).

If the total load $\sigma_\infty + \delta\sigma$ were applied all at once, the equilibrium condition in Eq. 2.7 states that,

$$\int_0^{a+\delta a} (\sigma_\infty + \delta\sigma) m(x, a+\delta a) dx + \int_{a_c}^{a+\delta a} \sigma_{closure}(x) m(x, a+\delta a) dx = 0. \tag{2.32}$$

If we instead apply σ_∞ and $\delta\sigma$ in successive steps we find that the equilibrium for σ_∞ is,

$$\int_0^{a} \sigma_\infty m(x, a) dx + \int_{a_c}^{a} \sigma_{closure}(x) m(x, a) dx = 0. \tag{2.33}$$

Next we apply the load $\delta\sigma$,

$$\delta\sigma \int_0^{a+\delta a} m(x, a+\delta a) dx + \int_{a}^{a+\delta a} \sigma_{eff}(x) m(x, a+\delta a) dx = 0, \tag{2.34}$$

where $\sigma_{eff}(x)$ is the combined stress field ahead of the crack tip. To enforce uniqueness in the application of loading, we should find that solving Eqs. 2.32 and 2.34 both produce the same incremental crack opening δa. Using the definition $\sigma_{eff} = \sigma'_{yy} + \sigma_{closure}$ and solving for $\delta\sigma$ in Eq. 2.32, Eq. 2.34 becomes,

$$\int_a^{a+\delta a} \sigma'_{yy} m(x, a+\delta a)dx = \sigma_\infty \int_0^{a+\delta a} m(x, a+\delta a)dx + \int_{a_c}^{a} \sigma_{closure}(x) m(x, a+\delta a)dx. \tag{2.35}$$

Since 2.35 has the form of a Volterra integral equation, we can solve for σ'_{yy} the same way we solved for the unknown closure stress distribution in Sect. 2.4.1. For the uniform bin method,

$$\sigma'_{yy}(a_i) = \frac{\sigma_\infty \int_0^{a_i} m(x, a_i)dx + \int_{a_c}^{a} \sigma_{closure}(x) m(x, a_i)dx - \sum_{j=1}^{i-1} \sigma'_{yy}(a_j) \int_{a_{j-1}}^{a_j} m(x, a_i)dx}{\int_{a_{i-1}}^{a_i} m(x, a_i)dx}, \tag{2.36}$$

where $a = a_0$ is the effective crack length and a_i varies from a to a_o. It may be apparent that this full solution is in fact analogous to the approximate solution, except it is solved in steps of length δa.

In the approximate solution, we made the observation that applying either σ_{op}, or σ_∞ and then $\sigma_{op} - \sigma_\infty$ will cause the crack to fully open from a_c to a_o. We know that after applying σ_∞ the stress field present on the crack surface is no longer $\sigma_{closure}$. We use the proposed form of σ'_{yy} to define the new apparent closure stress field on the crack surface.

With the more rigorous solution, we are decreasing this step size and are instead stating that applying either $\sigma_\infty + \delta\sigma$, or σ_∞ and then $\delta\sigma$ will cause the crack to open from a_c to $a + \delta a$. Since we are not assuming any global form for the apparent closure stress field (as we did with σ'_{yy} in the approximate solution), we instead assume that the stress field σ'_{yy} is constant between a and $a + \delta a$. This allows us to solve for the apparent closure stress field incrementally, using forward substitution of the results from the previous steps.

The full solution is identical to Eq. 2.31 if $a + \delta a = a_o$ and if the proposed form of σ'_{yy} is used (Eq. 2.29).

2.5 Validation and discussion

The preceding fracture mechanics based model builds off of the work of N. A. Fleck [18]. The most important aspect of Fleck's work which we borrow is the idea that when a partially closed crack is loaded by a remote load and closure stresses, the equilibrium point along the crack is where the effective stress intensity factor is zero. This is not unlike the pioneering work by Elber on crack closure during fatigue crack growth, where the crack is considered to be fully open when the stress intensity factor due to remote loading is equal to some threshold stress intensity factor K_{op}.

The model is validated in three steps: i) validation of the effective crack length, ii) validation of the opening displacement prediction, and iii) validation of the proposed apparent stress field. The model is validated with numerical simulation using the finite element method package Abaqus/CAE. A through crack with known geometry and closure stress is simulated.

2.5.1 FEM model

A 2D planar, deformable body is used to simulate the cracked plate. The material model is restricted to elastic to more closely follow the analytical model. The crack itself is modeled as a seam with a frictionless contact condition preventing it from overlapping. The remote load is applied as a surface traction on the outer faces, while the closure stresses are crack surface tractions of the form,

$$\sigma(x) = \sigma_c \left(\frac{x - a_c}{a_o - a_c} \right)^n . \tag{2.37}$$

The form of the closure stress is chosen based on work by Davidson [11] and Fleck [18], and on predictions made by the Dugdale strip-yield model [13; 48; 47], both of which find monotonically increasing closure stresses.

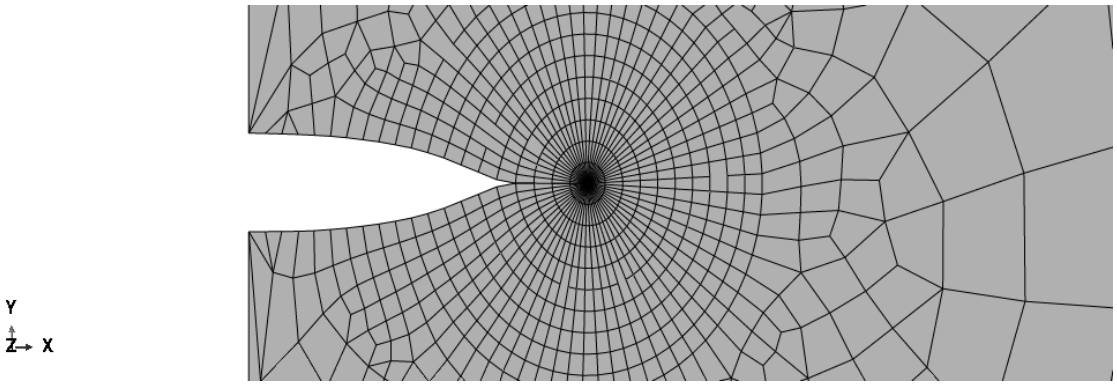

Figure 2.8 Finite element mesh showing the focused mesh around the physical crack tip

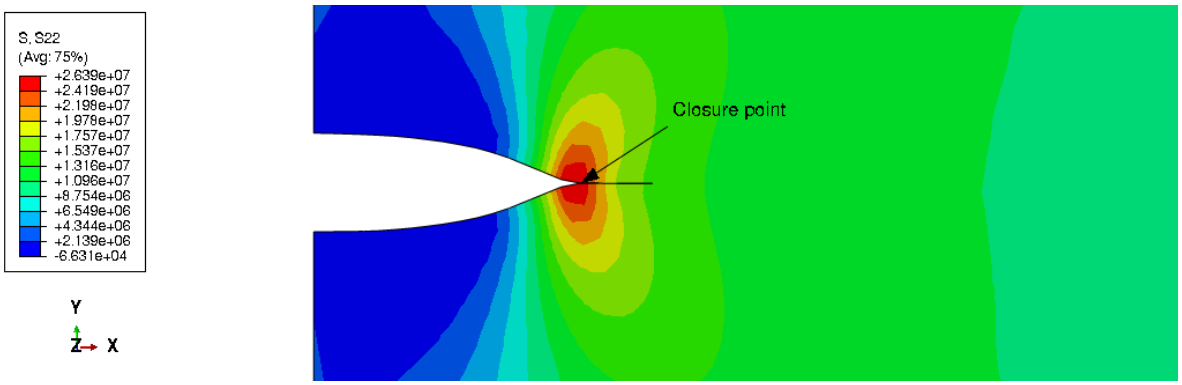

Figure 2.9 Stress distribution for a through crack

Near the crack tip a focused mesh is used, with a collapsed element at the tip to simulate the $1/\sqrt{r}$ singularity (Fig. 2.8).

Figure 2.9 shows the crack tip stresses, σ_{yy}, for a single remote load. The closure point has propagated towards the physical crack tip under the influence an opening load. It appears that the stress "singularity" (although we know it is not mathematically singular) is located at the closure point, which corroborates the hypothesis that the closure point acts as a virtual crack tip.

Furthermore, it appears that the singularity is located slightly behind the closure point, which would corroborate the proposed form of the redistributed stress field in Section 2.4.2 (the shift manifests as Δ in Eq. 2.29).

2.5.2 Effective crack tip

With a simple equilibrium statement and the utilization of weight functions, we are able to describe the evolution of the effective crack length in response to a non-uniform closure stress and remote tensile or bending load by Eq. 2.7. This is a powerful result which is fundamental to developing the closure stress inversion algorithm (Sect. 2.4.1) and the apparent stress field formulation (Sect. 2.4.2). For the through crack considered here, the weight function (Eq. 2.3) coefficients are given in [19].

Using the assumed closure stress field (Eq. 2.37), Eq. 2.7 is solved for a range of external loads σ_∞. For each external load, there is a unique crack length a which satisfies the equilibrium of Eq. 2.7. The integrals are computed numerically using an adaptive Gaussian quadrature algorithm, such as QAGS [50]. A root finding algorithm, such as Brent's method or bisection [51; 4], is used to find the optimal value of a.

The externally applied load is plotted against the effective crack tip in Fig. 2.10. The numerical results are compared to the model prediction for three different closure stress distributions: i) uniform closure stress, ii) linear closure stress, and iii) quadratic closure stress. The closure point is normalized by the physical crack length. A closure point of one corresponds to a fully opened crack. In these results, the closure stress distribution starts at $a_c/a_o = 0.5$.

From Fig. 2.10, we see that FEM and the closure model agree within the expected error (the error bars associated with the numerical results arise from the finite resolution mesh). In the finite element model, the closure point is chosen as the first mesh node where the contact stress is nonzero.

This simple but effective method for determining the equilibrium between the remote load and closure stresses makes it possible to quantitatively determine the evolution of the partially closed crack tip. As discussed in Sect. 2.4.1, this statement of equilibrium leads directly to an inversion algorithm which can be used solve for the closure stresses if the effective crack length is measured as a function of the external load. This is a

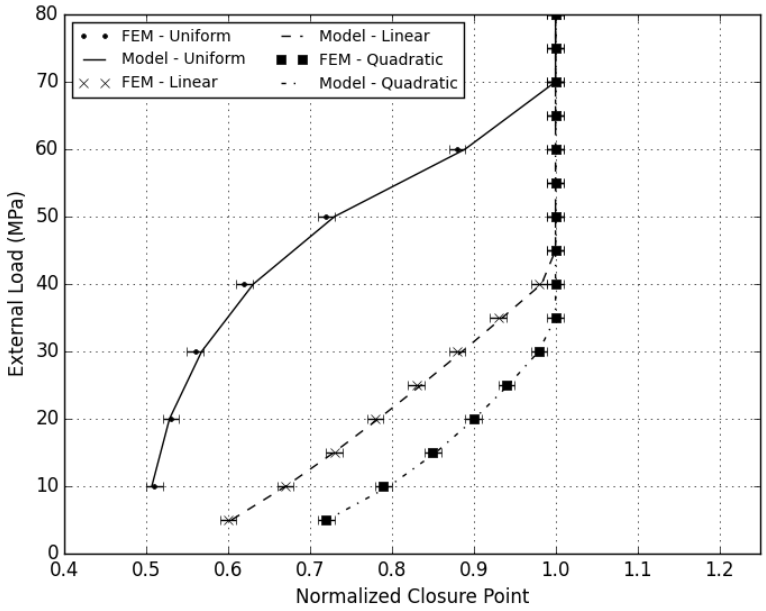

Figure 2.10 Externally applied load vs. equilibrium crack tip position for uniform, linear, and quadratic closure stress distributions.

powerful algorithm, as no prior assumptions are required to evaluate the stresses, aside from an estimate of the crack geometry.

2.5.3 Crack opening displacement

Since the closure stresses are non-uniform, we use the weight function approach to represent the total crack opening displacement. The crack opening displacement is found directly from the definition of the weight function [58], which by definition is only valid close to the tip. For this reason, when validating the results against FEM, we only consider the displacements close to the crack tip.

