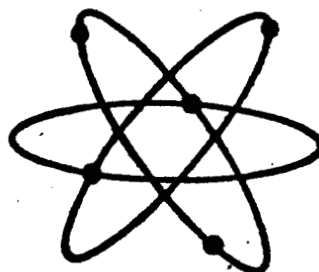


NASA
FACILITY FORM 002
N65-20969
(ACCESSION NUMBER)
109
(PAGES)
CB-57779
(NASA CR OR TMX OR AD NUMBER)
(THRU)
(CODE)
32
(CATEGORY)

MIDWEST APPLIED SCIENCE CORPORATION



SCIENTIFIC RESEARCH

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

FINAL REPORT
ON
STATISTICAL MODELS
OF
CUMULATIVE DAMAGE

by

Dr. Frank Kozin

for

Midwest Applied Science Corp.
1220 Potter Drive
McClure Research Park
West Lafayette, Indiana

December 1964

This research was conducted by MASC
for the
National Aeronautics & Space Administration
under
Contract No. NAS8-11180

ABSTRACT

214.7
Statistical Models of Cumulative Damage. A cumulative damage theory dealing with the failure of specimens in fatigue is presented. The damage is treated as a random variable, and the average damage is related to the irreversible work input to the specimen. The theory is compared to other theories of fatigue failure. Recommendations for further work are presented.

Author ↗

INTRODUCTION

This report represents an attempt on the part of Midwest Applied Science Corp. of West Lafayette, Indiana, to set forth a theoretical basis upon which one may develop a cumulative damage theory that is associated with the work input into a material specimen. The motivation for the theory is to take into account at once, the statistical variability that is inherent in fatigue tests, the changing material properties and to account for random as well as pure sinusoidal load inputs. It may appear to be a somewhat over-ambitious task to attempt a theory incorporating all of these factors. However, a very straight forward phenomenological approach yields the cornerstone needed to accomplish all of these points above. This cornerstone is the irreversible work input into the specimen as measured by the area of the stress strain hysteresis loop. This phenomenon has only relatively recently come back into favor in the study of the material properties of metals. It has, for example, already provided the basis for unification of many phenomena studied by materials properties researchers in dislocation theory, etc. The irreversible

work concept has also begun to permeate the fatigue and cumulative damage literature as seen in some recent papers, (Gatts [1], Feltner and Morrow [1]). However, we feel that the present report constitutes the first attempt to systematize a unification of the various statistical aspects of cumulative damage based upon this work input. The statistical properties of the present formulation are based upon the fact that cumulative damage may be viewed as a continuous birth process. That is, in the simplified case of a pure sinusoidal input, each cycle of the input stress function or load yields a non-negative accumulation of damage.

This damage is considered to be a random variable representing a random amount of birth or population growth. Assuming that there is a fixed upper limit of damage beyond which the specimen is said to fail, the basic question is; When does the total damage, or total population, first exceed this given value of the upper limit of accumulated damage? Furthermore, what is the average damage at any given time? Thus, the birth process concept underlies any cumulative damage theory. Statistically speaking, the problem of determining the first time for which the total damage exceeds any given level is referred to as the first passage time problem or, perhaps, it might be referred to as the first

passage cycle problem when referring to load cycles. In any case it is this point that was first recognized independently by Parzen (Parzen [1]) and Bogdanoff and Kozin (Bogdanoff and Kozin [1]).

Parzen was interested in the cumulative damage of a material specimen according to a level crossing model. However, he introduced the very clever idea of applying renewal ideas to the failure problem. Bogdanoff and Kozin were essentially interested in the failure of any "brittle" system. That is, given a state variable of the system, the system is said to have failed if the state variable passes outside of a given region having started within the region. Actually the formulation of Bogdanoff and Kozin is sufficiently general to include every dynamical situation since no restriction was placed upon the nature of the state variable. Therefore, when the state variable is the cumulative damage, then the first passage above the upper total damage limit constitutes a "brittle" failure. The question is; What is the mechanism of failure?, or from the engineering point of view, What is a suitable phenomenological indication of the mechanism of failure? Since the time of Palmgren (Palmgren [1]) and more recently the work of Miner (Miner [1]), the stress alone has been the indication via the so-called

Palmgren-Miner hypothesis. Based upon the Palmgren-Miner hypothesis, many inconsistencies have been witnessed.

