PREDICTING THE FAILURE RATE OF ELECTRICAL CONTACTS BY MODELING THE CONTACT RESISTANCE DISTRIBUTION

by
Robert W. Caven, Jr.

Submitted in Partial Fulfillment of the Requirements
for the Degree of
Master of Science in Engineering
in the
Electrical Engineering
Program

M. Malaki 11/22/89
Adviser Date

Sally M. Hotchkiss November 23, 1989
Dean of the Graduate School Date

YOUNGSTOWN STATE UNIVERSITY
December, 1989

ABSTRACT

PREDICTING THE FAILURE RATE OF ELECTRICAL CONTACTS BY MODELING THE CONTACT RESISTANCE DISTRIBUTION

Robert W. Caven, Jr.

Master of Science in Engineering

Youngstown State University, 1989

Several versions of a computer model were developed to predict contact resistance distributions for electrical contacts. In the models CR1, CR2, and CR3, it was assumed that the surface asperities of a contact metal are spherically shaped and always undergo purely elastic deformation. A comparison with actual gold contacts showed that CR1 and CR3 predicted the resistances of clean contacts reasonably well. The largest percent error between the model prediction and an actual measurement was 78%. The absolute error in this case was only 0.097mΩ. CR2 has not been tested.

For the model CR1P, it was assumed that the spherically-shaped asperities can undergo plastic deformation if the deformation is large enough. CR1P is an important step in modeling the degradation of contacts: thin films that can contaminate a contact surface are broken, and electrical contact made, only where an asperity is plastically deformed.

A failure rate equation and the method of uniform residuals are two methods described in this thesis that can predict the failure rate of contacts based on the contact resistance distribution. Examples of each method are provided.

ACKNOWLEDGEMENTS

The author would like to thank Paul DiLiello and Brian Sabo of the Electrical Contacts Research Lab at Packard Electric Division for their help in making contact resistance measurements. The author would also like to thank Bill Phillips and Dave Bailey of Packard for their technical support, and Phil Engler for his fine drawings in Figure 26. Finally, many thanks are due Dr. J. Jalali, thesis adviser, for his suggestions, help and encouragement.

TABLE OF CONTENTS

I. INTRODUCTION

1.1 Objective 1

1.2 Scope 1

1.3 Historical Background 4

1.4 The Importance of Studying Electrical Contacts 4

1.5 Contact Failure and the Effects of Degraded Contacts 6

1.5.1 Very high contact resistance 7
1.5.2 Noisy contacts 7

1.6 Frequency Distribution of Contact Resistance 8

1.6.1 Constriction resistance 10
1.6.2 Frequency distribution of degraded contacts 11
1.6.3 Currence and Rhoades: Laboratory testing 12
1.6.4 Grau: Field testing 13
1.6.5 Yamada and Rowlands: Quantification of the distribution 15
1.6.6 Contact resistance as a general failure criterion 16

1.7 Future Directions in Contact Modeling 18

II. CONTACT RESISTANCE DISTRIBUTION OF CLEAN CONTACTS USING AN ELASTIC MODEL OF THE INTERFACE 20

2.1 Model of a Metal Surface 20

2.1.1 Description of a metal surface 20
2.1.2 Formula for avg(r) 21
2.1.3 Distribution of the number of asperities 22

2.2 Elastic Contact of Two Solid Spheres 23

2.3 Bumpy Surface Contacting a Smooth Flat Surface 24

2.3.1 Description of the model 24
2.3.2 Algorithm for calculating the separation of the contact members 26

2.4 Two Bumpy Surfaces in Contact with Each Other 27
2.5 Dimple in Contact with a Bumpy Flat Surface 31
2.6 Calculating the Constriction Resistance 32
2.6.1 The contact spots 32
2.6.2 Solution using Laplace's equation 32
2.6.3 Approx. constriction resistance calculation 39
2.7 Algorithms for Plotting the Cumulative Distribution of Contacts 40
2.8 Results from a Few Sample Runs 42
2.9 Verification of the Model for Gold Contacts 48

III. PREDICTING FAILURE RATE BASED ON CONTACT RESISTANCE DISTRIBUTION 56
3.1 Method I: Threshold Resistance 56
3.1.1 Threshold resistance as a failure criterion 56
3.1.2 Designed experiment for obtaining equations for μ and σ 60
3.1.3 Example: Fraction failed using threshold resistance as failure 68
3.2 Method II: Uniform Residuals 69

IV. PLASTIC DEFORMATION OF ASPERITIES 75
4.1 The Significance of a Plastic Deformation Model 75
4.2 Description of Plastic Deformation 76
4.3 Huber-von Mises Equation: Criterion for Plastic Deformation 78
4.3.1 Principal stresses 78
4.3.2 Development of the basic equation 78
4.3.3 Principal shear stresses 80
4.4 Contact Pressure and Deformation Required for Plastic Yielding 81
4.4.1 Hardness as a function of yield stress 81
4.4.2 Maximum shear stress under an elastic indentation 83
4.4.3 Direct yield stress and contact pressure 85

4.4.4 Contact pressure required for plastic deformation 86
4.4.5 Greenwood and Williamson's derivation of the formula for δ^* 87

4.5 Greenwood and Williamson's Elastic Deformation Model 89

4.5.1 Surface model 89
4.5.2 Plasticity Index 92
4.5.3 Load supported by a plastically deformed asperity 94
4.5.4 Contact area vs. load for plastic and elastic surfaces 95
4.5.5 Contact resistance 97
4.5.6 Summary of Greenwood and Williamson's model 100

4.6 Results from CR1, CR1P, and Greenwood and Williamson 100

V. SUMMARY AND CONCLUSIONS 107

5.1 Summary 107

5.2 Conclusions 108

5.3 Recommendations 111

Appendix A: Failure Rate and System Reliability 113

Appendix B: Generation of Random Numbers having a Specific Distribution 117

Appendix C: Derivation of Greenwood's Potential Equation 122

Appendix D: Normal Probability Plot 125

Appendix E: Principal Shear Stress Example 130

Appendix F: Calculation of Gamma-Function Related Integral) 134

Notes 136

LIST OF FIGURES

1. A typical connector used in an automobile 2
2. (a) Actual vs. apparent contact area
   (b) Constriction of current through a contact spot 10
3. Histograms of contact resistance 12
4. Distribution plots of tin-tin contacts after fretting 13
5. NPP of contact resistance 14
6. Standard deviation of a segment of an NPP 15
7. Control charts for contact resistance 16
8. Model of a metal surface 20
9. Slopes on surfaces and angle α 21
10. Elastic contact between two spheres 24
11. Deformation of an asperity 26
12. Two asperities in contact 28
13. Deformation of contacting asperities 30
14. Vertical component of asperity load 31
15. Contact spots, smooth flat surface on bumpy surface 33
16. Contact spots, smooth dimple on bumpy surface 34
17. Grid point in the two-dimensional contact problem 35
18. An insulating boundary 36
19. Potential contour plots for a 2D contact 37
20. A box drawn on the contour plot of ϕ 38
21. Contact resistance NPP, smooth flat vs. bumpy flat 44
22. Contact resistance NPP, 1.0mm radius dimple vs. flat 45
23. Contact resistance NPP, 3.0mm radius dimple vs. flat 46
24. Contact resistance NPP, 5.0mm radius dimple vs. flat 46
25. Average resistance vs. force, dimples and flats 47

26. Experimental hardware for gold contact resistance measurements 49
27. NPP of gold contacts, 1.6mm round rod 52
28. Mean and median contact resistance of gold, round rod 53
29. NPP of gold contacts, flat rod vs. flat 54
30. Mean and median contact resistance of gold, flat rod 55
31. Average contact resistance at different levels of force 64
32. Average contact resistance at different levels of ρ 64
33. Average contact resistance at different levels of r/h 65
34. Average contact resistance at different levels of "shape" 65
35. Average contact resistance at different levels of d 66
36. Average contact resistance at different levels of E 66
37. Average contact resistance at different levels of avg(h) 67
38. Average contact resistance at different levels of A_{app} 67
39. Normal distribution 69
40. Student's "t"-distribution 72
41. Graphical representation of Method of Uniform Residuals 74
42. Edge dislocation and the slip vector 77
43. Principal stresses and reduced stresses on a cube 79
44. Plastic deformation along the x-axis 80
45. Indentation made by a hard ball in a metal slab 82
46. Coordinate system and stresses in Davies' problem 84
47. Principal shear stresses at ζ 86
48. Full plastic deformation of an asperity 87
49. Height and separation 89
50. Plastic deformation of an asperity 91
51. Stress vs. strain 92
52. Metal surface with large σ_h/r 93
53. Metal surface with small σ_h/r 93
54. Contact area and no. of contact points vs. contact force 104

55. Contact resistance vs. contact force 105

56. Contact resistance vs. force: CR1 model vs. Greenwood and Williamson's model 106

LIST OF SYMBOLS

a = radius of a contact spot, m

avg(···) = average of ···

A_{app} = apparent contact area, m^2

A_t = total actual contact area, m^2

A_{tp} = total contact area under plastic deformation, m^2

$\underline{b}$ = slip vector, m

D = (x,y) component of distance between the centers of two asperites, m

d = width of a rectangular "box" on a potential contour plot, m

d = distance between the centers of contact spots, m

d = bump density, number of bumps per mm^2

E = Young's modulus, Pa

E' = $E/2(1-\mu^2)$, Pa

F = applied normal contact force, N

h = height of an asperity, m

H = hardness, Pa (published tables use units of kg/mm^2)

I = electrical current, A

K = unknown constant

l = length of a rectangular "box" on a potential contour plot, m

length = length of a rectangular contact surface, m

max(···) = maximum value of ···

n = a specific value of the random variable N

n_b = total number of asperities on a surface (subscript "b" = "bumps")

n_c = number of asperities in contact

n_{pc} = number of plastically deformed asperities in contact

N = number of asperities on a base, a random variable

p, prob = probability

P = contact load or contact force supported by asperities

P_t = total load supported by asperities; at equilibrium, $P_t = F$.

p = principal stress, Pa

p_h = hydrostatic pressure

q = pressure, Pa

r = distance in Davies' coordinate system (see Fig. 46), m

r = radius of curvature of an asperity, m

r/h = avg(r)/avg(h), used as an abbreviation in Chapter III.

R = resistance, contact resistance, Ω

Rci = initial contact resistance, Ω

Rcmin = theoretical minimum constriction resistance, Ω

ΔR = change in (contact) resistance, Ω

s = separation of the bases of opposing contact surfaces, m

s = sample standard deviation

s' = standardized value of s, $s' = s/\sigma_h$, (dimensionless)

U = number of (x,y) locations on a base

V = applied voltage across contact members, V

var(···) = variance of ···

width = width of a rectangular contact surface, m

Xx, Yy, Zz = normal stress components parallel to x, y, z axes, Pa

Xp = number of standard deviations from the mean, assuming gaussian distribution

Y = yield stress, Pa

Yz = shear stress component parallel to y-axis and acting in x-y plane, Pa

Z = Bernoulli random variable, equal to 0 or 1

z' = standardized value of z, $z' = z/\sigma_h$, (dimensionless)

α = contact angle between an asperity and its base surface

β = contact angle between two opposing asperities (see Fig. 12)

δ = deformation, m

δ^* = deformation at which plastic deformation commences, m

Δx, Δy, Δz = change in x, y, z

ϵ = error

ζ = point of maximum difference between principal stresses

θ = angle in Davies' coordinate system (see Fig. 46)

μ = mean

μ = Poisson's ratio, dimensionless

ρ = electrical resistivity, $\Omega \cdot m$

σ = standard deviation

σ_h = standard deviation of asperity heights, m

ϕ = electric potential, V

$\phi(z)$ = probability density function of z

ψ = plasticity index, dimensionless

subscripts e, w, n, s = the grid point right, left, up, down respectively

subscript m = mean or average

subscript p = plastic deformation

superscript v = vertical component

∇ = del operator, $(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$

CHAPTER I.

INTRODUCTION

1.1 Objective

A reliable electrical or electronic system requires reliable components, the most numerous of which are electrical contacts. Because electrical and electronic systems are becoming more complex, the need for a tool to aid in the design of reliable electrical contacts is crucial. Thus the author proposed the development of a computer model to predict the failure rate (see Appendix A) of electrical contacts. This thesis presents one phase of the model development: predicting the contact resistance distribution of butt contacts.

1.2 Scope

An electrical contact is a releasable junction which is apt to carry electric current. [1] A connector is a type of contact designed for no relative movement between the contact members while the contact members are mated (i.e. connected), and the contact members are designed to be separated only when there is no current flowing. [2] An example is shown in Fig. 1. One type of connector, called a butt contact, is designed so that there is no sliding of the contact surfaces while the contact members are being mated. Butt contacts are the simplest to model, and thus are the main focus of this thesis.

Electrical contacts that are not classified as connectors include switches, relays, slip rings, commutator/brush assemblies, circuit breakers, etc. These devices require modeling of such phenomena as

arcing and material transfer, and are beyond the scope of this thesis. Therefore, with the exception of some isolated examples in the introduction, they are generally excluded from the discussions.

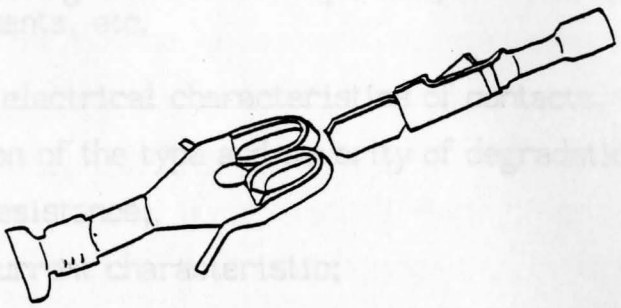


Fig. 1.--A typical connector used in an automobile.

If a contact is damaged during assembly, it will probably fail in service. However, this research is concerned with the more interesting and more common cause of contact failure: failure brought on by some sort of degradation. Degradation can be classified into two groups: (1) mechanical degradation, such as a reduction in contact force due to relaxation of the metal; and (2) chemical degradation, resulting in contamination of the contact interface. Degradation does not always lead to contact failure; the reader must keep in mind that the influential factors that constitute contact failure depend on the circuit in which the contact functions.

The phases in predicting failure rate of a contact design are:

1. Understanding the mechanisms of contact degradation, such as:
   a. stress relaxation of metal, which can reduce the contact normal force;
   b. growth of tarnish films on the contact interface, which affects the flow of current through the interface;
   c. growth of organic films on the contact interface, which also affects the flow of current
2. Predicting the probability and severity of each type of degradation versus age of the contact, such as:

a. stress relaxation versus age and temperature;

b. metal corrosion versus age, temperature, metal composition, corrosive agents, and humidity;

c. organic film growth versus age, temperature, type of organic contaminants, etc.

3. Predicting the electrical characteristics of contacts, such as a, b, c below, as a function of the type and severity of degradation:

a. contact resistance;

b. voltage-current characteristic;

c. number and duration of intermittent open circuit "events."

4. Correlating the electrical characteristics to the age of the contact for each type of degradation. This is a problem in probability and statistics because of the probabilistic nature of Phase 2.

5. Given the circuit requirements, defining failure criteria in terms of the electrical characteristics of the contact; then predicting the probability of failure given the type of degradation and the age of the contact. From these data, failure rate can be calculated. Failure depends on the circuit in which the contact functions. However, "generic failure" can be defined in terms of: increase in contact resistance; duration of intermittent open circuit events; or noise produced by a nonlinear voltage vs. current characteristic.

This thesis addresses Phase 3, prediction of electrical characteristics. Among these, contact resistance is the most fundamental and important. It is used to predict the onset of other abnormal behavior; some authors even define failure in terms of contact resistance irrespective of the application. Predicting the contact resistance of "new" or "clean" contacts became the primary focus of this thesis. Prediction of failure rate is an area of research suggested for future study. Chapter III provides a starting point for this research.

1.3 Historical Background

The pioneer of the study of electrical contacts was the Swedish physicist Dr. Ragnar Holm (b.1879-d.1970), who worked at Siemens-Berlin and later at the Stackpole Carbon Company in Pennsylvania. His 1967 book, Electric Contacts (which is a revised version of his 1958 book, Electric Contacts Handbook), is a classic work covering the basics of contact phenomena; it will be referenced many times in this thesis. An annual meeting, The IEEE Holm Conference on Electrical Contacts, developed in part by Holm in 1953, is the forum for the exchange of new technical information in the field of electrical contacts. It was named in honor of Holm in 1968, and fell under the jurisdiction of the IEEE Components, Hybrids, and Manufacturing Technology Society in 1984. The growth and success of this conference is proof that, as electrical and electronic systems become more sophisticated, the study of electrical contacts becomes more important. [3]

1.4 The Importance of Studying Electrical Contacts

Now one may well ask "why put so much effort into studying such a simple device as an electrical contact?" The answer is that electrical contacts are not simple, and designing a reliable one is not a trivial task.

In 1987, Dr. Brian Williamson, formerly of Cambridge University, England, and the recipient of the 1981 Ragnar Holm Scientific Achievement Award, called electrical contacts "the most failure-prone devices in electrical engineering" and added, "if Star Wars ever gets going, I guarantee it will be an electrical connector that fails." He backed up his claims by quoting from an article written by an automotive expert: "ten percent of all failures in automobiles are ultimately traced to connectors." Then, with his characteristically-

British dry sense of humour, Williamson remarked, "You press two metals together and—by the grace of God—they will continue to pass electrons...Any fool can make a connective device work; what separates a good vendor from a not-so-good one is: will it continue to work?" [4]

Improving performance and reliability of electrical and electronic equipment is obviously the benefit of understanding contacts. But to convince an engineer of the need for a comprehensive study, it is necessary to make known the extent to which electrical and electronic systems depend on contacts. In 1987, Dr. Williamson stated that "there are more connective devices than any other discrete electrical component." [5] The following examples help to substantiate his claim:

In 1957, Professor Frank Llewellyn-Jones of the University of Swansea, Wales reported that a local telephone call in the London area involved as many as 500 contacts, and that the British Post Office telephone system contained "something like 300 million contacts." [6]

Harold Wagar, of Bell Telephone Laboratories (Madison, NJ) presented a more recent example (1976) which hints at how electrical systems are becoming more sophisticated.

> ...during one single typical interoffice [telephone] call, some 1200 relays, or about 12,000 contacts, are called into action. In addition, hundreds of other contacts are in the subscriber's path, or in paths which will be needed on another call. [7]

He listed the number of contacts in use in the Bell System in 1975. The bottom line was: about 1100 contacts per subscriber, with 2400 being added to the system for every new subscriber. An abridged version of his table is shown in Table 1. [8]

Table 1 is based on 67 million subscribers in service in 1975. (The heading "connectors" is apparently a less general category than the one defined in the Objective and Scope section of this thesis.

According to that definition, twisted wires, plug and jack, wire splicing sleeves, and solderless wire wrap would all fall under the category "connectors.")

In 1975, Jack Wimsey of Wright-Patterson Air Force Base in Dayton, Ohio described the electrical system of a portable, deployable-anywhere air base capable of accommodating five squadrons and 4500 personnel. The base required, among other things, 20,550 Class L 60-amp connectors. Each connector contains five contact members. [9]

TABLE 1

TOTAL NUMBER OF CONTACTS, BY TYPE, IN THE BELL SYSTEM

Twisted Wires [a]	1.5 Billion
Plug and Jack [a]	1.8 Billion
Wire Splicing Sleeves [a]	5.0 Billion
Connectors [b]	9.0 Billion
Solderless Wire Wrap	30.0 Billion
Relays	16.9 Billion

[a] The author would classify these as connectors.

[b] Does not include the devices covered in note (a).

Closer to home, a 1948 Chevrolet Sedan contained 75 automotive terminals; a fully equipped 1986 Buick Riviera contains 1610 terminals. These numbers do not include switch contacts, motor brushes/commutators, nor relay contacts.

1.5 Contact Failure and the Effects of Degraded Contacts

Knowing exactly what has caused a contact to fail can help the systems engineer prevent the same problem in the future. This

section will present several examples of contact failure, and the degradation mechanism to which failure was attributed.

1.5.1 Very High Contact Resistance

Contact resistance is the resistance added to an electrical path by the presence of an electrical contact interface. Two obvious sources of contact resistance are: mechanical defects which result in a loss of intimate contact; and contamination of the contact surfaces by corrosion, organic films, dirt, and dust.

Everyone has experienced equipment failure which could be traced directly to a corroded or dirty electrical contact. For example, the author recently lost water pressure in his home because the water pump switch had severely corroded contacts.

Sometimes high-resistance contact failure is catastrophic; for example one may have heard that aluminum wiring may cause fires in buildings. The problem, however, is not with the aluminum wires themselves; rather, it has been traced to aluminum wire splices made with twist-on connectors. [10] Aluminum, as it turns out, develops a very tough oxide on its surface when exposed to air. The oxide, which reduces the metal-to-metal contact area, is difficult to break and causes the contact resistance to be quite high. The result is overheating of the contact and the potential for a fire. [11]

1.5.2 Noisy Contacts

In the October, 1984 issue of Communication International E.M. Edwards of Celwave Corp. wrote a brief article titled "Connector Problems in Two-Way Antennas." He wrote

> Like any other system, a two-way radio system is susceptible to degradation. When it is checked by the technician the underlying reason is often found to be the

same old problem; the dodgy connection.

> Symptoms of connector problems are noisy duplex operation, loss of receiver sensitivity, medium to high transmitter standing wave ratio. The problem: water gets into a connector interface and contaminates the interface or may cause the connectors to rupture if they are subjected to extremely cold temperatures. [12]

The connector in question was the u.h.f. connector, or PL259, which was chosen as the industry standard in the early days of two-way radio. Failure of the connector could be defined quantitatively in terms of a noise level, a sensitivity level, or a SWR level; Edwards did not go into this. He explained that the apparently simple solution, keeping the connector dry, is not at all easy to do, and went on to explain how to do it.

Another description of the noise problem is presented in a 1957 book by Llewellyn-Jones, who wrote

> Metallic contact...may be hampered by tarnish films, which may be conducting, semi-conducting, or insulating. A semi-conducting film may give rise to a rectifying action which can distort the wave-form of the current passed. [13]

Unfortunately, he did not give any specific examples.

1.6 Frequency Distribution of Contact Resistance

The first and most important model developed for this project was a computer prediction of the frequency distribution of contact resistance as a function of age of the contact for different contact designs and applications. Strictly speaking, a frequency distribution is "a function which measures the relative frequency or probability that a [random] variable can take on a set of values." [14] It gives the probabilities of a single contact having certain resistances; it also tells how the contact resistances of a population of contacts will be

distributed. In practice, the frequency distribution is approximated by "sampling," i.e., by measuring the contact resistance of each of several contacts in a sample group, then plotting their contact resistances as either a histogram or a normal probability plot. (A normal probability plot (NPP) is a cumulative frequency distribution scaled such that data following a gaussian distribution fall on a straight line; see Appendix D.)

The "sampled" distribution is a valid approximation of a true frequency distribution only if all contacts in the sample group are of identical design and have nearly the same age and application. In the author's opinion, the "sampled" distribution is only useful insofar as it approximates the true frequency distribution.

Contact resistance is a useful diagnostic tool for measuring the progress of degradation of a contact, and the likelihood that the contact will soon fail. It will be seen in this section that a high resistance 'tail' in the frequency distribution is a sign of degradation. The symptoms —which include unstable contact resistance, nonlinear voltage vs. current behavior, intermittent open circuits, and others—may be a sign of imminent failure if they do not themselves constitute failure.

The computer prediction spoken of in this thesis takes the practical approach by simulating a "sample group" of identical contacts and "measuring" the contact resistance of each.

A Monte Carlo approach was used for the simulation because many of the variables affecting contact resistance are random; for it is well established that among a sample of real contacts, no two will have identical contact resistances — even if they are all of the same design and used in the same application. It is also well known that contact resistance can randomly vary with time.

This section will present some examples of how researchers have used contact resistance distribution to detect contact pathologies and to quantify the health of contacts.

1.6.1 Constriction Resistance

First it is necessary to appreciate the fact that even clean contacts have some contact resistance; it is called constriction resistance, and it generally varies from contact to contact, even if the contacts have identical designs.

Constriction resistance is due to the fact that, when two contact members are in contact, they are actually in contact in only a few spots. That is because even the smoothest polished surface is quite bumpy on the micron scale. Were it not for this fact, the second basic law of friction--that frictional resistance is independent of the apparent area of contact of the sliding surfaces [15] --would not hold. Referring to Fig. 2(a), one can see that the actual contact area is a very small fraction (perhaps 1/1000) of the apparent contact area. [16,17] When contact is only at a few discrete spots, the current flowing from one contact member to the other must bend its path to get through these spots. The bending (causing a lengthening of the current path) and the reduced conduction area are what cause the resistance; Fig. 2(b). [18] The value of the constriction resistance can be calculated using Laplace's equation $\nabla^2\phi = 0$; constriction resistance depends on the number, sizes, and locations of the contact spots and on the material of the contact members.

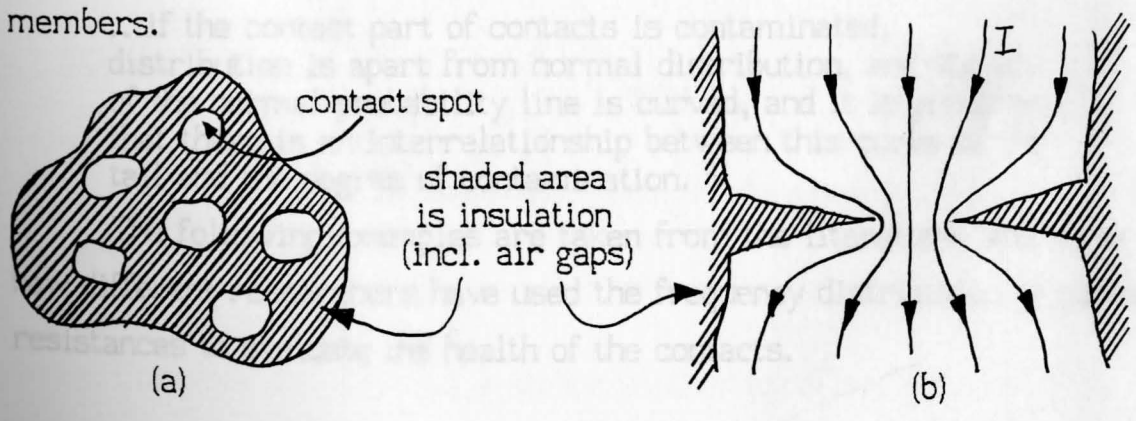


Fig. 2.--(a) Actual vs. apparent contact area.
(b) Constriction of current through a contact spot.

Constriction resistance varies from contact to contact because no two metal surfaces are identical; each has its own topography and therefore each contact interface has its own arrangement of contact spots.

Yamada et al. described the frequency distribution of clean contacts:

> If the contact part of the contacts is clean, distribution of contact resistances approaches normal distribution and the normal probability curve approximates a straight line. [19]

1.6.2 Frequency Distribution of Degraded Contacts

The total contact resistance of a "dirty" contact is the sum of constriction resistance and film resistance. [20] Films are of two types: tarnish films such as oxides and sulfides; and organic films which are produced at the contact interface from organic contaminants such as processing lubricants, splashed oil, and outgassed polymers. Films grow over the life of the contact depending on the environment and any contaminating substances (including humidity in the air) present.

Yamada explained the effect of films on the distribution of contact resistance:

> ...if the contact part of contacts is contaminated, distribution is apart from normal distribution, and the tail of the normal probability line is curved, and it is presumed that there is an interrelationship between this curve of the tail and the degree of contamination. [21]

The following examples are taken from the literature, and show how different researchers have used the frequency distribution of contact resistances to evaluate the health of the contacts.

1.6.3 Currence and Rhoades: Laboratory testing

Figure 3 shows how "aging" affects contact resistance. Currence and Rhoades (1983) measured the contact resistance of a sample of identical contacts, the design being a tin-plated member mated to a gold-plated member. The contact resistance of each contact was measured twice: The first measurement was made immediately after the initial mating of the contact members; the distribution is plotted in Fig. 3(a). The second measurement was made after unmating and remating the contact members nine more times; the distribution is plotted in Fig. 3(b).

In Fig. 3(a), the distribution is tight, and the contact resistance relatively low. The slight dispersion can be attributed to the small differences in constriction resistance (no two surfaces are exactly the same) and the unavoidable contamination (albeit small amounts) that appears even on "new" contact surfaces. In Fig. 3(b) the distribution is more spread out, and this may be attributed to increased corrosion as the protective plating wears out and possibly other phenomena.

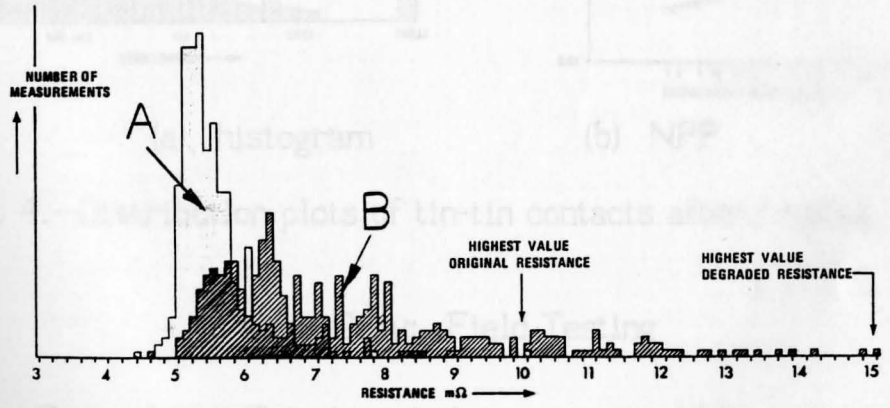


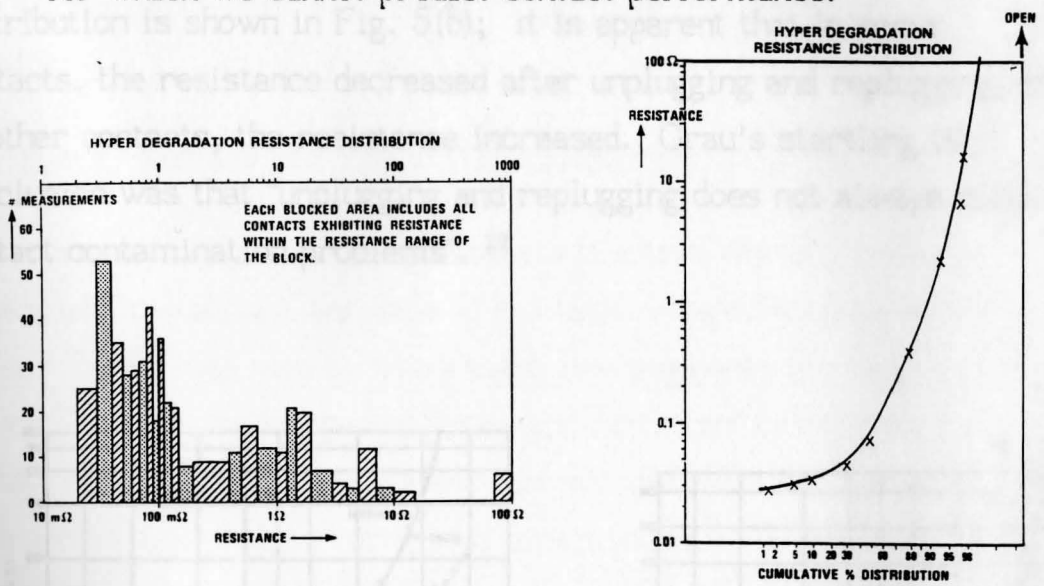
Fig. 3.--Histograms of contact resistance. [22]

Currence and Rhoades describe another experiment in which a sample of tin-plated vs. tin-plated contacts were subjected to simulated fretting. Fretting is a small amplitude (a few microns) cyclic rubbing

action between the two contact members. After perhaps several thousand fretting cycles, the contact resistance of each contact was measured. The histogram is shown in Fig. 4(a). A normal probability plot (NPP) of the same data is shown in Fig. 4(b).

A high-resistance "tail" is evident in the aged contacts. According to Currence and Rhoades:

> The "tail" is an indication that the contact system is strongly affected by a substantial degradation mechanism such as film growth, corrosion effects or surface contamination. This "tail" identifies a performance region for which we cannot predict contact performance. [23]



(a) histogram (b) NPP

Fig. 4.--Distribution plots of tin-tin contacts after fretting. [24]

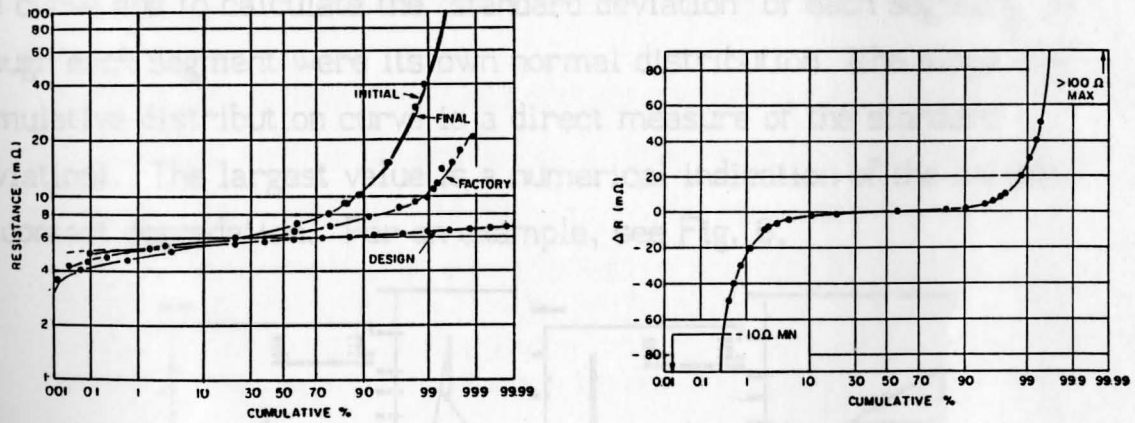
1.6.4 Grau: Field Testing

Tom Grau of Bell Telephone Laboratories in 1978 described a field study of a gold-plated blade mated to a tuning fork terminal used in office electronic equipment. He visited several cities to make measurements on contacts that were still in the equipment. In his words

> When we arrived at the field site, [equipment] power was turned off. We made...resistance measurements without disturbing the contacts, and obtained the distribution called 'initial' [Fig. 5(a)]...Then we unplugged the circuit pack...and remeasured resistance to obtain the distribution called 'final'. [25]

Notice the high-resistance tail on the initial and final curves. Grau attributed the tail to contamination from a manufacturing lubricant and a printed circuit board lacquer. [26]

A distribution called "delta R" was obtained by subtracting the final resistance of a contact from its initial resistance. This distribution is shown in Fig. 5(b); it is apparent that in some contacts, the resistance decreased after unplugging and replugging, and in other contacts, the resistance increased. Grau's startling (sic) conclusion was that "unplugging and replugging does not always solve contact contamination problems". [27]



(a) Contact Resistance (b) Delta R

Fig. 5.--NPP's of contact resistance. [28]

Grau also measured the contact resistance of a batch of factory-fresh contacts. He states

> These connectors were manufactured, placed on printed circuit boards, and were ready to be shipped to the fields to be installed in new equipment. [29]

The distribution is labeled 'factory' in Fig. 5(a).