Figures 2.11, 2.12, and 2.13 show the crack opening displacement at the two evaluation points as a function of the applied remote load. The displacement is evaluated at $0.8a$ and $0.9a$, where a is the effective crack length or closure point at that particular load

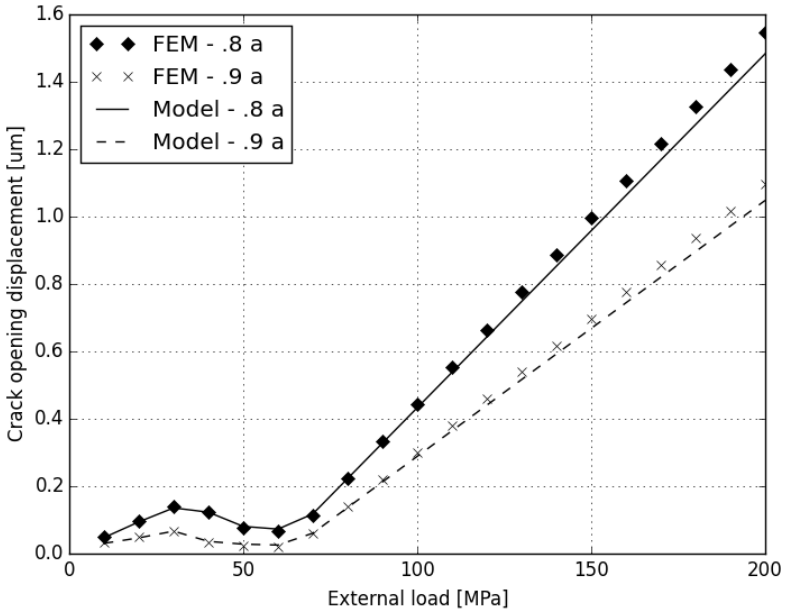

Figure 2.11 COD versus remote load (uniform closure stress)

step. These points are chosen based on the limitations of the crack opening displacement equation, which is only valid near the crack tip.

Comparing these results to Fig. 2.10, we see that the opening displacement calculations are accurate even after the closure point has fully saturated. With these results, we are confident that the partially closed crack can be described in terms of its effective crack length and crack opening displacement.

2.5.4 Apparent closure stress field

After determining the effective crack length and opening displacement, we wish to represent the apparent contact stresses which remain in the partially closed region of the crack. We have proposed a redistributed crack tip stress field. This stress field accounts for the fact that crack faces can only support stresses equal to that of the closure stress.

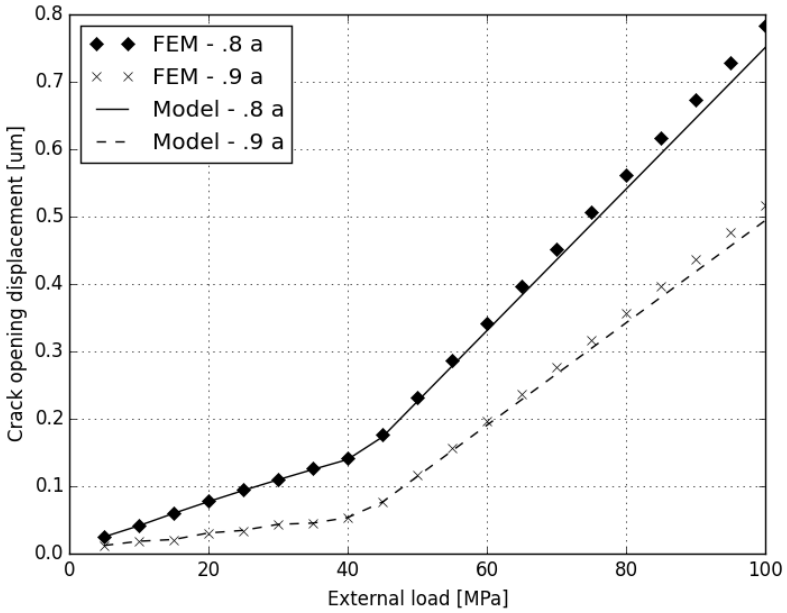

Figure 2.12 COD versus remote load (linear ramp closure stress)

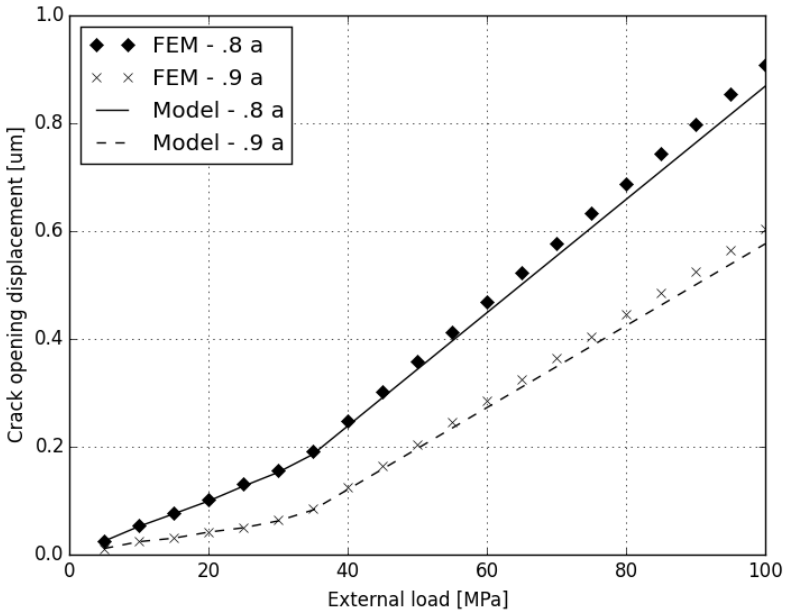

Figure 2.13 COD versus remote load (quadratic closure stress)

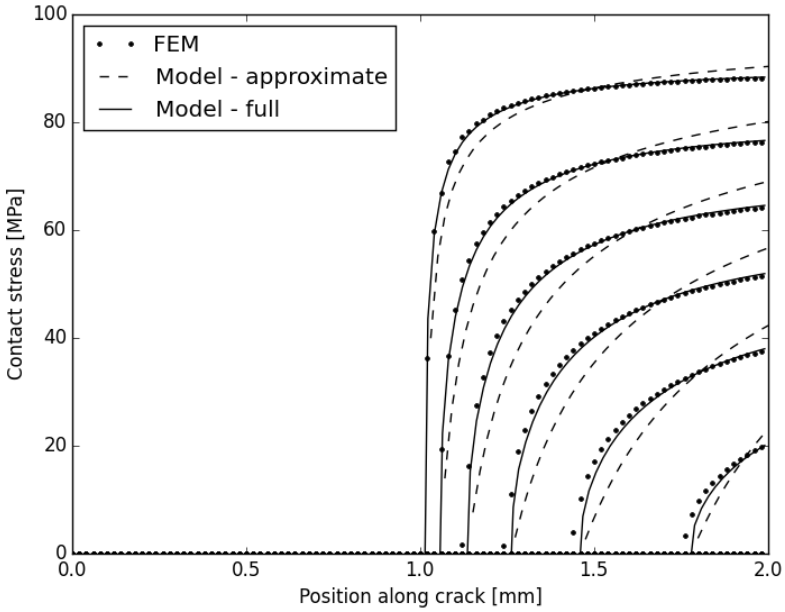

Figure 2.14 Contact stress vs. location along crack (uniform closure stress). From left to right, $\sigma_\infty = 10, 20, 30, 40, 50, 60\ MPa$.

We employ two methods: i) an approximate model which takes the form of the elastic stress field and ii) a more rigorous model which considers the compatibility of loading.

For the approximate model two constraints are proposed: i) the apparent stress field at the effective crack tip is zero and ii) the redistributed stress field produces the same crack opening stress.

For the more rigorous treatment, the apparent stress field is chosen such that loading a single load step versus a series of small load steps produces the same result (see Sect. 2.4.2 for more details).

The FEM results in Figs. 2.14, 2.15, and 2.16 are predicted very accurately by the full apparent stress model. The approximate model does capture the overall characteristics of the stress field, but is not as accurate.

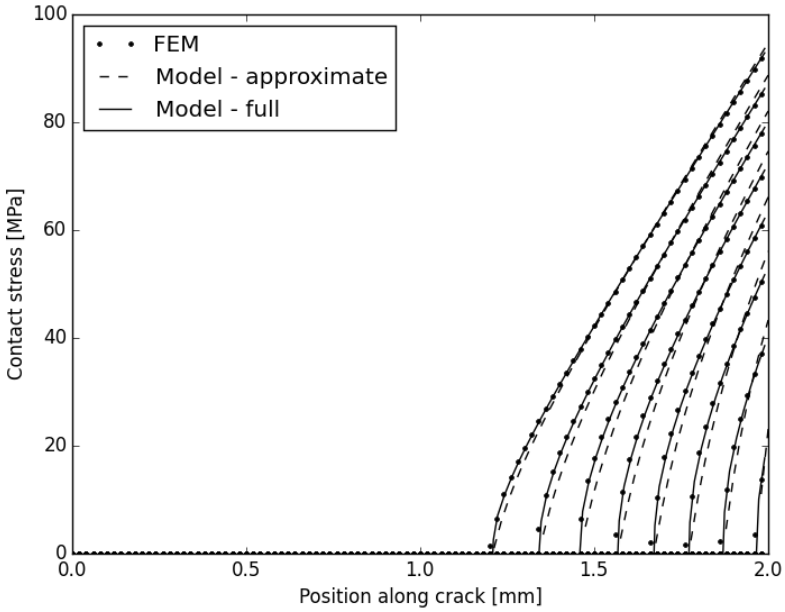

Figure 2.15 Contact stress vs. location along crack (linear closure stress). From left to right, $\sigma_\infty = 5, 10, 15, 20, 25, 30, 35, 40\ MPa$.

2.6 Conclusion

The model in this work describes the mechanics of a partially closed, non-propagating fatigue crack. This is an area of research which has had little attention, but is crucial in understanding the mechanics of a fatigue crack after growth. The model is based on Fleck's work [18] and outlines a framework for describing the evolution of the closure state of a surface crack under external loading and non-uniform closure stresses.

The equilibrium criteria of the crack takes into account the closure and remote stress fields. The equilibrium point under given loading corresponds to the point where there is a transition from fully opened to fully closed, at which the stress intensity factors are in equilibrium. The crack opening is determined by the superposition of the displacement due to the remote load and the displacement due to the closure stress, evaluated using the weight function point load solution. When compared with FEM, the one dimensional

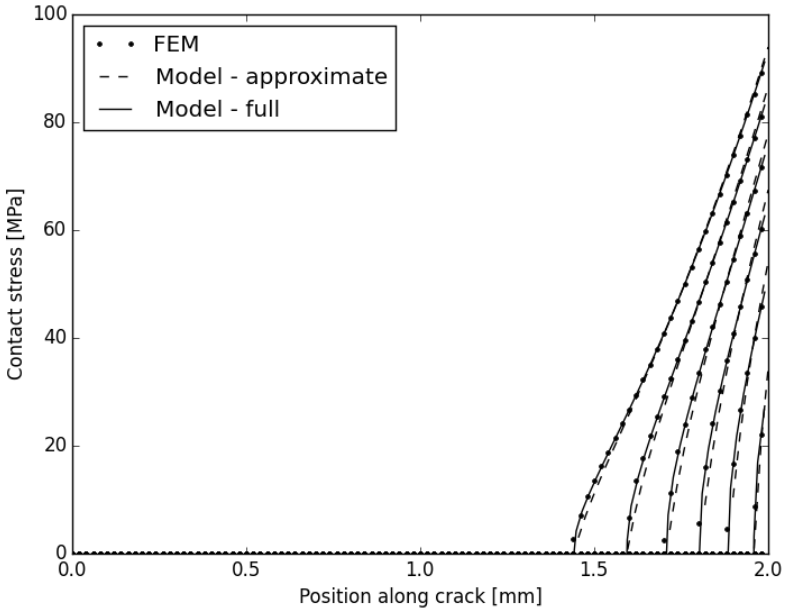

Figure 2.16 Contact stress vs. location along crack (quadratic closure stress). From left to right, $\sigma_\infty = 5, 10, 15, 20, 25, 30\ MPa$.

through crack predictions are within the expected FEM error. This is true for both the closure point and opening displacement predictions. The agreement between the model and simulation gives us confidence that the model correctly captures the mechanisms governing closure and its evolution.

An inversion algorithm is developed based on the equilibrium criteria to evaluate an unknown closure stress based on measurement of the effective crack length in response to a known remote load. Finally, a redistributed crack tip stress field is proposed which takes into account the fact that the crack surfaces are already fractured. The tensile load that can be supported is bounded by the closure stress holding the surfaces together. Both an approximate model and rigorous model are introduced. The approximate model performs well near the crack tip, while the full model agrees with FEM over the entire crack.

2.7 Acknowledgment

This material is based on work supported by the Air Force Research Laboratory Contract #FA8650-10D-5210, Task Order #023 and performed at Iowa State University.

CHAPTER 3. A QUANTITATIVE AND QUALITATIVE ANALYSIS OF VIBROTHERMOGRAPHY CRACK HEATING MECHANISMS

Modified from a paper to be submitted to *NDT&E International*.