For example, upon subjecting a specimen to two different stress levels yields a non-commutativity in the accumulated damage that cannot be accounted for by the Palmgren-Miner hypothesis. There have been a number of attempts to try to derive analytical devices to do away with this problem such as non-linear laws or the work of Freudenthal and Heller (Freudenthal and Heller [1]) using a modified linear accumulation hypothesis. However, such modifications of the Palmgren-Miner hypothesis in general are unrealistic and contain the basic faults that are contained in the hypothesis itself. Various attempts have been made to understand the nature of fatigue life of a specimen under random loadings. A very recent work along these lines is the paper of Leybold and Naumann (Leybold and Naumann [1]), in which a very interesting study was made of the relative failure times predicted from computed statistics (e.g. number of peaks, magnitude of absolute maxima and minima, etc.) as compared with the failure time under the actual load function, where the time is taken as unity. Their results are, indeed, most interesting and their work clearly represents one of the first sound studies of the

failure phenomenon under random loading. However, as the authors themselves state, their work does not yield any underlying analytical basis upon which to predict the results for other environments.

In fact it is conceivable that their results may change from sample to sample, thus not giving useful information for the random case. Our feeling here is that only average damage or failure properties can make sense in the random case, since we can never know with probability one what the properties of the materials are and hence never know the damage per cycle accumulated by the specimen or even its distributional properties. Hence, we shall probably have to accept the fact that moments will be the best information we can obtain.

In this report, we shall first present a review of various statistical theories of cumulative damage and state their valid points as well as their weaknesses. We shall then present our approach to the problem and analyze a few examples. We shall terminate this report with a comparison of the approach presented and the theories described in the review. We shall, furthermore, make suggestions for future experimental research work that is urgently required

in order to provide the necessary data concerning hysteresis loop areas that are included as basic quantities in our formulas.

CHAPTER II

A REVIEW OF SOME PERTINENT STATISTICAL THEORIES OF CUMULATIVE DAMAGE

Those theories of cumulative damage in which probabilities and statistics have played a basic role can be roughly broken into two categories: Theories in which damage itself either implicitly or explicitly has been considered to be a random variable, and theories in which damage has been presumed deterministic but the input or stress history is assumed random. In only one case (Parzen [1]) has the random damage and random input problems been simultaneously modeled. In almost all theories, except for a few, (Freudenthal and Heller [1], Gatts [1]) the changing material properties have not been taken into account. Furthermore, a large number of the theories, especially those with deterministic damage and random inputs are merely variations of the Palmgren-Miner approach. In the present section we shall consider first models for which the damage is implicitly or explicitly assumed to be a random variable.

The underlying idea for all such models is that a material is subjected to successive load applications of a sinusoidal nature at a fixed stress amplitude s . One then studies the number, $N(s)$, of load applications at this stress level that are required to produce "failure" of the specimen. It is a well known fact, frequently observed (Gumbel [1]), that the number of stress applications required to produce fatigue failure is a random variable. That is, for any given collection of apparently identical specimens of the same material it is observed that the number of load cycles required to produce fatigue failure varies in an unpredictable way from specimen to specimen. Furthermore, such statistical spread as has been observed is too great to explain away by mere experimental procedure. Indeed, high quality controlled experiments still yield a spread wide enough to conclude that this is a physical phenomenon that is being observed. Hence, one must consider the inherent variation in the number of cycles to failure, as Gumbel puts it, "...the very essence of the problem". As a result various models of materials have been proposed in order to attempt to predict a theoretical distribution function for the number of cycles to failure, $N(s)$, at a given stress level s .

That is, these models have been constructed to determine the failure distribution.

$$2.1) \quad P_s(n) = \text{Prob} \{N(s) \leq n\},$$

or, equivalently, the survival function

$$2.2) \quad L_s(n) = 1 - P_s(n) = \text{Prob} \{N(s) > n\}.$$

Several models have been proposed in the past and we wish to describe a few of them along with their assumptions (Murphy [1]).