Finally, Grau used a Monte Carlo technique to generate a simulated distribution of factory-fresh contact resistance (labeled 'design' in Fig. 5(a)). He states

> To use the method, each part dimension is assumed to come from some distribution; e.g., normal, uniform, log normal, etc., as are material resistivities. All variables are mathematically combined to produce one...resistance value. [30]

The Monte Carlo method is basic to the modeling described in this thesis.

1.6.5 Yamada and Rowlands: Quantification of the Distribution

Yamada et al. of the Matsushita Electric Works, Japan, attempted to quantify the slope of the tail in the cumulative distribution curve. Their idea was to fit straight-line segments through the tail of the curve and to calculate the "standard deviation" of each segment, as though each segment were its own normal distribution (the slope of a cumulative distribution curve is a direct measure of the standard deviation). The largest value is a numerical indication of the severity of contact degradation. For an example, see Fig. 6.

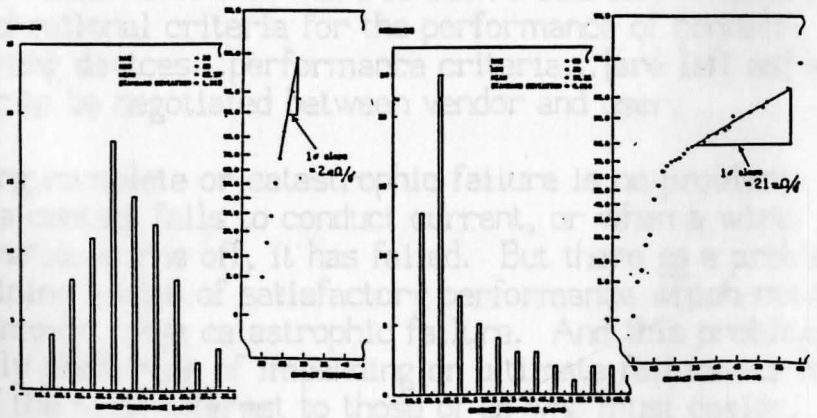
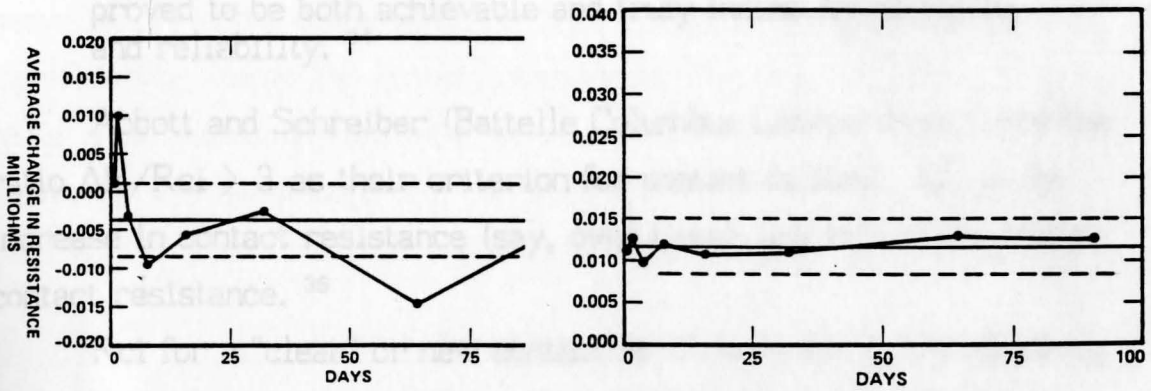


Fig. 6.--Standard deviation of a segment of an NPP. [31]

E. Rowlands of TRW applied statistical process control (SPC) techniques to monitor contact resistance distributions. He created control charts, plotting average contact resistance and standard deviation as a function of time. He also applied the criterion of an upper control limit to find out when the sample goes "out of control." For an example, see Fig. 7.



(a) mean (b) standard deviation

Fig. 7.--Control charts for contact resistance. [32]

1.6.6 Contact Resistance as a General Failure Criterion

Some authors have even used contact resistance as a general failure criterion, even if failure is not a direct result of an increase in contact resistance. Whitley and Malucci (AMP Inc.) said

> ...it is remarkable that there is such a scarcity of clearly-defined rational criteria for the performance of contact-containing devices...performance criteria...[are left as] a matter to be negotiated between vendor and user.
>
> Defining complete or catastrophic failure is no problem. When a contact fails to conduct current, or when a wire termination burns off, it has failed. But there is a problem in defining limits of satisfactory performance which may be far removed from catastrophic failure. And this problem—of early prediction of impending or ultimate failure—is the one of the most interest to those of us who must design, test, and use contacts and connectors. [33]

They used the ratio R/Rcmin > 3 as a criterion for contact failure. Here, Rcmin is the theoretical minimum constriction resistance (see section 1.6.1), and R is the actual contact resistance. They said

> This criterion, which was at first vigorously opposed by both customer and producer (for opposite reasons), has proved to be both achievable and truly indicative of quality and reliability. [34]

Abbott and Schreiber (Battelle Columbus Laboratories) used the ratio ΔR/Rci > 3 as their criterion for contact failure. ΔR is the increase in contact resistance (say, over time), and Rci is the initial contact resistance. [35]

Rci for a "clean" or new contact is of the order 1-3 milliohms, meaning that a contact resistance increase of about 10 milliohms is a sign of impending failure.

Abbott said

> Milliohm level increases...can seldom account for an actual contact or device failure. This is particularly true in modern, low-current digital circuitry. It is often argued that since (1) circuits will often tolerate hundreds or thousands of ohms, and/or (2) high open circuit voltages, even 5 volts, may cause films to be electrically destroyed, small (milliohm) increases are not significant.
>
> Field experience, however, has shown that perhaps the most common mode of contact failure fall in the intermittent category characterized by random, infrequent "events." [36]

Abbott and Schreiber found from lab testing that ΔR/Rci > 3 implies a "sharply increasing probability for observing short-duration 'opens' at typical TTL logic levels." [37] Abbott found a direct correlation between ΔR/Rci and the number and duration of the intermittent open circuit events. [38]

From a frequency distribution of contact resistances, one can see how many contacts among the group have "failed," using a criterion such as one of those summarized above. Possibly, one can quantify the probability of failure for the given design, application, and age of the contact. However, one must be cautious when measuring real contacts: a contact exhibiting abnormally high contact resistance is usually unstable. A slight disturbance such as attaching ohmmeter clips can significantly change the contact resistance; meaning that several low resistance measurements may have come from failed, but disturbed, contacts. One advantage of a computer simulation is that there is no need to worry about disturbing the contacts.

1.7 Future Directions in Contact Modeling

The models described in this thesis provide a steppingstone for more complete and elaborate models. Later models will attempt to simulate:

1) Loss of contact force with time, due to stress relaxation, etc.
2) Corrosion and organic films on the contact interface.
3) Dust particles and other debris on the interface.
4) Contact fretting (cyclic rubbing of the contact surfaces).

Also, a further study must be made of alternatives to the surface model outlined in section 2.1 of this thesis. Especially important is the modeling of contacts in which mating involves sliding of one contact surface over another (non-butt contacts). When two contact members slide over each other, the tops of the asperities fracture and the contact surface topography changes. And, films grown on the surfaces are pushed away.

Modeling of electrical characteristics other than contact

resistance is another area to be considered in future contact models. Any nonlinear effects that may degrade the performance of electronic circuitry must be predictable by the contact model.

Verification of all models is necessary as well. This will require measurements of actual contacts in various stages of degradation.

CHAPTER II.

CONTACT RESISTANCE DISTRIBUTION OF CLEAN CONTACTS USING AN ELASTIC MODEL OF THE INTERFACE

This chapter describes a series of computer programs, called "CR1," "CR2," and "CR3," that were written to predict the distribution of contact resistances for a group of "clean" electrical contacts. The models assume that the asperities or "bumps" on the contacting surfaces are spherically shaped and subject to pure elastic deformation. In Chapter IV, plastic deformation of the asperities is considered.

2.1 Model of a Metal Surface

2.1.1 Description of a Metal Surface

The "bumpy" surface of a metal was modeled as a flat smooth "base" dotted with spherical asperities whose heights and radii of curvature are random (Fig. 8). The number of bumps and location of bumps are also random.

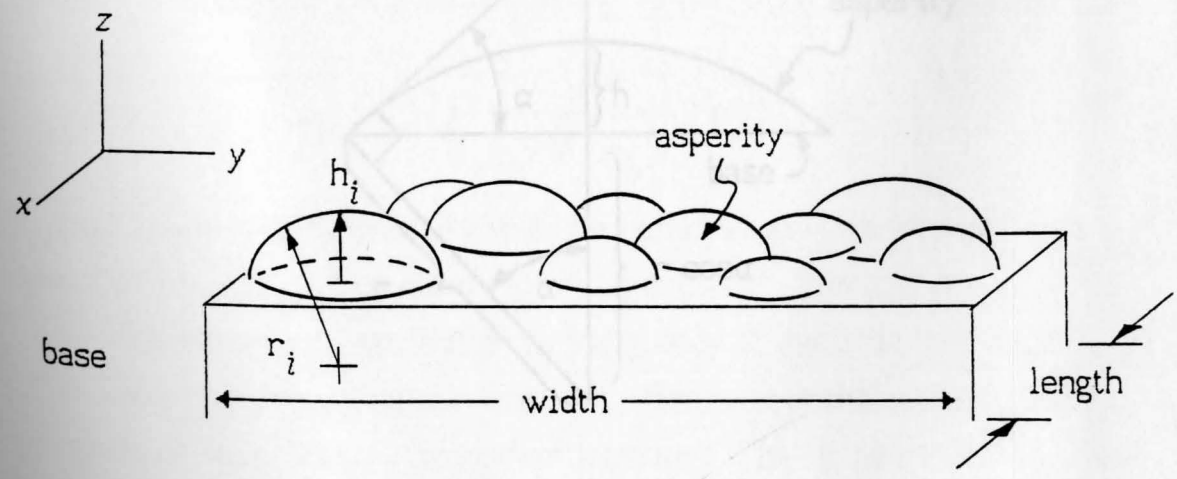


Fig. 8.--Model of a metal surface

The height of an asperity is a gaussian random number [39,40] with mean avg(h) and variance var(h). The radius of curvature of an asperity is a gaussian random number with mean avg(r) and variance var(r).

2.1.2 Formula for avg(r)

According to Williamson and Hunt, "the slopes on surfaces rarely exceed ten degrees." [41] To approximate this surface feature, α, the maximum angle which the bottom of the asperity makes with the flat base on which the asperity sits, was chosen as ten degrees (Fig. 9). This leads to a formula for avg(r):

$$h = r - r\cos\alpha = r(1-\cos\alpha)$$

$$h/r = 1-\cos\alpha$$

$$r = h/(1-\cos\alpha)$$

for $\alpha \leq 10^0$:

$$\text{let} \qquad \text{avg}(r) \geq \frac{\text{avg}(h)}{1-\cos\alpha} = \frac{\text{avg}(h)}{1-\cos 10} \simeq 66\,\text{avg}(h) \qquad (1)$$

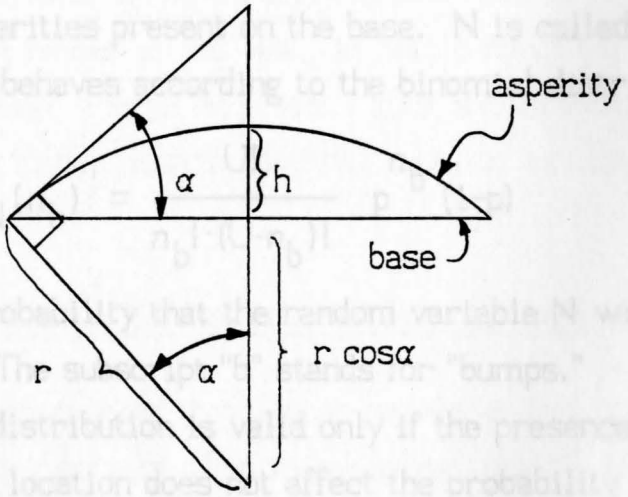


Fig. 9.--Slopes on surfaces and angle α.

2.1.3 Distribution of the Number of Asperities

The author assumed that the number of asperities on the bumpy surface was a Poisson random number with a "density" of d asperities per mm^2. The reasoning for using a Poisson distribution was as follows.

The presence of an asperity at a location (x_i, y_i) on the flat base is given a probability p, $0 \le p \le 1$. The absence of an asperity at that location is given the probability 1-p. The random variable Z, "presence of an asperity," is called a Bernoulli random variable. Now pick U locations on the base, (x_1,y_1), (x_2,y_2),..., (x_U,y_U).

At each location, the presence of an asperity is modeled by a Bernoulli random variable Z. Assign to Z the value "one" if an asperity is present at (x_i, y_i), and "zero" if an asperity is not present there. The probability of an asperity being present is p at all locations. Now define a random variable N such that

$$N = \sum_{i=1}^{U} Z_i$$

N will have an integer value between zero and U, and is a record of the number of asperities present on the base. N is called a binomial random variable and behaves according to the binomial distribution:

$$p_N(n_b) = \frac{U!}{n_b! \cdot (U-n_b)!} \; p^{n_b} (1-p)$$

$p_N(n_b)$ reads "the probability that the random variable N will equal the specific value n_b." The subscript "b" stands for "bumps."

The binomial distribution is valid only if the presence or absence of an asperity at one location does not affect the probability of the presence of asperities at any other location. The bumps were assumed to be "spread out" enough so that they were more or less independent.

A base having a finite area contains an infinite number of locations (x_i, y_i). Since this base will contain a finite number of asperities, then the probability of an asperity being present at any specific location approaches zero. Letting

$$U \rightarrow \infty$$
$$p \rightarrow 0$$

but allowing $U \cdot p$ to be finite, then

$$p_N(n_b) \rightarrow \frac{(U \cdot p)^{n_b} \cdot e^{-(U \cdot p)}}{n_b!} \qquad (2)$$

Equation (2) is called the Poisson distribution. [42]

$U \cdot p$ is the average or "expected" number of asperities on the base. It can be calculated from the formula

$$U \cdot p = d \cdot (\text{area of surface})$$

where d is the average or expected density of asperities.

The locations of the N asperities are uniform random numbers, because there is no reason to believe that the asperities will be clustered in any peculiar way. That is, one should not expect the probability of the existence of an asperity to decrease as its location moves away from the center of the base, as would happen if the locations were gaussian random numbers. One should expect the density of asperities to be uniform over the base, in the long run at least.

2.2 Elastic Contact of Two Solid Spheres

The formulas describing the elastic contact of two solid spheres were used to model the contact of opposing asperities. These formulas were derived by Hertz and are quoted in Holm and Roark. [43,44] They are given in equations (3) and (4) with reference to Fig. 10.

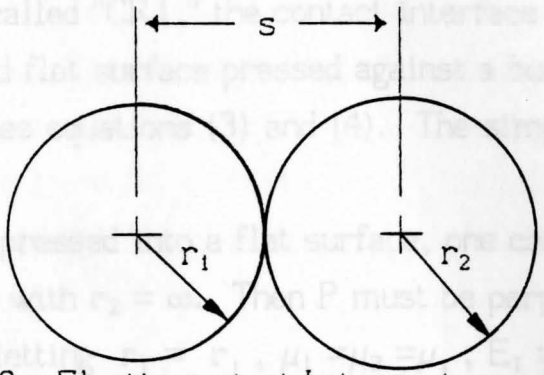


Fig. 10.--Elastic contact between two spheres.

$$a = \left[\frac{3}{4} \cdot P \cdot \left(\frac{1-\mu_1^2}{E_1} + \frac{1-\mu_2^2}{E_2} \right) \cdot \left(\frac{1}{r_1} + \frac{1}{r_2} \right)^{-1} \right]^{1/3} \qquad (3)$$

$$\delta = a^2 \cdot \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \qquad (4)$$

where: P is the force pressing the two spheres together. For eq. (3) and (4) to be valid, P must be parallel to the line connecting the centers of the two spheres.

a is the radius of the circular contact area.

μ is Poisson's ratio (a property of the material).

E is Young's modulus of elasticity.

r is the radius of the sphere.

δ is the *deformation* defined as

$$\delta = \begin{cases} s-r_1-r_2 & \text{if } s<r_1+r_2 \\ 0 & \text{if } s \geq r_1+r_2 \end{cases}$$

s is the distance separating the centers of the two spheres

2.3 Bumpy Surface Contacting a Smooth Flat Surface: Program CR1

2.3.1 Description of the Model

In the first and simplest model, implemented in a BASIC

computer program called "CR1," the contact interface was modeled as a perfectly smooth and flat surface pressed against a bumpy surface. This assumption simplifies equations (3) and (4). The simplifications are as follows:

For a sphere pressed into a flat surface, one can let the flat surface be sphere 2 with $r_2 = \infty$. Then P must be perpendicular to the flat surface. Also letting $r_1 = r_i$, $\mu_1 = \mu_2 = \mu_i$, $E_1 = E_2 = E_i$, and $P_1 = P_i$, one can solve for the force supported by a spherical asperity at location i. One obtains

$$a_i = \left[\frac{3}{4} \cdot P_i \cdot 2 \cdot \left(\frac{1-\mu^2}{E_i} \right) \cdot \left(\frac{1}{r_i} \right)^{-1} \right]^{1/3}$$

Solving for P_i:

$$P_i = \frac{(4/3) \cdot a_i^{\,3}}{2 \cdot \left(\frac{1-\mu^2}{E_i} \right) \cdot r_i}$$

But from eq. (4), $$a_i = (\delta_i \cdot r_i)^{1/2} \qquad (5)$$

so that $$P_i = (4/3) \cdot (\delta_i \cdot r_i)^{3/2} \cdot (1/2) \cdot \left(\frac{E_i}{1-\mu^2} \right) \cdot \frac{1}{r_i}$$

$$= (2/3) \cdot \delta_i^{\,3/2} \, r_i^{\,1/2} \left(\frac{E_i}{1-\mu^2} \right) \qquad (6)$$

The deformation δ_i for an asperity of height h_i (Fig. 11) is defined as

$$\delta_i = \begin{cases} h_i - s & \text{if } s < h_i \\ 0 & \text{if } s \geq h_i \end{cases} \qquad (7)$$

where s is the separation between the base of the bumpy surface and the smooth flat surface that comprises the opposite contact member. The form of eq. (7) assures that if an asperity is not making contact, then δ is zero and, by equation (6), P_i is also zero.

Equation (7) may be more conveniently written as

$$\delta_i = (h_i - s) \cdot \frac{1}{2} \cdot \left[1 + \frac{h_i - s}{|h_i - s|} \right] \qquad (8)$$

Each asperity was modeled this way.

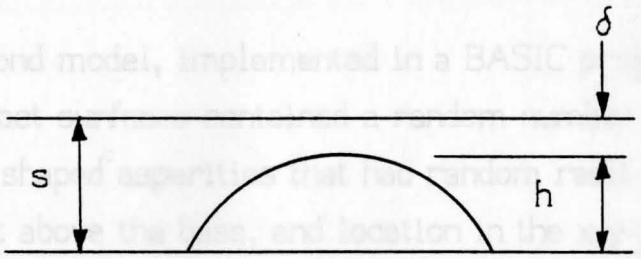


Fig. 11.--Deformation of an asperity.

2.3.2 Algorithm for calculating the separation of the contact members

Final contact separation is that which satisfies the equilibrium force equation:

$$F = \sum P_i \qquad (9)$$

where F is the applied contact force.

The secant method was used to find the contact separation. The algorithm is:

1) Guess starting values of s.

perform steps 2 through 4 for all asperities

2) Calculate δ_i, using equation (8).

3) Calculate P_i using δ_i and equation (6)

4) Calculate a_i from equation (5) and store for later use.

5) Calculate $\epsilon = F - \Sigma P_i$ from eq. (9).

6) If ϵ is less than the allowable error then stop, else calculate new s from the secant method formula and go to step 2.

2.4 Two Bumpy Surfaces in Contact with Each Other: Program CR2

In the second model, implemented in a BASIC program called "CR2," both contact surfaces contained a random number of hemispherically shaped asperities that had random radii of curvature, height above the base, and location in the x-y plane.

Contact of two opposing asperities is shown in Fig. 12. Assume that the top asperity is centered at $p_i = (x_i, y_i, z_i)$ and the bottom asperity is centered at $p_j = (x_j, y_j, z_j)$, and the asperities are barely touching. The x-y component of the distance between centers is

$$D_{ij} = \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]}$$

The directional vector from (x_i, y_i) to (x_j, y_j) is

$$(x_j - x_i)\,\underline{x} + (y_j - y_i)\,\underline{y}$$

where $\underline{x}$ and $\underline{y}$ are unit vectors.

The unit directional vector is

$$\frac{(x_j - x_i)\,\underline{x} + (y_j - y_i)\,\underline{y}}{D_{ij}}$$

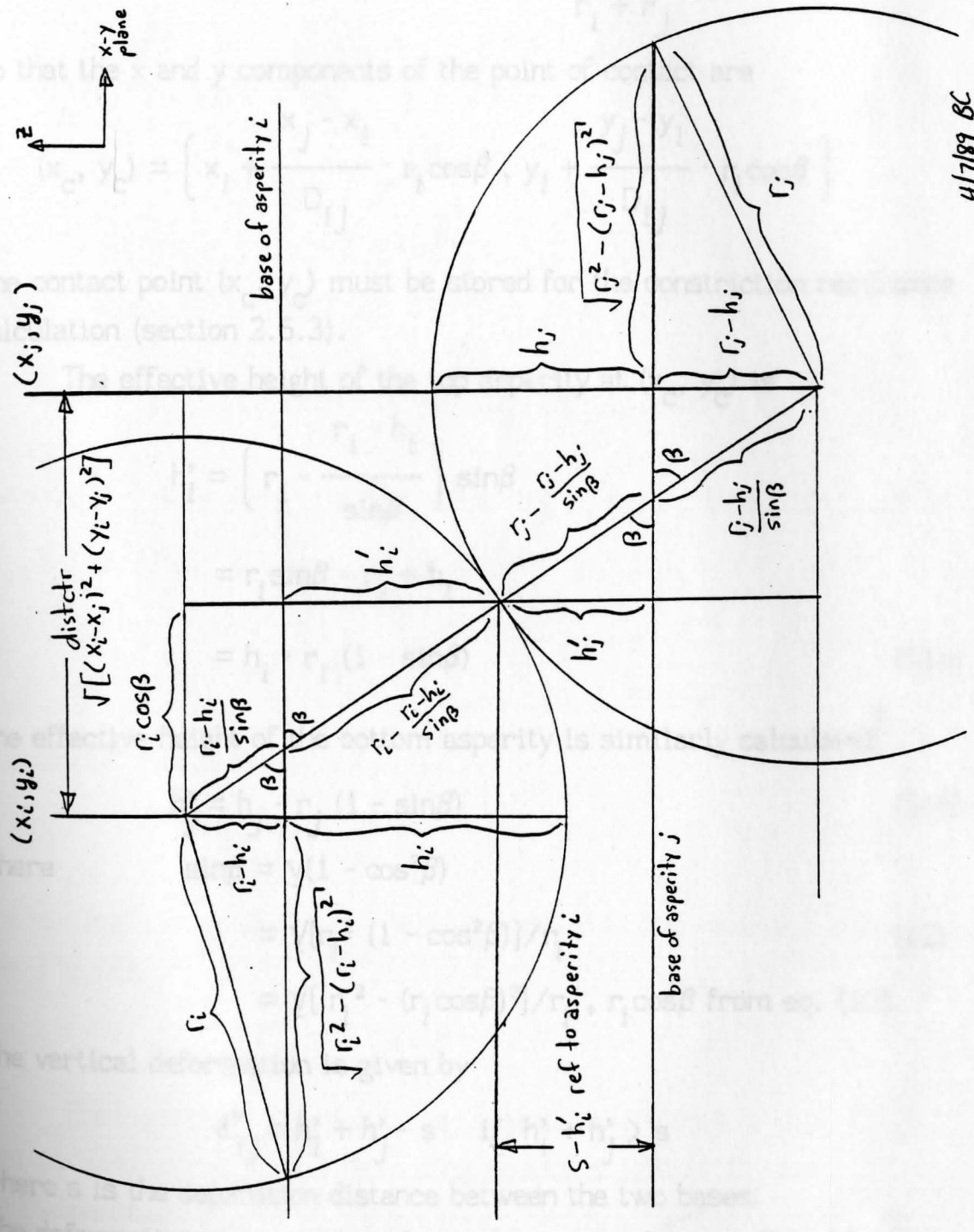


Fig. 12.--Two asperities in contact.

The x-y direction from (x_i, y_i) to the point of contact is

$$r_i \cos\beta \qquad \text{where } \cos\beta = \frac{D_{ij}}{r_i + r_j} \tag{10}$$

So that the x and y components of the point of contact are

$$(x_c, y_c) = \left(x_i + \frac{x_j - x_i}{D_{ij}} \cdot r_i\cos\beta \, , \, y_i + \frac{y_j - y_i}{D_{ij}} \cdot r_i\cos\beta \right)$$

The contact point (x_c, y_c) must be stored for the constriction resistance calculation (section 2.6.3).

The effective height of the top asperity at (x_c, y_c) is

$$h_i' = \left(r_i - \frac{r_i - h_i}{\sin\beta} \right) \sin\beta$$

$$= r_i\sin\beta - r_i + h_i$$

$$= h_i - r_i\,(1 - \sin\beta) \tag{11a}$$

The effective height of the bottom asperity is similarly calculated

$$h_j' = h_j - r_j\,(1 - \sin\beta) \tag{11b}$$

where $\sin\beta = \sqrt{(1 - \cos^2\beta)}$

$$= \sqrt{[r_i^{\,2}\,(1 - \cos^2\beta)]}/r_i \tag{12}$$

$$= \sqrt{[\,r_i^{\,2} - (r_i\cos\beta)^2]}/r_i \;,\; r_i\cos\beta \text{ from eq. (10).}$$

The vertical deformation is given by

$$\delta^v_{ij} = h_i' + h_j' - s \quad \text{if } h_i' + h_j' > s$$

where s is the separation distance between the two bases.

The deformation along the line from p_i to p_j is shown in Fig. 13.

$$\delta_{ij} = \delta^{V}_{ij} \sin\beta \tag{13}$$

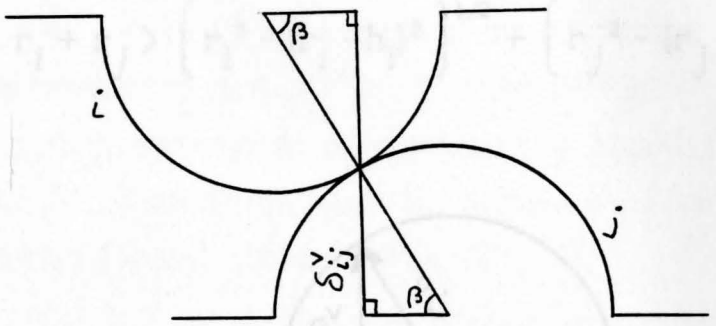


Fig. 13.--Deformation of contacting asperities.

The load P_{ij} acts along the line connecting p_i and p_j, thus it depends on δ_{ij}, not on δ^{V}_{ij}.

$$P_{ij} = (2/3)\, \delta_{ij}^{3/2} \left(\frac{1}{r_i} + \frac{1}{r_j} \right)^{-1/2} \left(\frac{E}{1-\mu^2} \right) \tag{14}$$

The contact radius is

$$a_{ij} = \left(\delta_{ij} \left(\frac{1}{r_i} + \frac{1}{r_j} \right)^{-1} \right)^{1/2} \tag{15}$$

The vertical component of P_{ij} helps support the total contact force (see Fig 14).

$$P^{V}_{ij} = P_{ij} \sin\beta \tag{16}$$

The algorithm for calculating separation distance is similar to that outlined in sec. 2.3.2, but using equations (13), (15), and (16).

One must, of course, calculate a_{ij} and P_{ij} for every combination of i and j. This will require $n_{bi} \times n_{bj}$ calculations for every iteration of separation distance. But since asperity i will only contact its nearest

neighbors on the opposite surface, one can save calculation time by using a "proximity check": asperity i cannot contact asperity j on the opposite surface if

$$r_i + r_j > \left[r_i^2 - (r_i - h_i)^2 \right]^{1/2} + \left[r_j^2 - (r_j - h_j)^2 \right]^{1/2}$$

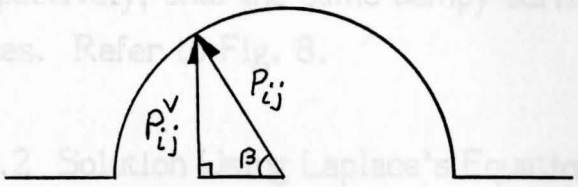


Fig. 14.--Vertical component of asperity load.

2.5 Dimple in Contact with a Bumpy Flat Surface: Program CR3

This is a special case of bumpy flat vs. bumpy flat: the top surface contains a single "asperity" called a dimple. The dimple's height is much greater than a regular asperity's (1 or 2 mm vs. 1 to 15μm). The dimple's surface is assumed to be smooth. The equations for bumpy flat vs. bumpy flat are used. All one needs to do is make the following substitutions into equations (10) through (16):

$$r_i = r_{dimp}$$

$$\left.\begin{aligned} x_i &= \text{length}/2 \\ y_i &= \text{width}/2 \end{aligned}\right\} \quad \text{The dimple is centered at the center of the top base (Fig. 8).}$$

where "length" and "width" are the length and width of the nominally flat bumpy surface. This model was implemented in a BASIC program called "CR3."

2.6 Calculating the Constriction Resistance

2.6.1 The Contact Spots

Figures 15 and 16 show plots of actual contact spots. These plots were created as part of the computer simulation, and represent to scale the contact spots made when pressing a smooth flat surface and a smooth dimple, respectively, onto the same bumpy surface, at different applied contact forces. Refer to Fig. 8.

2.6.2 Solution Using Laplace's Equation

The "exact" value for the constriction resistance is solved for by using Laplace's equation in three dimensions:

$$\nabla^2\phi = 0$$

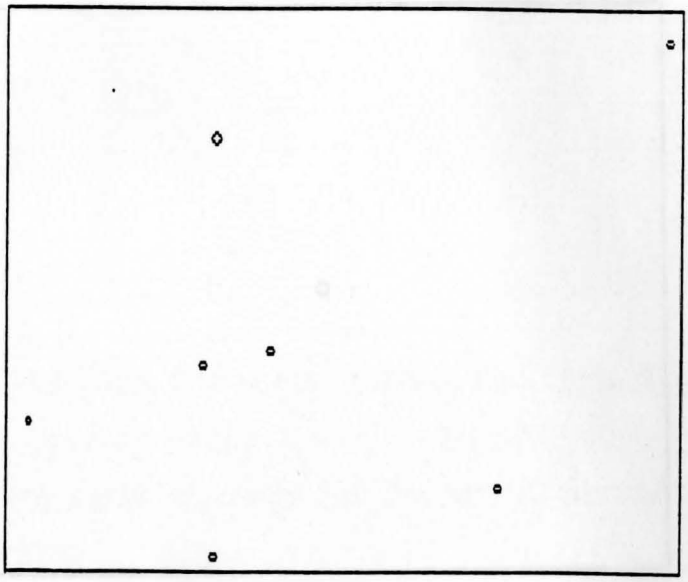
where ϕ is the potential at a point (x,y,z) within the contact members or at the interface. A numerical approximation of Laplace's equation is recommended for this type of problem.

A numerical solution to a two-dimensional contact interface problem was worked out using the usual finite-difference method. To summarize this method, recall Laplace's equation in two dimensions:

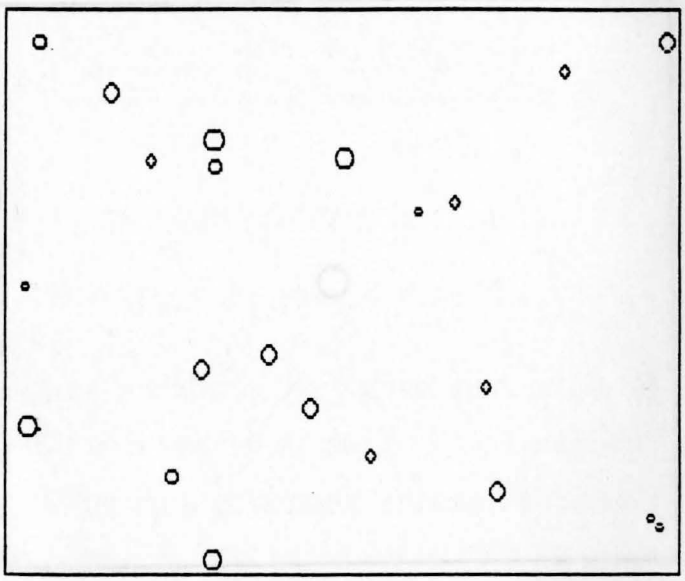
$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial z^2} = 0$$

If one breaks up the two-dimensional area by grids, as in Fig. 17, one can approximate Laplace's equation at a grid point by the formula

$$\frac{\partial^2\phi}{\partial x^2} \simeq \frac{\left(\frac{\phi_e - \phi}{\Delta x}\right) - \left(\frac{\phi - \phi_w}{\Delta x}\right)}{\Delta x}$$

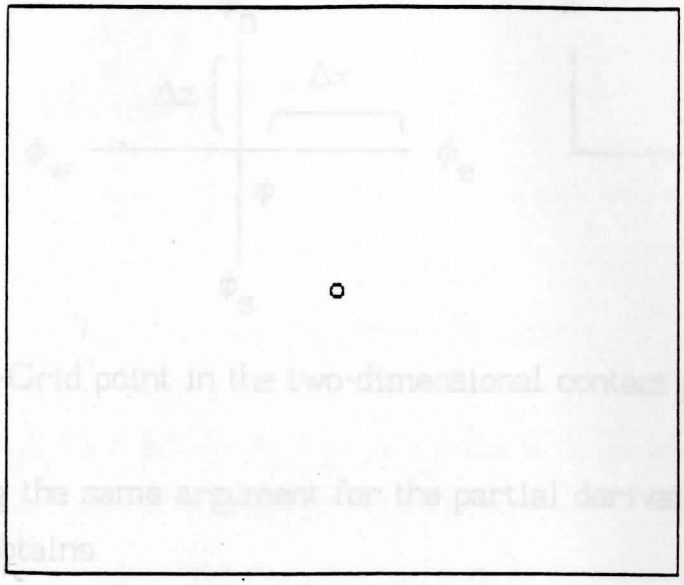


(a) 1N applied contact force

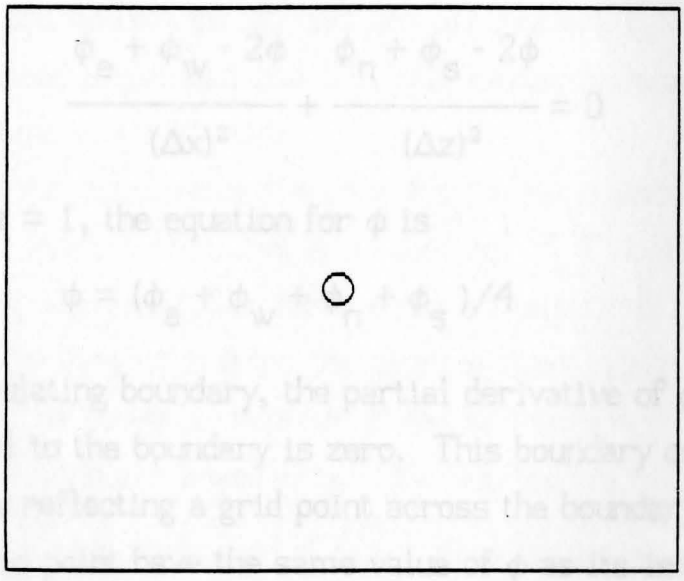


(b) 15N applied contact force

Fig. 15.--Contact spots, smooth flat surface on bumpy surface.