Bryan E. Schiefelbein, Tyler Lesthaeghe, Stephen D. Holland, Ashraf Bastawros

3.1 Abstract

Vibrothermography is a form of NDE that utilizes vibration induced crack heating for defect detection. Currently, there is a lack of consensus regarding the mechanisms that lead to crack heating [31]. In this work, experimental observations and theory are used to evaluate a list of plausible heating mechanisms, each of which have been suggested in literature to be the dominant cause of vibrothermography heating. This work concludes that adhesion/friction is at the root of crack heating, although the distinction between the two is not clear at a fundamental level. This work provides evidence against plastic flow, thermoelasticity, and linear absorption as dominant heating mechanisms. More work is needed to understand the role of friction and adhesion, with the goal of better predicting heat generation in real NDE applications.

3.2 Introduction

Vibrothermography or Sonic IR is a form of NDE in which mechanical vibrations are converted to heat at crack surfaces. Vibrothermography was first pioneered by Ed

Henneke in the late 1970's [28; 53], and has since gained popularity as a fast, full-field inspection of metal or composite parts [15]. Specimens are vibrated with an ultrasonic transducer, and any near surface cracks experience mechanical vibrations. These vibrations lead to crack face interference and heat generation. This heat generation is measured by a thermal imaging system and used to quantify possible defects [28; 15; 31].

Most NDE techniques rely on material discontinuities or inhomogeneities to detect possible defects. Vibrothermography is unique in that it works best when there is some crack closure, a phenomenon that hinders detection in other NDE modalities. The probability of detection and quantification of the location and size of defects in vibrothermography is directly related to the magnitude and location of heating at the crack surfaces. In order to better understand and predict this phenomenon, we take a two-fold approach: i) understand the evolution of crack closure state in response to external loading [40; 55; 60] and ii) understand how crack flank interference and contact is tied into the conversion of vibrational energy to heat. The latter is investigated in this work.

Currently, the understanding of vibrothermography heating mechanisms is limited. Many researchers simply label heating as friction induced [41; 57; 15; 70], implying tangential sliding friction between the crack flanks. Other researchers have suggested plastic flow or thermo-plastic heat generation at the crack flanks or near the crack tip [42; 32; 29]. Others yet have suggested that thermoelasticity [33], viscoelasticity and anelasticity [54; 33; 45], or a combination of mechanisms [31; 54] play a role in crack heating. In this paper, we analyze a number of these heating mechanisms through experimental observations and theory.

3.3 Background

To investigate the possible heating mechanisms, we take experimentally observed phenomena and construct a set of criteria with which to evaluate them. These experimental

observations are derived from work by Holland et al. [56; 31; 39]. The experimental observations are as follows:

1. Cracks heat at the closure point, not at the crack tips.
2. Heating does not occur when there is: i) no closure and ii) too much closure.
3. Linear dependence of heat generation of excitation frequency [70].
4. Power law dependence on vibrational strain amplitude between one and two [31].
5. Little dependence on crack opening mode [70].

Expressed mathematically,

1, 2	$\lim_{P\to 0} Q = 0 \quad \text{and} \quad \lim_{P\to\infty} Q = 0.$
3	$Q \propto f.$
4	$Q \propto \epsilon^n \quad 1 \geq n \geq 2.$
5	$Q_{Mode\ I} \approx Q_{Mode\ II} \approx Q_{Mode\ III}.$

where Q is the thermal power generated, f is the excitation frequency, σ_∞ is the applied load, and ϵ is the local strain amplitude at the crack surface. Statements 1 and 2 are important, as they illustrate the closure conditions needed for vibrothermography heating. In reality, heat generation follows a pattern as shown in Fig. 3.1, where heating occurs at locations of tenuous contact between the opposing crack faces.

3.4 Heating mechanisms

The heating mechanisms investigated in this work are:

- Sliding friction

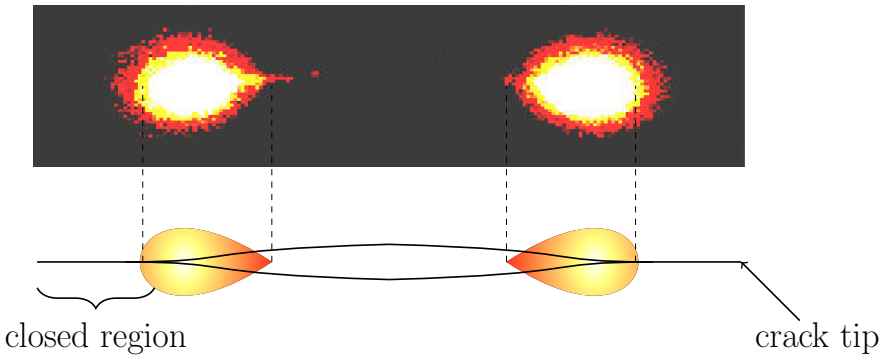

Figure 3.1 Heating zones are located at regions of tenuous crack flank contact.

- Adhesion hysteresis
- Plastic flow
- Linear absorption
- Thermoelastic absorption

Sliding friction: Energy is dissipated as heat when two surfaces are moved tangentially relative to one another. The adhesion and deformation of asperities resist motion and dissipate energy.

Adhesion hysteresis: When two surfaces come into close contact, it is sometimes energetically favorable for them to adhere and destroy the original surfaces. The surfaces can be pulled apart, from which there is a net energy loss.

Plastic flow: Plasticity is an inherently irreversible process. When a material is plastically deformed, there is heat generation [52].

Linear absorption: When a material is deformed elastically it is implied to be fully reversible, when in reality there is hysteresis. This phenomenon is also known as anelasticity, viscoelasticity, bulk hysteresis, or internal friction.

Thermoelastic absorption: The mechanical and thermal states of a material are coupled through the thermoelastic effect. When a material is loaded cyclically, there is a

thermal loading cycle as well. If there is a phase lag between the two cycles, there will be a irreversible heat loss to the surroundings.

3.4.1 Linear absorption (strain hysteresis)

Elastic deformations are typically assumed to be entirely reversible. In reality, even when a material is in the elastic regime, some fraction of the deformation is irreversible. This phenomenon is referred to as strain hysteresis, bulk hysteresis, linear absorption, viscoelasticity, or anelasticity and simply represents a deviation from Hooke's law and an irreversible loss of energy during elastic deformation [67; 38; 35; 23].

In dynamic loading, there is energy loss (possibly as heat) in each cycle of compression and relaxation. The linear absorption phenomenon is analyzed using the Hertzian contact model [35]. This model assumes two smooth, elastic spheres in contact at a point, neglecting adhesion. As a load P is applied to the two spheres of radii R_1 and R_2, with corresponding elastic modulii E_1 and E_2, they will experience elastic deformation,

$$\delta = \left(\frac{9}{16R^*(E^*)^2}\right)^{1/3} P^{2/3}, \tag{3.4.1}$$

where delta is measured at a remote point on the sphere, $R^* = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)^{-1}$ is the effective radius, $E^* = \left(\frac{(1-\nu_1)^2}{E_1} + \frac{(1-\nu_2)^2}{E_2}\right)^{-1}$ is the effective modulus, and P is the load applied to both spheres.

Strain hysteresis is the fraction of strain energy lost during elastic deformation (represented by the hysteresis loss factor). In order to quantify hysteresis in the asperity compression cycle there must first be a description of the total energy in the cycle. The work performed by the load on the spheres is,

$$U_e = \int P d\delta. \tag{3.4.2}$$

Using the load-displacement relation given in Eq. 3.4.1 and a change of variables from $d\delta$ to dP, we find that,

$$U_e = \frac{2}{5}\left(\frac{9}{16R^*(E^*)^2}\right)^{1/3} P^{5/3}. \tag{3.4.3}$$

This reveals a five-thirds dependence between the strain energy and the applied load, which is in line with observation 3. However, consider a static load P and a dynamic load dP. The deformation is,

$$\delta = \left(\frac{9}{16R^*(E^*)^2}\right)^{1/3} (P + dP)^{2/3}. \tag{3.4.4}$$

Using Eq. 3.4.2,

$$dU_e = \frac{2}{5}\left(\frac{9}{16R^*(E^*)^2}\right)^{1/3} P^{5/3}\left[(1 + \frac{dP}{P})^{5/3} - 1\right]. \tag{3.4.5}$$

Using L'Hospital's rule, we find that as P goes to infinity, so does the change in strain energy. This is in direct conflict with observation 2. We can analyze this problem in another way by considering the macroscopic contact of the surfaces. In section 3.4.4 we will show that for randomly rough surfaces, even when contact is dominated by elasticity, the load-displacement relationship can be proportional $\delta \propto P$. With this assumption, the strain energy from P to $P + dP$ according to Eq. 3.4.2 is,

$$dU_e \propto 2PdP + dP^2. \tag{3.4.6}$$

Once again, we find that an increasing preload leads to a monotonically increasing strain energy, inconsistent with observation 2.

3.4.2 Plastic flow

The energy loss due to plastic flow is rooted in irreversible deformations of the material and has been suggested as a possible source of heating in vibrothermography [42; 27; 32; 29]. The location where this plastic flow occurs will be primarily at i) the crack flanks in tight closure and ii) the crack tips. The first case will occur when there is sufficient crack closure where the contacting surfaces are primarily in a state of plastic flow. Since the material is assumed to have yielded, any additional load will contribute to the plastic deformation. This load-displacement relationship will be linear, such that the stress supported by the surface is the yield stress. This local yielding defines the true contact area,

$$A_{contact} = \frac{P}{c\sigma_y}, \tag{3.4.7}$$

where c is a constant that ranges between $1.1 < c < 3$ and σ_y is the material yield stress [6]. The energy dissipated due to plasticity is found by integrating the load times the incremental displacement (as with Eq. 3.4.2),

$$U_p = \int P d\delta = \frac{1}{\pi R c \sigma_y} \int P dP. \tag{3.4.8}$$

The energy dissipation due to plasticity between loads P and $P + dP$ is,

$$dU_p = \frac{1}{\pi R c \sigma_y} \left[(P + dP)^2 - P^2 \right] = \frac{2PdP + dP^2}{\pi R c \sigma_y}. \tag{3.4.9}$$

Note that as $P \to \infty$, $dU_p \to \infty$. This implies that tightly closed regions will generate more heat (a contradiction of observation 2).

The second case will occur when the loading is large enough that there is yielding at the crack tips. This is the precursor to crack growth and represents permanent damage to the material. This heating mechanism is implausible for two reasons: i) heating is not

observed at the crack tips (in conflict with observation 1) and ii) the vibration amplitude is kept low enough that it does not induce crack growth.

3.4.3 Thermoelastic absorption

Thermoelasticity is a phenomenon that describes the coupling of the mechanical and thermal states of a material. This phenomenon in solids is analogous to the behavior of a gas; compress the gas and it heats up, allow it to expand and it cools down. When an asperity is brought into contact with another asperity, or a flat surface, it is compressed. The compression causes a localized increase in temperature, while the release of pressure decreases the temperature. This process is mostly reversible, but there is an irreversible component where energy is lost to the surroundings.

Goodman et al. [21] worked through a detailed analysis of energy dissipation in the thermoelastic cycling of an elastic solid. When the internal stresses are inhomogeneous, it leads to temperature gradients and irreversible heat flow. In a thermodynamic system, this increases the entropy of the system and leads to irreversible energy loss. In the case of a dynamic stress, the asperity stress state is cyclic. The same can be said of the thermal state; the asperity heats up and cools down during each cycle, causing energy to flow in and out of the system. The buildup of heat will be caused by the phase lag between the cyclic stress and the cyclic heat flow in and out of the asperity volume.

The foundation of thermoelasticity is the coupled partial differential equation linking heat flow and the strain field [21],

$$k\nabla^2 T = \rho c_p \frac{\partial T}{\partial t} + (3\lambda + 2\mu)\alpha T \frac{\partial \epsilon}{\partial t}, \tag{3.4.10}$$

where k is the thermal conductivity, α is the coefficient of thermal expansion, λ is Lame's constant, μ is the shear modulus, T is temperature, ρ is density, c_p is the specific heat capacity, and ϵ is strain. This equation can be linearized by decomposing the temperature into the sum of the ambient temperature and the deviation from ambient,

$$T = T_o + \theta, \tag{3.4.11}$$

where we assume that the deviation from ambient temperature, θ, is small. The strain is decomposed into the dilatational strain, ϵ", and thermal strain, ϵ',

$$\epsilon = \epsilon'(\theta) + \epsilon''. \tag{3.4.12}$$

Plugging the decomposed temperature and strain into Eq. 3.4.10,

$$k\nabla^2\theta = \rho c_p \frac{\partial \theta}{\partial t} + (3\lambda + 2\mu)\alpha T_o \frac{\partial \epsilon''}{\partial t}. \tag{3.4.13}$$

Goodman et al. show extensively that this linearization technique is a good approximation to the 'exact' solution to the nonlinear system. The only disclaimer is that, in the case of the linearized system, the work done on the system and the heat flow out of the system are not equivalent. In reality, they both represent the energy loss in a cycle due to irreversible heat flow and should be identical.