A.) The Log-Normal Distribution

The Probability density in this case is

$$2.3) \quad p(x) = \frac{dP(x)}{dx} = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp \left[\frac{-1}{2\sigma^2} (\log x - \mu)^2 \right] & x > 0 \\ 0 & x < 0. \end{cases}$$

A random variable X is log-normal if and only if $\log X$ is normally distributed with mean μ and variance σ^2 .

In such a case, one finds the statistics to be (Parzen [1])..

$$2.4) \left\{ \begin{array}{ll} \text{mode} & - e^{\mu - \sigma^2} \\ \text{median} & - e^{\mu} \\ \text{mean} & - e^{\mu + 1/2 \sigma^2} \\ \text{variance} & - e^{2\mu + \sigma^2} (e^{\sigma^2} - 1). \end{array} \right.$$

The model yielding the log-normal distribution for the number of cycles to failure at a given stress level s is based upon the model of proportional effects first advanced by Kapteyn in 1903.

Let $D_1(s), \dots, D_n(s)$ be a sequence of random variables that represent the amount of damage accumulated at each successive load application at stress level s . The basic postulate is that the damage at the n^{th} load application is related to its predecessor as

$$2.5) \quad D_n(s) - D_{n-1}(s) = \epsilon_n D_{n-1}(s)$$

where ϵ_n is the random effect due to the n^{th} load application and $\{\epsilon_i\}$ is a sequence of independent random variables. It follows immediately from 2.5) that

$$2.6) \quad D_n(s) = (1 + \epsilon_n) D_{n-1}(s) = (1 + \epsilon_n) \dots (1 + \epsilon_1) D_0.$$

Hence, the logarithm of $D_n(s)$ is the sum of a large number of independent terms for n large. If the assumption is made that the independent random variables are distributed in such a manner that the central limit theorem applies, this will imply that $\log D_n(s)$ is approximately normal for n large. Hence, $D_n(s)$ is approximately log-normal for n large.

It may then be argued (Freudenthal and Gumbel [1]) that from 2.5) and the fact that ϵ_n is independent of $D_{n-1}(s)$, it follows that the rate of increase of the average total damage is proportional to the average total damage. Furthermore, the number of cycles to yield a given damage will be approximately inversely proportional to this rate of change of damage, and hence approximately inversely proportional to the total damage. That is,

$$2.7) \quad N_D(s) = \frac{1}{D_N(s)},$$

where $N_D(s)$ is the number of cycles required to yield the total damage of magnitude $D_N(s)$.

However,

$$2.8) \quad \log \left[\frac{1}{D_N(s)} \right] = -\log D_N(s) ,$$

hence $\log N_D(s)$ is approximately normal, that is, $N_D(s)$ is log-normal for any D . Therefore, assuming a fixed level of damage for failure, it follows that the number of cycles to failure $N(s)$ is log-normal.

We may point out here that since $\{\epsilon_i\}$ is a sequence of independent random variables, it follows that $\{D_i(s)\}$ is a Markov chain implying all the features of the Markov property for the sequence of random variables. In particular the distribution desired is the so-called steady state or ergodic distribution, which is log-normal for the equation 2.5).

However, it is difficult to justify the very particular form 2.5) out of the general possibilities given by

$$2.9) \quad D_n(s) = A(D_{n-1}(s)) + \epsilon_n B(D_{n-1}(s)) ,$$

which in general will yield distributions different from the log-normal.

B.) Extreme Value Distribution-First Asymptotic Form.

The probability distribution in this case is

$$2.10) \quad P(x) = 1 - \exp \left\{ -\exp \left(-\frac{x-\mu}{\beta} \right) \right\}, \quad -\infty < x < \infty,$$

and the probability density function is given as

$$2.11) \quad p(x) = \frac{dP(x)}{dx} = \frac{1}{\beta} \exp \left\{ -\left(\frac{x-\mu}{\beta} \right) - \exp \left(-\frac{x-\mu}{\beta} \right) \right\}, \quad -\infty < x < \infty$$

The parameter ranges are $\mu \in (-\infty, \infty)$, $\beta \in [0, \infty)$. A random variable distributed according to the extreme value distribution above possesses the statistics (Parzen [1])

$$2.12) \quad \begin{cases} \text{mode} & - & \mu \\ \text{median} & - & \mu + (.36657) \beta \\ \text{mean} & - & \mu + (.57722) \beta \\ \text{variance} & - & \frac{\pi^2}{6} \beta^2 \end{cases}$$