(a) 1N applied contact force



(b) 15N applied contact force

Fig. 16.--Contact spots, smooth 1mm radius dimple on bumpy surface.

or $$\partial^2\phi/\partial x^2 \simeq (\phi_e + \phi_w - 2\phi)/(\Delta x)^2 \qquad (17)$$

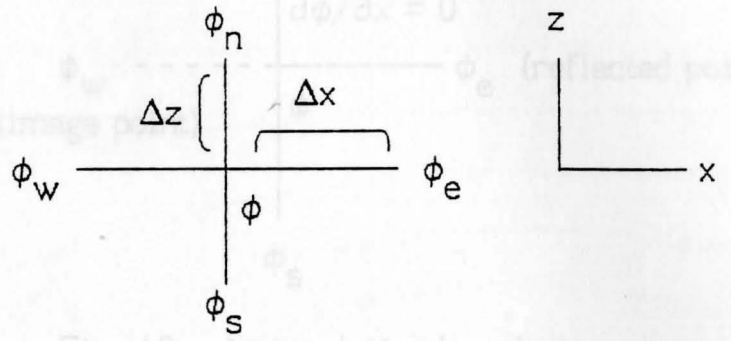


Fig. 17.--Grid point in the two-dimensional contact problem.

Following the same argument for the partial derivatives in the z-direction, one obtains

$$\partial^2\phi/\partial z^2 \simeq (\phi_n + \phi_s - 2\phi)/(\Delta z)^2 \qquad (18)$$

Combining eq. (17) and (18) to form Laplace's equation, one gets

$$\frac{\phi_e + \phi_w - 2\phi}{(\Delta x)^2} + \frac{\phi_n + \phi_s - 2\phi}{(\Delta z)^2} = 0$$

Letting $\Delta x = \Delta z = 1$, the equation for ϕ is

$$\phi = (\phi_e + \phi_w + \phi_n + \phi_s)/4$$

At an insulating boundary, the partial derivative of ϕ in the direction normal to the boundary is zero. This boundary condition is approximated by reflecting a grid point across the boundary and making the reflected grid point have the same value of ϕ as its image. For example, an insulating boundary parallel to the z-axis is shown in Fig. 18.

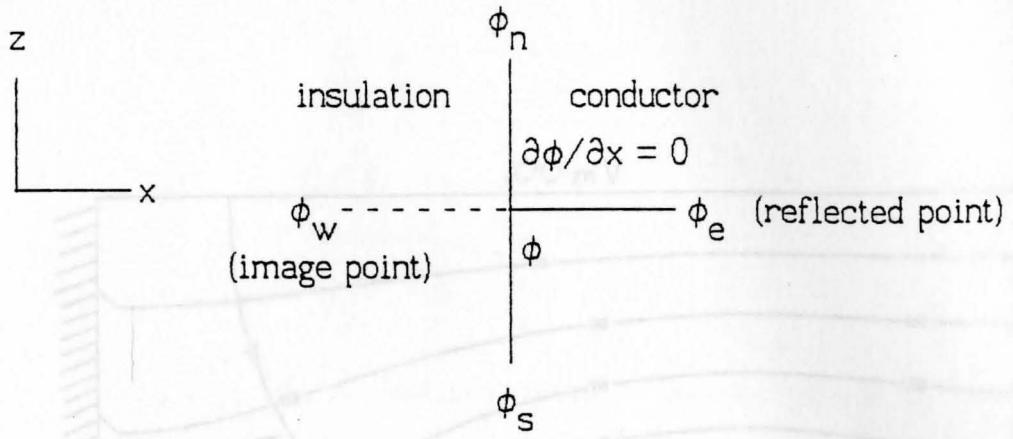


Fig. 18.--An insulating boundary.

If $\phi_w = \phi_e$, then $\partial\phi/\partial x \simeq (\phi_e - \phi_w)/\Delta x = 0$, which is the requirement. At this point then, the equation is

$$\phi = (2\phi_e + \phi_n + \phi_s)/4$$

Figures 19(a) and 19(b) are contour plots of the potential for simple two-dimensional contact regions containing three contact spots. The left, right, and bottom boundaries are all insulating, the contact spots were assigned a constant ϕ = 0V, and the top boundary was assigned a constant ϕ = 100 mV.

A method for calculating the constriction resistance from a 2D contour plot was given by Francyk and Kulikjan. [45] The disadvantage is that this method, when modified for the more realistic 3D case, is very cumbersome. The idea is to draw the current lines (always perpendicular to the potential contour lines) so that one creates approximately rectangular boxes (Fig. 20). Define the following :

d_{ki} = the dimension of box k along potential contour ϕ_i

d_{kj} = the dimension of box k along potential contour ϕ_j

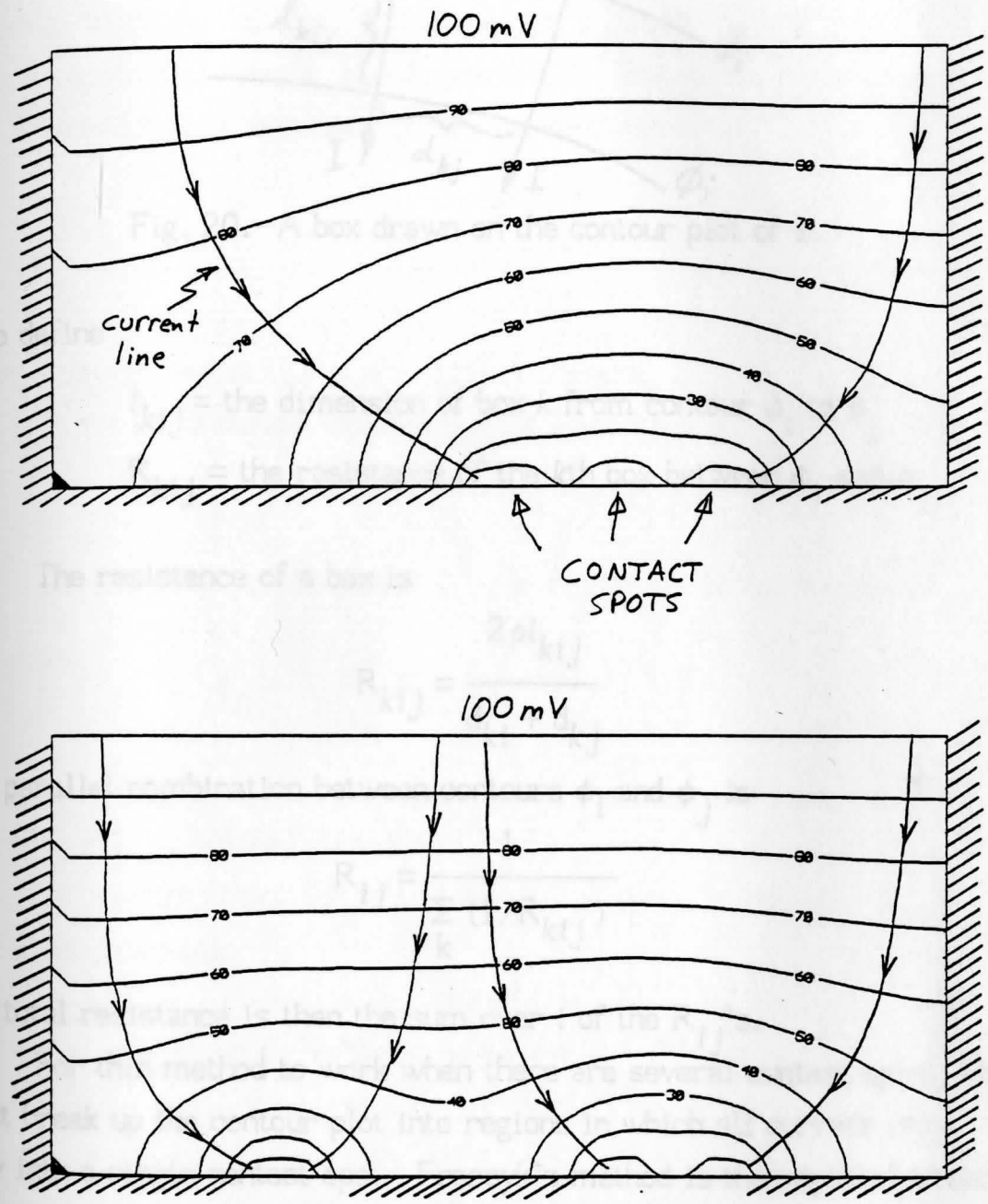


Fig. 19.--Potential contour plots for a 2D contact.

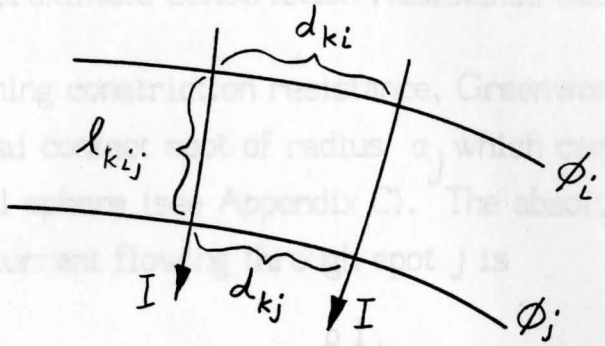


Fig. 20.--A box drawn on the contour plot of ϕ.

Also define

l_{kij} = the dimension of box k from contour ϕ_i to ϕ_j

R_{kij} = the resistance of the kth box between ϕ_i and ϕ_j

The resistance of a box is

$$R_{kij} = \frac{2\rho l_{kij}}{d_{ki} + d_{kj}}$$

The parallel combination between contours ϕ_i and ϕ_j is

$$R_{ij} = \frac{1}{\sum_k (1/R_{kij})}$$

The total resistance is then the sum over i of the R_{ij}'s.

For this method to work when there are several contact spots, one must break up the contour plot into regions in which all current lines flow into a single contact spot. Francyk's method is then applied to each region, and the total constriction resistance is a parallel combination.

The author decided that a Laplace's equation solution was not practical for implementation in CR1-CR3, and that a simpler method would be better. This method will now be described.

2.6.3 Approximate Constriction Resistance Calculation [46]

In determining constriction resistance, Greenwood assumed that each individual contact spot of radius a_j which carries current I_j is an equipotential sphere (see Appendix C). The absolute potential at spot i caused by current flowing through spot j is

$$\phi_{ij} = \frac{\rho I_j}{2\pi d_{ij}}$$

where ρ is the resistivity of the contact metals;

d_{ij} is the distance separating spot i and spot j,

$$d_{ij} = \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2]}$$

The potential at contact spot i due to current through spot i is

$$\phi_{ii} = I_i \frac{1}{2} R_{ci} = I_i \frac{\rho_i}{4a_i}$$

The quantity R_{ci} is called the *constriction resistance* of spot i and its formula $\rho/2a_i$ is widely quoted and used.

By superposition, the total potential at spot i is the sum of all individual potentials:

$$\phi_i = \phi_{ii} + \sum_{j \neq i} \phi_{ij} = \frac{\rho I_i}{4 a_i} + \frac{\rho}{2\pi} \sum_j \frac{I_j}{d_{ij}}$$

Since all the contact spots are in parallel, their potentials must all be equal. This gives rise to a set of simultaneous equations:

$$\phi_i = V/2$$

where V is an applied contact voltage across the two contact members.

In matrix form:

$$\begin{bmatrix} \rho/4a_1 & \rho/2\pi d_{12} & \rho/2\pi d_{13} & \cdots \\ \rho/2\pi d_{21} & \rho/4a_2 & \rho/2\pi d_{23} & \cdots \\ \rho/2\pi d_{31} & \rho/2\pi d_{32} & \rho/4a_3 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ \vdots \\ I_{n_c} \end{bmatrix} = \begin{bmatrix} V/2 \\ V/2 \\ V/2 \\ \vdots \\ V/2 \end{bmatrix} \qquad (19)$$

where n_c is the number of contact spots,

n_c << number of asperities, a fact which has been shown many times in the literature. [47]

The constriction resistance is then

$$R_c = \frac{V}{\left(\sum_{j=1}^{n_c} I_j \right)} \qquad (20)$$

2.7 Algorithms for Plotting the Cumulative Distribution of Contacts

To generate a cumulative distribution plot of a group of contacts, the user of the programs CR1, CR2, or CR3 must select:

a) the number of contacts in the group;

b) the nominal contact force, F;

c) the tolerance, "tol," as a fraction of the nominal contact force;

d) the radius of the dimple (CR3 only).

The algorithm for calculating constriction resistance was as follows:

1) Generate a random bumpy surface (Appendix B).
2) Calculate equilibrium separation distance using equation (9), the Hertz equations (5)-(6) or (14)-(16), and the secant method as described in section 2.3.2.
3) Calculate constriction resistance using equations (19) and (20).
4) Repeat steps 1-3 for the next contact.
5) Plot resistance values on normal probability paper.

Normal probability paper is scaled in such a way that, if the data being plotted follow a gaussian distribution, then the data will fall on a straight line. See Appendix D. The procedure for constructing a plot is

1) Sort the data in descending order. Assume N data points.
2) Assign an integer i to each data point, the largest data point will be assigned i=1, and the smallest data point will be assigned i=N.
3) Calculate for each data point

$$edf_i = i/(N+1)$$

$$t_i = \sqrt{(\ln(1/edf_i^2))}$$

$$Xp_i = t_i - \frac{2.515517 + 0.802853 \cdot t_i + 0.010328 \cdot t_i^2}{1 + 1.432788 \cdot t_i + 0.189269 \cdot t_i^2 + 0.001308 \cdot t_i^3} \qquad (21)$$

Xp_i is the number of standard deviations from the mean corresponding to the percentile of the data point, assuming a gaussian distribution. [52]

For example, the 73rd lowest data point (i = 73) out of a group of 100 will have a percentile of 0.73 and an Xp_i of 0.613. Xp_i is such that the area under the gaussian distribution curve from $-\infty$ to Xp_i, obtained from a standard statistical table, is 0.73.

Xp_i calculated from equation (21) differs slightly from the value

in the published statistical tables. The author is helping to develop a computer program that accurately matches Xp_i with table values.

2.8 Results from a few Sample Runs

Cumulative distribution plots were made for different contact forces and different diameter dimples (including flat). Several values kept constant are shown in Table 2.

TABLE 2

DATA FOR CONTACT RESISTANCE SIMULATION [48,49]

Surface data:		
average bump density	d	$100/mm^2$
apparent contact area	A_{app}	$4\ mm^2$
Asperity data:		
average height	avg(h)	$5.0\mu m$
variance of height	var(h)	$(avg(ht)/3)^2$
avg. radius of curvature	avg(r)	0.0005 m
variance of radius	var(r)	$(avg(r)/3)^2$
Material properties (brass, 70% Cu, 30% Zn):		
Young's modulus	E	1.1×10^{11} Pa
Poisson's ratio	μ	0.3
electrical resistivity	ρ	$6.2\times10^{-8}\ \Omega\cdot m$

The justification for these values is as follows: Uppal and Probert used 100 asperities/mm^2 in experiments in which they bead-blasted smooth metal surfaces to create an artificial rough contact surface. [50] An apparent contact area of $4mm^2$ is typical for a small connector. Williamson, Pullen, and Hunt showed that typical metals

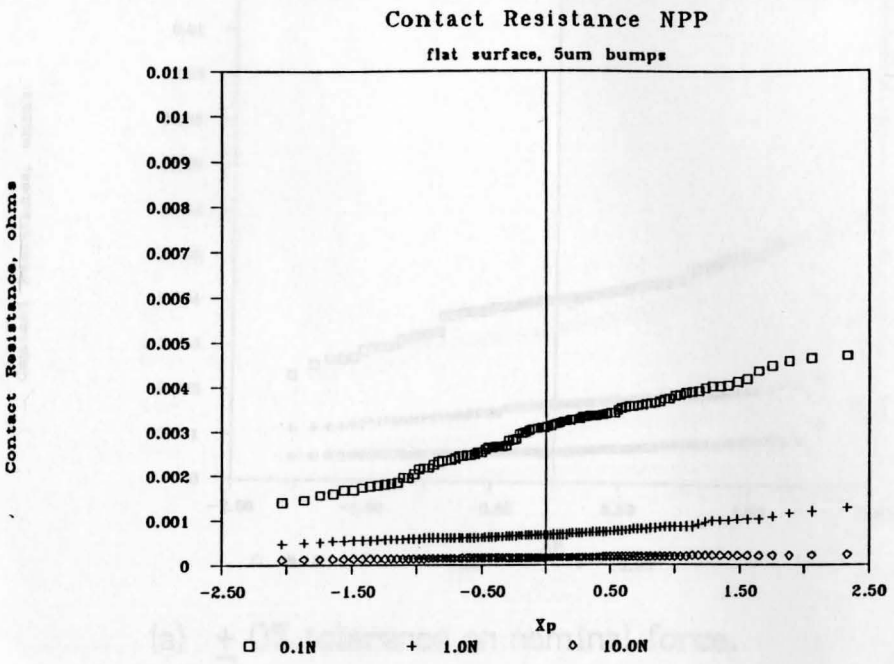
have mean asperity heights in the range 1-5μm [51], but CR1 had trouble with avg(h) $<$ ~ 3μm at higher forces because too many asperities came into contact and the number of contact spots made equation (19) too large for the personal computer. The radius of curvature was chosen such that it was 100 times larger than avg(h), in accordance with equation (1). The variance of asperity height and radius of curvature were chosen such that avg - $3\cdot(\text{var})^{1/2} = 0$.

Applied contact forces ranged from 0.1N to 10.0N. In some groups, the force was normally distributed such that $\pm$ tol = $\pm$ 3σ, where σ is the standard deviation. In other groups, tol was chosen as zero, and the force was identical for all contacts in the group.

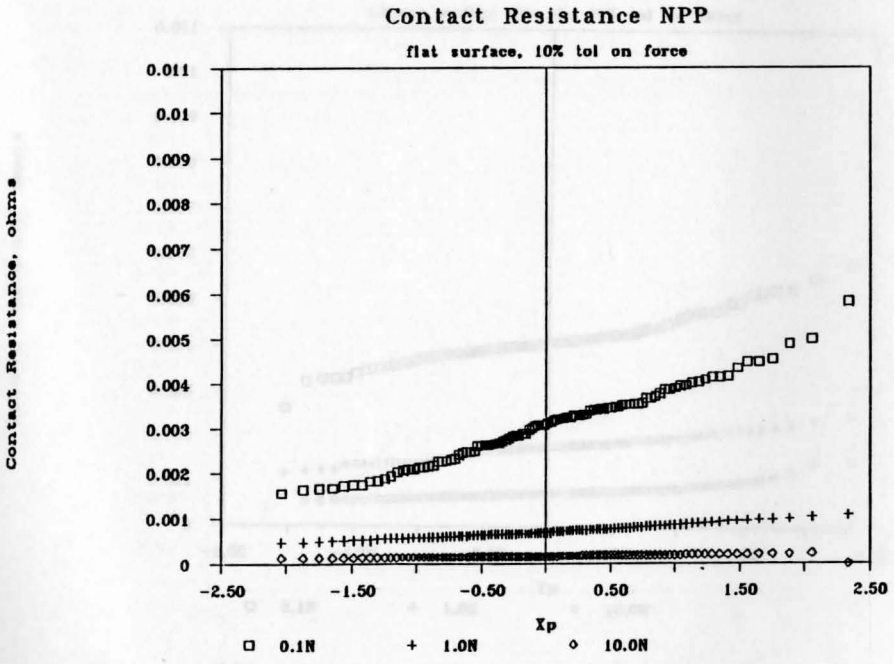
The graphs in Figs. 21 through 24 plot resistance on one axis and Xp as calculated from eq. (21) on the other, for the cases flat vs. flat, 1.0mm radius dimple vs. flat, and 3.0mm radius dimple vs. flat. In each case, one contact member is a bumpy flat and the other member is smooth. (That is, the models CR1 and CR3, but not CR2, were used.)

Fig. 25 is a plot of average contact resistance for each of the three cases and for each force. Notice that the flat vs. flat contact design has the lowest resistance for each force; and for the dimple vs. flat design, the smaller the radius of the dimple, the higher the resistance. That is because the flatter the mating interface, the more "spread out" the contact spots are, and the less that the current must bend its path to pass through the contact spots. Less bending of current leads to lower constriction resistance. Thus, the smaller the dimple's radius, the more constriction, and the higher the constriction resistance.

Then why are dimpled contacts used in practice? Because they concentrate the contact force over a smaller apparent area, leading to higher pressure on the contacting asperities and a greater chance of breaking contaminating films. Films were not considered in the data of Figs. 21-25. Thin films will be considered in Chapter IV.

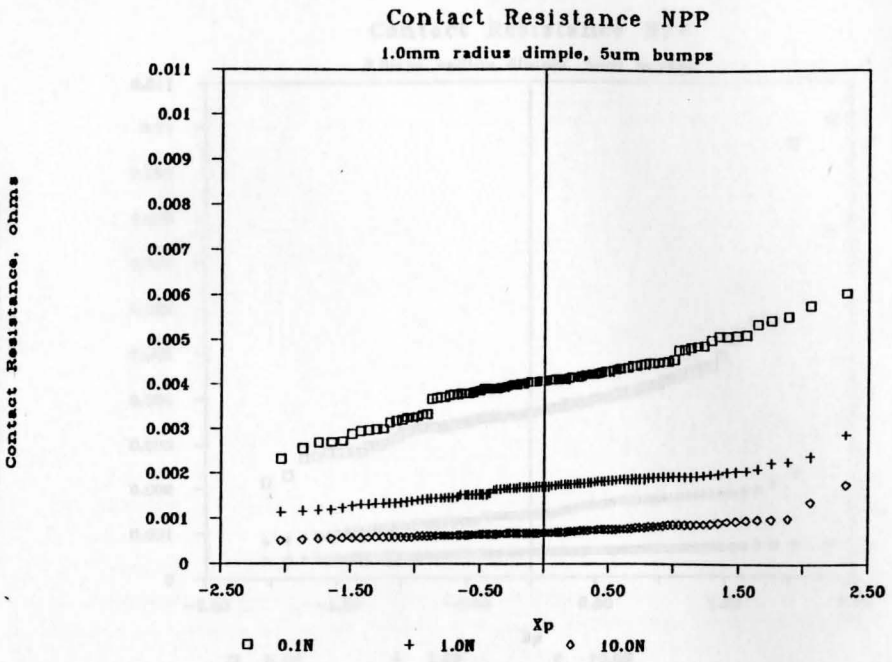


(a) ± 0% tolerance on nominal force.

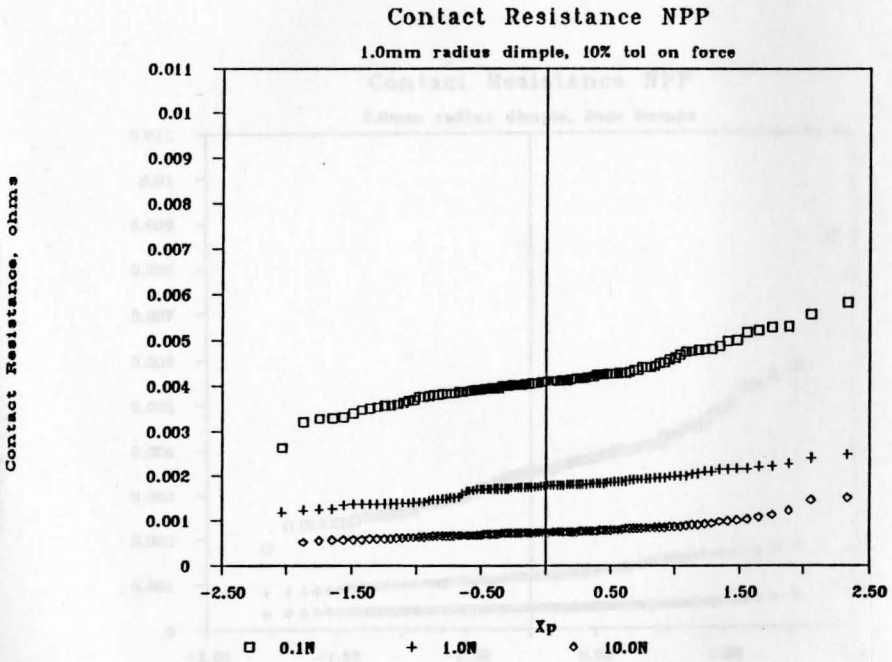


(b) ± 10% tolerance on nominal force.

Fig. 21.--Contact resistance NPP; smooth flat vs. bumpy flat (CR1).



(a) ± 0% tolerance on nominal force.



(b) ± 10% tolerance on nominal force.

Fig. 22.--Contact resistance NPP; 1.0mm radius dimple vs. bumpy flat.

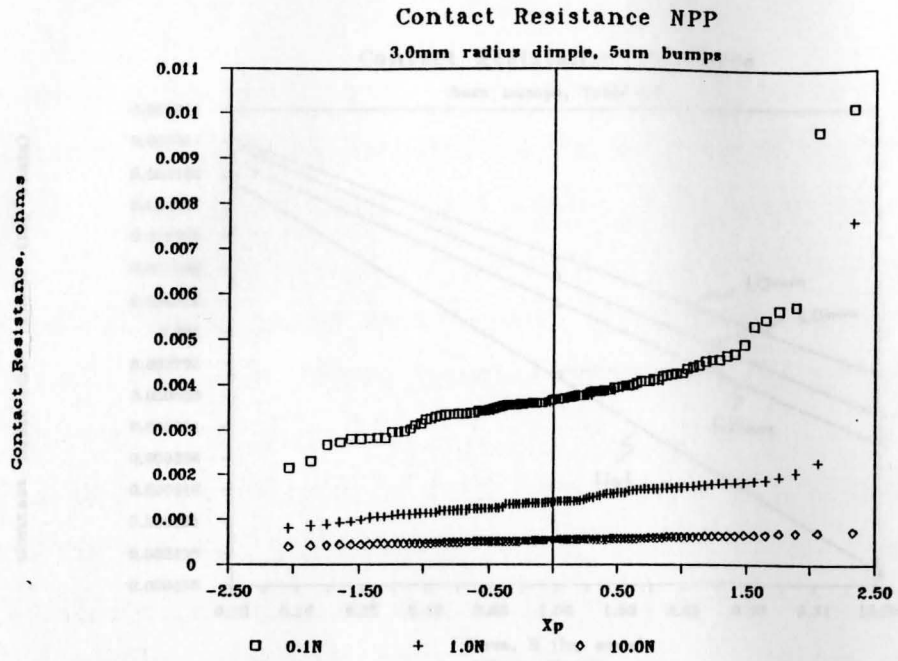


Fig. 23.--Contact resistance NPP; 3.0mm radius dimple vs. bumpy flat. ± 0% tolerance on nominal force.

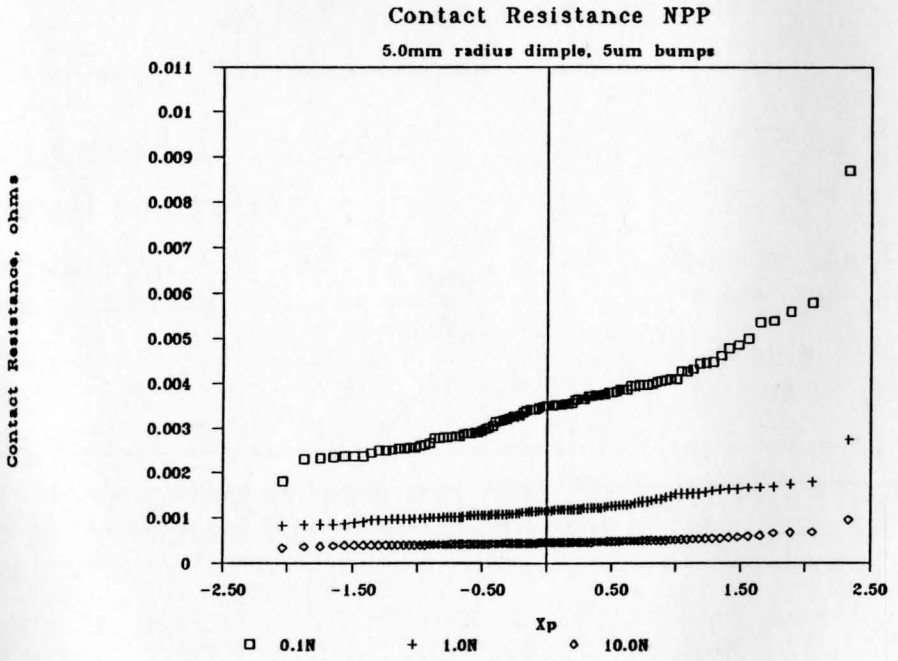


Fig. 24.--Contact resistance NPP; 5.0mm radius dimple vs. flat. ± 0% tolerance on nominal force.

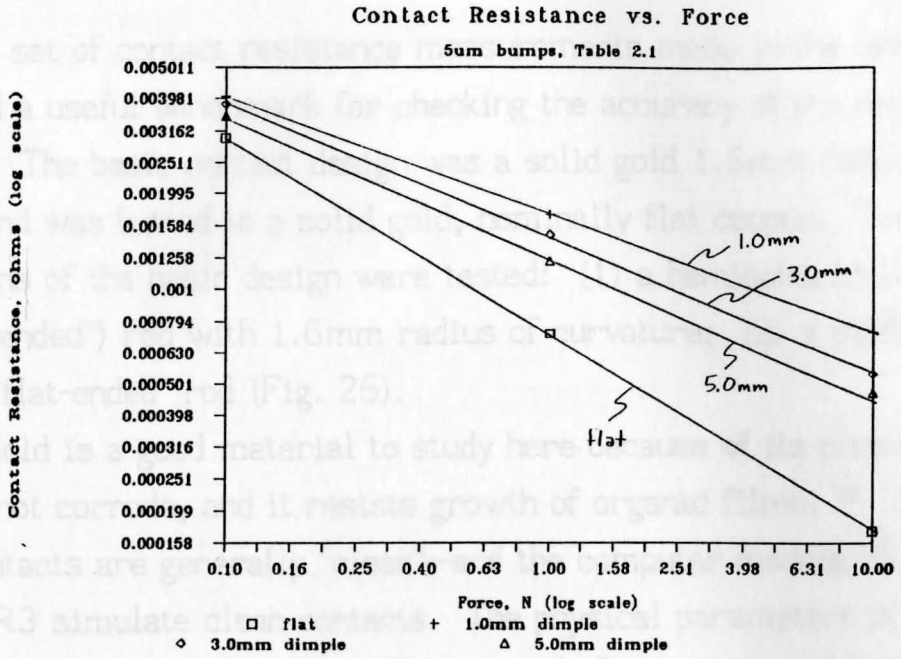


Fig. 25.--Average resistance vs. force, dimples and flats.

2.9 Verification of the Model for Gold Contacts

A set of contact resistance measurements made in the laboratory provided a useful benchmark for checking the accuracy of the computer models. The basic contact design was a solid gold 1.6mm radius rod whose end was butted to a solid gold, nominally flat coupon. Two variations of the basic design were tested: (1) a hemispherically-ended ("round-ended") rod with 1.6mm radius of curvature; (2) a machine-ground "flat-ended" rod (Fig. 26).

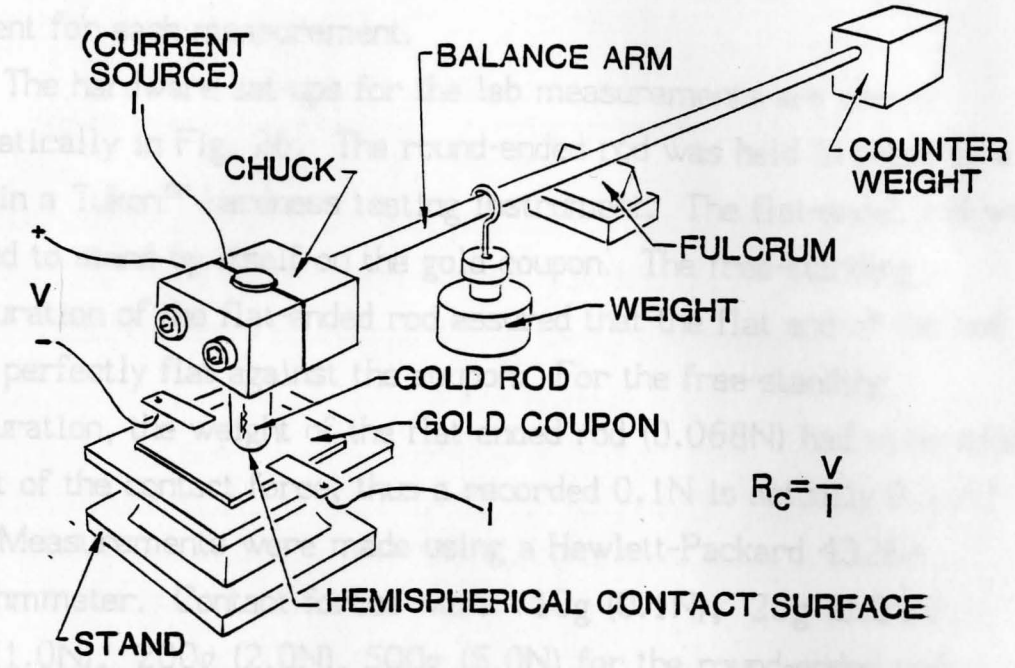
Gold is a good material to study here because of its properties: it does not corrode, and it resists growth of organic films. [53] Thus gold contacts are generally "clean"—and the computer models CR1, CR2, CR3 simulate clean contacts. The physical parameters of Table 2 remained the same for the simulation, with the exception of Table 3.