To apply this analysis to the case of vibrothermographic crack heating, we investigate the energy dissipation in a single asperity contact, subject to a cyclic load. First, we assume some shape for the asperity. We expect the cross sectional area to be small at the tip of the asperity, and increase with depth,

$$A = A_o e^{\frac{z}{c}}, \tag{3.4.14}$$

where A_o is the area at the tip of the asperity ($z = 0$), and c is a geometric constant controlling the rate of area increase as z increases. Imagine a normal force applied to the tip of the asperity, with the form,

$$F(t) = F_A e^{-i\omega t}, \tag{3.4.15}$$

where F_A is the amplitude of the stress and ω is the frequency of loading. As a direct comparison to a typical vibrothermography experiment, we use an exciting frequency of $f \approx 20kHz$, or $\omega \approx 3183rad/s$. At this frequency, the vibration wavelength is, to a first approximation, $\lambda = \sqrt{\frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)}} f^{-1} \approx 32cm$. The wavelength is on the order of centimeters, while our asperity height is on the order of microns. This allows us to make the assumption that the forces in the asperity are uniform and unchanging with depth.

The dilatational strain is given by (see Appendix A.2),

$$\epsilon'' = \frac{\Delta V}{V} = -2\nu\epsilon_z. \tag{3.4.16}$$

Let us assume that temperature rise at the tip of the asperity is large, and that it decays into the bulk,

$$\theta(z,t) = T_1 e^{-i\omega t} e^{-\frac{z}{k_3}}, \tag{3.4.17}$$

where k_3 represents the thermal decay length, and again should be on the order of nanometers or microns. The coefficient T_1 represents the temperature at the tip of the asperity at time $t = 0$. By specifying the form of θ with Eq. 3.4.17, we are restricting Eq. 3.4.13 to a particular solution. If we assume that $k_3 \approx c$ [1], the thermoelastic equation (Eq. 3.4.13) becomes,

$$\frac{kT_1}{k_3^2} + i\omega\rho c_p T_1 = \frac{2F_A\nu(i\omega)}{EA_o}(3\lambda + 2\mu)\alpha T_o, \tag{3.4.18}$$

or,

[1] This assumes that the thermal decay length and the geometric increase factor are of the same order of magnitude.

$$T_1 = \frac{\beta \frac{k}{k_3^2} i\omega + \beta \omega^2 \rho c_p}{(\frac{k}{k_3^2})^2 + (\omega \rho c_p)^2}, \tag{3.4.19}$$

where $\beta = \frac{2F_A \nu}{EA_o}(3\lambda + 2\mu)\alpha T_o$ (see Appendix A.3).

In a vibrothermography inspection, the excitation frequency is relatively low. In other words, the wavelength is large relative to the asperity size and the ratio $(\frac{k}{k_3^2})^2/(\omega \rho c_p)^2 >> 1$. Using the equation for energy dissipation derived by Goodman (see Appendix A.4),

$$W_o = \frac{k}{T_o} \int_t^{t+\frac{2\pi}{\omega}} \int_V (\frac{\partial \theta}{\partial z})^2 dV dt = \frac{\pi k_3^3 \omega T_o}{A_o k} \left(\frac{2F_A \nu}{E}(3\lambda + 2\mu)\alpha \right)^2. \tag{3.4.20}$$

Note that this solution is energy loss for a single period. The energy loss per unit time is found by multiplying by the excitation frequency. This yields a frequency squared dependence, which is contradictory to observation 3. Figure 3.2 shows the energy dissipated over the entire crack surface for a 1 second excitation (with and without the assumption $(\frac{k}{k_3^2})^2/(\omega \rho c_p)^2 >> 1$). If we assume an excitation time of $t = 1\ sec$ and a crack contact area of $\approx 4^{-6}\ m^2$, the energy dissipation is borderline insignificant for typical vibrothermography excitation frequencies (≈ 20 kHz). So while thermoelasticity may play some role, it is likely not a dominant mechanism.

3.4.4 Friction/Adhesion

3.4.4.1 Rough Contact Models

Let us revisit the assumption that rough surfaces in contact experience localized yielding or complete plastic flow (used in Sect. 3.4.4.2 and 3.4.1). This assumption leads to a load-displacement proportionality (and consequently load-area proportionality), but this assumption is not universally valid. There are circumstances where the load-displacement and load-area relationships are dominated by either elastic deformation or plastic flow. Greenwood and Williamson [25] developed a widely used contact model, where a randomly rough surface is brought into contact with a flat surface. They

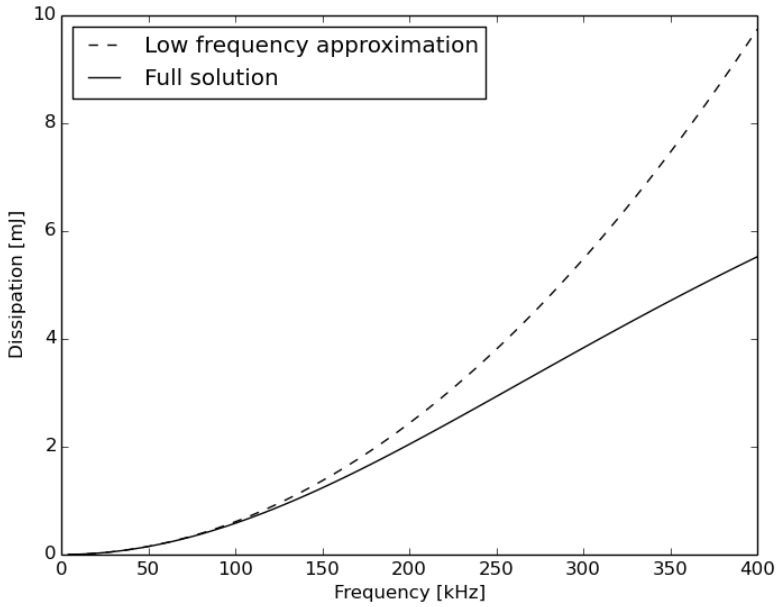

Figure 3.2 Energy dissipation per asperity per cycle. The reduced solution assumes that the wavelength of vibration is large relative to the asperity geometry.

describe the transition from elastic to plastic driven contact by a quantity referred to as the *plasticity index*. A low plasticity index ($<< 1$) represents a particularly smooth surface, where the real contact area is a significant fraction of the apparent contact area. In this scenario, the force holding the surfaces together is distributed over a large area, allowing for elastic contact. The scenario seen in most real surfaces, however, is one in which there is significant roughness and a real contact area which is a small fraction (a few percent) of the apparent contact area. The applied load is supported by a small fraction of the area, leading to plastic flow at relatively small loads.

While in reality the two contacting surfaces (crack flanks) are both rough, the assumption of one rough surface and one smooth surface is a valid approximation. Greenwood and Tripp [24] found that whether one or both surfaces are rough does not change the important relationships (e.g. load-area proportionality). In fact, one can always find a combination of a rough surface and a smooth surface that behaves the same as two rough surfaces in contact.

With a slightly different approach, Archard [3] was able to show that, regardless of whether elastic or plastic deformation dominates, the load-area relationship can approach linear. The one requirement is that the increasing load primarily results in an increase in the number of contact points, as opposed to increasing the area of existing contacts. This representation of rough surfaces is in contrast to the Greenwood-Williamson model, where the asperity heights are modeled as a statistical distribution. Archard represents the surface roughness as a distribution of hemispherical protuberances with fractal characteristics.

3.4.4.2 Sliding friction

The fundamental mechanism of friction has been studied extensively over the years, but the basic theory of dry friction (the Coulomb model) is based on experimentally observed relationships [6]. In the Coulomb model of friction, the friction force opposes tangential motion of two surfaces in contact,

$$F = \mu P, \tag{3.4.21}$$

where F is the friction force, μ is the coefficient of friction, and P is the normal force holding the two surfaces in contact. Based on work by Holland et al. [31], we can approximate the frictional energy dissipation over a dynamic excitation,

$$Q = \mu(P + dP)duf, \tag{3.4.22}$$

where P is the static normal force, dP is the dynamically induced normal force, and du is the relative motion between the surfaces induced by the dynamic loading. This relationship satisfies the power law dependence on vibration amplitude since both dP and du are assumed to be proportional to the vibration amplitude. It also satisfies the linear dependence on excitation frequency. One might assume that, since frictional dissipation only occurs for lateral motion, that an open/close mode vibration would not lead to

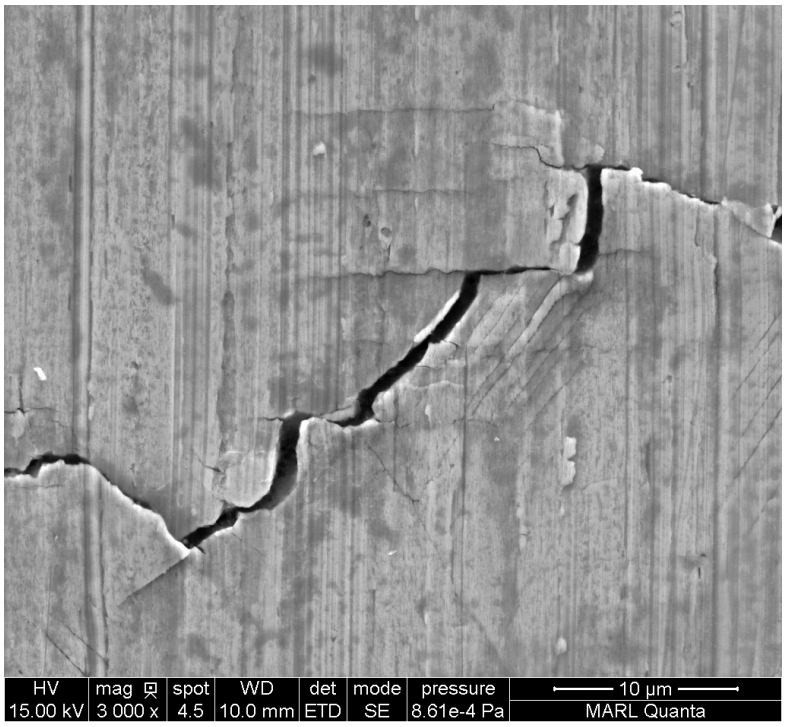

Figure 3.3 SEM image of a surface fatigue crack. Contact occurs at discrete points, often along ledges which are at an angle with the macroscopic crack path.

significant frictional losses. In reality, the morphology of fatigue crack surfaces (namely, the tortuosity of the crack path as shown in Fig. 3.3) is such that even an open/close mode vibration will induce shear or relative lateral motion between the opposing crack faces. This explains the experimental observation that heating does not depend on vibration mode (observation 5).

Equation 3.4.22 implies that heat generation increases linearly with static load. This equation, however, neglects the concept of a static friction force. It is expected that above some static load, the dynamic strain will not be sufficient to overcome the static friction force. This will result in a truncated heat generation. So while Eq. 3.4.22 does not explicitly take into account the behavior at large static loads, the physics tells us that frictional dissipation will indeed drop off above a certain static load, satisfying all of the necessary experimental observations.

3.4.4.3 The adhesion model of friction

The description of frictional dissipation in Sect. 3.4.4.2 focuses on a classical, macroscopic view crack surface sliding. If we instead adopt an atomistic view of surface contact, the distinction between adhesion hysteresis and friction becomes blurred [43; 63; 64]. Both phenomena are rooted in the creation and destruction of surface area; a sliding asperity must destroy a free surface at the front and create a free surface in it's wake. This process is irreversible, and could be a source of vibrothermography heat generation.

Imagine two surfaces in contact, supported at discrete asperity contact points, where the true contact area is much less than the apparent contact area. When the surfaces are pressed together, the discrete asperity contacts experience local yielding where the contact area behaves according to Eq. 3.4.7. This implies that the normal force holding the surfaces in contact is proportional to the real contact area. If the two surfaces are moved tangentially against one another, the friction force opposes the motion. If we assume that the friction force derives from an adhesion mechanism, friction force will be proportional to the real contact area. Since we have shown that normal force is also proportional to contact area, we can conclude that the normal force is proportional to the friction force. This result is analogous to Coulomb's law of friction, but invokes the concept of adhesion in its derivation.