The physical assumptions leading to such an extreme value distribution may be stated briefly in the following fashion (Murphy [1]). Let us suppose that the material is made up of fibers or very thin rods. Furthermore, suppose that damage to this material is equivalent to the snapping of fibers in the bundle. The properties of the fibers are

assumed to be independent of one another and each distinct fiber can withstand a number, $N(s)$, of load applications at stress level s to the entire bundle. The number $N(s)$ is a random variable distributed according to the distribution $F_s(n)$. That is,

$$2.13) \quad \text{Prob } \{N(s) \leq n\} = F_s(n).$$

One postulates that the system fails when all fibers have failed. The question becomes, "What is the survival function?" That is, what is the probability that the material survives n cycles at stress level s . This is just the probability that after n cycles there is at least one fiber in tact. This is essentially a strongest link theory. If there are M fibers in the bundle, then this probability is simply given as

$$2.14) \quad 1 - [F_s(n)]^M.$$

The associated density function is

$$2.15) \quad f_s(n) = \frac{d(1 - [F_s(n)]^M)}{dn} = -M [F_s(n)]^{M-1} \frac{dF_s(n)}{dn},$$

which is merely the probability density for an extreme value in a collection of M independent samples from a given distribution. Under a wide set of conditions (Gumbel [2], Gnedenko [1]), the limiting distribution of $[F_S(n)]^M$ as $M \rightarrow \infty$ is of the form 2.10).

C.) Extreme Value Distribution - Weibull Distribution

The probability distribution in this case is

$$2.16) \quad P(x) = \begin{cases} 1 - \exp \left[-\left(\frac{x-\epsilon}{v-\epsilon} \right)^\kappa \right] & x \geq \epsilon \\ 0 & x < \epsilon \end{cases}$$

where $\kappa, v > 0$.

The probability density is given as

$$2.17) \quad p(x) = \frac{dP(x)}{dx} = \begin{cases} \frac{\kappa}{v-\epsilon} \left(\frac{x-\epsilon}{v-\epsilon} \right)^{\kappa-1} \exp \left[-\left(\frac{x-\epsilon}{v-\epsilon} \right)^\kappa \right] & x \geq \epsilon \\ 0 & x < \epsilon \end{cases}$$

A random variable distributed according to 2.16) possesses the statistics

$$2.18) \left\{ \begin{array}{lcl} \text{mode} & - & \epsilon + (v-\epsilon) \left(1 - \frac{1}{\kappa}\right)^{1/\kappa} \\ \text{median} & - & \epsilon + (v-\epsilon) (\log 2)^{1/\kappa} \\ \text{mean} & - & \epsilon + (v-\epsilon) \Gamma\left(1 + \frac{1}{\kappa}\right) \\ \text{variance} & - & (v-\epsilon)^2 \left[\Gamma\left(1 - \frac{2}{\kappa}\right) - \Gamma^2\left(1 + \frac{1}{\kappa}\right) \right], \end{array} \right.$$

where $\Gamma(y)$ is the Gamma function. For, ϵ, v fixed, the mode, median, and mean all approach v as $\kappa \rightarrow \infty$. Therefore, v is usually considered the parameter of location of the Weibull distribution.

Physical assumptions leading to the above distribution may be briefly stated in the following fashion (Murphy [1]).

We assume the material and statistical properties to hold in exactly the same fashion as assumed in the previous case. We only change the postulate for failure. We shall assume that the system fails if only one fiber fails. This is tantamount to saying that the material is only as strong as its weakest link. The probability that none of the fibers has failed in n applications of the load at stress level s

is given by

$$2.19) \quad [1 - F_s(n)]^M,$$

where $F_s(n)$ was defined in the previous case and M is the number of fibers making up the material.

The density function for survival is therefore

$$2.20) \quad f_s(n) = -M [1 - F_s(n)]^{M-1} \frac{dF_s(n)}{dn}.$$

Again, under a wide set of conditions given in the references cited above, $1 - [1 - F_s(n)]^M$ will approach the Weibull distribution 2.16), as $M \rightarrow \infty$.