TABLE 3
DATA FOR GOLD CONTACTS [54]

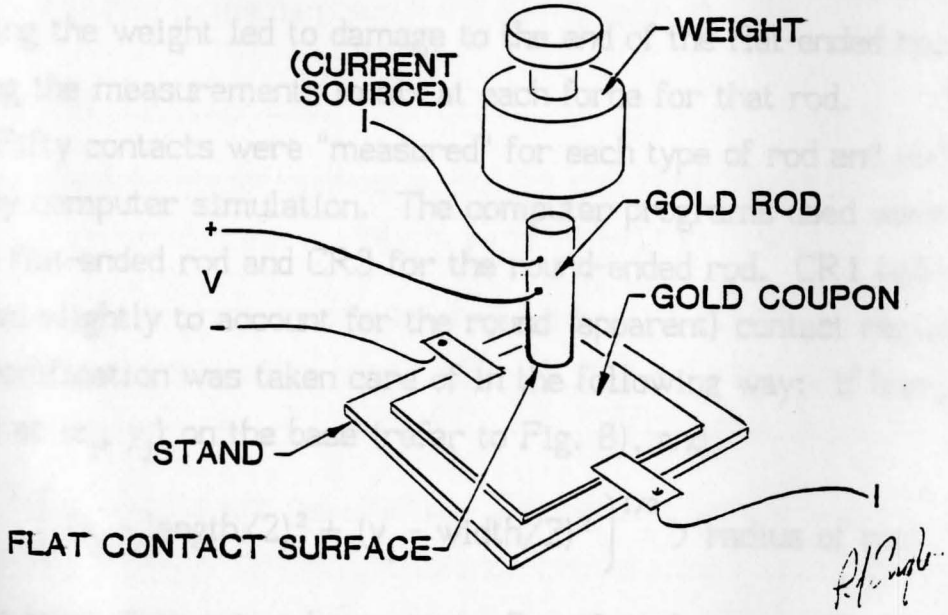
Young's modulus	E	8×10^{10} Pa
Electrical resistivity	ρ	2.25×10^{-8} Ω·m
Apparent area of bumpy surface	A_{app}	2mm × 2mm* round rod 3.2mm × 3.2mm flat ended rod

* note: since contact is made only near the center of the bumpy surface by the rounded rod, the apparent area of the bumpy surface is not important.

All lab measurements used the same gold coupon, and only one of each type of rod was used. The rod and coupon were cleaned prior to



(a) Set-up for round-ended rod.



(b) Set-up for flat-ended rod.

Fig. 26.--Experimental hardware for gold contact resistance measurements.

each measurement, and the location of contact on the flat coupon was different for each measurement.

The hardware set-ups for the lab measurements are shown schematically in Fig. 26. The round-ended rod was held in place by a chuck in a Tukon™ hardness testing instrument. The flat-ended rod was allowed to stand by itself on the gold coupon. The free-standing configuration of the flat-ended rod assured that the flat end of the rod butted perfectly flat against the coupon. For the free-standing configuration, the weight of the flat-ended rod (0.068N) had to be added as part of the contact force, thus a recorded 0.1N is actually 0.168N.

Measurements were made using a Hewlett-Packard 4328A milliohmmeter. Contact forces were: 10g (0.1N), 25g (0.25N), 100g (1.0N), 200g (2.0N), 500g (5.0N) for the round-ended rod; and 10g (0.1N), 50g (0.5N), and 100g (1.0N) for the flat-ended rod. Twenty lab measurements were made, all at different locations on the coupon, at each force for the round-ended rod. Extreme difficulties in balancing the weight led to damage to the end of the flat-ended rod, limiting the measurements to six at each force for that rod.

Fifty contacts were "measured" for each type of rod and each force by computer simulation. The computer programs used were CR1 for the flat-ended rod and CR3 for the round-ended rod. CR1 had to be modified slightly to account for the round (apparent) contact region. This modification was taken care of in the following way: if bump i is located at (x_i, y_i) on the base (refer to Fig. 8), and

$$\left[(x_i - \text{length}/2)^2 + (y_i - \text{width}/2)^2 \right]^{1/2} > \text{radius of rod}$$

then the bump does not make contact: $P_i = 0$ (refer to eq. 6). It was assumed that the center of the rod was at $(x, y) = (\text{length}/2, \text{width}/2)$.

The graphs in Figs. 27 through 30 show how the models compare to the lab measurements. [55]

The percent error between the model values and the lab values were calculated as follows:

$$\text{percent error} = \frac{\text{model value - lab value}}{\text{lab value}} \times 100\%$$

The maximum absolute and percent errors in the mean and the median for both the round-ended rod and flat-ended rod are shown in Table 4. Each error's counterpart is shown in parentheses.

TABLE 4

MAXIMUM ERRORS IN CONTACT RESISTANCE MEASUREMENTS

Round-ended rod (Fig. 26a)

Max absolute error in mean = -0.20 mΩ (-13.3%) @ 0.1N
Max % error in mean = 77.7% (0.097mΩ) @ 10N

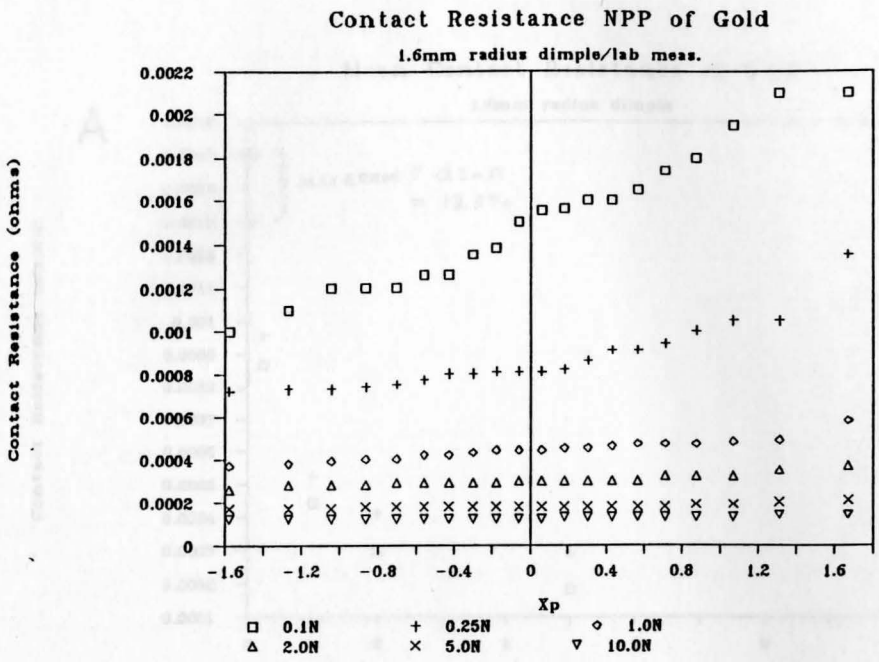
Max absolute error in median = -0.23 mΩ (-15.7%) @ 0.1N
Max % error in median = 75.8% (0.090mΩ) @ 10N

Flat-ended rod (Fig. 26b)

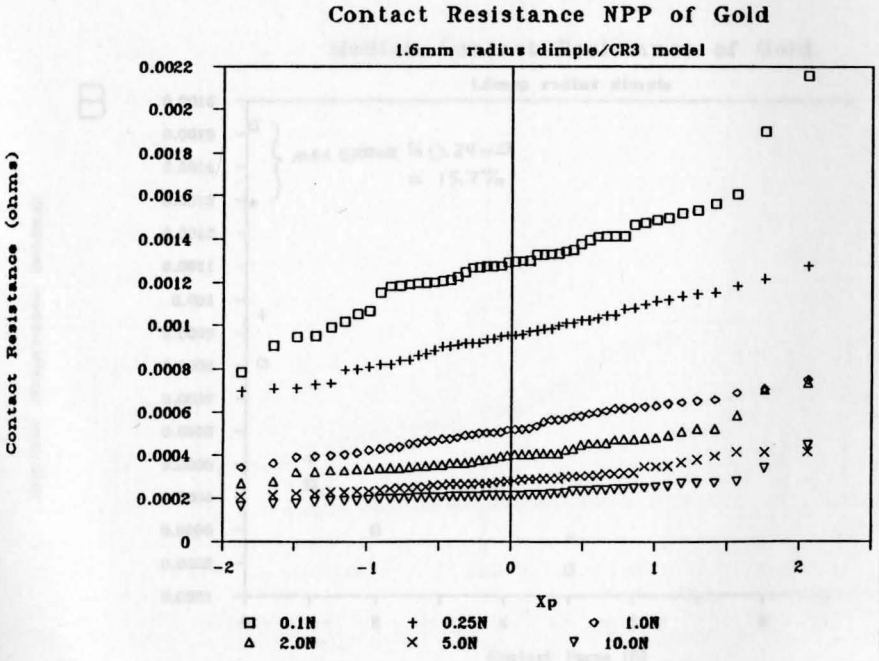
Max absolute error in mean = -0.205 mΩ (-24.0%) @ 0.1N
Max % error in mean = -38.9% (-0.179 mΩ) @ 0.5N

Max absolute error in median = -0.264 mΩ (-30.5%) @ 0.1N
Max % error in median = -43.65% (-0.212 mΩ) @ 0.5N

These results are reasonable. The percent errors are large only because the values of contact resistance are so small—note how small the absolute errors are. The author feels that the absolute error is a better measure of the "goodness" of the model. Figures 28 and 30 are very promising and show that the assumptions for the contact surface model are in the ballpark.



(a) Results from laboratory measurements.



(b) Results from CR3 computer model.

Fig. 27.--Contact resistance NPP of gold contacts, 1.6mm radius round rod.

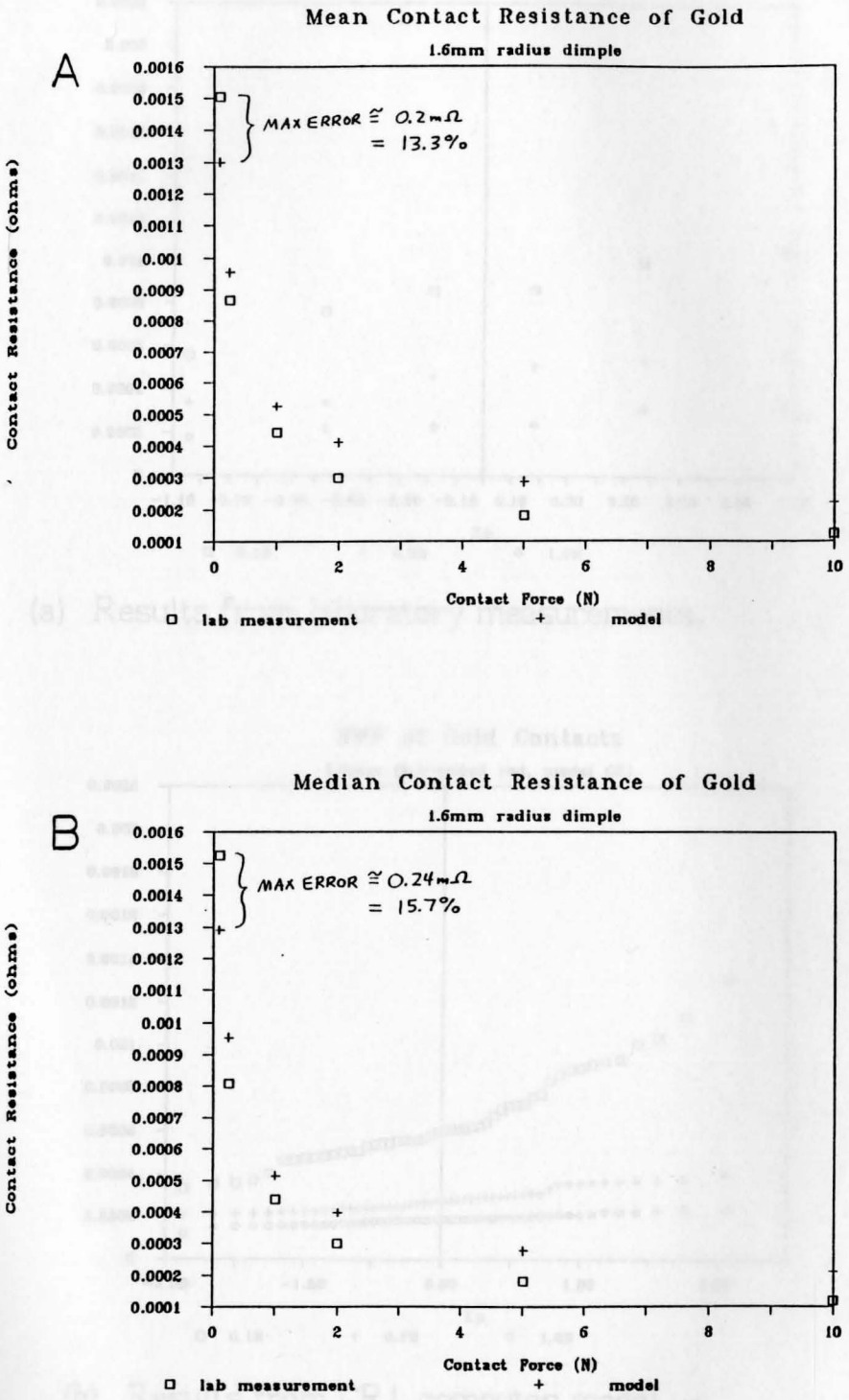
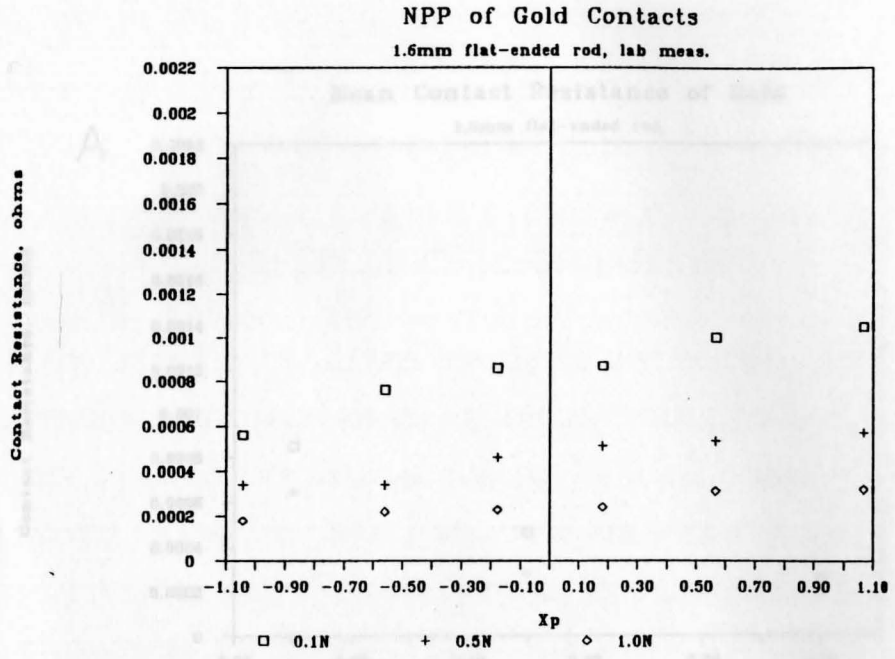
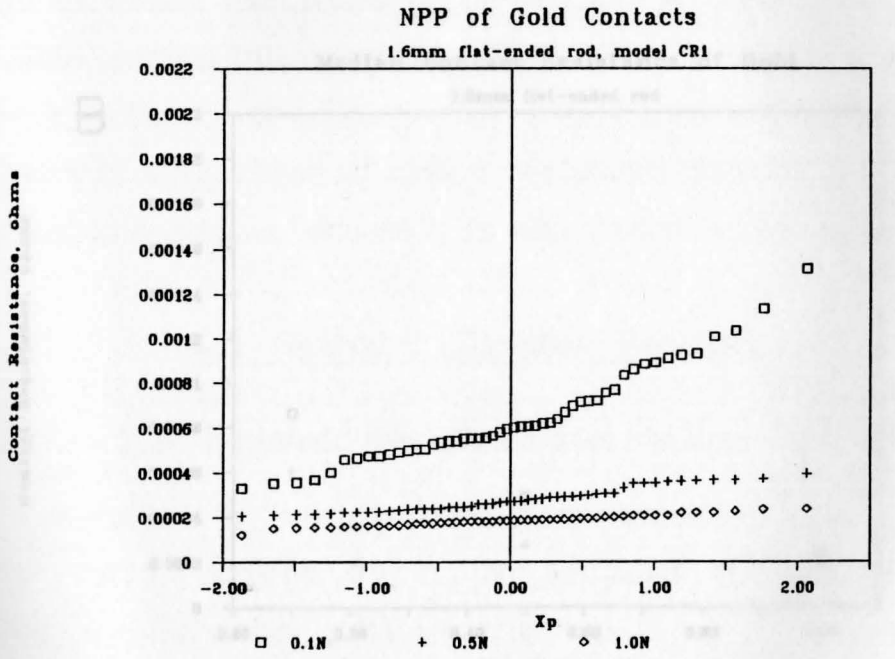


Fig. 28.--(a) Mean and (b) Median contact resistances of gold, model and lab 1.6mm radius round rod vs. flat.



(a) Results from laboratory measurements.



(b) Results from CR1 computer model.

Fig. 29.--Contact resistance NPP of gold contacts, 1.6mm flat probe vs. flat.

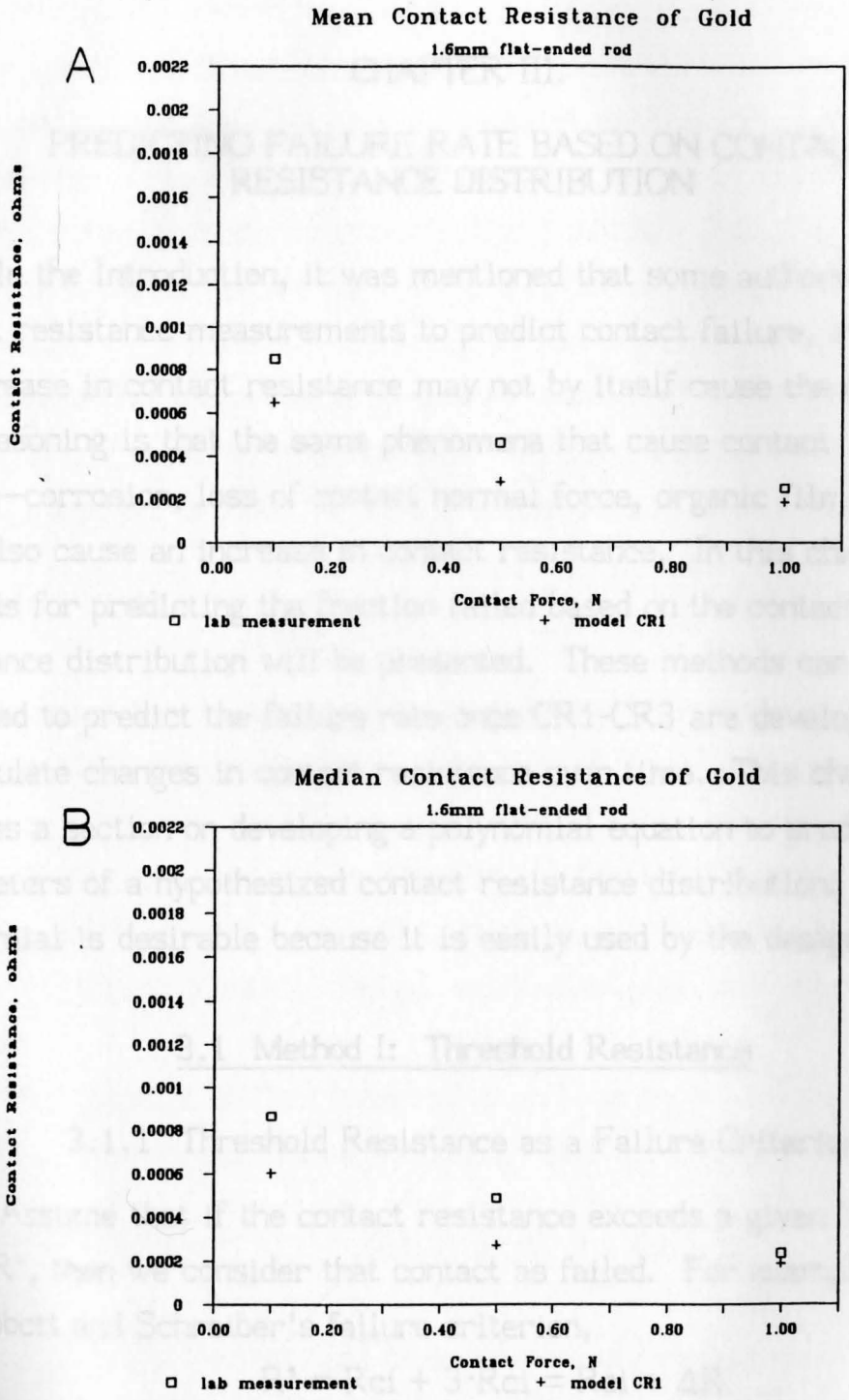


Fig. 30.--(a) Mean and (b) Median contact resistance of gold. 1.6 mm radius flat rod.

CHAPTER III.

PREDICTING FAILURE RATE BASED ON CONTACT RESISTANCE DISTRIBUTION

In the Introduction, it was mentioned that some authors have used contact resistance measurements to predict contact failure, even though an increase in contact resistance may not by itself cause the failure. The reasoning is that the same phenomena that cause contact failure—corrosion, loss of contact normal force, organic film growth, etc.—also cause an increase in contact resistance. In this chapter two methods for predicting the fraction failed based on the contact resistance distribution will be presented. These methods can be expanded to predict the failure rate once CR1-CR3 are developed enough to simulate changes in contact resistance over time. This chapter also includes a section on developing a polynomial equation to predict the parameters of a hypothesized contact resistance distribution. A polynomial is desirable because it is easily used by the designer.

3.1 Method I: Threshold Resistance

3.1.1 Threshold Resistance as a Failure Criterion

Assume that if the contact resistance exceeds a given "threshold" value R', then we consider that contact as failed. For example, one may use Abbott and Schreiber's failure criterion,

$$R' = Rci + 3 \cdot Rci = Rci + \Delta R$$

where Rci is the initial or "clean" contact resistance calculated from CR1 or one of its offspring; or one may use Whitley and Malucci's

failure criterion,

$$R' = 3 \cdot Rcmin$$

where Rcmin is a theoretical or ideal minimum contact resistance. One may also choose R' based on the application if one knows what value of contact resistance may lead to problems; it may be known, for example, that a certain contact in a certain application begins to show problems when the contact resistance exceeds *R* milliohms. Then R' = *R*.

The expected failure rate of a population of "aging" contacts is the fraction of good (R<R') contacts that fail per unit time interval. Now assume that the probability density function of the frequency distribution of contact resistances for this population of contacts is f(r,t), where t is time. Then the failure rate at time t is

$$\text{failure rate} = \frac{d}{dt} \left\{ \int_{R'}^{\infty} f(r,t)\, dr \right\} / \left\{ \int_{0}^{R'} f(r,t)\, dr \right\}$$

One estimates the derivative if f(r,t) is known only at discrete times.

Usually, one does not know the form of f(r,t). However, given a sample group of resistance measurements of either real or computer-generated contacts, one can hypothesize the form of f(r,t) and estimate its parameters. For a meaningful estimate, all contacts in the sample group should have identical design, application, age, and length of service.

At this point of the research, time, t, is not a variable in the models. However, the "infant mortality" at t = 0 can be predicted:

$$\text{fraction failed} = \int_{R'}^{\infty} f(r,0)\, dr \equiv \int_{R'}^{\infty} f_0(r)\, dr \qquad (22)$$

It is desirable that $f_0(r)$ be a polynomial function of the integrating variable r. This way, eq. (22) is easy to evaluate. Most practical probability density functions, however, are not polynomials, so one must seek a good polynomial approximation for whatever

distribution the contact resistances follow.

For illustration, let us hypothesize that the contact resistances at t = 0 follow a normal (gaussian) distribution. (A commonly occurring contact resistance distribution for any time t is the lognormal distribution. According to this distribution, the log of contact resistance is normally distributed. Thus the gaussian distribution is useful to look at here.) The parameters of the gaussian distribution, μ and σ, mean and standard deviation, are well known and easy to predict from a sample of measurements. The density function is

$$f_0(r)_{\text{gaussian}} = \frac{1}{\sigma\sqrt{(2\pi)}} \exp\left[-\frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2\right]$$

which can be approximated by a polynomial given by Chhabra and Wenning [56], which the author generalized:

$$f_0(r) \simeq A\,(r-a)^2\,(r-b)^2$$

where $a = \mu - 2.63\sigma$

$b = \mu + 2.63\sigma$

so that $f_0(r) \simeq A\,(r-\mu+2.63\sigma)^2\,(r-\mu-2.63\sigma)^2$

valid for $\mu - 2.63\sigma \leq r \leq \mu + 2.63\sigma$

Thus equation (22) becomes

$$\text{fraction failed} = \int_{R'}^{\mu+2.63\sigma} A\,(r-\mu+2.63\sigma)^2\,(r-\mu+2.63\sigma)^2\,dr \qquad (23)$$

If the density function is found to be lognormal, μ and σ must be the mean and standard deviation of log(R). Then, one would substitute log(r) for r and log(R') for R' in equation (23). The introduction of the logarithmic function does not make eq. (23) a non-polynomial if log(r) is treated as a variable unto itself.

The constant "A" must satisfy

$$\int_0^{\infty} f_0(r)\, dr = 1$$

which in this example becomes

$$\int_{\mu - 2.63\sigma}^{\mu + 2.63\sigma} A\,(r - \mu + 2.63\sigma)^2\,(r - \mu - 2.63\sigma)^2\, dr = 1 \qquad (24)$$

It is not necessary to compute "A" because eq. (23) can be divided by (24) which is the same as dividing (23) by 1. When this is done, the "A" is cancelled out of the computation:

$$\text{fraction failed} = \frac{\displaystyle\int_{R'}^{\mu + 2.63\sigma} A\,(r - \mu + 2.63\sigma)^2\,(r - \mu - 2.63\sigma)^2\, dr}{\displaystyle\int_{\mu - 2.63\sigma}^{\mu + 2.63\sigma} A\,(r - \mu + 2.63\sigma)^2\,(r - \mu - 2.63\sigma)^2\, dr} \qquad (25)$$

Note that, by eq. (24), the denominator of eq. (25) equals 1.

The numerator and denominator of eq. (25) are simple polynomials of the integrating variable r. The key now is to find functions that relate μ and σ to the variables that describe contact design, application, and time or length of service. Otherwise, the designer will have to generate several contact resistance values from a contact resistance simulation computer program, and then calculate μ and σ from the data generated. The designer, without doubt, has better things to do. To obtain a simple polynomial function, the method of least squares was used. The data for the least squares fit can be obtained by running the computer simulation using different values of the variables for each run. An example will now be given.

3.1.2 Designed Experiment for Obtaining Equations for μ and σ

Even though the simulation program CR1 predicts contact resistance distributions for "clean" contacts, and therefore is not expected to produce a "failed" contact (except, maybe, in the case of an extremely critical application requiring a very low contact resistance), it can be used in an example to clarify the preceding discussion and to demonstrate the use of equation (25). In this example, μ and σ were modeled in terms of the contact design parameters only. Application variables (such as temperature, moisture, environment, et al.) and time were not a part of the simulation program CR1.

An experimental design array based on G. Taguchi's techniques [57] was used to set values of all variables for each trial of the experiment. The values are shown in Tables 5(a) and 6. The ranges of the variables were carefully chosen after a review of the literature and some initial trial-and-error to make sure the program would run for all desired combinations of the variables. The range for average height is due to Williamson, Pullen, and Hunt, who showed that many metals have average asperity heights in the range of 1-5μm (CR1 had trouble converging with ht $< 3\mu$m at higher forces such a 10N, so the range 5-15μm was selected); the range for r/h is due to Williamson and Hunt (refer to Chapter II of this report for their criterion r/h $\geq$ 66); number of asperities per mm^2 is due to Uppal and Probert (who claimed 100/mm^2 is typical); the contact normal force covers a range typical for small automotive terminals; Young's modulus covers the range from silver (7×10^{10} Pa) to brass (11×10^{10} Pa); also resistivity: silver (1.8×10^{-8} $\Omega\cdot$m) to brass (6.2×10^{-8} $\Omega\cdot$m).

For each of the 27 trials dictated by Taguchi's array, 30 "random" contact resistances were generated by CR1. The mean and standard deviation were calculated for each trial; they are shown in the

last two columns of Table 6. The average contact resistance at each level of each variable was also calculated; plots are shown in Figures 31-38. Based on these plots, an analysis of variance on the mean contact resistance was conducted to determine which of the variables were significant (Table 5b), then a least squares curve fit was made to obtain polynomial equations for μ and σ, which can be substituted into equation (25). The polynomial equations for μ and σ are as follows:

$$\hat{\mu} = 1.22\times10^{-3} + 6.94\times10^{-5}\cdot(\text{avg(h)}) \; 4.27\times10^{-4}\cdot(r/h) + 1.35\times10^{-5}\cdot(r/h)^2 + 1.35\times10^{-4}\cdot(\text{shape}) - 1.42\times10^{-4}\cdot(\text{shape})^2 + 6.57\times10^{-5}\cdot(\text{area}) + 2.08\times10^{-4}\cdot(n_b) - 1.21\times10^{-3}\cdot\log(F) + 2.29\times10^{-4}\cdot(E) + 6.54\times10^{-4}\cdot(\rho) - 2.77\times10^{-4}\cdot\log(F)\cdot(\rho) \qquad (26)$$

$$\hat{\sigma} = 2.69\times10^{-4} + 2.84\times10^{-5}\cdot(\text{avg(h)}) - 1.09\times10^{-4}\cdot(r/h) + 3.18\times10^{-5}\cdot(r/h)^2 + 2.53\times10^{-5}\cdot(\text{shape}) - 5.09\times10^{-5}\cdot(\text{shape})^2 + 2.38\times10^{-5}\cdot(\text{area}) + 5.96\times10^{-5}\cdot(n_b) - 2.65\times10^{-4}\cdot\log(F) + 6.24\times10^{-5}\cdot(E) + 1.29\times10^{-4}\cdot(\rho) - 1.49\times10^{-5}\cdot\log(F)\cdot(\rho) \qquad (27)$$

It is worthwhile to discuss why the terms $(r/h)^2$, $(\text{shape})^2$, and $(\log(F)\cdot\rho)$ were included in eq. (26) and (27). The first of these terms, $(r/h)^2$ and $(\text{shape})^2$, were included because of the shape of the graphs "resistance vs. r/h" (Fig. 33) and "resistance vs. shape" (Fig. 34), both of which look significantly parabolic. No other variable showed such a great parabolic tendency. The interaction term, $(\log(F)\cdot\rho)$, was included because, by inspection of Table 6, the three "best" contacts (lowest μ and σ,) occurred at the three combinations of high force and low ρ: trials 5, 18, 19 in column 1 of Table 6. Two of the three worst contacts (highest μ and σ) occurred at two of the three combinations of low force and high ρ: trials 12 and 22. Thus, force and ρ appeared to show interaction.

TABLE 5

(a) VARIABLE NAMES

avg(h) = average height of asperities, units are μm;
the variance is $(ht/3)^2$.

r/h = (average radius of curvature)/(average height) ratio;
the variance of curvature is $(\text{avg radius of curv}/3)^2$

shape = length/width ratio of rectangular apparent contact area.

A_{app} = apparent contact area in mm^2.

d = asperities per mm^2. $d = n_b/A_{app}$, n_b is not random.

F = contact normal force in Newtons.

E = Young's modulus of elasticity in 10^{10} Pa.

ρ = electrical resistivity in 10^{-8} $\Omega \cdot m$.

(b) ANALYSIS OF VARIANCE OF MEAN CONTACT RESISTANCE

variable	sum of squares	df	mean square	f	
avg(h)	8.682×10^{-8}	1	8.682×10^{-8}	0.181	
r/h	3.282×10^{-6}	1	3.282×10^{-6}	6.853	*
$(r/h)^2$	9.797×10^{-7}	1	9.797×10^{-7}	2.045	
shape	3.287×10^{-7}	1	3.287×10^{-7}	0.686	
$(shape)^2$	2.976×10^{-6}	1	2.976×10^{-6}	6.214	*
A_{app}	7.785×10^{-8}	1	7.785×10^{-8}	0.162	
d	2.162×10^{-6}	1	2.162×10^{-6}	4.513	
F	2.646×10^{-5}	1	2.646×10^{-5}	55.239	**
E	9.479×10^{-7}	1	9.479×10^{-7}	1.979	
ρ	7.701×10^{-6}	1	7.701×10^{-6}	16.077	**
total	9.303×10^{-5}	27			
residual	7.664×10^{-6}	16	4.790×10^{-7}		

* significant at 95% ** significant at 99%

TABLE 6

TAGUCHI ARRAY FOR CR1 VARIABLES [57]

no.	avg(h)	r/h	shape	A_{app}	d[a]	F	E	ρ	$\mu(m\Omega)$[b]	$\sigma(m\Omega)$[b]
1	5	5	1	1	50	0.1	7	1.8	1.915	0.524
2	10	50	5	2	50	1	9	4	0.791	0.219
3	15	100	10	3	50	10	11	6.2	0.354	0.087
4	5	5	1	2	75	1	9	6.2	1.312	0.191
5	10	50	5	3	75	10	11	1.8	0.086	0.019
6	15	100	10	1	75	0.1	7	4	1.438	0.209
7	5	5	1	3	100	10	11	4	0.175	0.010
8	10	50	5	1	100	0.1	7	6.2	2.766	0.459
9	15	100	10	2	100	1	9	1.8	0.276	0.097
10	5	50	10	1	100	1	11	1.8	0.369	0.116
11	10	100	1	2	100	10	7	4	0.142	0.026
12	15	5	5	3	100	0.1	9	6.2	5.744	1.372
13	5	50	10	2	50	10	7	6.2	0.192	0.025
14	10	100	1	3	50	0.1	9	1.8	0.779	0.136
15	15	5	5	1	50	1	11	4	1.580	0.435
16	5	50	10	3	75	0.1	9	4	2.253	0.474
17	10	100	1	1	75	1	11	6.2	1.228	0.267
18	15	5	5	2	75	10	7	1.8	0.107	0.019
19	5	100	5	1	75	10	9	1.8	0.070	0.013
20	10	5	10	2	75	0.1	11	4	4.349	0.995
21	15	50	1	3	75	1	7	6.2	0.992	0.282
22	5	100	5	2	100	0.1	11	6.2	3.577	0.766
23	10	5	10	3	100	1	7	1.8	0.397	0.083
24	15	50	1	1	100	10	9	4	0.261	0.076
25	5	100	5	3	50	1	7	4	0.508	0.126
26	10	5	10	1	50	10	9	6.2	0.479	0.064
27	15	50	1	2	50	0.1	11	1.8	0.870	0.181

[a] n_b is a constant and not Poisson distributed as in Chapter 2.