3.4.4.4 Surface energy driven friction/adhesion

While the Sections 3.4.4.2 and 3.4.4.3 imply a frictional dissipation mechanism, we can also analyze the dissipation in terms of adhesion hysteresis. When two surfaces are brought in close contact, there is a point at which it is energetically favorable for the surfaces to annihilate. The free surfaces are destroyed and the free surfaces adhere. Similarly, if a force is applied to pull the surfaces apart, there is a point at which they will 'snap' apart. If this is repeated, there is a net energy loss as the surfaces are brought together and pulled apart. This energy loss is termed adhesion hysteresis. Similar to

strain hysteresis, the energy loss in adhesion is represented as a fraction of the total surface energy which is lost.

Therefore, if we assume that heat generation is primarily due to a change in contact area, we can represent the energy loss as a fraction of the total change in surface energy attributed to the change in contact area,

$$dU_S = -2\gamma dA_c, \tag{3.4.23}$$

where γ is the surface energy density. If we consider an asperity undergoing elastic deformation (i.e. Hertzian contact), we can relate the contact area to the normal force with Eq. 3.4.1,

$$dU_S = -2\gamma\pi \left(\frac{3R}{4E^*}\right)^{2/3} dP^{2/3}. \tag{3.4.24}$$

Thus, there is a two-thirds dependence on vibration amplitude, which is not in agreement with observation 4. If we consider the case where there is a static load,

$$\lim_{P_i \to \infty} dU_S = -2\gamma\pi \left(\frac{3R}{4E^*}\right)^{2/3} ((P + P_i)^{2/3} - P_i^{2/3}) = 0. \tag{3.4.25}$$

As expected, heat generation vanishes as the static loading tends to infinity. We can also analyze the system under the assumption that contact area is proportional to normal force, as discussed in Sect. 3.4.4.1 (Eq. 3.4.7). The change in surface energy is,

$$dU_S \propto -2\gamma dP. \tag{3.4.26}$$

Now there is a proportionality between heating and vibration amplitude. If we consider static loading which tends to infinity, the proportionality will hold. Thus, while heat generation is bounded, this analysis implies heat generation will continue even at large static loads, which is in contradiction to observation 2. In summary, for contact

dominated by elasticity observation, 2 holds, but observation 4 is violated. For contact dominated by plasticity, observation 4 holds, but 2 is violated.

However, we have neglected one important aspect of adhesion, namely the distinction between a change in contact area of an existing contact point and the creation of an entirely new contact point. Research has shown that hysteretic losses in surface adhesion is greater near the point of contact (or point of instability) [71]. The motivation for this stems from the idea that adhesion is dynamic in nature. The unstable snapping of asperities into and out of contact results in energy being carried away by high frequency thermal vibrations, or phonons. While this same phenomenon happens when contacting asperities are pressed more tightly in contact, it is not as significant.

Let us consider then an alternate hypothesis, that heat generation is related to the number of asperities that just make and break contact. Consider a rough surface with asperities with randomly distributed heights, as in the Greenwood-Williamson model [25]. The asperity heights are normally distributed, $h_i \sim N(\mu, \sigma^2)$. Another surface is used to press down on the rough surface and flatten the asperities. The height of the flat surface is z, and the interference between the flat surface and an individual asperity is $\delta_i = h_i - z \sim N(\mu - z, \sigma^2)$, where $\delta_i > 0$.

Assume that the asperity, when depressed, has a circular cross section and that the radius of curvature of the asperity is a constant, R. The radius of the contact area is given by a and is related to the interference by $\delta = \frac{a^2}{R}$. The contact area for a single asperity is given by,

$$A_i = \pi a^2 = \pi R \delta. \tag{3.4.27}$$

If we assume that the asperities experience local yielding, Eq. 3.4.7 holds (recall that contact dominated by elasticity may still have a load-area proportionality, as discussed in Sect. 3.4.4.1). This area should be equivalent to the sum of all of the individual asperity contact areas supporting the normal force,

$$A_{contact} = \frac{P}{c\sigma_y} = \sum_i \pi R \delta_i. \tag{3.4.28}$$

Solving for normal force we find,

$$P = \pi R c \sigma_y \sum_i (h_i - z) \qquad \text{such that } (h_i - z) > 0. \tag{3.4.29}$$

To convert this to an integral, we multiply the height by the probability of this asperity height occurring. We also scale it by multiplying by the total number of asperities per unit area, R_α.

$$P = \pi R R_\alpha c \sigma_y \int_z^\infty (h - z) f(h) dh = \pi R R_\alpha c \sigma_y \left[(\mu - z)(1 - \Phi(\alpha)) + \sigma \phi(\alpha) \right], \tag{3.4.30}$$

where $\alpha = \frac{z-\mu}{\sigma}$ [22] and ϕ and Φ are the standard normal pdf and cdf, respectively. See Appendix A.5 for details. We can rearrange this,

$$P^* = \frac{P}{\pi R R_\alpha c \sigma_y \sigma} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha^2} - \frac{\alpha}{2} erfc\left(\frac{\alpha}{\sqrt{2}}\right). \tag{3.4.31}$$

We are interested in evaluating heat generation, specifically as a function of the applied load. If heating is related to the number of asperities coming into or leaving contact, we can evaluate the rate at which asperity contact is created,

$$dQ(P^*) \propto \frac{\partial}{\partial P^*}(1 - \Phi) \propto -\frac{\partial \Phi}{\partial \alpha} \frac{\partial \alpha}{\partial P^*} \propto -\phi \frac{\partial \alpha}{\partial P^*}. \tag{3.4.32}$$

The cdf of the standard normal, Φ, represents the total fraction of asperities in contact at height z and below. Since positive z represents moving the surface out of contact with the asperities, the total number of asperities in contact at a given surface height z (or α) is $R_\alpha(1 - \Phi(\alpha))$.

The derivative $\frac{\partial}{\partial P^*}(1 - \Phi)$ represents the change in total fraction of asperities in contact as a function of load. If we scale by the total number of asperities, this derivative

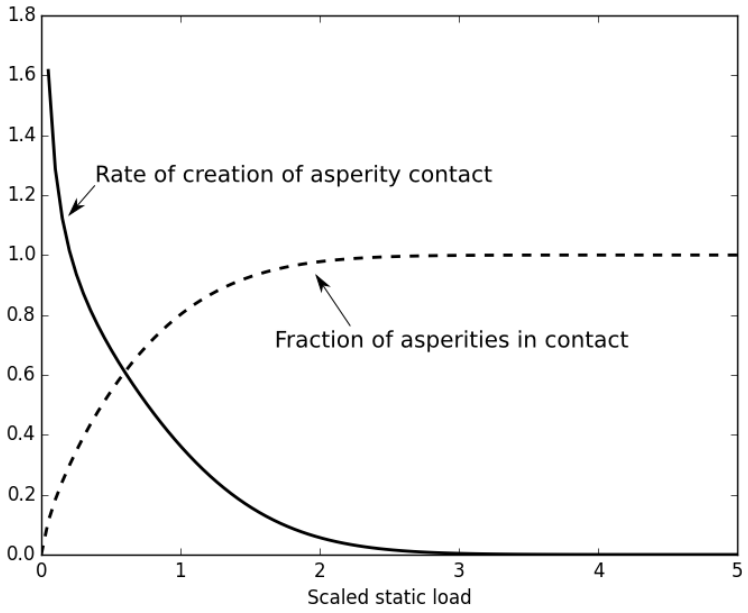

Figure 3.4 Rate of asperity contact creation/destruction as a function of static load.

represents the change in the total number of asperities per unit load. The only unknown is the surface height z as a function of normal force P (equivalently, α as a function of P^*). It turns out that the inverse relationship can be represented well by the following relation,

$$\alpha(P^*) \approx 3.62 - 0.72P^* - 3.78(P^*)^{0.172}, \tag{3.4.33}$$

where the coefficients are fit using a least squares approach.

The rate of asperity contact creation is plotted as a function of scaled load P^* in Fig. 3.4. We can see that at low static loading, the rate at which asperity contacts are created is relatively high. At elevated static loading, this rate of asperity contact creation drops off (eventually vanishing) as the fraction of asperities in contact approaches one. If total thermal power is related to the number of asperities in coming into and out of contact, this is the behavior expected from observation 2. If the total heat generation for a cyclic

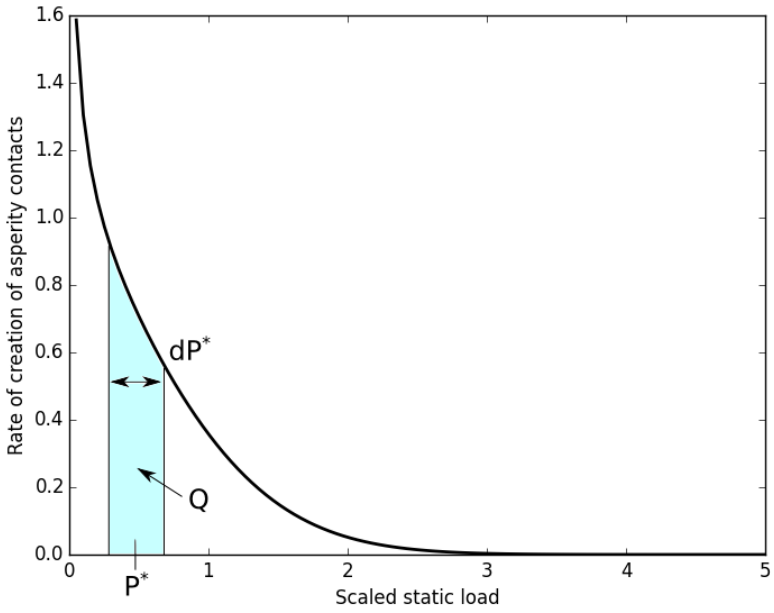

Figure 3.5 Heat generation, Q, for static load P^* and dynamic load dP^*.

load between P_1^* and P_2^* is proportional to the total number of asperities coming into and out of contact over the loading cycle,

$$Q(P^*) \propto \int_{\alpha_1}^{\alpha_2} dQ = [1 - \Phi(\alpha_2)] - [1 - \Phi(\alpha_1)]\,. \tag{3.4.34}$$

Plugging in the CDF,

$$Q \propto \frac{1}{2}\left| erf\left(\frac{\alpha(P_2^*)}{\sqrt{2}}\right) - erf\left(\frac{\alpha(P_1^*)}{\sqrt{2}}\right)\right|. \tag{3.4.35}$$

This quantity represents the area under the rate of asperity contact creation curve. For a given static load P^*, heat generation Q is a function the dynamic load dP^* as shown in Fig. 3.5. Since this is the energy lost in a single cycle, the energy lost per unit time will be Qf, satisfying observation observation 3. The only requirement we haven't yet satisfied is observation 4. To evaluate this observation, Q is plotted against dP^* for three different static load levels: i) zero static load, ii) median static load, and iii) high static load (see Fig. 3.6).

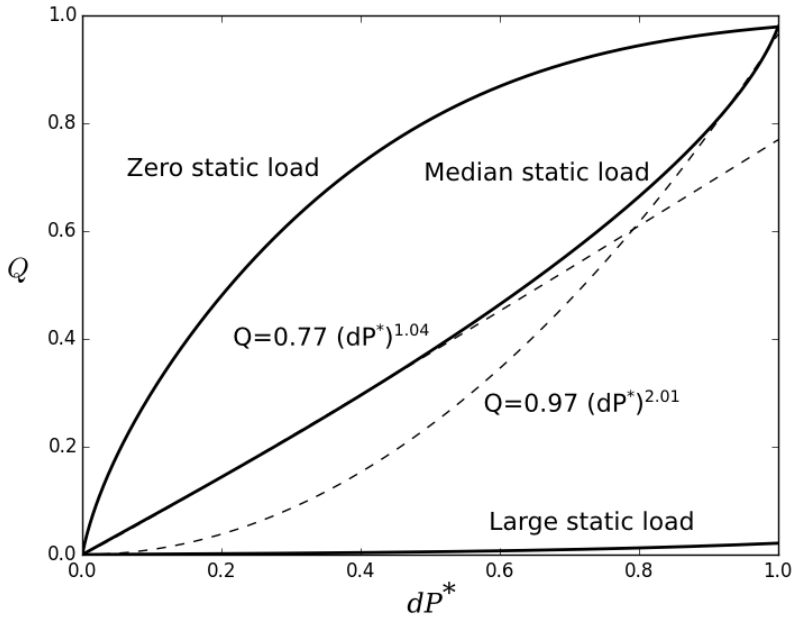

Figure 3.6 Heat generation (qualitative) as a function of dynamic load amplitude for three different static load levels using Eq. 3.4.35.

As expected, at high static loading heat generation is reduced. At zero static load, heat generation increases sharply and then levels off as the rate of asperity contact creation drops. Finally, at a median static load heat generation starts as roughly linear and gradually increases until it is quadratic in dP^*. While the prediction is qualitative, this is the exact behavior expected based on experimental observation.