The three failure distributions above constitute the most commonly used distributions in the study of failure under load applications at a given stress level. There are arguments both for and against the models used above along with the distributions derived from them. Before we state such pros and cons it will be beneficial to discuss the very useful concept of the hazard function (Parzen [2]).

We define $\mu(x)$, referred to as the hazard function, to be the conditional failure density function. Thus,

$$2.21) \quad \mu(x)dx = \text{Prob} \{ \text{failure for } t \in (x, x + dx) \mid \text{no failure for } t \in (t_0, x) \}.$$

From the definition of conditional probabilities, one has

$$\begin{aligned} & \text{Prob} \{ \text{failure for } t \in (x, x + dx) \mid \text{no failure for } t \in (t_0, x) \} \\ & \quad \times \text{Prob} \{ \text{no failure for } t \in (t_0, x) \} \end{aligned}$$

$$\begin{aligned} 2.22) \quad & \\ & = \text{Prob} \{ \text{failure for } t \in (x, x + dx), \text{ no failure for } t \in (t_0, x) \} \\ & = \text{Prob} \{ \text{failure for } t \in (x, x + dx) \}. \end{aligned}$$

The last equality follows since failure for $t \in (x, x + dx)$ implies no failure for $t \in (t_0, x)$.

Therefore, from 2.22) assuming a failure distribution $F(x)$, we have

$$2.23) \quad \mu(x)dx [1-F(x)] = f(x)dx.$$

Therefore, $\mu(x)$, the hazard function, is given as

$$2.24) \quad \mu(x) = \frac{f(x)}{1-F(x)} .$$

If $F(\epsilon) = 0$, we easily solve 2.24) to yield

$$2.25) \quad F(x) = 1 - \exp \left[- \int_{\epsilon}^x \mu(y) dy \right] , \quad x > \epsilon .$$

From 2.25), one may derive the failure distribution if the hazard function, that is the conditional rate of failure, is known. Hence, on the basis of assumptions of the hazard function, failure laws can be derived. For example, we may suppose that the hazard rate is a constant μ . Thus, by the definition of hazard as a conditional probability density we have a stationary chance failure independent of the past. A simple application of 2.25) for this case yields the well-known exponential failure function commonly used in reliability studies.

It is of interest to determine the hazard functions associated with the three failure probabilities derived above. They may be found to be

A.) Log-Normal Distribution

$$\mu(x) = \frac{\frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2} (\log x - \mu)^2\right]}{\int_x^\infty dy \frac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2} (\log y - \mu)^2\right]}, \quad x > 0.$$

B.) Extreme Value Distribution - First Asymptotic Form

$$\mu(x) = \frac{1}{\beta} \exp\left[-\frac{x-\mu}{\beta}\right], \quad -\infty < x < \infty.$$

C.) Extreme Value Distribution - Weibull Distribution

$$\mu(x) = c\kappa (x-\epsilon)^{\kappa-1}, \quad \kappa > 1, \quad x > \epsilon.$$

Based upon a physical picture that describes fatigue as a process of progressive damage terminated by actual failure, it follows that the risk function shall be an increasing function of x , or of N , the number of cycles sustained by the material. That is, if a phenomenon such as coxing is not prevalent, we should expect the larger the number of load applications that a given specimen survives, the greater should be the probability of failure on the very next cycle. This property is possessed by the risk functions associated with the extreme value distributions B.), C.) as may easily be seen. However, the risk function for the Log-normal distribution possesses a very slow asymptotic approach to zero after a reasonable sharp rise to a maximum value. On the basis of the risk function Freudenthal (Freudenthal [1]) strongly rules out the Log-normal distribution by the theoretical arguments put forth by Freudenthal. However, he points out that these arguments do not invalidate the use of the Log-Normal distribution as an approximation to the failure density for purposes other than extrapolation, especially for the 5% - 95% range. Freudenthal concurs with Corten's statements on this point. Indeed Weibull has shown that one cannot distinguish between the Extreme Value distributions and the Log-Normal distribu-

tion in the 5% - 95% range.