[b] based on 30 resistances generated at the given settings.

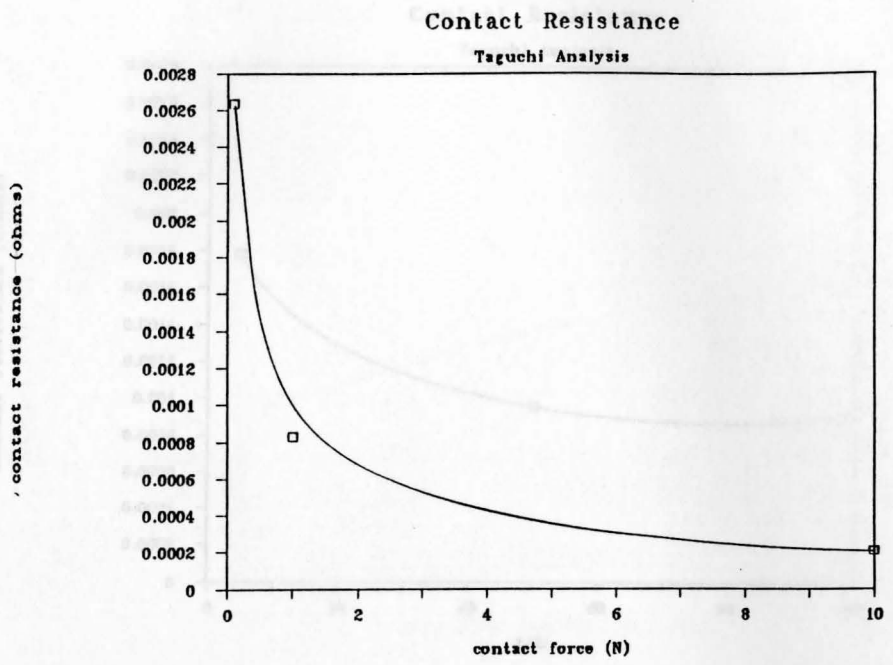


Fig. 31.--Average contact resistance at different levels of force.

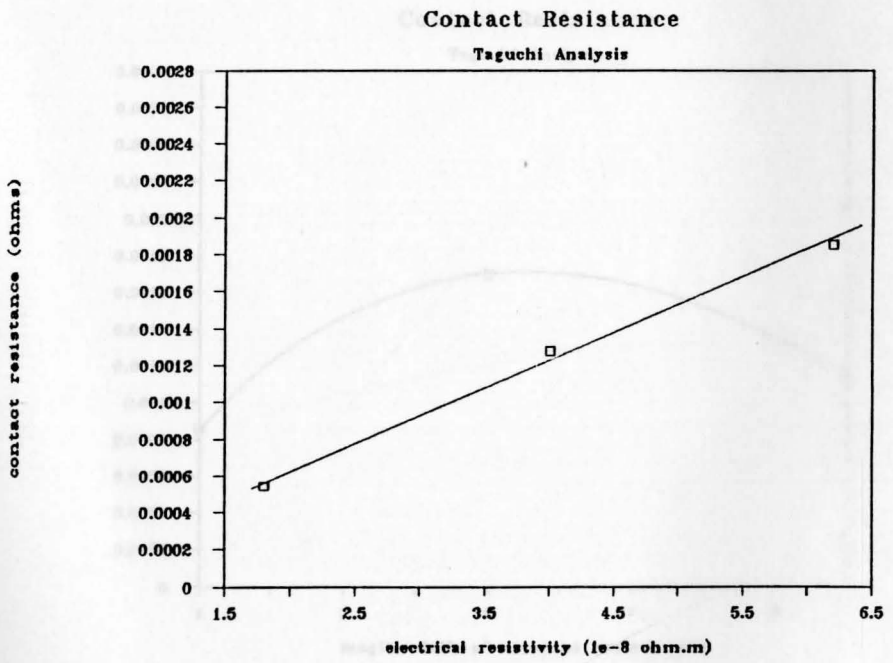


Fig. 32.--Average contact resistance at different levels of ρ.

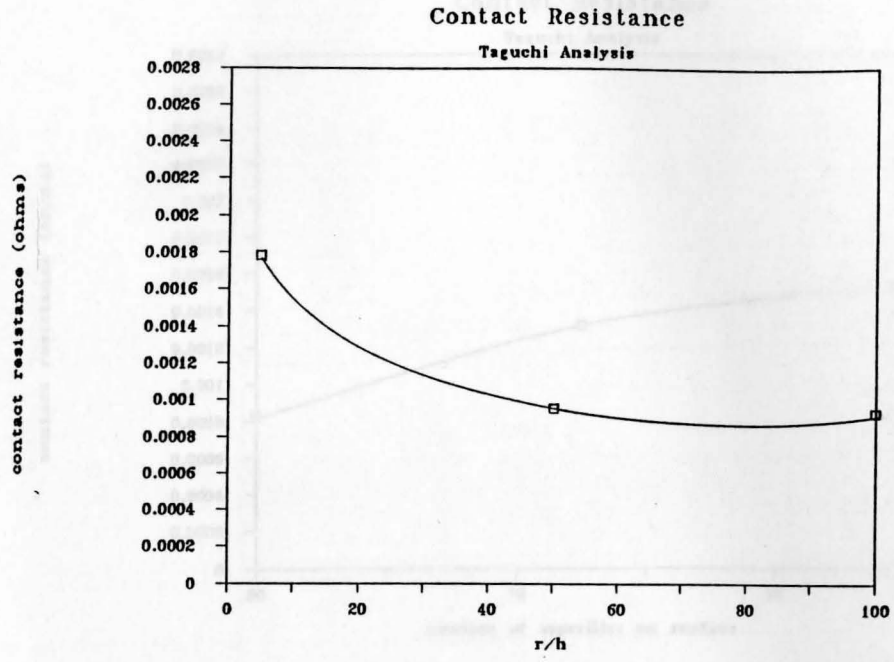


Fig. 33.--Average contact resistance at different levels of r/h.

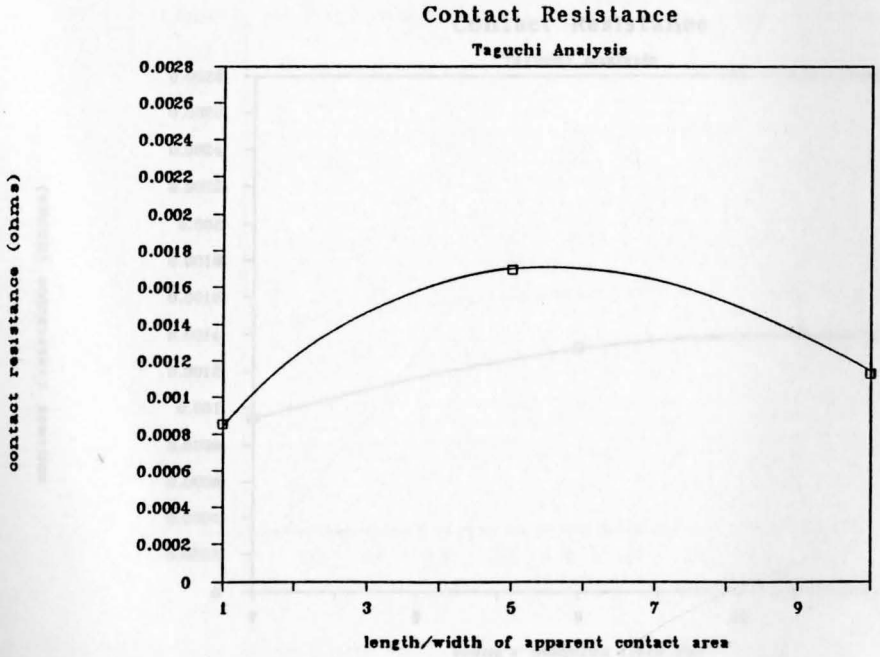


Fig. 34.--Average contact resistance at different levels of "shape."

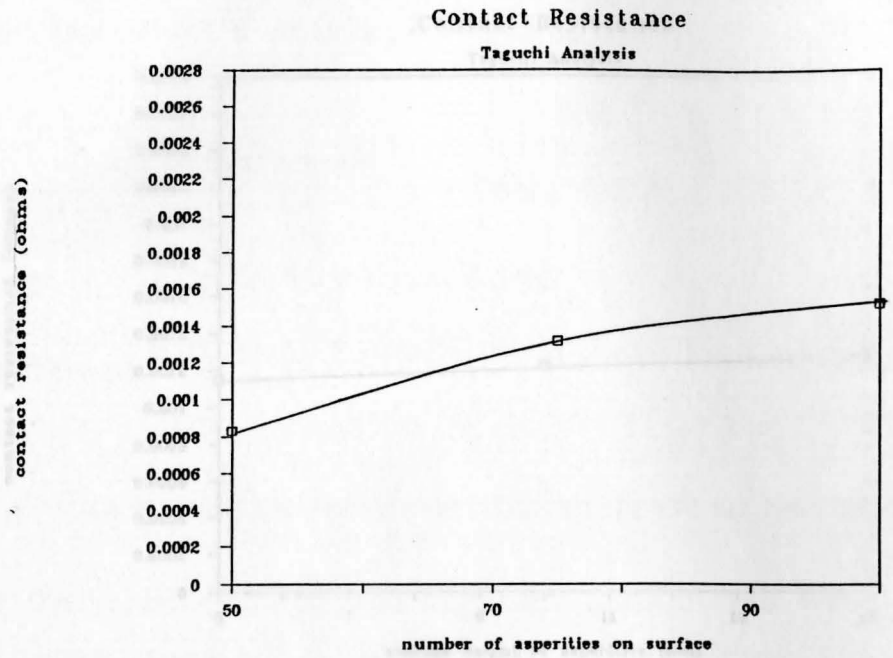


Fig. 35.--Average contact resistance at different levels of *d*.

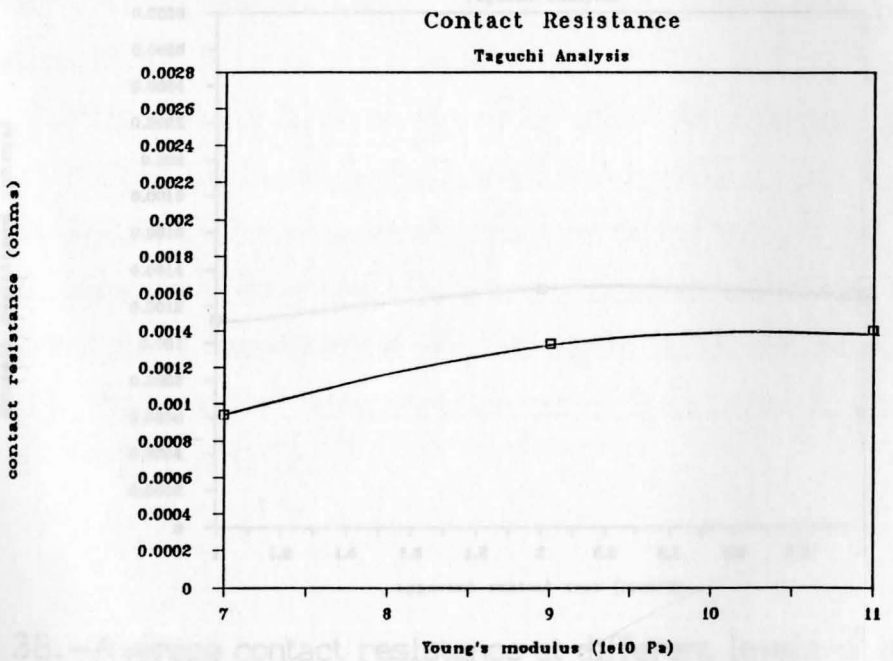


Fig. 36.--Average contact resistance at different levels of E.

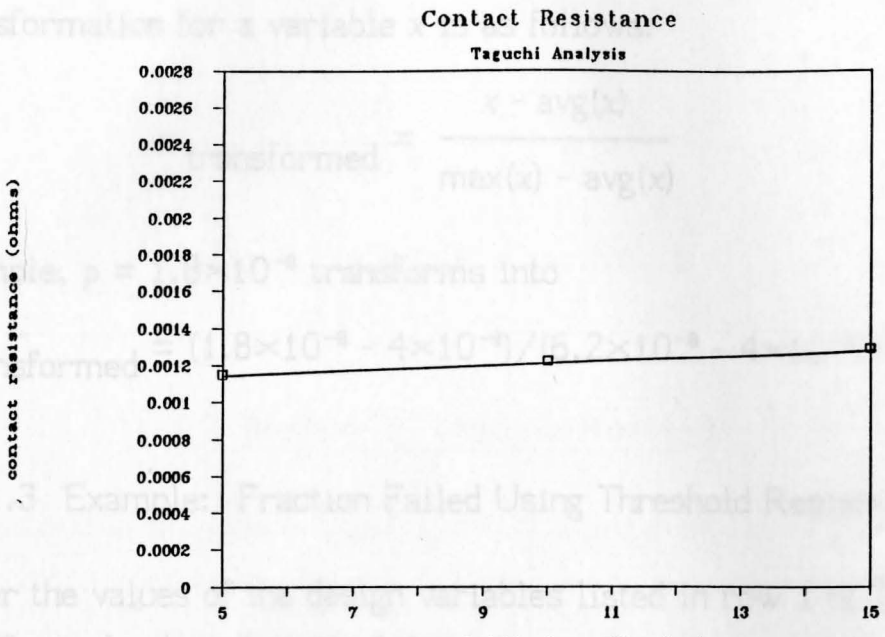


Fig. 37.--Average contact resistance at different levels of avg(h).

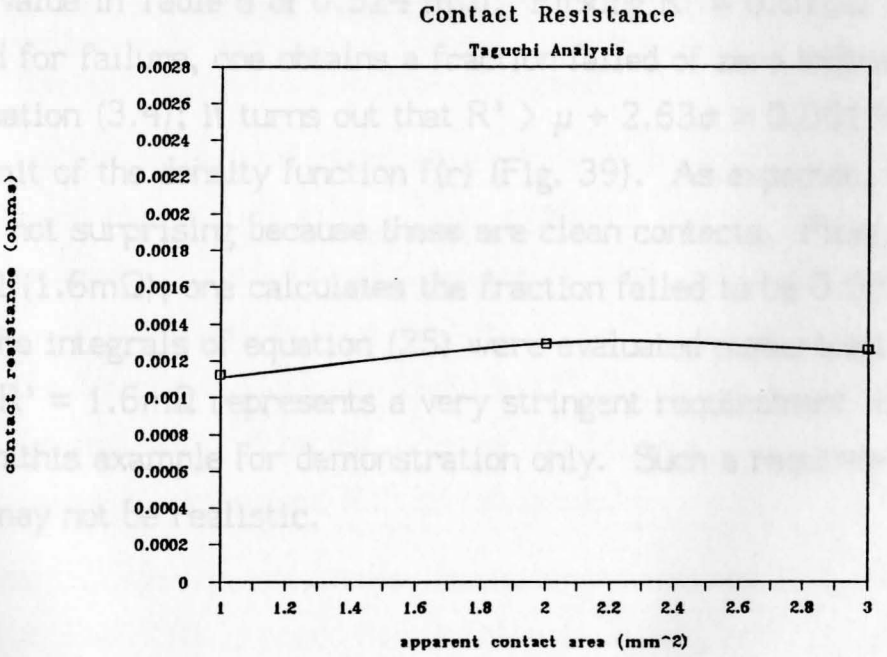


Fig. 38.--Average contact resistance at different levels of A_{app}.

All of the variables, with the exception of force, F, must be transformed before being substituted into equations (26) and (27). The transformation for a variable x is as follows:

$$x_{\text{transformed}} = \frac{x - \text{avg}(x)}{\max(x) - \text{avg}(x)}$$

For example, $\rho = 1.8\times10^{-8}$ transforms into

$$\rho_{\text{transformed}} = (1.8\times10^{-8} - 4\times10^{-8})/(6.2\times10^{-8} - 4\times10^{-8}) = -1.$$

3.1.3 Example: Fraction Failed Using Threshold Resistance

For the values of the design variables listed in row 1 of Table 6, the transformed values for all of them is -1. Substituting -1 into equations (26) and (27), one obtains $\mu = 1.22$ mΩ (as compared to the data value in Table 6 of 1.915 mΩ) and $\sigma = 0.281$ mΩ (as compared to the data value in Table 6 of 0.524 mΩ). Picking R' = 0.010Ω as a threshold for failure, one obtains a fraction failed of zero indirectly from equation (3.4); it turns out that $R' > \mu + 2.63\sigma = 0.00196$, the upper limit of the density function f(r) (Fig. 39). As expected, this result is not surprising because these are clean contacts. Picking R' = 0.0016Ω (1.6mΩ), one calculates the fraction failed to be 0.097 or 9.7%. The integrals of equation (25) were evaluated numerically. A value of R' = 1.6mΩ represents a very stringent requirement and was chosen in this example for demonstration only. Such a requirement may or may not be realistic.

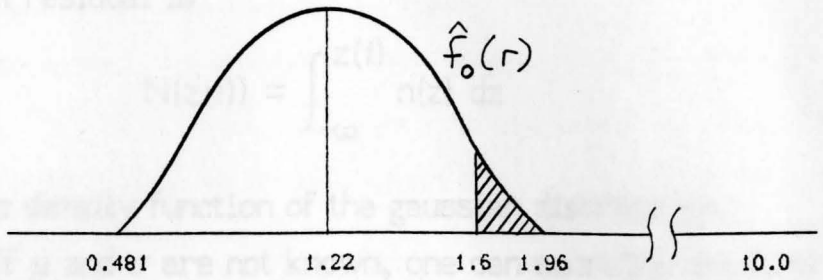


Fig. 39.--Normal distribution.

3.2 Method II: Uniform Residuals

Yamada et al. give a clue for determining fraction failed based directly on the contact resistance distribution of a sample group. Repeating their quotes from sections 1.6.1-1.6.2 of this thesis:

> If the contact part of the contacts is clean, distribution of contact resistance approaches normal distribution and the normal probability curve approximates a straight line.
> ...if the contact part of contacts is contaminated, distribution is apart from normal distribution, and the tail of the normal probability line is curved...[58]

A method is desired to determine how many contacts in a group fall off the normal distribution. These contacts, according to Yamada, can be considered as failed or doomed for failure. Quesenberry's Method of Uniform Residuals is the type of method required.

In the Method of Uniform Residuals [59], one calculates either the value of the statistic $z(i)$ or the value of the Student's t statistic, $t(i)$, for every value of resistance $R(i)$ in the group.

Case 1. Ideally, one knows the parameters μ and σ of the normal distribution that $R(i)$ is presumed to follow. In this case, one calculates

$$z(i) = \frac{R(i) - \mu}{\sigma}$$

Then the uniform residual is

$$N(z(i)) = \int_{-\infty}^{z(i)} n(z)\, dz$$

where n(z) is the density function of the gaussian distribution.

Case 2. If μ and σ are not known, one can sacrifice the first two values from the data to estimate μ and σ, then update the estimate with each successive value of R(*i*). In this case, one calculates

$$t(i) = \frac{i - 1}{i} \cdot \frac{R(i) - R_{avg(i-1)}}{s_{i-1}} \qquad (28)$$

Here, $R_{avg(i-1)}$ is the average, and s_{i-1} the standard deviation, of the preceding (*i*-1) values of R.

$$R_{avg(i-1)} = \frac{1}{i - 1} \sum_{j=1}^{i-1} R(j) \qquad (29)$$

$$s_{i-1} = \frac{1}{i - 2} \sum_{j=1}^{i-1} \left(R(j) - R_{avg(i-1)} \right)^2 \qquad (30)$$

From the value of t(*i*), one looks up the "uniform residual" in the statistical tables. The uniform residual is

$$G_\nu(t(i)) = \int_{-\infty}^{t(i)} g_\nu(t)\, dt$$

where $g_\nu(t)$ is the density function of the t-distribution with ν degrees of freedom, $\nu = i - 2$. $G_\nu(t)$ is the cumulative distribution function.

Case 2 will be more common, because given a set of resistance data, one usually has no knowledge of μ and σ. Therefore, the rest of this discussion will use the terminology of case 2, although the

arguments translate easily to case 1.

$G_\nu(t(i))$ is, loosely translated, a "percentile," meaning that $100 \cdot G_\nu(t(i))\%$ of the resistances following the presumed normal distribution are expected to be smaller than $R(i)$. $G_\nu(t(i))$ increases as $R(i)$ increases.

An unreasonably large $G_\nu(t(i))$ means that $R(i)$ is probably too large to be part of the normal distribution that it was presumed to follow; the probability that $R(i)$ can be part of the presumed normal distribution is $1-G_\nu(t(i))$, which will be very small. Taking a hint from Yamada, one can say that contact i has failed; $R(i)$ is part of the "tail" of the NPP. This is consistent with the knowledge that high resistances indicate failure.

How large $G_\nu(t(i))$ must be before one says contact i has failed depends on the expected fraction failed. If one of every N_f contacts is expected to fail, then the expected reliability is defined as

$$1 - \text{expected fraction failed} = 1 - 1/N_f = \frac{N_f - 1}{N_f} = \beta \qquad (31)$$

β is the fraction of contacts expected to not fail. β can be thought of as

$$\beta = \int_{-\infty}^{t^*} g_\nu(t)\, dt$$

or as

$$1/N_f = 1 - \beta = \int_{t^*}^{\infty} g_\nu(t)\, dt.$$

See Figure 40.

If $G_\nu(t(i)) > \beta$, then contact i is expected to fail (Fig. 40), and any other contact with a resistance greater than or equal to $R(i)$ is also

expected to fail. On the contrary, if $G_\nu(t(i)) > \beta$ and contact i is not expected to fail, then β was underestimated, because if contact i is not expected to fail, neither is any contact whose resistance is less than $R(i)$, and the expected reliability is at least equal to $G_\nu(t(i))$.

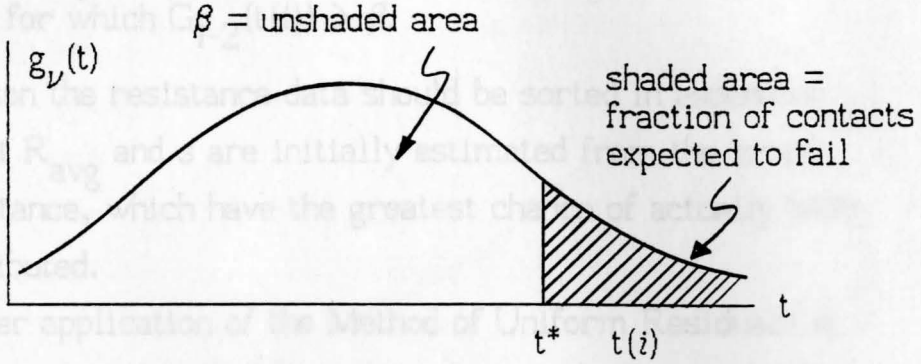


Figure 40.--Student's "t"-distribution.
Note that $t(i)$ is just a linear transformation of $R(i)$.

The procedure recommended for the Method of Uniform Residuals is as follows:

1) List the resistance data in ascending order. Resistance data can be obtained through measurements of actual contacts or through simulation by a computer program. Assign a value i to each resistance, the lowest resistance assigned i=1 and the highest resistance, i=N. Assume N resistance values.

2) Using the first two resistance data R(1) and R(2), calculate $R_{avg(2)}$ and $s_{(2)}$, using equations (29) and (30).

3) Calculate t(3) for the resistance value R(3) using equation (28), then look up and record $G_1(t(3))$.

4) Use R(3) (along with R(1) and R(2)) to calculate $R_{avg(3)}$ and $s_{(3)}$, then calculate t(4) for the resistance value R(4) and look up $G_2(t(4))$.

5) Continue in this fashion for all i, using $R(1)...R(i-1)$ to calculate $R_{avg(i-1)}$ and $s_{(i-1)}$; then calculate $t(i)$ and look up $G_{i-2}(t(i))$ for the resistance value $R(i)$.

6) Calculate β from equation (31) and compare $G_{i-2}(t(i))$ to β for each i. The estimated fraction failed is that fraction of data for which $G_{i-2}(t(i)) > \beta$.

The reason the resistance data should be sorted in ascending order is so that R_{avg} and s are initially estimated from the lowest values of resistance, which have the greatest chance of actually being normally distributed.

A simpler application of the Method of Uniform Residuals is inspection of the normal probability plot. A straight line is drawn to fit the lowest values of resistance on the plot. The contacts that fail are those with high values of resistance that fall off the line. An example is shown in Fig. 41, where the contact surface data is the same as that in Table 2, and the contact force is nominally 1N with an "out-of control" variation of $\pm$ 30% (standard deviation of the force is such that $\pm 3\sigma$ = 0.3 times the nominal force = 0.3N, or σ = 0.1N). In the example, there are six "outliers" (six high resistance values that fall off the straight line) for an estimated fraction failed of 6/100 = 0.06 or 6%.

In this example, it was assumed that the high resistance of the outliers was due to a manufacturing defect that led to an abnormally low contact force; and that this defect will worsen as the contacts age, eventually leading to contact failure.

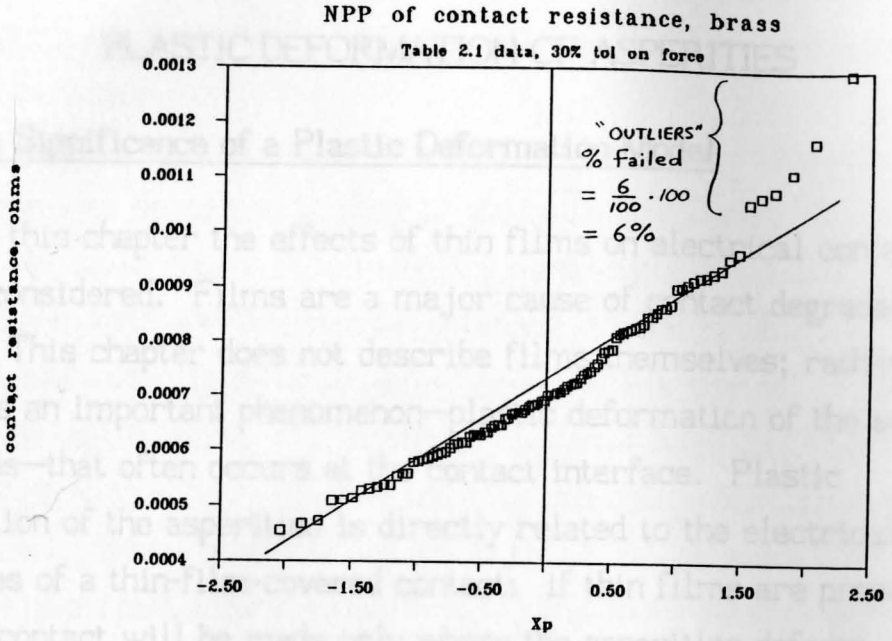


Fig. 41.--Graphical representation of the Method of Uniform Residuals.

CHAPTER IV.

PLASTIC DEFORMATION OF ASPERITIES

4.1 The Significance of a Plastic Deformation Model

In this chapter the effects of thin films on electrical contacts will be considered. Films are a major cause of contact degradation and failure. This chapter does not describe films themselves; rather, it describes an important phenomenon—plastic deformation of the surface asperities—that often occurs at the contact interface. Plastic deformation of the asperities is directly related to the electrical properties of a thin-film-covered contact: if thin films are present, electric contact will be made only where the asperities deform plastically--this is where the films are broken. [60]

In the computer models CR1, CR2, and CR3, described in Chapters II and III, the asperities are made of solid metal and are subject to purely elastic deformation as given by the Hertz equations. However, in a real contact, the stresses may be high enough to cause plastic deformation of the individual asperities. This chapter addresses the issue of plasticity on a contact surface and describes a modified computer model, CR1P, which accounts for plastic deformation of the asperities.

Historically it was thought that at small contact forces, a contact would undergo purely elastic deformation, and at higher contact forces, that same contact would undergo purely plastic deformation. However, the transition force (the force at which the contact's deformation changes from elastic to plastic) has been found to be either extremely

high or negligibly low.

A negligibly low transition force means that, even at the smallest force, the contact will undergo significant plastic deformation. An extremely high transition force means that no matter how hard the contact members are pressed together, elastic deformation dominates at the contact points. [61]

A probabilistic model of the contact interface by Greenwood and Williamson was developed to predict the transition force. Their model outlines one basic approach to modeling plasticity. In developing their model, Greenwood and Williamson obtained a relation which tells at what value of asperity deformation (δ) plastic flow begins. It is called the critical deformation, δ^*. Another relation, quoted by Chhabra and Wenning, gives the load which an asperity can bear while under plastic deformation. These relations are the basis of the program CR1P.

An important intermediate relationship in Greenwood and Williamson's model tells us how large the maximum contact pressure (q_0) must be to cause plastic deformation. Several steps from different sources were pieced together in this chapter to reach their conclusion that the contact pressure must reach a certain fraction of the hardness (H) of the metal: $q_0 > 0.6H$. Following a necessary background section on plasticity (sec. 4.2-4.3), this chapter is devoted to first deriving this criterion (sec. 4.4), then describing Greenwood and Williamson's model (sec. 4.5).

4.2 Description of Plastic Deformation

First it is necessary to define two terms: (1) normal stress is the stress perpendicular (normal) to a plane; (2) shear stress is the stress tangential or parallel to a plane.

Holm says that "plastic deformation of crystalline solids

proceeds by slips". [62] A slip, also called a plastic shear by Van Vlack, is the sliding of one plane of atoms over an adjacent plane of atoms. [63] The interface between the two planes of atoms is called the "slip plane." A shear force along a slip plane will cause shear: the displacement of one plane of atoms over the other.

If after the removal of the shear stress, the two planes of atoms remain displaced relative to each other, then the shear is called a slip or a plastic shear. If the two planes of atoms return to their original positions relative to each other, the shear is called elastic.

The "strength" of a metal depends on how well it resists plastic shear. Real metals are much weaker than metals made up of ideal crystal lattices. That is because real metals contain defects called "dislocations" that weaken the real metal by producing slip planes that slip easily. Shearing perpendicular to an edge dislocation or parallel to a screw dislocation is much easier than shearing in any direction in an ideal defect-free crystal lattice.

Depending on the type of dislocation, one defines a "slip vector" $\underline{b}$, which is the displacement of atoms from their hypothetical position in an ideal lattice due to a dislocation. [64] $\underline{b}$ for an edge dislocation is illustrated in Fig. 42.

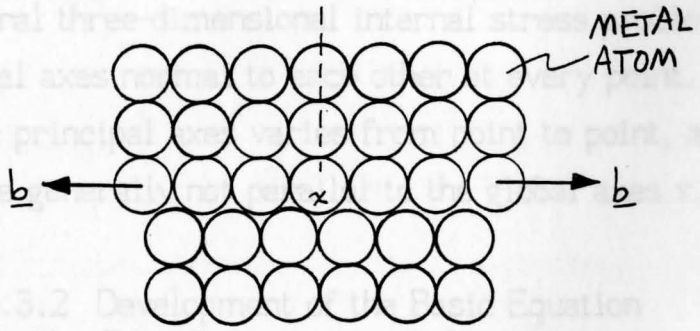


Fig. 42.--Edge dislocation and the slip vector.

For the edge dislocation, $\underline{b}$ is perpendicular to the dislocation. The edge dislocation is the extra plane of atoms; its edge is a line into the paper at point x.

Slip occurs most easily in a direction parallel to $\underline{b}$, so that plastic deformation depends on the shear stress parallel to $\underline{b}$. In a real metal, there are enough dislocations in enough directions so that shearing of the metal lattice can occur in any direction. That is why macroscopic properties like hardness are practically isotropic.

4.3 Huber-von Mises Equation: Criterion for Plastic Deformation

4.3.1 Principal Stresses

The amount of stress required for plastic deformation was empirically determined by Huber in 1904 and independently by von Mises in 1913. This equation is fundamental to any discussion of plastic deformation, and since it is used extensively in the ensuing calculations, it is worthwhile to derive and explain it now. The derivation is adapted from Tabor.

Before entering the following discussion, three terms must be defined. (1) A principal plane (of stress) is a plane in an elastic body across which the shearing stress is zero. [65] (2) A principal axis is a line normal to a principal plane. (3) A principal stress is a stress acting normal to a principal plane.

For a general three-dimensional internal stress problem, there are three principal axes normal to each other at every point. The orientation of the principal axes varies from point to point, and the principal axes are generally not parallel to the global axes x, y, and z. [66]

4.3.2 Development of the Basic Equation

According to Tabor

> Hydrostatic pressure will not of itself produce plastic deformation. [67]

Hydrostatic pressure is the pressure that presses on all sides of a body equally, such as that when a body is submersed in a pressurized water tank. This pressure is described in Pascal's law.

Tabor continues:

> ...if a metal is subjected to combined stresses e.g. components in the x,y,z directions], the only part of the stresses which will be effective in producing plastic deformation will be the part left after the hydrostatic component has been subtracted. [68]

Take a small cube of metal whose sides are principal planes of stress, subjected to principal stresses p_1, p_2 and p_3 (Fig. 43(a)). The hydrostatic component is

$$p_h = (p_1 + p_2 + p_3)/3$$

The hydrostatic component p_h is subtracted the principal stresses to get the reduced stresses (Fig. 43(b)).

$$p_1 - p_h \ , \ p_2 - p_h \ , \ p_3 - p_h$$

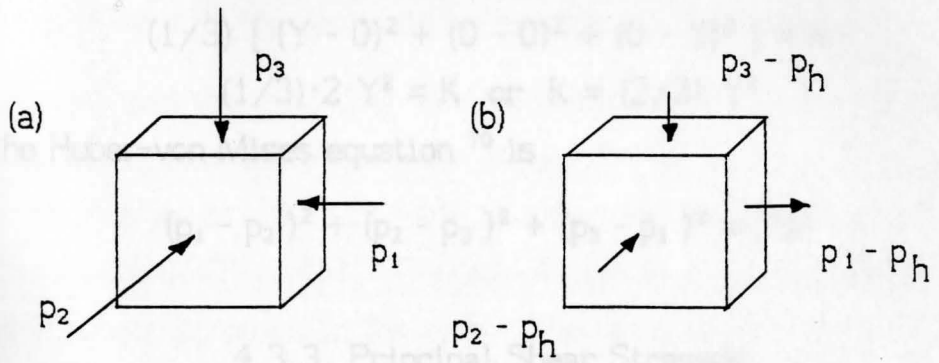


Fig. 43.--Principal stresses and reduced stresses on a cube.