3.5 Conclusion

There is a lack of agreement in the NDE community regarding the heating mechanisms that dominate vibrothermography crack heating. In this work, we have given experimental and theoretical evidence to disprove a number of mechanisms, namely linear absorption, thermoelasticity, and plasticity. Furthermore, we have introduced a framework with which to predict contact area driven heating through adhesion hysteresis, and frictional heating through relative surface motion. Both approaches satisfy the exper-

imental criteria set forth by observation. Within the classical view of sliding friction, heating comes from the static normal force, the dynamic normal force, and the relative surface motion induced by the dynamic strain. It is assumed that there is a static load level above which the dynamic strain is not sufficient to overcome the static friction force. With adhesion hysteresis, energy loss is first attributed to a change in contact area (and consequently surface energy). In reality, hysteresis losses are dominant at the point of instability when the asperity is brought completely into and out of contact. This satisfies the observation that heating vanishes at high static loads. At atomistic scales, the distinction between adhesion and friction is unclear and more work is needed to distinguish the true source of heat generation.

3.6 Acknowledgements

This material is based on work supported by the Air Force Research Laboratory Contract #FA8650-10D-5210, Task Order #023 and performed at Iowa State University.

CHAPTER 4. CONCLUSION

The crack closure model developed in this work is novel in both the fields of NDE and fracture mechanics. The one-dimensional crack closure model describes the mechanics of a partially closed fatigue crack below the crack opening stress in terms of stresses and displacements in response to an externally applied tensile load and a local closure stress distribution. The fatigue process produces a wake of plastically stretched material surrounding the crack flanks. This residual strain manifests itself as contact stresses at the interface between the opposing crack surfaces. When an external tensile load is applied, it is balanced by these compressive stresses, resulting in an equilibrium closure state. For a 1D crack, this equilibrium state can be described by the effective crack length, the apparent contact stress, and the opening displacement. These predictions are validated numerically and shown to be within expected FEM error. An algorithm to solve for an unknown closure stress is developed based on the foundations of the aforementioned closure mechanics model. This algorithm estimates the closure stress acting on a partially closed crack, provided measurements of the effective surface length as a function of remote load.

The application of these fracture and closure mechanics concepts to vibrothermography NDE requires an understanding of the mechanisms by which heat generation occurs. Given the uncertainty which surrounds the topic in the scientific community, an analysis of plausible mechanisms is carried out. We provide theoretical evidence which suggests that linear absorption, plastic flow, and thermoelasticity are not plausible heat genera-

tion mechanisms. The likely mechanism is a form of friction or adhesion, both of which are influenced by the real area in contact (as opposed to the apparent contact area).

This work is the first attempt at understanding and modeling the fundamental connection between crack closure and vibrothermography heat generation analytically. Past attempts to model vibrothermography heat generation have focused on empirical or finite element models, ignoring the complexities which arise from the interplay between fracture mechanics, contact mechanics, and heat generation. Aside from the applicability to the field of NDE, the crack closure modeling is novel in the field of fracture mechanics for its description of a partially closed fatigue crack below the crack opening stress.

BIBLIOGRAPHY

[1] T. L. Anderson. *Fracture mechanics: fundamentals and applications*. CRC press, 2005.

[2] P. Anselone. Singularity subtraction in the numerical solution of integral equations. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, 22(04):408–418, 1981.

[3] J. F. Archard. Elastic Deformation and the Laws of Friction. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 243(1233):190–205, Dec. 1957.

[4] K. E. Atkinson. *An introduction to numerical analysis*. John Wiley & Sons, 2008.

[5] S. Banerjee. A review of crack closure. Technical report, DTIC Document, 1984.

[6] F. P. Bowden and D. Tabor. *The friction and lubrication of solids*, volume 1. Oxford university press, 2001.

[7] O. Buck, B. R. Tittmann, and C. J. Beevers. Advances in crack length measurement. *Beevers, C. J. eds*, page 413, 1982.

[8] B. Budiansky and J. W. Hutchinson. Analysis of closure in fatigue crack growth. *Journal of Applied Mechanics*, 45(2):267–276, 1978.

[9] H. Bueckner. Novel principle for the computation of stress intensity factors. *Zeitschrift fuer Angewandte Mathematik & Mechanik*, 50(9), 1970.

[10] R. Clark, W. D. Dover, and L. J. Bond. The effect of crack closure on the reliability of NDT predictions of crack size. *NDT international*, 20(5):269–275, 1987.

[11] D. L. Davidson. Plasticity induced fatigue crack closure. In *Mechanics of Fatigue Crack Closure*. ASTM International, 1988.

[12] L. Delves and J. Mohamed. *Computational Methods for Integral Equations*. Cambridge University Press, 1988.

[13] D. S. Dugdale. Yielding of steel sheets containing slits. *Journal of the Mechanics and Physics of Solids*, 8(2):100–104, May 1960.

[14] W. Elber. The significance of fatigue crack closure. *Astm Stp*, 486:230–243, 1971.

[15] L. D. Favro, X. Han, Z. Ouyang, G. Sun, H. Sui, and R. L. Thomas. Infrared imaging of defects heated by a sonic pulse. *Review of Scientific Instruments*, 71(6):2418–2421, June 2000.

[16] T. Fett. *Stress Intensity Factors, T-stresses, Weight Functions*, volume 50. KIT Scientific Publishing, 2008.

[17] T. Fett. *Stress Intensity Factors, T-stresses, Weight Functions: Supplement Volume*, volume 55. KIT Scientific Publishing, 2009.

[18] N. A. Fleck. *An investigation of fatigue crack closure*. PhD dissertation, University of Cambridge, 1984.

[19] G. Glinka. Development of weight functions and computer integration procedures for calculating stress intensity factors around cracks subjected to complex stress fields. *Stress and Fatigue-Fracture Design, Petersburg Ontario, Canada, Progress Report*, 1(1):1, 1996.

[20] G. Glinka and G. Shen. Universal features of weight functions for cracks in mode I. *Engineering Fracture Mechanics*, 40(6):1135–1146, 1991.

[21] L. E. Goodman, C. S. Chang, and A. R. Robinson. Thermoelastic damping. Technical report, DTIC Document, 1962.

[22] W. H. Greene. *Econometric analysis*. Pearson Education India, 2003.

[23] J. Greenwood, H. Minshall, and D. Tabor. Hysteresis losses in rolling and sliding friction. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 259, pages 480–507. The Royal Society, 1961.

[24] J. A. Greenwood and J. H. Tripp. The Contact of Two Nominally Flat Rough Surfaces. *Proceedings of the Institution of Mechanical Engineers*, 185(1):625–633, June 1970.

[25] J. A. Greenwood and J. B. P. Williamson. Contact of Nominally Flat Surfaces. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 295(1442):300–319, Dec. 1966.

[26] A. A. Griffith. The phenomena of rupture and flow in solids. *Philosophical transactions of the royal society of london. Series A, containing papers of a mathematical or physical character*, 221:163–198, 1921.

[27] X. Han, M. S. Islam, G. Newaz, L. D. Favro, and R. L. Thomas. Finite element modeling of the heating of cracks during sonic infrared imaging. *Journal of Applied Physics*, 99(7):074905, Apr. 2006.

[28] E. G. Henneke, K. L. Reifsnider, and W. W. Stinchcomb. Thermography - An NDI Method for Damage Detection. *JOM*, 31(9):11–15, Sept. 1979.

[29] I. Henneke, K. L. Reifsnider, and W. W. Stinchcomb. Vibrothermography: Investigation, Development, and Application of a New Nondestructive Evaluation Technique. Technical report, Nov. 1986.

[30] S. D. Holland, L. Koester, J. Vaddi, T. Lesthaeghe, W. Q. Meeker, B. Schiefelbein, D. E. Chimenti, and L. J. Bond. Vibrosim: A hybrid computational/empirical model of vibrothermography nondestructive evaluation. In *AIP Conference Proceedings*, volume 1706, page 100008. AIP Publishing, 2016.

[31] S. D. Holland, C. Uhl, Z. Ouyang, T. Bantel, M. Li, W. Q. Meeker, J. Lively, L. Brasche, and D. Eisenmann. Quantifying the vibrothermographic effect. *NDT & E International*, 44(8):775–782, 2011.

[32] C. Homma, M. Rothenfusser, J. Baumann, R. Shannon, D. O. Thompson, and D. E. Chimenti. Study of the heat generation mechanism in acoustic thermography. In *AIP Conference Proceedings*, volume 820, pages 566–573. AIP, 2006.

[33] C. Ibarra-Castanedo, A. Saboktakin, X. Maldague, A. Bendada, P. Bison, E. Grinzato, and S. Marinetti. Analysis of ultrasonic elastic waves in vibrothermography using fem. In *6th International Workshop - NDT Signal Processing*, 2009.

[34] G. R. Irwin. Onset of fast crack propagation in high strength steel and aluminum alloys. Technical report, May 1956.

[35] K. L. Johnson and K. L. Johnson. *Contact mechanics*. Cambridge university press, 1987.

[36] R. Kemp. Fatigue crack closure-a review. Technical report, DTIC Document, 1990.

[37] A. Kiciak, G. Glinka, and D. Burns. Calculation of stress intensity factors and crack opening displacements for cracks subjected to complex stress fields. *Journal of pressure vessel technology*, 125(3):260–266, 2003.

[38] A. L. Kimball and D. E. Lovell. Internal Friction in Solids. *Physical Review*, 30(6):948–959, Dec. 1927.

[39] T. Lesthaeghe. Evaluation of some parameters influencing vibrothermographic crack heating. *Graduate Theses and Dissertations*, Jan. 2015.

[40] J. Lu, X. Han, G. Newaz, L. D. Favro, and R. L. Thomas. Study of the effect of crack closure in Sonic Infrared Imaging. *Nondestructive Testing and Evaluation*, 22(2-3):127–135, June 2007.

[41] F. Mabrouki, M. Thomas, M. Genest, and A. Fahr. Frictional heating model for efficient use of vibrothermography. *NDT & E International*, 42(5):345–352, July 2009.

[42] F. Mabrouki, M. Thomas, M. Genest, and A. Fahr. Numerical modeling of vibrothermography based on plastic deformation. *NDT & E International*, 43(6):476–483, Sept. 2010.

[43] L. Makkonen. A thermodynamic model of sliding friction. *AIP Advances*, 2(1):012179, Mar. 2012.

[44] B. Masserey. *Ultrasonic surface crack characterization using Rayleigh waves*. PhD dissertation, Diss., Technische Wissenschaften, Eidgenössische Technische Hochschule ETH Zürich, Nr. 16646, 2006, 2006.

[45] N. G. H. Meyendorf, P. B. Nagy, and S. Rokhlin. *Nondestructive Materials Characterization: With Applications to Aerospace Materials*. Springer Science & Business Media, Nov. 2013. Google-Books-ID: IUbtCAAAQBAJ.

[46] M. Morbidini, P. Cawley, T. Barden, D. Almond, and P. Duffour. The relationship between vibration level and minimum detectable defect size in sonic-ir inspection. In *Proceedings of the 9th European Conference on Non-Destructive Testing*, 2006.

[47] J. Newman. A crack-closure model for predicting fatigue crack growth under aircraft spectrum loading. In *Methods and models for predicting fatigue crack growth under random loading*. ASTM International, 1981.

[48] J. C. Newman. A finite-element analysis of fatigue crack closure. In *Mechanics of Crack Growth*. ASTM International, 1976.

[49] P. Paris and F. Erdogan. A critical analysis of crack propagation laws. *Journal of basic engineering*, 85(4):528–533, 1963.

[50] R. Piessens, E. de Doncker-Kapenga, C. W. Überhuber, and D. K. Kahaner. *QUADPACK: A subroutine package for automatic integration*, volume 1. Springer Science & Business Media, 2012.

[51] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C*, volume 2. Cambridge Univ Press, 1982.

[52] G. Ravichandran, A. J. Rosakis, J. Hodowany, and P. Rosakis. On the conversion of plastic work into heat during high-strain-rate deformation. 2002.

[53] K. Reifsnider, E. G. Henneke, and W. Stinchcomb. The mechanics of vibrothermography. In *Mechanics of nondestructive testing*, pages 249–276. Springer, 1980.

[54] J. Renshaw, J. C. Chen, S. D. Holland, and R. B. Thompson. The sources of heat generation in vibrothermography. *NDT & E International*, 44(8):736–739, 2011.

[55] J. Renshaw, S. Holland, R. Thompson, and C. Uhl. The Effect of Crack Closure on Heat Generation on Vibrothermography. *Aerospace Engineering Conference Papers, Presentations and Posters*, July 2008.

[56] J. Renshaw, S. D. Holland, and R. B. Thompson. Measurement of crack opening stresses and crack closure stress profiles from heat generation in vibrating cracks. *Applied Physics Letters*, 93(8):081914, 2008.