In the same paper Freudenthal presents an argument in favor of the Weibull distribution, assuming that one does not expect failure below some initial finite number of load applications and the monotone non-decreasing properties required of the risk function. He puts forth the argument that the simplest such risk function is a power function, and hence is lead to the Weibull distribution. It is interesting to note that Freudenthal has been studying statistical models of fatigue since 1946 (Freudenthal [2]). However, it is our opinion that in many of the arguments presented the physical picture seems to be secondary to arithmetical simplicity. In fact, the arguments appear arithmetical rather than physical. We certainly agree that tractability is important for applications, however it should aid the theoretical development and not lead it. Concerning the models presented above we agree with the comments of Epstein (Epstein [1]), who essentially states that any theoretical argument that leaves out the history of the stress function and neglects the fact that each time a material is cycled its properties change, implying a non-stationarity of the distribution of the fiber strengths, leaves out the fundamental physical phenomenon of fatigue

failure. Such is the case with the "static" theories presented above, as well as almost all models posed so far.

We shall now turn to a model which in our opinion is the most significant statistical model yet posed that is physically based upon the Palmgren-Miner motivated stress exceedence principles. Statistically this model, due to Parzen (Parzen [1]), offers an approach that is sure to yield important results in the future. Unfortunately, as far as we know, the model is not too well known in the general cumulative damage literature.

Parzen assumes that damage is a non-negative random variable associated with the application of a load at stress level s , referred to as $D(s)$. For the i^{th} application, he assumes the damage done to be represented as $D_i(s)$. His basic assumption is that the damage done at each successive application is independent of any other application, and furthermore the successive damages at each application are identically distributed. Thus, $\{D_i(s)\}$, for fixed parameter s , is a sequence of independent identically distributed random variables. Again, since the foundation of the idea of cumulative damage is that damage is additive, it is assumed that the total damage accrued after N -stress appli-

cations to be given by the sum

$$2.26) \quad D_1(s) + D_2(s) + \dots + D_N(s)$$

of non-negative, independent identically distributed random variables. Now let us assume that the material specimen can only withstand a given amount of damage, D , before it is said to have failed. Let $N_D(s)$ be the number of cycles to failure of a specimen of strength D at stress level s .

Obviously, $N_D(s)$ is the smallest integer N for which

$$2.27) \quad D_1(s) + \dots + D_{N_D(s)}(s) > D.$$

Thus, we see that we are looking here at a simple first passage situation.

Parzen's basic contribution here is to recognize that his assumptions lead to a model of cumulative damage that is a stationary renewal counting process, (Smith [1]).

If one defines,

$$2.28) \begin{cases} F_s(D) = \text{Prob} \{D(s) \leq D\} \\ M_s(D) = E \{N_D(s)\}, \end{cases}$$

where $E \{ \}$ is the expectation operator, it follows that $F_s(D)$, $M_s(D)$ are related through the "renewal equation", (Parzen [2]),

$$2.29) \quad M_s(D) = F_s(D) + \int_0^D M_s(D-D) dF_s(D).$$

Theoretically speaking, in order to determine the probability distribution of the damage random variable for a single load application, one would only require a knowledge of the expected numbers of cycles to the various damage levels and not the probability distributions of the cycles to failure. However, this is only of theoretical interest and of no practical significance. But, on the other hand, it follows from the basic limit theorems of Renewal Theory, (Smith [1]), that

$$2.30) \left\{ \begin{array}{l} \lim_{D \rightarrow \infty} \frac{E \{N_D(s)\}}{D} = \frac{1}{E \{D(s)\}} \\ \lim_{D \rightarrow \infty} \frac{\text{Var} \{N_D(s)\}}{D^2} = \frac{\text{Var} \{D(s)\}}{(E \{D(s)\})^3} \end{array} \right.$$

which is of important practical significance, for these formulas yield the method by which one may obtain the mean and variance of the damage on any given cycle and hence, because of the assumed independence, 2.30) will also yield the mean and variance after any given number of load applications. This is all given in terms of a readily accessible experimentally determinable statistic, namely, the average number of cycles to failure. Hence, for large D , 2.30) yields

$$2.31) \quad \begin{array}{l} E \left\{ \frac{D(s)}{D} \right\} = \frac{1}{E \{N_D(s)\}} \\ \text{Var} \frac{D(s)}{D} = \frac{\text{Var} \{N_D(s)\}}{(E \{N_D(s)\})^3} \end{array}$$

where $\frac{D(s)}{D}$ represents the proportion of damage on a given cycle.