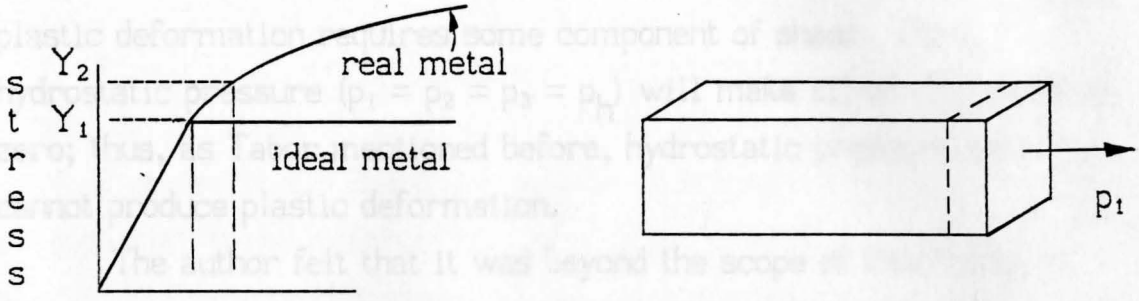
The criterion for producing plastic deformation was determined experimentally as

$$(p_1 - p_h)^2 + (p_2 - p_h)^2 + (p_3 - p_h)^2 = K, \quad K \text{ is a constant}$$

This equation can be written as [69]

$$(1/3)\ [\ (p_1 - p_2)^2 + (p_2 - p_3)^2 + (p_3 - p_1)^2\ ] = K$$

In pure tension, $p_1 = Y$, $p_2 = p_3 = 0$ produces plastic deformation along the x-axis only (Fig. 44). Y is the yield stress of the metal. (For an ideal metal, Y equals the elastic limit. For a real metal, Y is slightly larger than the elastic limit.)



Y_1 = elastic limit, Y_2 = yield stress

Fig. 44.--Plastic deformation along the x-axis.

substituting,

$$(1/3)\ [\ (Y - 0)^2 + (0 - 0)^2 + (0 - Y)^2\ ] = K$$

$$(1/3)\cdot 2\cdot Y^2 = K \quad \text{or} \quad K = (2/3)\ Y^2$$

Then the Huber-von Mises equation [70] is

$$(p_1 - p_2)^2 + (p_2 - p_3)^2 + (p_3 - p_1)^2 = 2Y^2 \qquad (32)$$

4.3.3 Principal Shear Stresses

According to Roark

> The maximum shear stress [in the cube of metal] occurs on each of two planes inclined at 45° to the two principal stresses whose algebraic difference is greatest and is equal to one-half that algebraic difference. [71]

That is, the maximum shear stress in the cube is the largest of

the three stresses given in equations (33).

$$p_{12} = (p_1 - p_2)/2$$
$$p_{23} = (p_2 - p_3)/2 \qquad (33)$$
$$p_{31} = (p_3 - p_1)/2$$

The stresses p_{12}, p_{23} and p_{31} are called principal shear stresses by Holm and by Tabor. [72] Holm notes that, for plastic deformation, at least one of eq. (33) must be nonzero, else (32) will not be satisfied. That is, plastic deformation requires some component of shear. Pure hydrostatic pressure ($p_1 = p_2 = p_3 = p_h$) will make all of (33) equal to zero; thus, as Tabor mentioned before, hydrostatic pressure alone cannot produce plastic deformation.

The author felt that it was beyond the scope of this thesis to derive equations (33), however, some insight can be obtained by considering the example in Appendix E.

4.4 Contact Pressure and Deformation Required for Plastic Yielding

4.4.1 Hardness as a Function of Yield Stress

Holm used the following informal discussion to roughly show how hardness (H) is related to the yield stress (Y) of a metal.

Imagine an indentation made in a nominally flat slab of metal by a smooth, hard ball (Fig. 45). The indentation is a plastic deformation caused by shear stresses in the slab. The ratio of the depth to the radius of curvature of the indentation (called the "specific depth") is approximately 0.03 (an important fact, but not a number which will be used in any calculations here).

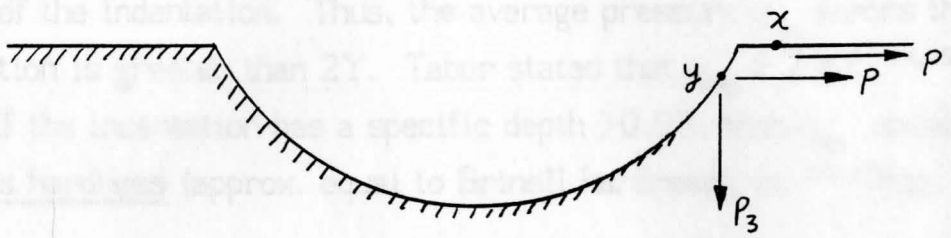


Fig. 45.--Indentation made by a hard ball in a metal slab.

Pick a point x near the rim of the indentation, the outer edge of plastic deformation. There is no vertical deformation at x, therefore no vertical stress at x. There are only horizontal stresses: one component is tangential to the mouth of the indentation, the other is normal to the mouth. The principal stresses p_1, p_2, p_3 at x must satisfy the Huber-von Mises equation.

If one assumes that the horizontal principal stresses are at least approximately equal, then one can let $p_1 = p_2 = p$, and $p_3 = 0$ in eq. (32):

$$(p - p)^2 + (p - 0)^2 + (0 - p)^2 = 2Y^2$$

$$2p^2 = 2Y^2 \quad \text{or} \quad p = Y$$

which makes intuitive sense: the principal stresses in the horizontal directions must equal the yield stress before plastic deformation can occur.

At point y, near x, the horizontal stresses are of the same order of magnitude as those at x. However, a vertical stress p_3 is also present. Its value is given by the Huber-von Mises equation (32). [73]

$$(Y - Y)^2 + (Y - p_3)^2 + (p_3 - Y)^2 = 2Y^2$$

$$2 \cdot (p_3 - Y)^2 = 2Y^2$$

$$2p_3^2 - 4p_3Y + 2Y^2 = 2Y^2$$

$$2p_3 \cdot (p_3 - 2Y) = 0$$

$$p_3 = 2Y$$

The vertical stress given by this ballpark calculation is twice the yield stress. It should be noted that p_3 increases as we move toward the center of the indentation. Thus, the average pressure q_m across the indentation is greater than 2Y. Tabor states that $q_m \simeq 2.8Y$. [74,75]

If the indentation has a specific depth >0.03, then q_m equals the Meyer's hardness (approx. equal to Brinell hardness), H. [76] Thus,

$$H = q_m \simeq 2.8Y \tag{34}$$

4.4.2 Maximum Shear Stress Under an Elastic Indentation

For the same ball indenter problem, Davies (1949) calculated the stresses inside the slab—as a function of the maximum pressure (q_0) on the surface of the slab—prior to the onset of plastic deformation. His result and Holm's (eq. 34) are used to find the pressure at which plastic deformation will begin.

He selected a global coordinate system x,y,z (Fig. 46). The plane perpendicular to the y-axis and passing through the center of the indentation is the "y_0 plane." At every point (x,0,z) in the y_0 plane there are normal stress components parallel to the x-, y-, and z-axes: respectively they are Xx, Yy, Zz. There is also one component of shear stress parallel to the x-axis and acting along the x-y plane: call it Xz. [77] If at (x,0,z) the shear stress Xz is zero, the principal stresses are simply Xx, Yy, and Zz. This follows from the definition of principal stress. In general, however, the principal stresses are given by the following equations: [78]

$$\begin{aligned} p_1 &= Xx \\ p_2 &= (1/2)[(Yy + Zz) - \{(Zz - Yy)^2 + 4Xz^2\}^{1/2}] \\ p_3 &= (1/2)[(Yy + Zz) + \{(Zz - Yy)^2 + 4Xz^2\}^{1/2}] \end{aligned} \tag{35}$$

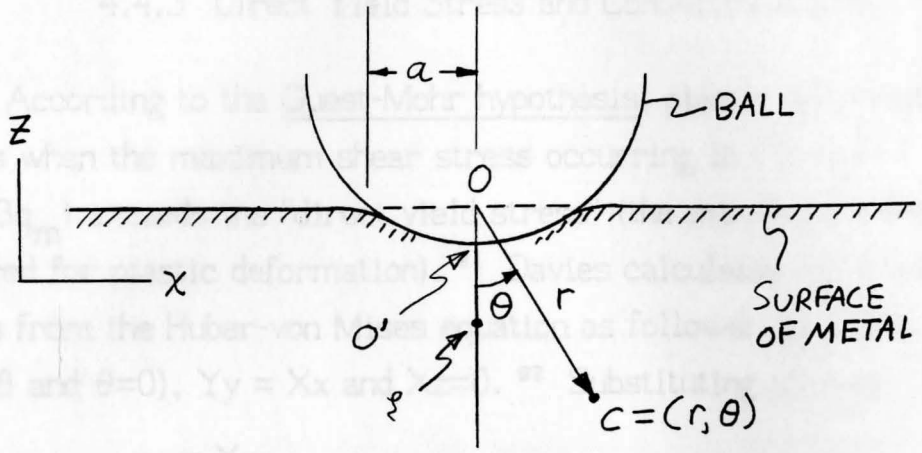


Fig. 46--Coordinate system and stresses in Davies' problem.

Davies plotted the "maximum difference between principal stresses" versus (r/a) and θ, a being the radius of contact. The overall maximum occurred at $(r/a) \simeq 0.48$ and $\theta = 0$. At this point, which will be called ζ, [79]

$$\frac{\max |p_i - p_j|}{q_0} = 0.63 \qquad i=1,2,3;\ j=1,2,3$$

where q_0 is the pressure at point O' in Fig. 46. q_0 is the maximum pressure on the surface of the slab.

According to Tabor

$$q_0 = (3/2)\, q_m \qquad \text{where } q_m \text{ is the average pressure over the contact area.} \quad [80] \qquad (36)$$

Thus, $$\max |p_i - p_j| = 0.63 \cdot (3/2) \cdot q_m$$

It follows from eq. (33) that the maximum shear stress in the slab is

$$\begin{aligned} \text{max shear} &= (1/2)\ \max |p_i - p_j| \\ &= (1/2) \cdot 0.63 \cdot (3/2) \cdot q_m \\ &= 0.473\ q_m \end{aligned}$$

4.4.3 Direct Yield Stress and Contact Pressure

According to the Guest-Mohr hypothesis, plastic deformation at ζ occurs when the maximum shear stress occurring in the slab ($0.473q_m$) exceeds the "direct yield stress" (the minimum shear stress required for plastic deformation). [81] Davies calculated the direct yield stress from the Huber-von Mises equation as follows: at ζ (i.e., $(r/a) = 0.48$ and $\theta=0$), $Yy = Xx$ and $Xz=0$. [82] Substituting into eq. (35)

$$p_1 = Xx$$

$$p_2 = (1/2)\,[(Xx + Zz) - \left((Zz - Xx)^2 + 4\cdot 0^2\right)^{1/2}]$$

$$= (1/2)\,[Xx + Zz - Zz + Xx]$$

$$= Xx$$

$$p_3 = (1/2)\,[(Xx + Zz) + \left((Zz - Xx)^2 + 4\cdot 0^2\right)^{1/2}]$$

$$= (1/2)\,[Xx + Zz + Zz - Xx]$$

$$= Zz$$

Thus one obtains the important result $p_1 = p_2$.

Letting $p_1 = p_2 = p$ in the Huber-von Mises equation (32), one obtains the criterion for plastic deformation at ζ

$$(p - p)^2 + (p - p_3)^2 + (p_3 - p)^2 = 2Y^2$$

$$2\cdot(p_3 - p)^2 = 2Y^2$$

that is,

$$|p_3 - p| = Y$$

In terms of the principal shear stresses (eq. 33) the criterion is

$$p_{12} = |p - p|/2 = 0$$

$$p_{23} = |p - p_3|/2 = Y/2$$

$$p_{31} = |p_3 - p|/2 = Y/2$$

These stresses are shown in Fig. 47. $Y/2$ is the "direct yield stress" mentioned in the Guest-Mohr hypothesis.

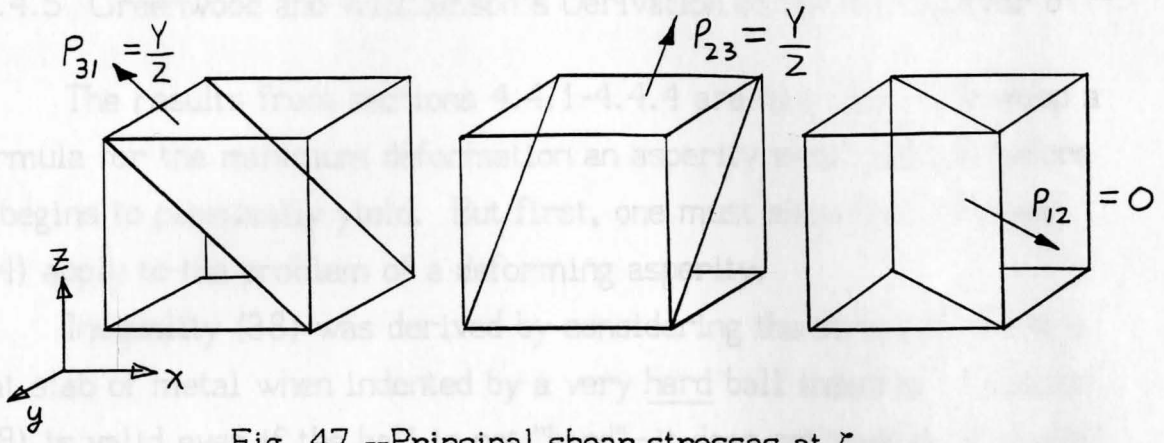


Fig. 47.--Principal shear stresses at ζ.

4.4.4 Contact Presssure Required for Plastic Deformation

Now the Guest-Mohr hypothesis can be used to obtain the value of contact pressure, in terms of the yield strength and the hardness, required for the onset of plastic deformation.

According to the Guest-Mohr hypothesis, plastic deformation should occur when [83,84]

$$0.473\, q_m \geq Y/2$$

$$\text{or } q_m \simeq 1.1Y \text{ for the onset of plastic deformation.} \qquad (37)$$

It was found that $H = q_m \simeq 2.8Y$ for full plastic deformation,

and $$q_0 = (3/2)\, q_m,$$

Taking the ratio,

$$\frac{q_0}{H} = \frac{(3/2)\, q_m}{H} = \frac{(3/2)\cdot(1.1)Y}{2.8Y} = 0.59$$

one gets the criterion

$$q_0 > \sim 0.6H \text{ for plastic deformation to begin.} \qquad (38)$$

[85]

4.4.5 Greenwood and Williamson's Derivation of the Formula for δ^* [86]

The results from sections 4.4.1-4.4.4 are now used to develop a formula for the minimum deformation an asperity must undergo before it begins to plastically yield. But first, one must show that (38) and (34) apply to the problem of a deforming asperity.

Inequality (38) was derived by considering the stresses inside a flat slab of metal when indented by a very hard ball indenter. However, (38) is valid even if the ball is not "hard"—it does not matter what kind of ball produces the pressure q_0, plastic deformation will begin in the slab as long as (38) is met. And if the flat slab is hard and the ball is soft, we expect plastic deformation to begin in the ball when (38) is met (H being the hardness of the ball). Thus by this informal argument, one can generalize (38) to apply to the onset of plastic deformation inside a spherically-shaped asperity, which is actually a "ball indenter."

Also, it is interesting to use a thought experiment similar to the one Holm used in deriving (34), but for the case of a hard flat slab plastically deforming a soft asperity (Fig. 48).

One uses the Huber-von Mises equation at x and then at y to show that the average pressure q_m on the contact surface of the asperity is ~ 3Y, so that eq. (34) applies to the problem shown in Fig. 48. [87]

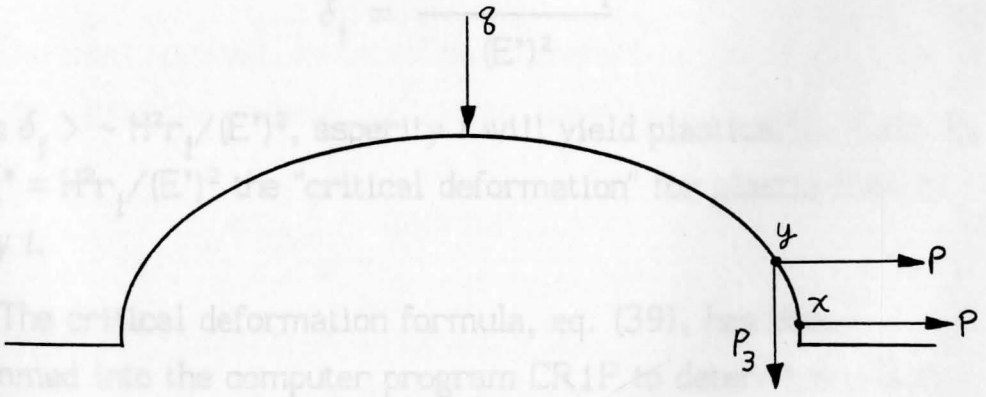


Fig. 48.--Full plastic deformation of an asperity.

Now recall the Hertz equation for the load supported by a single asperity under elastic deformation:

$$P = (4/3)\ \delta^{3/2}\ r^{1/2}\ E', \quad E' = (1/2)\ \frac{E}{1 - \mu^2} \qquad (6)$$

$$\delta^{3/2} = (3/4)\ P\ r^{-1/2}\ (E')^{-1}$$

$$= (3/4)\ q_m\ A\ r^{-1/2}\ (E')^{-1} \quad \text{where } q_m \text{ is the mean pressure on the contact spot; } q_m = P/A.$$

but from (36) it is known that

$$q_m = (2/3)\ q_0$$

$$\delta^{3/2} = (3/4)\ (2/3)\ q_0\ A\ r^{-1/2}\ (E')^{-1}$$

but $A = \pi r \delta$

$$\delta^{3/2} = (1/2)\ q_0\ \pi\ r^{1/2}\ \delta\ (E')^{-1}$$

$$\delta^{1/2} = (1/2)\ q_0\ \pi\ r^{1/2}\ (E')^{-1}$$

$$\delta = (1/2)^2\ \pi^2\ q_0^2\ r\ (E')^{-2}$$

It was found that (eq. 38) plastic deformation commences when

$q_0 > \sim 0.6H$ where H is the hardness of the metal

substituting: $\delta = (1/2)^2\ \pi^2\ (0.6H)^2\ r\ (E')^{-2}$

letting $r = r_i$ and $\delta = \delta_i$ for asperity i

$$\delta_i = \frac{0.888\ H^2\ r_i}{(E')^2} \qquad (39)$$

so when $\delta_i > \sim H^2 r_i/(E')^2$, asperity i will yield plastically. Call the term $\delta_i^* = H^2 r_i/(E')^2$ the "critical deformation" for plastic flow of asperity i.

(The critical deformation formula, eq. (39), has been programmed into the computer program CR1P to determine which contacting asperities are under plastic deformation.)

4.5 Greenwood and Williamson's Elastic Deformation Model [88]

4.5.1 Surface Model

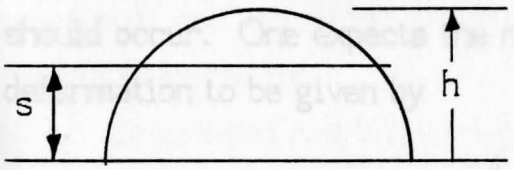
Assume a contact interface in which one surface is smooth and flat and the other is bumpy and nominally flat. The bumpy surface contains asperities which have identical radii of curvature, r. The probability that an asperity's height h is between z and z+dz above the base is $\phi(z)dz$.

Recall the Hertz equations:

$$P = (4/3)\ \delta^{3/2}\ r^{1/2}\ E' \qquad (6)$$

$$a = \delta^{1/2}\ r^{1/2} \qquad (5)$$

Let us take "s" as the separation between the face of the smooth flat surface and the base of the bumpy surface. Then the probability that an asperity will be in contact is the probability that its height will be greater than "s" (Fig. 49).



$$\text{prob}(h>s) = \int_{z=s}^{\infty} \phi(z)\ dz$$

Fig. 49.--Height and separation.

The total number of asperities in contact is the total number of asperities times the probability that an asperity is in contact.

$$n_c = n_b \int_{z=s}^{\infty} \phi(z)\ dz \qquad (40)$$

The area of contact of an asperity is

$$A = \pi\ a^2 = \pi\ (\delta^{1/2}\ r^{1/2})^2 = \pi\delta r$$

but $\delta = z\text{-}s$ from (14), using z for h_i.

so that $A = \pi\ (z\text{-}s)\ r$

The total area of contact depends on the "expected value" of $A = \pi \cdot r \cdot (z\text{-}s)$ taken over all asperities:

$$A_t = n_b \cdot \mu_A = n_b\ \pi\ r \int_{z=s}^{\infty} (z\text{-}s)\ \phi(z)\ dz \qquad (41)$$

Notice that eq. (41) implies that μ_A, the average area of contact of a single asperity, is A_t/n_b and not A_t/n_c. That is, the average area of contact is for all asperities, not just those asperities in contact.

The total value of the load supported by all asperities depends on the expected value of $(z\text{-}s)^{3/2}$ and is

$$P_t = n_b\ (4/3)\ E'\ r^{1/2} \int_{z=s}^{\infty} (z\text{-}s)^{3/2}\ \phi(z)\ dz \qquad (42)$$

This is an important relationship between P_t and separation, s.

If one looks at the stress vs. deformation diagram (Fig. 50) of a typical metal one sees that for $h\text{-}s > \delta^*$ (or $h > s+\delta^*$), plastic deformation should occur. One expects the number of asperities undergoing plastic deformation to be given by

$$n_{pc} = n_b \int_{z=s+\delta^*}^{\infty} \phi(z)\ dz$$

That is, any asperity having a height greater than $s+\delta^*$ should experience some plastic deformation. If an asperity is deformed by less than δ^*, it will experience only elastic deformation and will return to its original shape when the contact force is released.

From equation (39), one can write

$$n_{pc} = n_b \int_{z=s+H^2r/(E')^2}^{\infty} \phi(z)\ dz$$

The total contact area under plastic deformation is

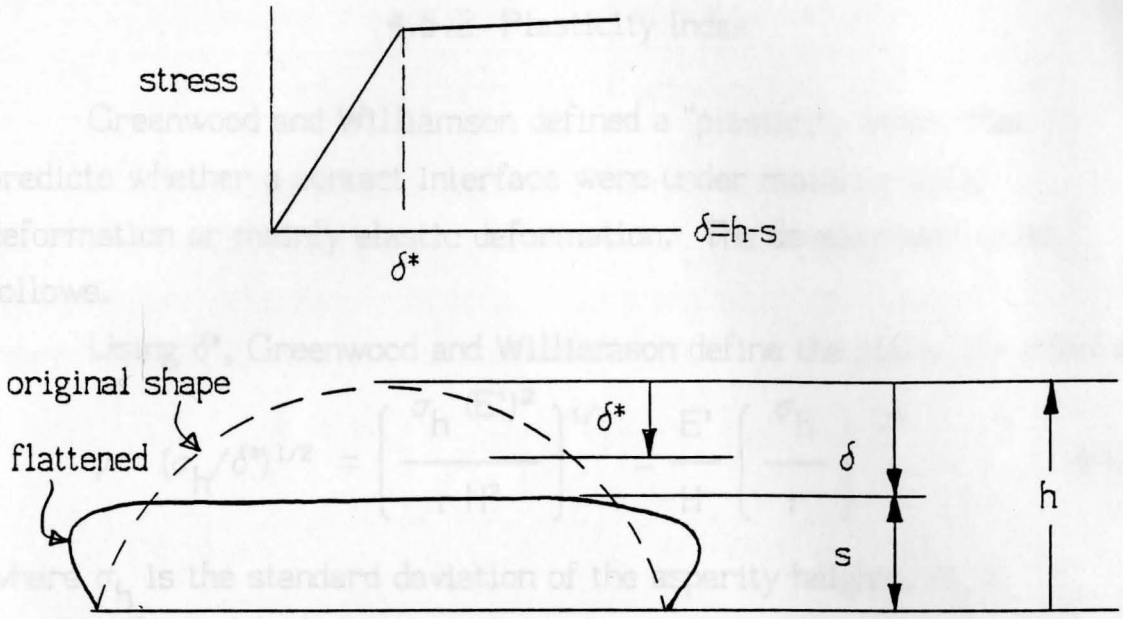


Fig. 50.--Plastic deformation of an asperity.

$$A_{tp} = n_b \pi r \int_{z=s+H^2r/(E')^2}^{\infty} (z-s)\, \phi(z)\, dz \qquad (43)$$

assuming, as Greenwood and Williamson do, that (5) still holds under plastic deformation.

Greenwood and Williamson arbitrarily pick $A_{tp}/A_{app} > 0.02$ as a criterion for "significant amount of plastic deformation." For a given surface, this inequality is solved for the separation s. From s and the equation for P_t, one can calculate the force required for the transition from elastic to plastic deformation. It turns out that the transition force is either extremely high (on the order of 50kg for a $1cm^2$ nominal contact area) or negligibly low (on the order of 1mg for a 1 cm^2 nominal contact area). This means that a given surface undergoes either elastic or plastic deformation at all practical loads with no transition.

4.5.2 Plasticity Index

Greenwood and Williamson defined a "plasticity index" that predicts whether a contact interface were under mainly plastic deformation or mainly elastic deformation. The development is as follows.

Using δ^*, Greenwood and Williamson define the plasticity index as

$$\psi = (\sigma_h/\delta^*)^{1/2} = \left(\frac{\sigma_h (E')^2}{r H^2} \right)^{1/2} = \frac{E'}{H} \left(\frac{\sigma_h}{r} \right)^{1/2} \qquad (44)$$

where σ_h is the standard deviation of the asperity heights, $\sigma_h = (\mathrm{var}(h))^{1/2}$. A detailed look at the terms in ψ is appropriate at this point.

a) E', Young's modulus of elasticity.

At strain s' (Fig. 51), metal "a" will deform plastically and metal "b" will deform elastically. A large value of E' (the slope of the stress vs. deformation curve) is more likely to result in plastic deformation. (A rubber band has a very small E', and is very elastic.)

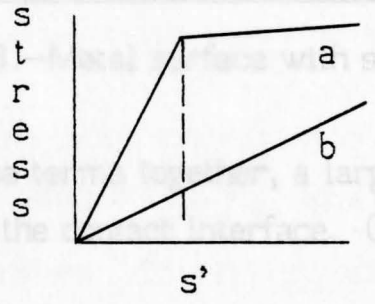


Fig. 51.--Stress vs. strain.

b) H, hardness.

Hardness is a measure of the size of the plastic indentation left by a diamond or ball shaped indenter. H = P/A, where P is the force of the indenter and A is the area of the mouth of the indentation. So more plastic deformation means larger A and smaller H.

c) $\sigma_h/r = \dfrac{\text{standard deviation of the heights of asperities}}{\text{radius of curvature of asperities}}$

If σ_h/r is large, the asperities tend to be tall and thin rather than short and stumpy. Large σ_h/r is more likely to result in plastic deformation. When σ_h/r is large, there are many tall asperities and many more occurrences of $\delta > \delta^*$ (Fig. 52).

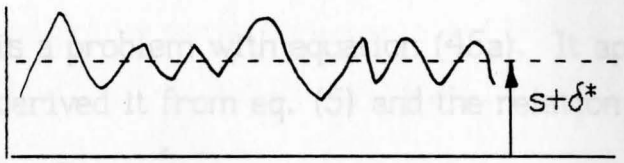


Fig. 52.--Metal surface with large σ_h/r.

When σ_h/r is small, the asperities are short and the occurrence of $\delta > \delta^*$ is rare (Fig. 53).

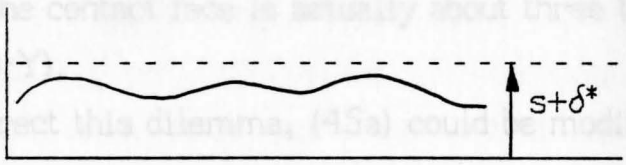


Fig. 53.--Metal surface with small σ_h/r.

Viewing all these terms together, a large value of ψ indicates plastic deformation at the contact interface. Greenwood and Williamson found that:

if $\psi > 1.0$, then the transition load is small ($<2mg/cm^2$ nominal pressure) and plastic deformation will occur at all practical loads (that is, all loads above 2mg for a $1cm^2$ apparent contact area);

if $\psi < 0.6$, then the transition load is large ($>50kg/cm^2$ nominal pressure) and elastic deformation will occur at all practical loads (that is, all loads below 50kg for a $1cm^2$ apparent contact area). [89]

4.5.3 Load Supported by a Plastically Deformed Asperity

According the Chhabra and Wenning, an asperity under plastic deformation supports a load given by [90]

$$P_{ipl} = \pi\, r_i\, Y\, \delta_i \qquad (45a)$$

where r is the radius of curvature of the asperity

Y is the elastic limit (or yield strength) of the metal

δ is the deformation.

There is a problem with equation (45a). It appears that Chhabra and Wenning derived it from eq. (5) and the relation

$$P_{ipl} = (\text{average pressure}) \cdot (\text{contact area})$$
$$= Y \cdot \pi a_i^2 = Y \cdot \pi \cdot (r_i\, \delta_i).$$

implying that the average pressure on the face of the plastically deformed asperity is equal to the yield stress. However, it was shown in equation (34) and section 4.4.5 (Figure 48) that the average pressure on the contact face is actually about three times the yield stress ($\sim 2.8 \cdot Y$).

To correct this dilemma, (45a) could be modified using eq. (34):

$$P_{ipl} = H \cdot \pi a_i^2 = (2.8 \cdot Y) \cdot \pi a_i^2 = (2.8 \cdot Y) \cdot \pi \cdot (r_i\, \delta_i) \qquad (45b)$$

The problem with making this modification to eq. (45a) is that the "constant" 2.8 actually changes with P_{ipl} in a nonlinear way between 1.1 (see eq. 37) and 2.8 (see eq. 34) until <u>full</u> plastic yielding takes place. Ignoring the nonlinearity may or may not be important in light of the other simplifications in CR1P (one must also consider the validity of eq. (5) within eq. (45a and b)). The author tried both (45a) and (45b) in the CR1P model.

Therefore, in the program CR1P, the load supported by a single asperity is

$$P_i = \begin{cases} (4/3)\ \delta_i^{3/2}\ r_i^{1/2}\ E/(1-\mu^2) & \text{if } \delta \leq \delta^* \quad (6) \\ \pi\ r_i\ (K \cdot Y)\ \delta_i\ ,\ K = 1 \text{ or } 2.8 & \text{if } \delta > \delta^* \quad (45) \end{cases}$$

4.5.4 Contact Area vs. Load for Plastic and Elastic Surfaces

This section is important because contained in it are some useful benchmarks for checking the adequacy of the elastic model CR1. This section is based on the same Greenwood and Williamson contact model described in the previous section, and uses results from their model development.

It has been shown that the asperity heights follow a gaussian probability distribution. [91] Greenwood and Williamson claim that the right tail of the gaussian distribution, which represents the tallest asperities—those which are likely to make contact—may be approximated by an exponential distribution. The exponential distribution in z must first be scaled properly before it can be substituted into the formulas (41) and (42) for total contact area and total contact load. It is most convenient to use the normalized variables $z'=z/\sigma_h$ and $s'=s/\sigma_h$ and the distribution of z':

$$\phi(z') = e^{-z'} = e^{-z/\sigma_h}$$

Letting $z = z' \cdot \sigma_h$ and $s = s' \cdot \sigma_h$ in eq. (41) and (42):

$$P_t = n_b \cdot (4/3) \cdot E' \cdot r^{1/2} \cdot \int_{z'=s'}^{\infty} \sigma_h^{3/2}\ (z'-s')^{3/2}\ e^{-z'}\ dz' \quad (42)$$

from Appendix F, $P_t = n_b \cdot (4/3) \cdot E' \cdot r^{1/2} \cdot \sigma^{3/2}\ (3/2)! \cdot e^{-s'}$ (46)

and $$A_t = n_b \cdot \pi \cdot r \cdot \int_{z'=s'}^{\infty} \sigma_h\ (z'-s') \cdot e^{-z'}\ dz' \quad (41)$$

from Appendix F, $A_t = n_b \cdot \pi \cdot r \cdot \sigma_h \cdot e^{-s'}$ (47)

dividing, $$\frac{A_t}{P_t} = \frac{n_b \cdot \pi \cdot r \cdot \sigma_h \cdot e^{-s'}}{(4/3) \cdot (3/2)! \cdot n_b \cdot E' \cdot r^{1/2} \cdot \sigma_h^{3/2} \cdot e^{-s'}},$$

$$= \frac{(3/4) \cdot \pi \cdot r^{1/2}}{(3/2)! \cdot \sigma_h^{1/2} \cdot E'} = \text{a constant}$$

therefore the relation obtained is

$$A_t \propto P_t \qquad (48)$$

— a proportional relationship that has been observed in real contacts.

This proportionality was derived using a purely elastic model, however, if one assumes full plastic deformation, one is led to the famous relation stated in Tabor [92]

$$\frac{A_t}{P_t} = \frac{1}{H} \qquad (49a)$$

which is the definition of hardness, H.

In fact, if one looks at equation (45a), one can also derive the proportionality between P_t and A_t for full plastic deformation:

$$P_{ipl} = \pi r_i Y \delta_i \qquad (45a)$$

$$P_t = \sum_i P_{ipl} = Y \sum_i \pi r_i \delta_i = Y \sum_i \pi [(r_i \delta_i)^{1/2}]^2$$

$$= Y \sum_i \pi (a_i)^2 \qquad \text{from eq. (5)}$$

$$= Y \cdot A_t \quad \text{(QED)}. \qquad (49b)$$

Thus, $A \propto P$ assuming either elastic or plastic deformation.