[57] J. Renshaw, S. D. Holland, R. B. Thompson, and J. Anderegg. Vibration-induced tribological damage to fracture surfaces via vibrothermography. *International Journal of Fatigue*, 33(7):849–857, July 2011.

[58] J. R. Rice. Some remarks on elastic crack-tip stress fields. *International Journal of Solids and Structures*, 8(6):751–758, 1972.

[59] R. O. Ritchie. Mechanisms of fatigue crack propagation in metals, ceramics and composites: role of crack tip shielding. *Materials Science and Engineering: A*, 103(1):15–28, 1988.

[60] B. E. Schiefelbein, S. D. Holland, and A. Bastawros. Precision measurement of crack closure state with vibrothermography. In *AIP Conference Proceedings*, volume 1706, page 020004. AIP Publishing, Feb. 2016.

[61] J. Schijve. Fatigue crack closure: observations and technical significance. In *Mechanics of fatigue crack closure*. ASTM International, 1988.

[62] C. Sun and Z.-H. Jin. Chapter 3 - the elastic stress field around a crack tip. In C. S.-H. Jin, editor, *Fracture Mechanics*, pages 25 – 75. Academic Press, Boston, 2012.

[63] R. Szoszkiewicz, B. Bhushan, B. D. Huey, A. J. Kulik, and G. Gremaud. Correlations between adhesion hysteresis and friction at molecular scales. *The Journal of Chemical Physics*, 122(14):144708, 2005.

[64] R. Szoszkiewicz, B. Bhushan, B. D. Huey, A. J. Kulik, and G. Gremaud. Adhesion hysteresis and friction at nanometer and micrometer lengths. *Journal of Applied Physics*, 99(1):014310, Jan. 2006.

[65] R. Thompson and C. Fiedler. The effects of crack closure on ultrasonic scattering measurements. In *Review of Progress in Quantitative Nondestructive Evaluation*, pages 207–215. Springer, 1984.

[66] R. B. Thompson, B. Skillings, L. Zachary, L. Schmerr, and O. Buck. Effects of crack closure on ultrasonic transmission. In *Review of Progress in Quantitative Nondestructive Evaluation*, pages 325–343. Springer, 1983.

[67] H. Walther. Internal Friction in Solids. *The Scientific Monthly*, 41(3):275–277, 1935.

[68] E. Wolf. Fatigue crack closure under cyclic tension. *Engineering Fracture Mechanics*, 2(1):37–45, 1970.

[69] A. T. Zehnder. *Fracture Mechanics*, volume 62 of *Lecture Notes in Applied and Computational Mechanics*. Springer Netherlands, Dordrecht, 2012. DOI: 10.1007/978-94-007-2595-9.

[70] W. Zhang. Frequency and load mode dependence of Vibrothermography. *Graduate Theses and Dissertations*, Jan. 2010.

[71] Zheng Wei, Meng-Fu He, and Ya-Pu Zhao. The Effects of Roughness on Adhesion Hysteresis. *Journal of Adhesion Science & Technology*, 24(6):1045–1054, May 2010.

APPENDIX ADDITIONAL MATERIAL

A.1 Solving Volterra integral equations

Volterra integral equations are a class of integral equations where the upper bound of integration is a variable. The Volterra integral equation of the first kind has the form,

$$f(t) = \int_a^t x(s)K(t,s)ds, \tag{A.1.1}$$

where $x(t)$ is an unknown function. This class of integral equations can typically be solved for using forward substitution. When discretized and placed in matrix form, the equation becomes a linear algebra problem with the form,

$$F = Kx, \tag{A.1.2}$$

where F is a vector containing values of $f(t_i)$, x is a vector with the unknown values $x(s_j)$, and K is a lower triangular matrix filled with values $K(t_i, s_j)$. Equation 2.19, used in the closure stress inversion, has the form of a Volterra integral equation of the first kind. To discretize the equation, we employ the trapezoidal integration rule,

$$\int_{a_c}^{a} \sigma_{closure}(x)m(x,a)dx = \frac{h}{2}\sum_{k=0}^{N-1}\left[\sigma_{closure}(x_{k+1})m(x_{k+1},x_N) + \sigma_{closure}(x_k)m(x_k,x_N)\right]. \tag{A.1.3}$$

Equivalently,

$$\int_{a_c}^{a} \sigma_{closure}(x)m(x,a)dx = \frac{h}{2}\left[\sigma_{closure}(x_0)m(x_0,x_N) + 2\sum_{j=0}^{N-1}\sigma_{closure}(x_j)m(x_j,x_N) + \sigma_{closure}(x_N)m(x_N,x_N)\right], \tag{A.1.4}$$

where $x_0 = a_0$, $x_N = a_i$, and $h = \frac{x_N - x_0}{N}$. It should be apparent from the form of $m(x,a)$ (Eq. 2.3), that evaluating $m(x_N, x_N) \to \infty$, which is problematic. To account for this singularity, one can apply a singularity subtraction method [2; 51],

$$-K_{op}(a) = \int_{a_c}^{a} (\sigma_{closure}(x) - \sigma_{closure}(a)) m(x,a) dx + \sigma_{closure}(a) \int_{a_c}^{a} m(x,a) dx. \quad \text{(A.1.5)}$$

When $x = a$, the first integral will vanish and the second integral can be evaluated numerically. Using the trapezoidal rule for the first integral,

$$\int_{a_c}^{a} \sigma_{closure}(x) m(x,a) dx = \frac{h}{2}\left[(\sigma_{closure}(x_0) - \sigma_{closure}(a_i)) m(x_0,a_i) + 2\sum_{j=0}^{N-1} (\sigma_{closure}(x_j) - \sigma_{closure}(a_i)) m(x_j,a_i)\right], \quad \text{(A.1.6)}$$

where the last term $(\sigma_{closure}(a_i) - \sigma_{closure}(a_i)) m(a_i,a_i)$ vanishes. Including the other terms in Eq. A.1.5,

$$-\left(K_{op}(a_i) + \sigma_{closure}(a_i) \int_{a_0}^{a_i} m(x,a_i) dx\right) = \frac{h}{2}\left[(\sigma_{closure}(x_0) - \sigma_{closure}(a_i)) m(x_0,a_i) + 2\sum_{j=0}^{N-1} (\sigma_{closure}(x_j) - \sigma_{closure}(a_i)) m(x_j,a_i)\right]. \quad \text{(A.1.7)}$$

Grouping terms that include $\sigma_{closure}(a_i)$,

$$\sigma_{closure}(a_i)\left(\frac{h}{2} m(x_0,a_i) + h\sum_{j=0}^{N-1} m(x_j,a_i) - \int_{a_0}^{a_i} m(x,a_i) dx\right) = K_{op}(a_i) + \frac{h}{2}\left[\sigma_{closure}(x_0) m(x_0,a_i) + 2\sum_{j=0}^{N-1} \sigma_{closure}(x_j) m(x_j,a_i)\right]. \quad \text{(A.1.8)}$$

Solving for $\sigma_{closure}(a_i)$,

$$\sigma_{closure}(a_i) = \frac{K_{op}(a_i) + \frac{h}{2}\left[\sigma_{closure}(x_0) m(x_0,a_i) + 2\sum_{j=0}^{N-1} \sigma_{closure}(x_j) m(x_j,a_i)\right]}{\frac{h}{2} m(x_0,a_i) + h\sum_{j=0}^{N-1} m(x_j,a_i) - \int_{a_0}^{a_i} m(x,a_i) dx}. \quad \text{(A.1.9)}$$

A.2 Dilatational strain

The linearized thermoelastic equation (Eq. 3.4.13) has the dilatational strain ϵ'' as a forced excitation. Consider the assumed asperity cross-sectional area given by Eq. 3.4.14. For a stress applied normal to the asperity tip (along the z-axis), the strain components are given by,

$$\epsilon_z = \frac{\sigma}{E}, \tag{A.2.1}$$

$$\epsilon_x = \epsilon_y = -\nu\frac{\sigma}{E}. \tag{A.2.2}$$

The dilatational strain is defined as the change in volume with respect to the original volume,

$$\epsilon'' = \frac{\Delta V}{V}. \tag{A.2.3}$$

Using Eq. 3.4.14, the total volume is given by,

$$V = \int_0^h A_o e^{\frac{z}{c}} dz = A_o c(e^{\frac{h}{c}} - 1), \tag{A.2.4}$$

where h is the asperity height. After applying a load along z, the new volume will be given by,

$$V_{new} = \int_0^{h_{new}} A_{new} dz = \int_0^{h(1+\epsilon_z)} (1-\nu\epsilon_z)^2 A dz, \tag{A.2.5}$$

$$V_{new} = \int_0^{h(1+\epsilon_z)} (1-\nu\epsilon_z)^2 A dz = \frac{A_o(1-\nu\epsilon_z)^2}{c}(e^{\frac{h}{c}(1+\epsilon_z)} - 1). \tag{A.2.6}$$

The change in volume due to stress is given by,

$$\Delta V = V_{new} - V = \frac{A_o(1-\nu\epsilon_z)^2}{c}(e^{\frac{h}{c}(1+\epsilon_z)} - 1) - A_o c(e^{\frac{h}{c}} - 1). \tag{A.2.7}$$

Plugging Eqs. A.2.4 and A.2.7 into Eq. A.2.3,

$$\epsilon'' = \frac{\Delta V}{V} = \frac{A_o c(1-\nu\epsilon_z)^2(e^{\frac{h}{c}}e^{\frac{h}{c}\epsilon_z} - 1) - A_o c(e^{\frac{h}{c}} - 1)}{A_o c(e^{\frac{h}{c}} - 1)}. \tag{A.2.8}$$

Since c represents the characteristic length of the asperity, it should be on the order of nanometers or microns. We are also assuming that strains are small. This means that, in the product $e^{\frac{h}{c}}e^{\frac{h}{c}\epsilon_z}$, the first term will dominate. This simplifies to,

$$\epsilon'' = \frac{A_o c(1-\nu\epsilon_z)^2(e^{\frac{h}{c}} - 1) - A_o c(e^{\frac{h}{c}} - 1)}{A_o c(e^{\frac{h}{c}} - 1)} = (1-\nu\epsilon_z)^2 - 1. \tag{A.2.9}$$

Expanding,

$$\epsilon'' = -2\nu\epsilon_z + (\nu\epsilon_z)^2. \tag{A.2.10}$$

Again, with the small strain assumption, we can neglect the second term. Finally,

$$\epsilon'' = -2\nu\epsilon_z. \tag{A.2.11}$$

A.3 Thermoelastic absorption

The time derivative of the dilatational strain (needed to solve Eq. 3.4.13) is found by Eq. 3.4.16, where ϵ_z is defined in Eq. A.2.1. The stress applied to the asperity is $\sigma(t) = F(t)/A$ where F is defined in Eq. 3.4.15 and A in Eq. 3.4.14,

$$\frac{\partial \epsilon''}{\partial t} = \frac{2F_A \nu(i\omega)}{EA_o} e^{-\frac{z}{c}} e^{i\omega t}. \tag{A.3.1}$$

Inserting this into the linearized thermoelastic equation (Eq. 3.4.13),

$$k\nabla^2\theta = \rho c_p \frac{\partial \theta}{\partial t} + \frac{2F_A \nu(i\omega)}{EA_o}(3\lambda + 2\mu)\alpha T_o e^{-\frac{z}{c}} e^{i\omega t}. \tag{A.3.2}$$

Inserting the particular form of θ given in Eq. 3.4.17 into Eq. A.3.2,

$$\frac{kT_1}{k_3^2} e^{-i\omega t} e^{-\frac{z}{k_3}} = -i\omega\rho c_p T_1 e^{-i\omega t} e^{-\frac{z}{k_3}} + \frac{2F_A \nu(i\omega)}{EA_o}(3\lambda + 2\mu)\alpha T_o e^{-\frac{z}{c}} e^{-i\omega t}. \tag{A.3.3}$$

If we assume that $k_3 \approx c$, we can cancel the term $e^{-i\omega t} e^{-\frac{z}{k_3}}$. Solving for T_1 yields,

$$T_1 = \frac{\beta i\omega}{\frac{k}{k_3^2} + i\omega\rho c_p}. \tag{A.3.4}$$

Multiplying the numerator and denominator by the complex conjugate of the denominator,

$$T_1 = \frac{\beta \frac{k}{k_3^2} i\omega + \beta\omega^2 \rho c_p}{(\frac{k}{k_3^2})^2 + (\omega\rho c_p)^2}. \tag{A.3.5}$$

Note that the two terms in the denominator are functions of the excitation frequency and the scale of the system (k_3). For this application, we take $\omega \approx 125 \times 10^3 rad/s$ for $f = 20kHz$ and $k_3 \approx 1\mu m$. This leads to the ratio,

$$(\frac{k}{k_3^2})^2/(\omega\rho c_p)^2 \approx 500. \tag{A.3.6}$$

This shows that the geometric term is dominant for this application, although for frequencies above 250 kHz, the terms are comparable.