There is an apparent contradiction between eq. (49a) and (49b). Eq. (49a) says that $P_t = H \cdot A_t$ whereas (49b) says that $P_t = Y \cdot A_t$. Solving, one obtains $H = Y$, which contradicts eq. (34): $H \simeq 2.8Y$. The

problem, as was mentioned before, is that eq. (45b) may be more correct than eq. (45a), since the average pressure on the contact face of a plastically deformed asperity is ~ 2.8Y. (Section 4.4.5, Fig. 48). Equation (49a) can now be used to support this claim.

Early attempts at using an elastic model failed to give the proportionality between A_t and P_t. That is because if one assumes that contact is made at one large single asperity, one is led to the result

$$P_t = (4/3)\ \delta^{3/2}\ r^{1/2}\ E'$$

but $A_t = \pi \cdot r \cdot \delta$ gives us $\delta = A_t/\pi r$, which when substituted yields

$$P_t = (4/3)\ (A_t/\pi r)^{3/2}\ r^{1/2}\ E'$$

$$= \text{constant} \cdot A^{3/2}$$

for the relation $A \propto P^{2/3}$.

The proportionality between A and P arises because new contact spots form as P is increased: a phenomenon not accounted for in the early models, but accounted for in the programs CR1-CR3 and CR1P and in Greenwood and Williamson's model.

4.5.5 Contact Resistance

The resistance of a single contact spot is given by Holm's formula

$$R_{spot} = \rho/2a$$

where a is the radius of the spot. For "n" spots having mean radius μ_a, the total parallel resistance is

$$R = \rho/2n\mu_a.$$

The mean value of the contact radius is derived from equation (5). Letting $\delta = z\text{-}s$,

$$\mu_a = r^{1/2} \int_s^\infty (z - s)^{1/2}\ \phi(z)\ dz \qquad (50)$$

Greenwood and Williamson say that the total contact resistance is [93]

$$R_t = \frac{\rho}{2\, n_b r^{1/2} \int_s^\infty (z - s)^{1/2}\, \phi(z)\, dz} = \frac{\rho}{2\, n_b\, \mu_a} \qquad (51a)$$

However, μ_a is the average contact radius of n_b asperities, not n_c contact spots (compare eq. 40 to 41). Resistance should be a function of n_c, not n_b. To circumvent this difficulty, the author suggests redefining μ_a using equations (40) and (41):

$$\mu_A' = \frac{n_b \pi r \int_s^\infty (z - s)\, \phi(z)\, dz}{n_b \int_s^\infty \phi(z)\, dz} = \frac{\text{total contact area}}{\text{number of contact spots}}$$

$$= A_t/n_c \text{ rather than Greenwood's apparent } A_t/n_b.$$

Then

$$\mu_a' = \left[\frac{\mu_A'}{\pi}\right]^{1/2} = \left[\frac{r \int_s^\infty (z - s)\, \phi(z)\, dz}{\int_s^\infty \phi(z)\, dz}\right]^{1/2}$$

and

$$R_2 = \frac{\rho}{2n\mu_a'} = \frac{\rho}{2\left[n_b \int_s^\infty \phi(z)\, dz\right]} \cdot \frac{\left[\int_s^\infty \phi(z)\, dz\right]^{1/2}}{\left[r \int_s^\infty (z - s)\, \phi(z)\, dz\right]^{1/2}}$$

$$= \frac{\rho}{2\, r^{1/2}\, n_b \left(\int_s^\infty \phi(z)\, dz \right)^{1/2} \left(\int_s^\infty (z - s)\, \phi(z)\, dz \right)^{1/2}} \qquad (51b)$$

R_1 and R_2 can be compared by assuming an exponential distribution for the height z. As before, one uses the normalized exponential $\phi(z') = e^{-z'}$, where $z'=z/\sigma_h$ and $s'=s/\sigma_h$.

Eq. (51a) becomes

$$R_1 = \frac{\rho}{2\, n_b\, r^{1/2} \int_{s'}^\infty \sigma_h^{1/2}\, (z' - s')^{1/2}\, e^{-z'} dz'}$$

$$= \frac{\rho}{2\, n_b\, r^{1/2}\, \sigma_h^{1/2}\, \Gamma(1/2 + 1)\, e^{-s'}} \qquad \text{(see Appendix F)}$$

$$= \frac{\rho}{2\, n_b\, (r \cdot \sigma_h)^{1/2}\, e^{-s'}\, (0.89)}$$

Eq. (51b) becomes:

$$R_2 = \frac{\rho}{2\, r^{1/2}\, n_b \left(\int_{s'}^\infty e^{-z'}\, dz' \right)^{1/2} \left(\int_{s'}^\infty \sigma_h\, (z' - s')\, e^{-z'}\, dz' \right)^{1/2}}$$

$$= \frac{\rho}{2\, r^{1/2}\, n_b \left(- e^{-z'} \Big|_{s'}^\infty \right)^{1/2} \left(\sigma_h\, e^{-s'}\, \Gamma(1+1) \right)^{1/2}}$$

$$= \frac{\rho}{2\, r^{1/2}\, n_b\, (e^{-s'})^{1/2}\, (\sigma_h\, e^{-s'} \cdot 1)^{1/2}} = \frac{\rho}{2\, n_b\, (r \cdot \sigma_h)^{1/2}\, e^{-s'}\, (1)^{1/2}}$$

Comparing, one gets $R_1 \simeq 1.13 \cdot R_2$, a difference which, considering all the assumptions and approximations, is practically negligible.

If thin insulating films are present on the metal surface, then only those asperities that yield plastically will be able to conduct current, because only at those asperities will the film be cracked. To account for this, one replaces s with $s+\delta^*$ in the lower integral limits—assuming, once again, that eq. (5) still holds for an asperity under plastic deformation.

4.5.6 Summary of Greenwood and Williamson's Model

Greenwood and Williamson's elastic deformation model is valid only if the asperities are far enough apart to be considered independent. If they are clustered closely together, mechanical and electrical interactions must be taken into account. This restriction was stated by Greenwood and Williamson themselves.

As a summary of their model, let us review the procedure for calculating contact resistance:

(1) Given the contact force F, let $P_t = F$ and use equation (42) to calculate the contact separation, s ($= s' \cdot \sigma_h$). It is assumed that all other data required for (42) are known.

(2) Having calculated s, use equation (51a) or (51b) to calculate the constriction resistance. This is an "expected value" of the constriction resistance.

(3) Having calculated s, one may, if desired, use equation (41) to calculate the contact area.

4.6 Results from CR1, CR1P, and Greenwood and Williamson

(For notational purposes, if the contact pressure of an asperity

under plastic deformation is assumed to be "Y," then the plastic deformation model will be called "CR1P$^{(1)}$"; if the contact pressure is assumed to be 2.8·Y, then the model will be called "CR1P$^{(2.8)}$". If the reference pertains to both models, the notation will be "CR1P" without a superscript.)

The results from CR1 and CR1P are shown in Figures 54 and 55. The surface model is the same as that in Table 2 with the exception that n_b = 400 asperities is constant and not a Poisson random number. And in CR1P, values for yield stress (Y = 260 MPa) and hardness (H = 74 kg/mm^2 = 726 MPa) were used [94] (fortunately, the data for H and Y support eq. (34): H = 2.79·Y).

There were 10 "measurements" made at each force. For the elastic deformation model CR1, P_i was calculated by eq. (6) for all asperities. For CR1P$^{(1)}$, P_i was calculated by eq. (45a) if $\delta_i \geq \delta^*$; and by eq. (6) if $\delta_i < \delta^*$. For CR1P$^{(2.8)}$, P_i was calculated by eq. (45b) if $\delta_i \geq \delta^*$ and eq. (6) if $\delta_i < \delta^*$. An interesting note: when CR1P was run using the values of Table 2, all contact spots were in plastic deformation at all contact forces.

In Fig. 54(a), note how the linear relationship between actual contact area and contact force holds for both elastic and plastic deformation models. These plots demonstrate equations (48), (49a), and (49b). Also from these plots, it is interesting to calculate the ratio of actual to apparent contact area; and compare it to the comment in section 1.6.1 (in the Introduction) that claimed the ratio was "perhaps 1/1000." At 5N, the ratios are [2×10^{-8} m^2/4×10^{-6} m^2] $\simeq$ 1/200 for the plastic deformation model CR1P$^{(1)}$ and [$0.8\times10^{-8}/4\times10^{-6}$] $\simeq$ 1/500 for both the plastic deformation model CR1P$^{(2.8)}$ and the elastic model CR1. These numbers are definitely in the ballpark.

Fig. 55(b) is a plot of contact resistance vs. contact force on a

log-log graph. Note the linear shape of the curves. By linear regression, the slopes of the lines were found to be -0.678 for the elastic model CR1; -0.728 for the plastic model CR1P$^{(1)}$; and 0.746 for the plastic model CR1P$^{(2.8)}$. Thus, one has the relations $R \propto F^{-0.678}$, $R \propto F^{-0.728}$, and $R \propto F^{-0.746}$, respectively, for the three models.

Upon inspection of Figures 54 (a) and (b) and 55(a), it appears CR1P$^{(2.8)}$ is closer in behavior to the elastic deformation model CR1 than to CR1P$^{(1)}$. However, the slopes in Fig. 55(b) show that the two plastic deformation models behave similarly (they have slopes that are nearly equal) and that the elastic deformation model behaves differently from these.

To compare the elastic model CR1 to Greenwood and Williamson's model, data from Greenwood and Williamson were used. Their data were not complete enough to use in CR1 (much of the data were given in dimensionless form, so some of their values had to be guessed at and the rest of the values were calculated to agree with their formulas.
The data finally used in CR1 are shown in Table 7.

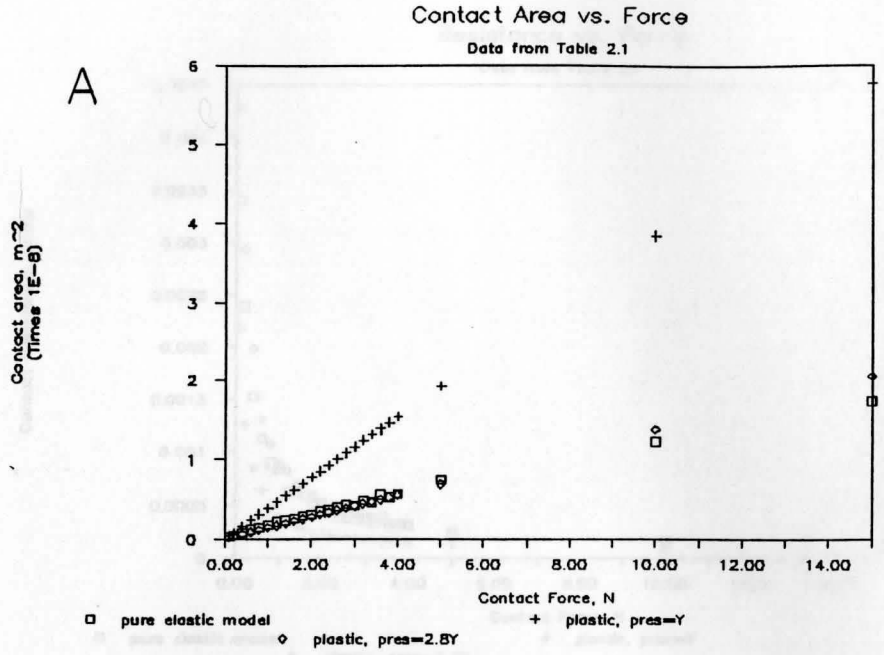
Greenwood and Williamson found a linear relationship between log(R) and log(F) for their elastic deformation model. Using equations (42) and (51a) developed in the preceding sections of this chapter, they found $R \propto F^{-0.9}$ for an apparent contact area of 1 cm^2 and a range of 0.1 to 1000 N contact force. [95] By comparison, data generated by CR1 (also using the values in Table 7) gave the relation $R \propto F^{-0.51}$ over the range 0.1 to 3.16 N and an apparent contact area of 1 mm^2. A comparison is shown in Fig. 56. CR1 could not calculate resistance when the force exceeded about 10N because of memory limitations (Note the large value of d in Table 7).

TABLE 7

VALUES USED IN GREENWOOD AND WILLIAMSON'S MODEL [95]

d[a]	bump density	300/mm^2
avg(h)	avg. asperity height	0.15 μm
r	avg radius of curvature	2.041 mm
σ_h	std dev of asperity heights	4.9×10^{-8} m
E	Young's modulus	9.1×10^{10} Pa
ρ	electrical resistivity	2.4×10^{-8} $\Omega\cdot$m

[a] n_b was not a random number in this experiment



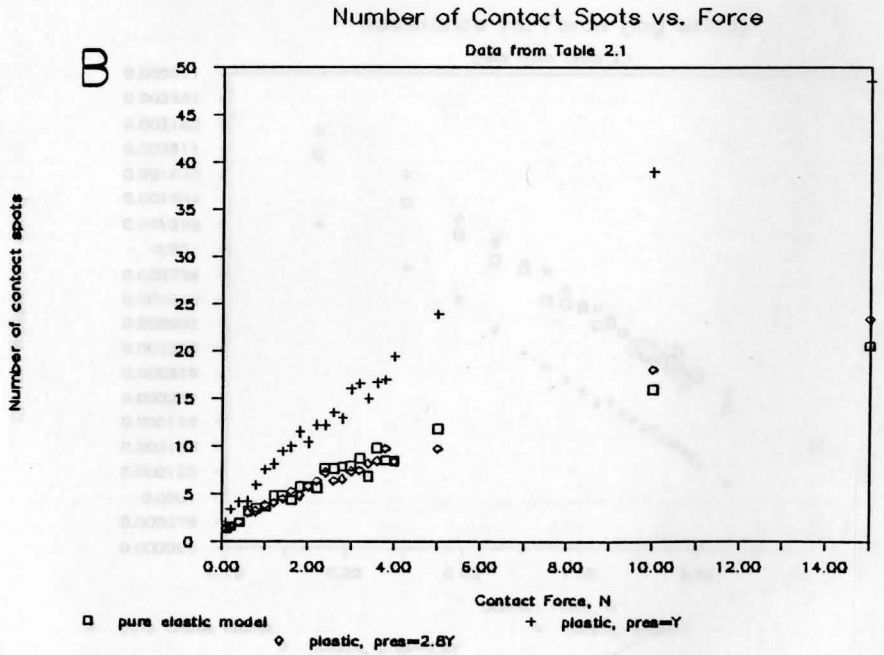
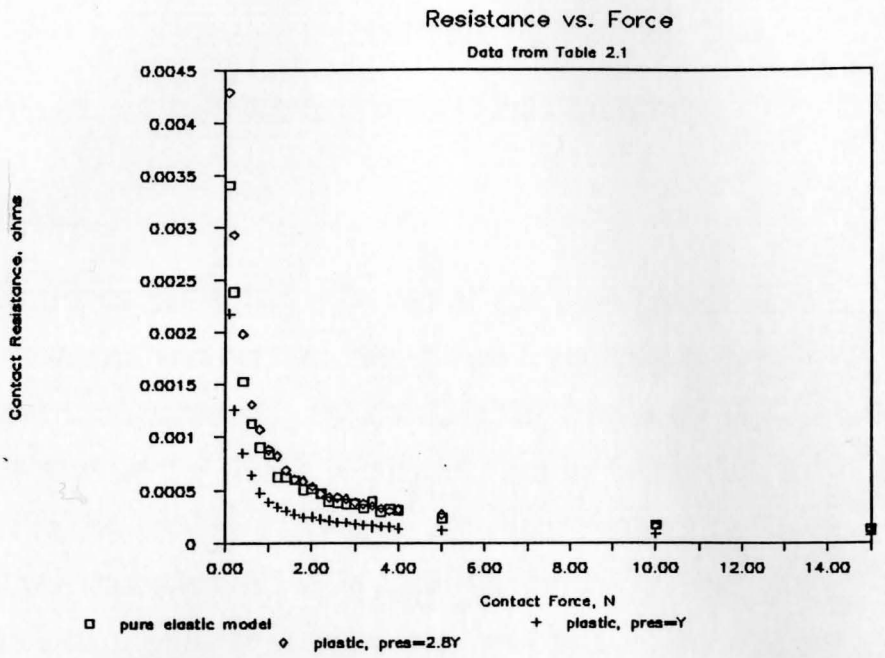


Fig. 54.--Contact area and no. of contact points vs. contact force.



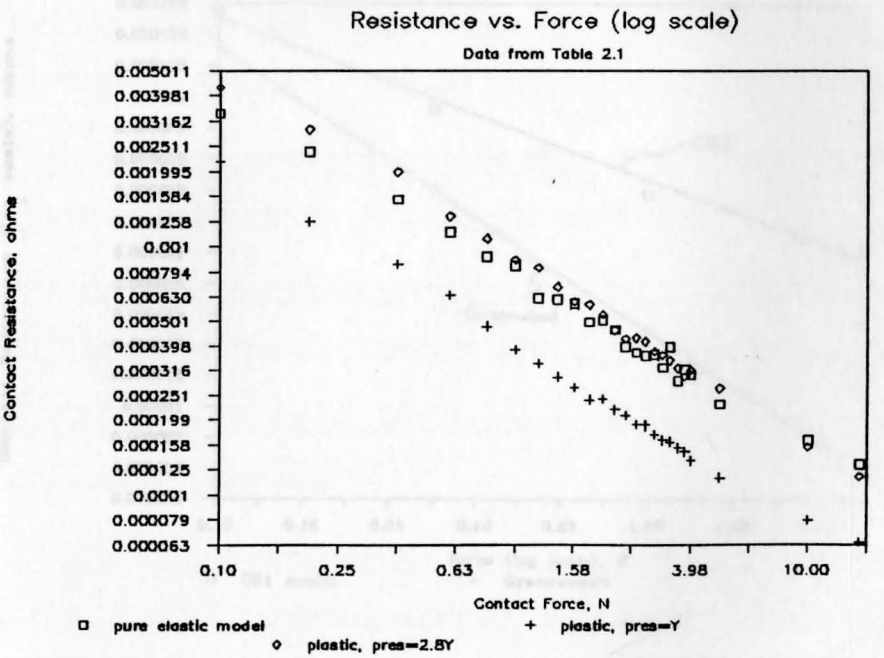


Fig. 55.--Contact resistance vs. contact force.

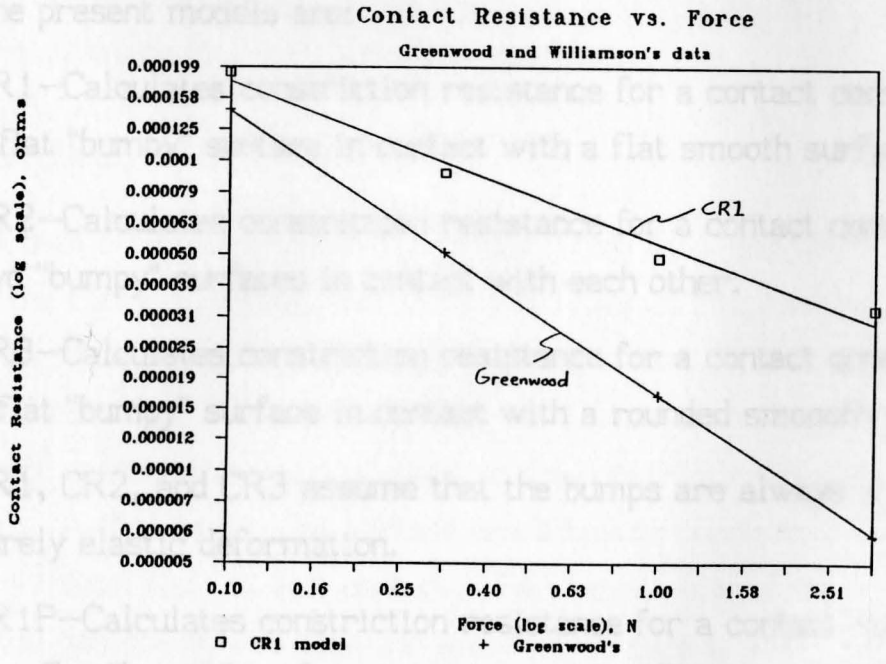


Fig. 56.--Contact resistance vs. force, CR1 model vs. Greenwood and Williamson's model. (log-log scale)

CHAPTER V.

SUMMARY AND CONCLUSIONS

5.1 Summary

This thesis describes a series of computer models that predict contact resistance and contact resistance frequency distributions of "clean" electrical contacts. By knowing the frequency distribution at different stages in the life of a contact, and by being able to predict the fraction of failed contacts based on a contact resistance failure criterion, one can predict the failure rate of the contact design at any time in its life. Thus, this work will lead to a model that can predict contact failure rate and contact reliability.

The present models are:

CR1--Calculates constriction resistance for a contact consisting of a flat "bumpy" surface in contact with a flat smooth surface.

CR2--Calculates constriction resistance for a contact consisting of two "bumpy" surfaces in contact with each other.

CR3--Calculates constriction resistance for a contact consisting of a flat "bumpy" surface in contact with a rounded smoooth surface.

CR1, CR2, and CR3 assume that the bumps are always under purely elastic deformation.

CR1P--Calculates constriction resistance for a contact consisting of a flat "bumpy" surface in contact with a flat smooth surface; if the deformation of a "bump" is great enough, then the bump undergoes plastic deformation, and the equations for plastic

deformation apply.

CR1P was written to predict the constriction resistance of thin-film-covered contacts. The assumption was made that the film is broken only at those asperities that are plastically deformed, and electrical contact is made only where the film is broken.

Two methods for predicting the fraction failed based on the contact resistance distribution were also described. (1) The threshold resistance equation assumes that if a contact resistance exceeds a certain value R', then the contact has failed. From a contact resistance distribution, the fraction failed at time t can be predicted knowing R'. (2) The method of uniform residuals assumes that the "failed" contacts at time t are those that are part of the high-resistance "tail" of a normal probability plot. In other words, failed contacts are those that do not follow a gaussian distribution.

To predict the failure rate, either of these two methods can be used on time-varying contact resistance distributions using a discrete-time approximation to the failure rate equation.

5.2 Conclusions

(Unless otherwise noted, the results from CR1 and CR1P are based on the contact surface data shown in Table 2; the hardness and yield strength data used in CR1P are those of brass, 70% Cu, 30% Zn.)

(1) From Fig. 25, it is seen that for "clean" contacts, the flatter the apparent mating interface, the lower the constriction resistance, all else equal. A dimple vs. flat contact has a higher constriction resistance than a flat vs. flat contact. This result makes sense because the dimpled design—by reducing the apparent contact area—induces a further constriction of current that is added on to the constriction produced by the individual contact spots. In practice, however, the

dimple vs. flat design is used to concentrate the contact normal force to break down the insulating films that appear on most real contact surfaces.

(2) Also from Fig. 25, it is seen that log(R) vs. log(F) is a straight line with a negative slope. Thus the relationship is $R \propto F^{-\alpha}$ where "R" is contact resistance (specifically, constriction resistance), "F" is contact normal force, and $-\alpha$ is the slope.

(3) From Figs. 28 and 30, it is seen that the models CR1 and CR3 can predict the contact resistance of actual gold contacts reasonably well. The percent errors are sometimes large, but the absolute errors are small. Gold contacts in real life are the closest things to "clean contacts" because gold does not corrode and it resists organic film growth. Contact surface data used in CR1 are shown in Tables 2 and 3.

(4) It is seen from Table 5b and Figs. 31-38 that contact normal force has the greatest influence on contact resistance. The next most influential parameters are, in order: electrical resistivity; the ratio of average radius of curvature to the average height of asperities; and the shape factor length/width of the apparent contact area. The other parameters listed in Table 5b have little or no influence on the contact resistance. This conclusion is valid for the ranges of parameters given in Table 6.

(5) From Fig. 54a, it is seen that, whether the elastic model CR1 or the plastic model CR1P is used, actual contact area vs. contact normal force is a linear relationship. The relationship is easily proven for the plastic deformation model by using eq. (45a) and eq. (5); and was proven for the elastic deformation model by Greenwood and Williamson in eq. (46)-(48).

(6) CR1P was written to predict the contact resistance of thin-film-covered contacts. The author expected that the number of

contacting asperities under plastic deformation would be less than the total number of asperities in contact. That is, CR1P was expected to predict a higher contact resistance for a film-covered contact than for a clean contact under the assumption that, for a film-covered contact, only those asperities under plastic deformation can conduct electrical current; while for a clean contact, all contacting asperities can conduct electrical current. This makes sense, because a film-covered contact should have a higher contact resistance than a clean contact.

However, in Section 4.6, it was pointed out that, under all contact forces studied (0.1-15N), CR1P predicted that all asperities would be under plastic deformation. Thus, it appears impossible for CR1P, in its present state, to predict a difference in contact resistance between clean and film-covered contacts. A new model may be necessary.

(7) Like CR1, CR1P predicts a negative linear correlation between log(R) and log(F). This is seen in Fig. 55. However, linear regression analysis showed that the slopes predicted by CR1 and CR1P are different. When the average contact pressure on a plastically deformed asperity was assumed to be equal to Y (yield strength), the slope predicted by CR1P was -0.728; when the pressure was assumed to be equal to 2.8Y, the slope was -0.746. These slopes are closer to each other than the slope predicted by CR1 for pure elastic deformation: -0.678.

(8) Fig. 56 shows how the elastic deformation model CR1 compares with Greenwood and Williamson's elastic deformation model. The data were taken from Greenwood and Williamson's paper. The discrepancy in the results is apparently due partially to the fact that some of the data required by CR1 were not supplied by Greenwood and Williamson, thus requiring guessing at one or two values of the parameters and forcing the rest of the parameters to follow their equations. The contact surface data used in CR1 are shown in Table 7.

5.3 Recommendations

The next step in this research may be to predict the change in the contact resistance distribution as a function of time (or age of the contacts). From the data obtained, one could then predict the failure rate, d(failures)/dt, using contact resistance as the failure criterion.

However, to predict the true failure of an electrical contact within a given circuit, one must know how the contact will affect the circuit. Knowledge of only the contact's resistance is not sufficient: electronic circuits can tolerate several hundreds of ohms of contact resistance. Thus, a more complete description of the contact's electrical characteristics are required. Two possible approaches are: modeling the voltage vs. current characteristic of a degraded contact; and predicting the probability, duration, and magnitude of intermittent open circuits at the contact. The models could then be used in a circuit analysis package such a SPICE to predict the effect of the contact on the circuit. Once this step is done, it may be possible to correlate the increase in contact resistance to probability of actual failure. Given this data and the time-varying contact resistance distributions, a true failure rate can be calculated.

Predicting these electrical characteristics will require some means of modeling a degraded contact surface. For example, predicting the voltage vs. current characteristic requires modeling of both the conduction and electrical breakdown properties of the contaminating films at the contact interface, and modeling of the distribution of the films over the metallic surfaces. Even the prediction of contact resistance distribution vs. time requires some means of modeling the insulating effects of those same films. Such modeling requires knowledge of chemistry (e.g. oxidation of metals, polymer film growth, diffusion of metal atoms, formation of metal phases, etc.), friction and

wear (and the consequent chemical reactions taking place), mechanical properties of thin films (such as oxides, tarnish, and polymers), and an intimate knowledge of electromagnetics (how is the electrical field disturbed by the films?).

The focus of the more advanced work should be on contacts designed for low voltage/low current/high data rate digital electronics.

The possibility of electromagnetic emissions from the contact should also be looked at, as dirty metallic junctions have been shown to act as mixers and antennas, [96] and because crosstalk between nearby circuits can be detrimental.

APPENDIX A

Failure Rate and System Reliability

The reliability of a series electrical circuit, such as a flashlight circuit (Fig. 1), is the product of the reliabilities of each of the components.

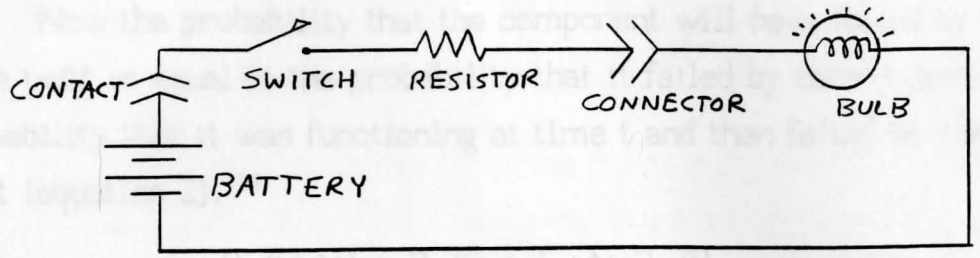


Fig. 1--Flashlight Circuit

Due to degradation of the components, the component reliabilities are generally functions of time; denote the reliability function of component i as $R_i(t)$. Then the circuit reliability (n components) is

$$R_{system}(t) = \prod_{i=1}^{n} R_i(t) \qquad (1)$$

The reliability function of a component is uniquely determined by its failure rate function $h_i(t)$. The failure rate of a component is the number of failures of that component per unit time. It turns out that the probability of a component failing during the interval t and $t+\Delta t$ is $h_i(t)\cdot\Delta t$.

In many practical engineering analyses, the failure rate function for a component is a constant, $h_i(t) = \lambda_i$. Then the probability that the component will fail between t and $t+\Delta t$ is $\lambda_i\cdot\Delta t$; this probability depends only on Δt and not on t.

The question to be asked now is: given a constant failure rate, what is the reliability function of component i? (The reliability function is not constant when the failure rate is constant.)

Define the probability that a component will have failed by

time t as $P_f(t)$. Also, assuming that the component has a constant failure rate λ_i, then the probability that the component will fail between t and t+Δt, assuming it was still functioning at time t, is $\lambda_i \cdot \Delta t$.

Now the probability that the component will have failed by the time t+Δt is equal to the probability that it failed by time t plus the probability that it was functioning at time t and then failed between t and t+Δt (equation 2). [97]

$$P_f(t+\Delta t) = P_f(t) + \lambda_i \cdot \Delta t \cdot R_i(t) \qquad (2)$$

where $R_i(t)$ is the probability that the component was functioning at t.

Equation (2) can be rearranged:

$$\frac{P_f(t+\Delta t) - P_f(t)}{\Delta t} = \lambda_i \cdot R_i(t)$$

In the limit as $\Delta t \rightarrow 0$, one gets the differential equation

$$dP_f(t)/dt = \lambda_i \cdot R_i(t) \qquad (3)$$

The probability that the component is still working at t+Δt is equal to the probability that it was working a t <u>and</u> does <u>not</u> fail between t and t+Δt (equation 4).

$$R_i(t+\Delta t) = R_i(t) \cdot (1 - \lambda_i \cdot \Delta t) \qquad (4)$$

Rearranged, this is

$$\frac{R_i(t+\Delta t) - R_i(t)}{\Delta t} = -\lambda_i R_i(t)$$

In the limit at $\Delta t \rightarrow 0$,

$$dR_i(t)/dt = -\lambda_i \cdot R_i(t) \qquad (5)$$

Equations (3) and (5) are a set of ordinary differential equations that can be solved using Laplace transform methods:

$$sP_f(s) - P_f(0) = \lambda_i R_i(s) \qquad (3')$$

$$sR_i(s) - R_i(0) = -\lambda_i R_i(s) \qquad (5')$$

Assuming that the component was functioning at t=0, one uses $P_f(0) = 0$ and $R_i(0) = 1$. Then one gets, in matrix form:

$$\begin{bmatrix} s & -\lambda_i \\ 0 & s+\lambda_i \end{bmatrix} \begin{bmatrix} P_f(s) \\ R_i(s) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The reliability $R_i(s)$ is solved for by Cramer's rule:

$$R_i(s) = \frac{\begin{vmatrix} s & 0 \\ 0 & 1 \end{vmatrix}}{\begin{vmatrix} s & -\lambda_i \\ 0 & s+\lambda_i \end{vmatrix}} = \frac{s}{s\,(s+\lambda_i)} = \frac{1}{s+\lambda_i}$$

Then $$R_i(t) = e^{-\lambda_i t} \qquad (6)$$

Thus if all components of the series circuit have constant failure rates, it follows from equations (1) and (6) that

$$R_{system}(t) = \prod_{i=1}^{n} e^{-\lambda_i(t)} = \exp\left[- t \sum_{i=1}^{n} \lambda_i \right]$$

The ultimate goal of this research is to find h(t). h(t) may or may not be a constant λ; the author guesses that h(t) will <u>not</u> be constant.

APPENDIX B

Generation of Random Numbers having a Specific Distribution

(1) Gaussian Random Numbers

Let $X(i)$ be a uniformly distributed random number, $0 \leq x(i) \leq 1$, such as that produced by a standard computer random number generator. (Fig. 1). By the Central Limit Theorem, a random number Y given by

$$Y = \frac{1}{n} \sum_{i=1}^{n} [\, a \cdot X(i) + b], \text{ where a and b are constants,}$$

has a distribution which approaches a normal (gaussian) distribution as n gets large.

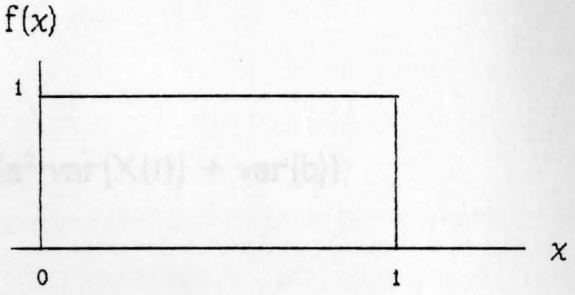


Fig. 1--Distribution of a Uniform Random Number

The mean or expected value of Y (denoted μ_Y or E(Y)) is

$$\mu_Y = E(Y) = E\left\{ \frac{1}{n} \sum_{i=1}^{n} [\, a \cdot X(i) + b] \right\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} [\, E\{a \cdot X(i)\} + E\{b\}]$$

since $E\{a \cdot X(i)\} = a \cdot E\{X(i)\} = a \cdot \mu_{X(i)}$ and $E(b) = b$, one can write

$$\mu_Y = \frac{1}{n} \sum_{i=1}^{n} (a \cdot \mu_{X(i)} + b) = \frac{1}{n} \cdot n \, (a \cdot \mu_{X(i)} + b)$$

$$= a \cdot \mu_{X(i)} + b \qquad (1)$$

The variance of Y is

$$\text{var}(Y) = \text{var}\left\{ \frac{1}{n} \sum_{i=1}^{n} [\, a \cdot X(i) + b \,] \right\}$$

One uses the property $\text{var}\{c \cdot X\} = c^2\, \text{var}(X)$ where c is a constant:

$$\text{var}(Y) = \frac{1}{n^2} \sum_{i=1}^{n} \text{var}\, \{\, a \cdot X(i) + b \,\}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} (a^2 \cdot \text{var}\{X(i)\} + \text{var}\{b\})$$

$$= \frac{1}{n^2} \cdot a^2 \cdot n\, \text{var}\{X(i)\}\,, \quad \text{var}\{b\} = 0 \text{ since b is a constant.}$$

$$= a^2 \cdot \text{var}\{X(i)\}/n \qquad (2)$$

Since $X(i)$ is uniformly distributed, its density function is

$$f(x(i)) = 1 \text{ for } 0 \leq x(i) \leq 1$$
$$= 0 \text{ elsewhere}$$

Then

$$\mu_{X(i)} = \int_{-\infty}^{\infty} x(i) \cdot f(x(i))\, dx(i) = \int_{0}^{1} x(i)\, dx(i) = \frac{x^2(i)}{2} \Bigg|_{0}^{1} = 1/2$$

$$\text{var}\{X(i)\} = \int_{-\infty}^{\infty} (x(i) - \mu_{X(i)})^2\, dx(i)$$

$$= \int_0^1 (x(i) - 0.5)^2 \, dx(i)$$

$$= \int_0^1 (x^2(i) - x(i) + 0.25) \, dx(i)$$

$$= x^3(i)/3 - x^2(i)/2 + 0.25x(i) \Big|_0^1$$

$$= 1/3 - 1/2 + 1/4 = 1/12$$

Substituting these values into equations (1) and (2):

$$E(Y) = \mu_Y = a(1/2) + b \qquad (1)$$

$$\text{var}(Y) = \sigma_Y^2 = a^2/12n \qquad (2)$$

We can generate a normally distributed random number Y using a uniform random number generator. We select a, b, and n to get suitable values of μ_Y and σ_Y^2.