For comparison, the work by Goodman et al. utilized a geometry where $h = 1$" and $f \approx 3MHz$. With steel we have $k = .5[W/mK]$, $c_p = 500[J/kgK]$, and $\rho = 7850[kg/m^3]$. For this scenario, the ratio is,

$$(\frac{k}{k_3^2})^2/(\omega\rho c_p)^2 \approx 10^-22. \tag{A.3.7}$$

This shows that Goodman et al. are analyzing a situation where the frequency is so high that the geometric effects are filtered out. This is indeed evident from the final solution for the energy dissipation,

$$W_o = \frac{(3\lambda + 2\mu)^2\alpha^2 T_o P^2 h\pi}{(\lambda + 2\mu)^2\rho c_p\sqrt{\frac{c_1'}{2}}}, \tag{A.3.8}$$

where $c_1' \propto h^2$.

A.4 Thermoelastic energy dissipation

With the assumed asperity cross-sectional area (Eq. 3.4.14) and applied load (Eq. 3.4.15), the coefficient T_1 in the assumed temperature increase (Eq. 3.4.17) can be solved for. Goodman et al. [21] derived an equation to describe the total energy dissipation for a loading cycle,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V (\frac{\partial\theta}{\partial z})^2 dV dt. \tag{A.4.1}$$

The coefficient T_1 is found to be (Eq. 3.4.19),

$$T_1 = \frac{\beta\frac{k}{k_3^2}i\omega + \beta\omega^2\rho c_p}{(\frac{k}{k_3^2})^2 + (\omega\rho c_p)^2}. \tag{A.4.2}$$

As stated in Appendix A.3, Goodman et al. consider a situation where the frequency is very high and the geometric terms are not important. However, in vibrothermography the frequency is relatively low, such that the vibration wavelength is large relative to the asperity height. We will solve for the thermoelastic dissipation first without making any approximations, and then by making a low frequency approximation.

A.4.1 Full solution

First, let us solve for the thermoelastic dissipation without simplifying the coefficient T_1. We can express T_1 from Eq. 3.4.19 as,

$$T_1 = C_1 + iC_2, \tag{A.4.3}$$

where $C_1 = \frac{\beta\omega^2\rho c_p}{(\frac{k}{k_3^2})^2+(\omega\rho c_p)^2}$ and $C_2 = \frac{\beta\frac{k}{k_3^2}\omega}{(\frac{k}{k_3^2})^2+(\omega\rho c_p)^2}$. Plugging these coefficients into the definition of the temperature increment (Eq. 3.4.17),

$$\theta(z,t) = (C_1 + iC_2)\, e^{-i\omega t} e^{-\frac{z}{k_3}}. \tag{A.4.4}$$

Using Eq. A.4.4,

$$\frac{\partial\theta}{\partial z} = -\frac{1}{k_3}\left(C_1 + iC_2\right)e^{-i\omega t}e^{-\frac{z}{k_3}}e^{-i\omega t}e^{-\frac{z}{k_3}}. \tag{A.4.5}$$

We define $\mathcal{C}_1$ and $\mathcal{C}_2$ to absorb the factor $-\frac{1}{k_3}e^{-\frac{z}{k_3}}$ in Eq. A.4.5,

$$\frac{\partial\theta}{\partial z} = \left(\mathcal{C}_1(z) + i\mathcal{C}_2(z)\right)e^{-i\omega t}, \tag{A.4.6}$$

where $\mathcal{C}_1 = -\frac{1}{k_3}\frac{\beta\omega^2\rho c_p}{(\frac{k}{k_3^2})^2+(\omega\rho c_p)^2}e^{-\frac{z}{k_3}}$ and $\mathcal{C}_2 = -\frac{1}{k_3}\frac{\beta\frac{k}{k_3^2}\omega}{(\frac{k}{k_3^2})^2+(\omega\rho c_p)^2}e^{-\frac{z}{k_3}}$.
Converting to the leading term $\mathcal{C}_1(z) + i\mathcal{C}_2(z)$ polar form,

$$\mathcal{C}_1(z) + i\mathcal{C}_2(z) = \sqrt{\mathcal{C}_1^2 + \mathcal{C}_2^2}e^{itan^{-1}(\frac{\mathcal{C}_2}{\mathcal{C}_1})}. \tag{A.4.7}$$

Equation A.4.6 becomes,

$$\frac{\partial\theta}{\partial z} = \sqrt{\mathcal{C}_1^2 + \mathcal{C}_2^2}exp\left(itan^{-1}(\frac{\mathcal{C}_2}{\mathcal{C}_1})\right)e^{-i\omega t}. \tag{A.4.8}$$

Plugging Eq. A.4.6 into Eq. A.4.1,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V\left[\sqrt{\mathcal{C}_1^2 + \mathcal{C}_2^2}exp\left(itan^{-1}(\frac{\mathcal{C}_2}{\mathcal{C}_1})\right)e^{-i\omega t}\right]^2 dVdt, \tag{A.4.9}$$

or,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V\left[\mathcal{C}e^{-i(\omega t-\phi)}\right]^2 dVdt, \tag{A.4.10}$$

where $\mathcal{C} = \sqrt{\mathcal{C}_1^2 + \mathcal{C}_2^2}$ and $\phi = tan^{-1}(\frac{\mathcal{C}_2}{\mathcal{C}_1})$. The integration in time is taken over one period and is analogous to finding the average power in an AC circuit. The average power is found by integrating the instantaneous power over one cycle,

$$P_{avg} = \frac{VI}{2}cos(\theta), \tag{A.4.11}$$

where V and I are the amplitudes for voltage and current, respectively, and θ is the phase angle between them. Since we are interested in the average power of a sinusoid squared, the phase angle will be zero,

$$P_{avg} = \frac{\mathcal{C}^2}{2}. \tag{A.4.12}$$

We can reach the same result by multiplying the phasor $\mathcal{C}e^{-i(\omega t-\phi)}$ by its complex conjugate. Since multiplying phasors means taking the product of the amplitudes and the sum of the phases, we will be left with,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V \mathcal{C}^2 e^{-i(\omega t+\phi)}e^{-i(\omega t-\phi)}dVdt. \tag{A.4.13}$$

In the time domain,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V \mathcal{C}^2\left[cos(\omega t+\phi)+i\sin(\omega t+\phi)\right]\left[cos(\omega t+\phi)-i\sin(\omega t+\phi)\right]dVdt. \tag{A.4.14}$$

We are just interested in the real part to find the average power,

$$W_o = \frac{k}{T_o}\int_t^{t+\frac{2\pi}{\omega}}\int_V \mathcal{C}^2 cos^2(\omega t+\phi)dVdt. \tag{A.4.15}$$

To find the energy, we multiply by the period as well. Since the integral of cos^2 over a period is $\frac{1}{2}$, we are left with,

$$W_o = \frac{\pi k}{\omega T_o}\int_V \mathcal{C}^2 dV = \frac{\pi k}{\omega T_o}\int_V \mathcal{C}_1^2+\mathcal{C}_2^2 dVdt. \tag{A.4.16}$$

Substituting for the constants,

$$W_o = \frac{\pi k}{\omega T_o k_3^2}\int_0^h A(z)\left(C_1^2+C_2^2\right)e^{-2\frac{z}{k_3}}dz, \tag{A.4.17}$$

where we have converted the integral over volume to an integral over z. With $A(z) = A_o e^{-\frac{z}{k_3}}$,

$$W_o = \frac{\pi k A_o}{\omega T_o k_3^2}\left(C_1^2 + C_2^2\right)\int_0^h e^{-\frac{z}{k_3}}dz \;\rightarrow\; W_o = \frac{\pi k A_o}{\omega T_o k_3}\left(C_1^2 + C_2^2\right)\left[1 - e^{-\frac{h}{k_3}}\right], \quad \text{(A.4.18)}$$

where $\beta = \frac{2F_A\nu}{EA_o}(3\lambda + 2\mu)\alpha T_o$, $C_1 = \frac{\beta\omega^2\rho c_p}{(\frac{k}{k_3^2})^2 + (\omega\rho c_p)^2}$, and $C_2 = \frac{\beta\frac{k}{k_3^2}\omega}{(\frac{k}{k_3^2})^2 + (\omega\rho c_p)^2}$.

If we integrate from 0 to ∞,

$$W_o = \frac{\pi k A_o}{\omega T_o k_3}\left(C_1^2 + C_2^2\right). \quad \text{(A.4.19)}$$

A.4.2 Low frequency approximation

As discussed in Sect. 3.4.3, vibrothermography leads to a scenario where the frequency is such that the geometric effects can be neglected (the vibration wavelength is large relative to the asperity height). This results in the updated constants $C_1 = \frac{\beta\omega^2\rho c_p}{(\frac{k}{k_3^2})^2} \approx 0$ and $C_2 = \frac{\beta\frac{k}{k_3^2}\omega}{(\frac{k}{k_3^2})^2} = \frac{\beta k_3^2\omega}{k}$. The total energy dissipated becomes,

$$W_o = \frac{\pi k A_o}{\omega T_o k_3}\left(\frac{\beta k_3^2\omega}{k}\right)^2 = \frac{\pi k_3^3\omega T_o}{A_o k}\left(\frac{2F_A\nu}{E}(3\lambda + 2\mu)\alpha\right)^2. \quad \text{(A.4.20)}$$

A.5 Asperity contact

The normal force supported by a distribution of asperities is given by Eq. 3.4.29,

$$P = \pi R c \sigma_y \sum_i (h_i - z) \qquad \text{such that } (h_i - z) > 0. \tag{A.5.1}$$

As discussed in Sect. 3.4.4.4, the summation can be converted to an integral,

$$P = \pi R R_\alpha c \sigma_y \int_z^\infty (h - z) f(h) dh, \tag{A.5.2}$$

where,

$$f(h) = \frac{1}{\sqrt{2\sigma^2 \pi}} e^{-\frac{(h-z)^2}{2\sigma^2}}, \tag{A.5.3}$$

is the probability distribution for a normally distributed random variable. The parameter σ is the standard deviation and μ is the mean. We can split this into two separate integrals,

$$P = \frac{\pi c \sigma_y \alpha}{k} \left[\int_z^\infty h f(h) dh - z \int_z^\infty f(h) dh \right]. \tag{A.5.4}$$

Note that the first integral in the bracket of Eq. A.5.4 resembles the expected value of h, $E(h)$, while the second integral corresponds to the product of the height of the flat surface and the cumulative distribution function of h, $z(1 - F(z))$.

Note that the first integral in the bracket has a lower bound of z. The expected value of h is,

$$E(h) = \int_{-\infty}^\infty h f(h) dh. \tag{A.5.5}$$

If the distribution is truncated at a lower bound z, the expected value changes to,

$$E(h; h > z) = \int_z^\infty h f(h; h > z) dh, \tag{A.5.6}$$

where,

$$f(h; h > z) = \frac{f(h)}{1 - \Phi(\alpha)}, \tag{A.5.7}$$

where $\alpha = \frac{z-\mu}{\sigma}$ [22]. This leaves,

$$P(z) = \pi R R_\alpha c \left[(1 - \Phi(\alpha)) E(h; h > z) - z(1 - F(z)) \right]. \tag{A.5.8}$$

The cumulative distribution is given by,

$$F(z) = \frac{1}{2} \left[1 + erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right]. \tag{A.5.9}$$

The expected value of the truncated normal distribution is [22],

$$E(h; h > z) = \mu + \sigma \frac{\phi(\frac{z-\mu}{\sigma})}{1 - \Phi(\frac{z-\mu}{\sigma})}. \tag{A.5.10}$$

The normal force evaluates to,

$$P(z) = \pi R R_\alpha c \left[\mu(1 - \Phi(\alpha)) + \sigma\phi(\frac{z-\mu}{\sigma}) - z\frac{1}{2} \left(1 - erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right) \right]. \tag{A.5.11}$$

Recall that $\Phi(\alpha) = \frac{1}{2} \left(1 + erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right)$.

$$P(z) = \pi R R_\alpha c \left[\mu\frac{1}{2} \left(1 - erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right) + \sigma\phi(\frac{z-\mu}{\sigma}) - z\frac{1}{2} \left(1 - erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right) \right]. \tag{A.5.12}$$

Grouping terms,

$$P(z) = \pi R R_\alpha c \left[\frac{1}{2}(\mu - z) \left(1 - erf(\frac{z-\mu}{\sigma\sqrt{2}}) \right) + \sigma\phi(\frac{z-\mu}{\sigma}) \right]. \tag{A.5.13}$$

Equivalently,

$$P(z) = \pi R R_\alpha c \sigma_y \left[(\mu - z) (1 - \Phi(\alpha)) + \sigma\phi(\alpha) \right]. \tag{A.5.14}$$