(2) Poisson Random Numbers

Let X be a uniformly distributed random number with density function $f(x)$. To generate Y, a random number with an arbitrary distribution, we need to find a transformation $y = g(x)$ (Fig. 2).

The function $y = g(x)$ must be monotonically increasing so that

$$p(X \leq x) = p(Y \leq y) \qquad (3)$$

will always be satisfied.

Now if X is uniformly distributed, $0 \leq x \leq 1$, then

$$p(X \leq x) = F(x) = \int_{-\infty}^{x} dx = \int_0^x dx = x \qquad (4)$$

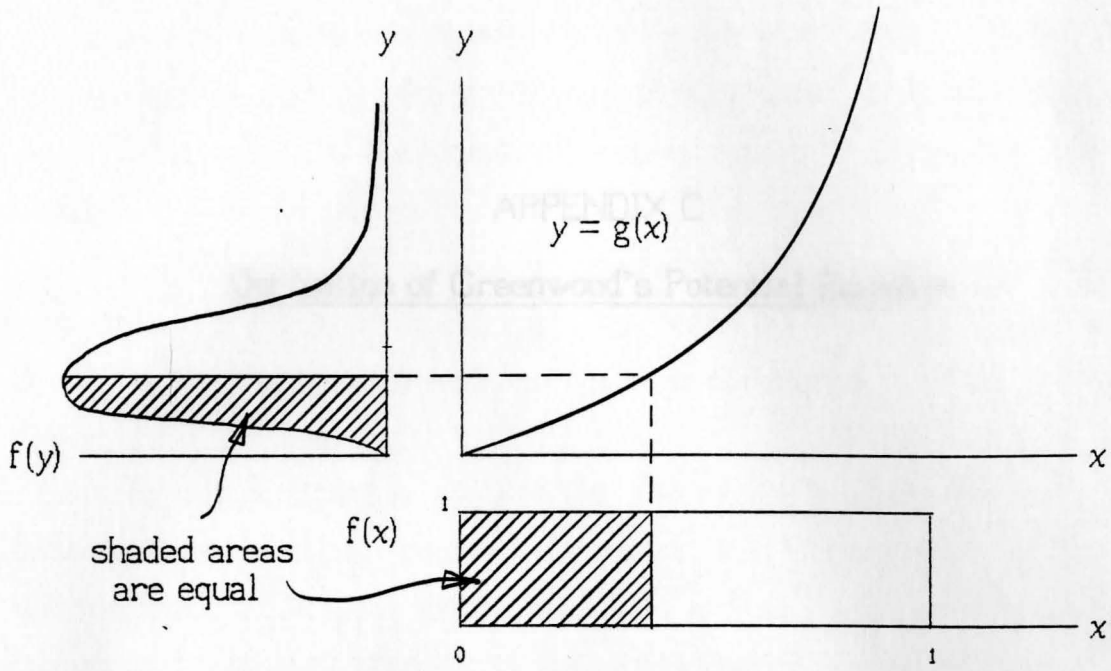


Fig. 2--A Transformation $y = g(x)$

Take Y as a Poisson-distributed random number.

$$P(Y \leq y) = \sum_{i=1}^{y} \frac{e^{-\mu} \mu^{i}}{i!} \tag{5}$$

Substituting (4) and (5) into (3), we get

$$x = \sum_{i=1}^{y} \frac{e^{-\mu} \mu^{i}}{i!} \tag{6}$$

which must be solved for y.

Since $y = g(x)$ is unique, equation (6) is the desired transformation. The easiest way to solve (6) is by iteration. Equation (6) cannot in general be satisfied exactly because Y, being a Poisson random number, is an integer. Thus the right side of (6) can only take on a discrete number of values once the parameter μ is chosen. The left side of (6), in contrast, can take on any value between zero and one.

APPENDIX C

Derivation of Greenwood's Potential Equation

In this appendix an equation will be derived for the potential at a contact spot caused by current flowing through another spot a distance d away. This equation was used by Greenwood and in Section 2.6.3 in this thesis.

Referring to Fig. 1, it is given that contact spot j conducts a current I_j. The goal is to calculate the potential at point i, which may or may not be another contact spot (it does not matter). Point i is a distance d_{ij} from spot j.

For simplification, assume that the region around spot j is a spherical equipotential volume. Current I_j will flow radially into the volume from the bottom contact member, and out of the volume into the top contact member. Hemispherical equipotential surfaces will appear in each of the contact member. Current lines are perpendicular to the equipotential surfaces everywhere.

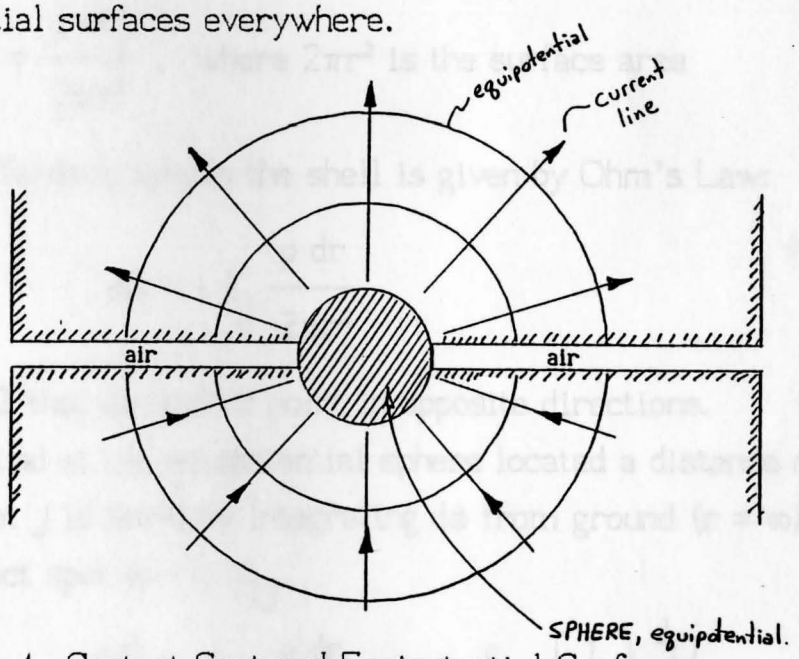


Fig. 1--Contact Spot and Equipotential Surfaces.

Consider the top contact member only (Fig. 2). Current passes through hemispherical "shells," normal to the surface of the shells. A shell has the incremental resistance

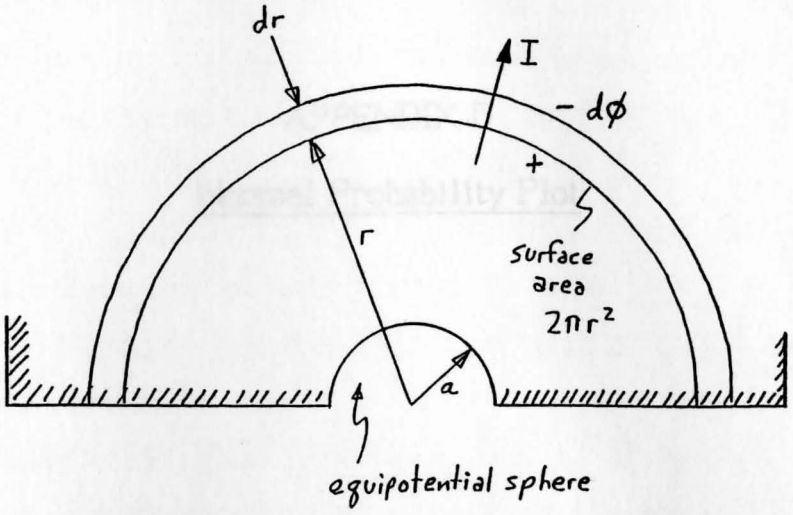


Fig. 2--Shell having incremental resistance and voltage drop.

$$dR = \frac{\rho\, dr}{2\pi r^2} \quad , \quad \text{where } 2\pi r^2 \text{ is the surface area}$$

The potential difference across the shell is given by Ohm's Law:

$$d\phi = - I_j \frac{\rho\, dr}{2\pi r^2}$$

Note from Fig. 2 that $d\phi$ and dr point in opposite directions.

The potential at the equipotential sphere located a distance d_{ij} from contact spot j is found by integrating $d\phi$ from ground ($r = \infty$), toward the contact spot to $r = d_{ij}$.

$$\phi_{ij} = \int_{r=\infty}^{d_{ij}} - I_j \frac{\rho\, dr}{2\pi r^2} = I_j \frac{\rho}{2\pi} \left[\frac{1}{r} \right]_{\infty}^{d_{ij}}$$

$$= \frac{\rho\, I_j}{2\pi} \left[\frac{1}{d_{ij}} - 0 \right] = \frac{\rho\, I_j}{2\pi d_{ij}} \qquad \text{Q.E.D.}$$

APPENDIX D

Normal Probability Plot

(This appendix will use the standard convention of naming a random variable with an uppercase letter--e.g., X--and denoting specific values of that random variable with the respective lowercase italicized letter--i.e., x. If the values of the random variable occur in a set, then the lowercase letter will include an index--e.g. $x(i)$.)

Let Xp be a gaussian-distributed random variable having a mean μ of zero and a variance σ^2 of one (written Xp = N(0,1)). If a plot of r versus x_p is a straight line, then R is also gaussian-distributed. [98] This will be proven here.

The distribution of Xp is shown in Fig. 1. If Xp = N(0,1) then the value $x_p = 1$ is one standard deviation from the mean. The probability density function of x_p is given by equation (1).

$$f(x_p) = \frac{1}{\sqrt{(2\pi)}} e^{-(1/2)\, x_p^2} \qquad (1)$$

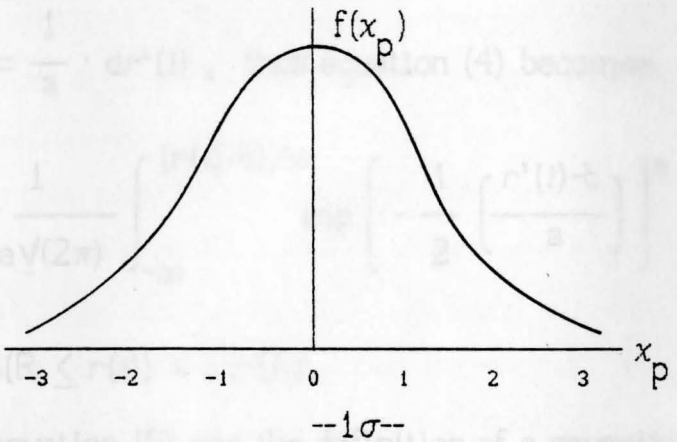


Fig. 1--Distribution of Xp = N(0,1)

The cumulative distribution is the integral of f(x) from $-\infty$ to x_p (equation (2)) and is the "percentile" of x_p divided by 100.

$$p(Xp \leq x_p) = F(x_p) \overset{\text{def}}{=} \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{x_p} e^{-(1/2)\,x^2}\, dx \qquad (2)$$

where $p(Xp \leq x_p)$ is read "the probability that the random variable Xp is less than or equal to the specific value x_p."

Now assume that $r(i)$ vs. $x_p(i)$ is a straight line (Fig. 2). This means that $R = a \cdot Xp + b$ where a and b are constants. One can write

$$r(i) = a \cdot x_p(i) + b$$

or
$$x_p(i) = \frac{r(i) - b}{a} \qquad (3)$$

Substituting equation (3) into equation (2):

$$F(x_p(i)) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{(r(i)-b)/a} \exp\left[-\frac{1}{2} \left(\frac{r'(i)-b}{a} \right)^2 \right] \cdot d\left(\frac{r'(i)-b}{a} \right) \qquad (4)$$

But $d\left(\frac{r'(i)-b}{a} \right) = \frac{1}{a} \cdot dr'(i)$, thus equation (4) becomes

$$F(x_p(i)) = \frac{1}{a\sqrt{(2\pi)}} \int_{-\infty}^{(r(i)-b)/a} \exp\left[-\frac{1}{2} \left(\frac{r'(i)-b}{a} \right) \right]^2 dr'(i)$$

$$= p(R \leq r(i)) = F(r(i)). \qquad (5)$$

It follows from equation (5) and the definition of a gaussian distribution that R is gaussian-distributed with variance a^2 (or standard deviation a) and mean b. We would write $R = N(b, a^2)$. Thus a line $r = a \cdot x_p + b$ has a slope equal to the standard deviation of R (Figure 2).

To create a normal probability plot (NPP) of data $r(i)$, one must satisfy (5) by finding corresponding values $x_p(i)$ such that

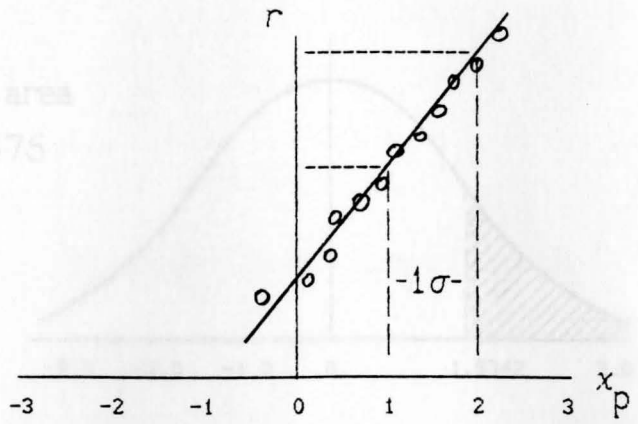


Fig. 2--r vs. x_p showing how the slope is the standard deviation of R.

$$F(x_p(i)) = F(r(i)) \tag{6}$$

for all i.

To do this, use the following steps:

1) List the n values of r in ascending order and index them such that the smallest r is assigned i=1. Then find the percentile of each $r(i)$. The percentile is $100i/n$. The percentile divided by 100 is $F(r(i))$.

2) Find the value of $x_p(i)$ from a standard normal probability table that satisfies equation (6).

As an example, assume one collected and ordered 160 resistance measurements in accordance with step 1 above; and $r(150)$ =1.35 mΩ. The percentile is $(150/160)\cdot 100\% = 93.75\%$ and $F(r(150)) = 0.9375$. The x_p that gives an area of 0.9375 under the normal curve is 1.5342 (Figure 3). Thus on an NPP one would plot a point at 1.35 on the resistance axis and 1.5342 on the x_p axis. Equation (21) in section 2.7 in the main body of this thesis gives an approximate value of x_p that is much easier to calculate in a computer program.

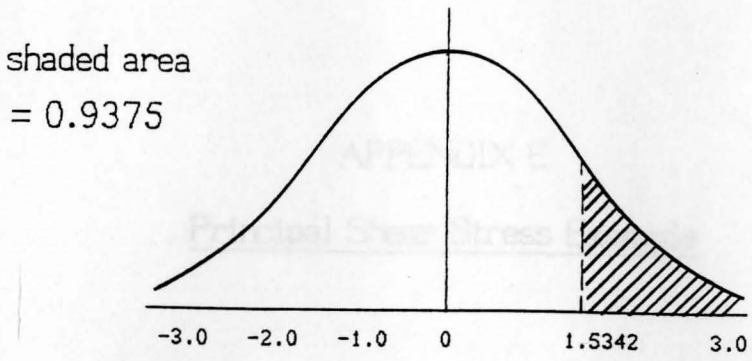


Fig. 3 Area under a Normal Curve

APPENDIX E

Principal Shear Stress Example

Fig. 1 shows a cubic block whose sides of area A are principal planes. The block is compressed by a force F_3 on the top plane Π_3. The plane of maximum shear stress (i.e., the angle θ) is found by maximizing

$$\sigma = \frac{F_3'}{A'} = \frac{F_3 \sin\theta}{A/\cos\theta} = \frac{F_3}{A} \sin\theta \cos\theta = p_3 \sin\theta \cos\theta$$

One shows that Π_1 and Π_3 are principal planes by letting $\theta = 90^0$ and 0^0 respectively and discovering that σ, the shear stress on these planes, is zero.

$$\frac{d}{d\theta}\left[\frac{F_3}{A}\sin\theta\cos\theta\right] = \frac{F_3}{A}\left[\cos^2\theta - \sin^2\theta\right] \overset{\text{set}}{=} 0$$

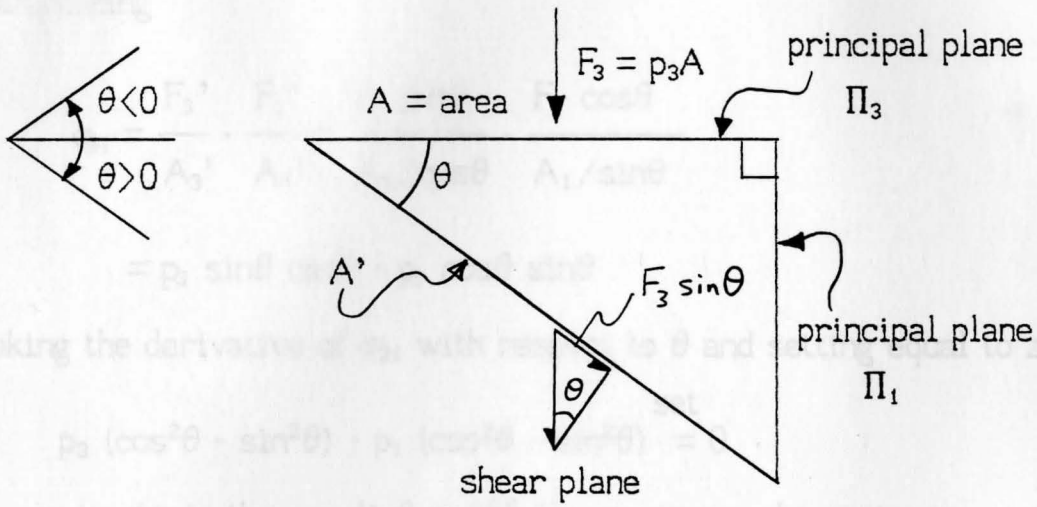


Fig. 1--Block subjected to a force F_3 .

$$\cos^2\theta = \sin^2\theta$$

$$\cos\theta = \pm \sin\theta \quad \text{or } \theta = \pm 45^0$$

The second derivative is

$$\frac{F_3}{A}\left[\, 2\cos\theta\,(-\sin\theta) - 2\sin\theta\,(\cos\theta)\,\right] < 0 \text{ for } 0 \leq \theta \leq 90^0.$$

Thus the maximum shear stress due to F_3 occurs on a plane oriented at $\theta = 45^0$; θ is positive in the clockwise direction from Π_3. This is the angle of the plane of maximum shear stress for the case of one principal stress.

The value of the maximum shear stress is obtained by letting $\theta = 45^0$:

$$\sigma = \frac{F_3}{A}\sin 45^0 \cos 45^0 = \frac{F_3}{A}\cdot\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} = \frac{1}{2}p_3$$

Fig. 2 shows the effect of two "principal" forces F_1 and F_2. The plane of the maximum shear stress (the angle θ) is found by maximizing

$$\sigma_{31} = \frac{F_3'}{A_3'} - \frac{F_1'}{A_1'} = \frac{F_3\sin\theta}{A_3/\cos\theta} - \frac{F_1\cos\theta}{A_1/\sin\theta}$$

$$= p_3\sin\theta\cos\theta - p_1\cos\theta\sin\theta$$

Taking the derivative of σ_{31} with respect to θ and setting equal to zero:

$$p_3(\cos^2\theta - \sin^2\theta) - p_1(\cos^2\theta - \sin^2\theta) \overset{\text{set}}{=} 0$$

we again obtain the result $\theta = 45^0$ for maximum shear stress.

The value of the shear stress acting along the 45^0 shear plane is

$$p_{31} = \sigma_{31}\Big|_{\theta=45^0} = p_3\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} - p_1\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} = \frac{1}{2}(p_3 - p_1)$$

which is the principal shear stress due to p_1 and p_3.

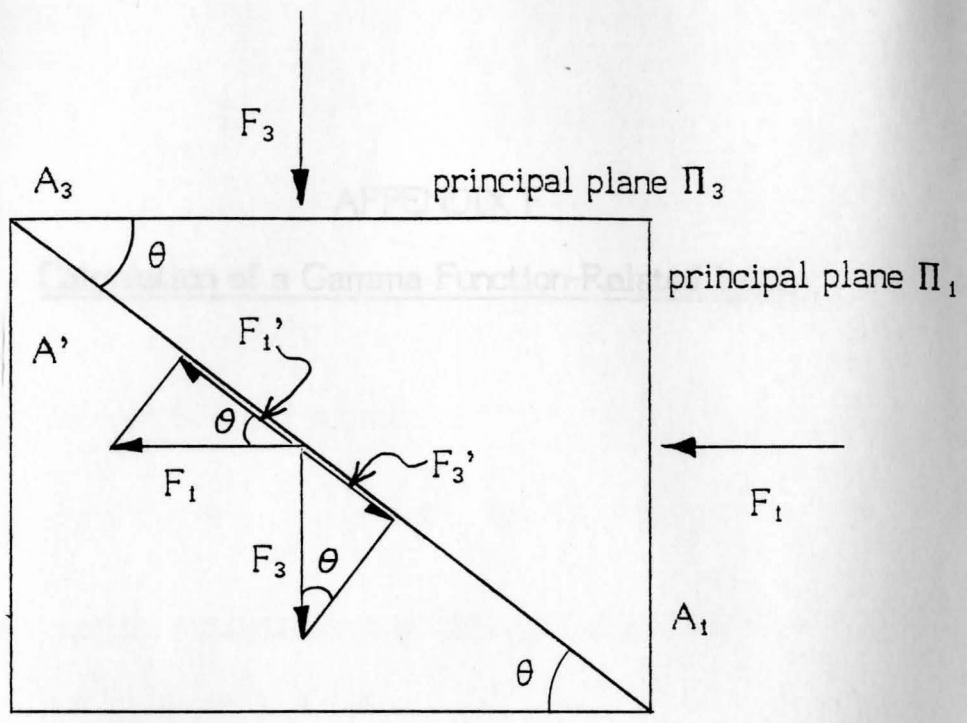


Fig. 2--Combined forces on a cube of principal planes.

APPENDIX F

Calculation of a Gamma Function-Related Integral

In this appendix the following definite integral will be calculated

$$\int_{z=s}^{\infty} (z-s)^{\alpha} e^{-z} dz \quad \text{for } s \text{ and } \alpha \geq 0 \tag{1}$$

To accomplish this, first recall the definition of the gamma function:

$$\Gamma(\alpha) = \int_{x=0}^{\infty} e^{-x} x^{\alpha-1} dx \tag{2}$$

let $\beta = \alpha-1$ or $\alpha = \beta+1$ and substitute

$$\Gamma(\beta+1) = \int_{x=0}^{\infty} e^{-x} x^{\beta} dx \tag{3}$$

now make a variable substitution so that (1) looks like (3)

$$\begin{aligned} x &= z - s \quad \text{or} \quad z = x + s \\ dz &= dx \end{aligned} \tag{4}$$

and the limits on the integral

$$\begin{aligned} &\text{when } z = s, \; x = 0 \\ &\text{when } z = \infty, \; x = \infty \end{aligned} \tag{5}$$

substituting (4) and (5) into (1)

$$\int_{x=0}^{\infty} x^{\alpha} e^{-(x+s)} dx = e^{-s} \int_{x=0}^{\infty} x^{\alpha} e^{-x} dx \tag{6}$$

now the right side of (6) looks like (3) except for the factor exp(-s).

So the right side of (6) is equal to

$$e^{-s} \Gamma(\alpha+1)$$

Therefore, $\int_{z=s}^{\infty} (z-s)^{\alpha} e^{-z} dz = e^{-s} \Gamma(\alpha+1)$ QED.

NOTES

1. Ragnar Holm, Electrical Contacts, 4th ed, Berlin, Springer-Verlag, 1981, pg. 1.

2. J.B.P. Williamson, seminar titled Intensive Course on Contact Theory and Application, Warren, MI, Oct. 12-15, 1987.

3. 1984 Holm Seminar, pg. xxi.

4. Williamson's seminar.

5. Ibid.

6. Frank Llewellyn-Jones, The Physics of Electrical Contacts, Oxford, Clarendon Press, 1957, pg. 1.

7. Harold N. Wagar, "Address of the Fifth Recipient of the Ragnar Holm Scientific Achievement Award," 1976 Holm Seminar, Illinois Institute of Technology (IIT), Chicago, pg. 3.

8. Wagar, pg. 4.

9. John E. Wimsey, "The Bare Base Electrical System," Eighth Annual Connector Symposium Proceedings, Camden, NJ, Electronic Connector Study Group Inc., 1975, pp. 426-449.

10. J. Aronstein and W.E. Campbell, "Failure and Overheating of Aluminum Wired Twist-on Connections," 1981 Holm Seminar, Chicago, IIT, pg. 259.

11. N.T. Bond, et al., "Fundamental Consideration of Aluminum Electrical Contact Interfaces," 1977 Holm Proceedings, Chicago, IIT, pg. 213.

12. E.M. Edwards, "Connector Problems in Two-Way Antennas," Communications International, Oct., 1984, pg. 54.

13. Llewellyn-Jones, pg. 9.

14. McGraw-Hill Dictionary of Scientific and Technical Terms, Third ed., New York, McGraw-Hill, 1984, pg. 646.

15. F.P. Bowden and D. Tabor, The Friction and Lubrication of Solids, Oxford, Clarendon Press, 1986, pg. 1.

16. J.B.P. Williamson, "The Microworld of the Contact Spot," 1981 Holm Seminar, Chicago, IIT, pp. 110.

17. J. Francyk and H. Kulikjan, "Mathematical Model of Constriction Resistance of Reed-Relay Contact," 1979 Holm Seminar, Chicago, IIT, pp. 235-239.

18. Else Holm, "Introductory Lecture on Contact Resistance," 1962 Holm Seminar, Orono, Maine, University of Maine, pp. 1-13.

19. K. Yamada et al., "Quantification of Distribution of Contact Resistance and Application Thereof," Osaka, Japan, Matsushita Electric Works, pg. 726.

20. R. Holm, pg. 10.

21. Yamada, pg. 726.

22. R. Currence and W. Rhoades, "Predicting, Modeling and Measuring Transient Resistance Changes of Degraded Electrical Contacts," 1983 Holm Seminar, Chicago, IIT, pg. 87.

23. Ibid.

24. ———, pg. 88.

25. Panel Discussion titled "Importance of Environment on Contact Performance," 1979 Holm Seminar, Chicago, IIT, pg. 243.

26. ———, pg. 249.

27. ———, pg. 243.

28. ———, pg. 242-243.

29. ———, pg. 243.

30. T.G. Grau, "A Field Study of the Electrical Performance of Separable Connectors," Proceedings of the 28th Electronic Components Conference, New York, IEEE, 1978, pg. 105.

31. Yamada, pg. 727.

32. E.R. Rowlands, "Long Range Reliability Prediction Study of Wire to Contact Junction," 7th Annual Connector Symposium Proceedings, Camden, NJ, Electronic Connector Study Group Inc., 1974, pp. 297-298.

33. J.H. Whitley and R.D. Malucci, "Contact Resistance Failure Criteria," 1978 Holm Seminar, Chicago, IIT, pg. 111.

34. ———, pp. 114-115.

35. W.H. Abbott and K.L. Schreiber, "Dynamic Contact Resistance of Gold, Tin, and Palladium Connector Interfaces During Low Amplitude Motion," 1981 Holm Seminar, Chicago, IIT, pg. 211.

36. W.H. Abbott, "Time Distribution of Intermittents versus Contact Resistance for Tin-Tin Connector Interfaces during Low Amplitude Motion," 1983 Holm Seminar, Chicago, IIT, pg. 55.

37. Abbott and Schreiber, pg. 211 (abstract).

38. Abbott, pg. 59.

39. J. Pullen and J.B.P. Williamson, "On the Plastic Contact of Rough Surfaces," Proc. R. Soc. Lond. A. 327, 1972, pp. 167.

40. J.B.P. Williamson, J. Pullen, and R.T. Hunt, "The Shape of Solid Surfaces," Surface Mechanics 24, New York, American Society of Mechanical Engineers, 1969, pg. 25.

41. J.B.P. Williamson and R.T. Hunt, "Asperity Persistence and the Real Area of Contact Between Rough Surfaces," Proc. R. Soc. Lond. A. 327, 1972, pg. 148.

42. J.R. Benjamin and C.A. Cornell, Probability, Statistics, and Decision for Civil Engineers, New York, McGraw-Hill, 1970, pp. 236-237.

43. R. Holm, 367-368.

44. R.J. Roark (W.C. Young, ed., 6th ed.), Roark's Formulas for Stress & Strain, 6th ed., New York, McGraw-Hill, 1989, pp. 647-651.

45. Francyk and Kulikjan, pg. 235.

46. J.A. Greenwood, "Constriction Resistance and the Real Area of Contact," Brit. J. Appl. Phys., Vol. 17, 1966, pp. 1621-1632.

47. R. Holm, pg. 8.

48. L.H. Van Vlack, Elements of Materials Science and Engineering, 4th ed., Reading, Mass., Addison-Wesley, 1980, pg. 522.

49. Llewellyn-Jones, pg. 9.

50. A.H. Uppal and S.D. Probert, "Deformation of Single and Multiple Asperities on Metal Surfaces," Wear, 20, 1972, pg. 389.

51. Williamson, Pullen, and Hunt, pp. 26-33.

52. Manual from Statistical Program written for Hewlett-Packard 9836 computer.

53. R. Holm, pg. 109.

54. Guide to Materials Engineering Data and Information, American Society for Metals, Metals Park, Ohio, 1986, pg. 2.6.

55. Thanks to Paul DiLiello, Electrical Contacts Research Lab, Packard Electric Division, General Motors Corporation, for making the measurements and providing the data.

56. D.S. Chhabra and G.T. Wenning, "The Effects of Surface Finish and Contact Geometry on Contact Resistance," 1965 Holm Seminar, Orono, Maine, University of Maine, pg. 327.

57. Genichi Taguchi, System of Experimental Design, Vol. 2, White Plains, NY, Kraus International Publications, 1987, pg. 1152.

58. Yamada, pp. 726-727.

59. Charles P. Quesenberry, "Screening Outliers in Normal Process Control Data with Uniform Residuals," GMR-4982, General Motors Research Laboratories, Warren, Mich., 1985.

60. Chhabra and Wenning, pg. 326.

61. Williamson and Hunt, pg. 147.

62. R. Holm, pg. 369.

63. Van Vlack, pg. 196.

64. ———, pg. 118.

65. McGraw Hill Scientific Dictionary, pg. 1271.

66. The proof requires showing that the stress tensor, S, which for any material is symmetric, can always be put into diagonal form by a proper rotation of the coordinate axes. The stress tensor is a 3×3 matrix whose diagonal elements are normal stresses and whose off-diagonal elements are shear stresses. The stress tensor is unique at a given point in the material; however, the components depend on the coordinate system to which it is referenced. If the coordinate axes are rotated to (x', y', z') such that the stress tensor matrix, referenced to this new coordinate system, is diagonal, then the principal planes are the planes normal to (x', y', z').

For references, see:

S. Timoshenko and J. Goodier, Theory of Elasticity, 3rd ed., New York, McGraw-Hill, 1970, pg. 221.

R.P. Feynman, et al., The Feynman Lectures on Physics, Vol. II, Reading, Mass., Addison-Wesley Publishing Co., 1964, pp. 31-1 thru 31-14.

67. D. Tabor, The Hardness of Metals, Oxford, Clarendon Press, 1951, pg. 28.

68. Ibid.

69. Ibid.

70. Ibid.

71. Roark, pg. 89.

72. Holm, pg. 371.

73. Holm, pg. 372.

74. Tabor, pg. 37.

75. Bowden and Tabor, pp. 11-13.

76. Holm, pg. 30.

77. R.M. Davies, "The Determination of Static and Dynamic Yield Stresses using a Steel Ball," Proc. Roy. Soc. of London, Series A, Vol. 197, 1949, pg. 422.

78. Davies, pg. 424.

79. Ibid.

80. Tabor, pg. 46.

81. Davies, pp. 424-425.

82. Davies, pg. 422.

83. Tabor, pg. 46.

84. Bowden and Tabor, pp. 11-13.

85. J.A. Greenwood and J.B.P. Williamson, "Contact of Nominally Flat Surfaces," 1966, pg. 307.

86. ———, pp. 307-308.

87. Tabor, pg. 141.

88. Greenwood and Williamson, pp. 301-303.

89. ———, pg. 309.

90. Chhabra and Wenning, pg. 320.

91. See references 39 and 40, also Greenwood and Williamson, pg. 303.

92. Tabor, pg. 141.

93. Greenwood and Williamson, pg. 302.

94. Metals Handbook, 9th ed., Vol. 2 Properties and Selection/ Nonferrous Alloys and Pure Metals, American Society for Metals, Metals Park, Ohio, 1979, pg. 324.

95. Greenwood and Williamson, pp. 305-307.

96. R.F. Elsner and M.J. Frazier, "Nonlinear Effects of Contacts in RF Systems Environments," 1967 Holm Seminar, pp. 155-161.

97. Daniel P. Siewiorek and Robert S. Swarz, The Theory and Practice of Reliable System Design, Digital Equipment Corporation, Digital Press, Bedford, Mass., 1982, pp. 201-259.

98. Wayne Nelson, Applied Life Data Analysis, John Wiley & Sons, New York, 1982, pp. 126-127.

99. Thanks to Mr. Bill Phillips of Packard Electric Division, General Motors Corporation, for this derivation.