It is important to note that Parzen's approach can take into account the complex load environments that a specimen may be subjected to, again by assuming the statistical independence of the damages produced by the various load applications. Further, it is assumed that all applications of the same stress amplitude, s , possess the same distributional properties and hence the same expected values as determined by 2.31). Thus, suppose in a given history of stress loads the levels $s_1, \dots, s_N$ are found in a total of M cycles. Let these be $M(s_i)$ cycles of load level s_i . Hence,

$$2.32) \quad M = M(s_1) + \dots + M(s_N).$$

For any given random environment, the variables $\{M(s_i)\}$ are also random variables. For a given stress level s_i , of which there are $M(s_i)$ cycles, we set

$$2.33) \quad D_M(s_i) = D_1(s_i) + D_2(s_i) + \dots + D_{M(s_i)}(s_i).$$

It easily follows that

$$2.34) \begin{cases} E \{D_M(s_i)\} = E \{M(s_i)\} E \{D(s_i)\} \\ \text{Var} \{D_M(s_i)\} = E \{M(s_i)\} \text{Var} \{D(s_i)\} + \text{Var} \{M(s_i)\} (E \{D(s_i)\})^2. \end{cases}$$

For the total damage, D , we find

$$2.35) \begin{cases} E \{D\} = \sum_i E \{M(s_i)\} E \{D(s_i)\} \\ \text{Var} \{D\} = \sum_i [E \{M(s_i)\} \text{Var} \{D(s_i)\}] + \text{Var} \left\{ \sum_i M(s_i) E \{D(s_i)\} \right\}. \end{cases}$$

Therefore, from the stated assumption and the renewal theory limits 2.30), we have for large D ,

$$2.36) \begin{cases} E \left\{ \frac{D}{\bar{N}} \right\} = \sum_i \frac{E \{M(s_i)\}}{E \{N_D(s_i)\}} \\ \text{Var} \left\{ \frac{D}{\bar{N}} \right\} = \sum_i E \{M(s_i)\} \frac{\text{Var} \{N_D(s_i)\}}{(E \{N_D(s_i)\})^2} + \text{Var} \left\{ \sum_i \frac{M(s_i)}{E \{N_D(s_i)\}} \right\} \end{cases}$$

Clearly one must know the statistics of the environment in order to determine the required mean values. However, such data is available in many cases such as in turbulent gust load data, various vibrational environments encountered from acoustical noise sources, etc.

Commenting upon Parzen's approach, we feel that it is the most significant statistical model yet posed in the theory of fatigue failure. Furthermore, his model allows one to directly obtain, through experiment, the quantities of interest in practical applications, that is, the mean and variance of the damage, without requiring assumptions concerning the nature of the distribution of the number of cycles to failure. It is fair to say that Parzen's approach possesses some weaknesses. The assumption of identical distribution for a damage of load level s independent of when the load occurs in the stress history as well as basing damage upon stress exceedence are the main weaknesses. However these weaknesses are shared by almost all theories so far presented of fatigue damage. Hence, Parzen's approach is certainly no worse, physically speaking, than those approaches put forward by experts in the field. In the case of a single amplitude, independence and identical distributions are probably not the poorest of assumptions. But, when various

stress amplitudes are present, these assumptions are already known to be incorrect, especially the assumption related to identical distributions.

The models discussed above have been models based upon the damage as a random variable. We wish now to discuss some results in cases for which the damage is a deterministic variable such as in the classical Palmgren-Miner approach, and in the very interesting approach of Gatts, but for which the complex stress histories have been studied.

Probably one of the most well known works along these lines was accomplished by Miles (Miles [1]). His work has been the subject of many subsequent investigations in the applied mechanics field. Using the Palmgren-Miner hypothesis, Miles determined the average damage when the stress history was the response of a single degree of freedom oscillation subjected to a Gaussian excitation and for which the S-N diagram for the material is of the form $NS^a = S_1^a$. Therefore, the randomness of the damage here is entirely due to the statistical properties of the environment. We shall present here a recent derivation of Miles' result (Crandall, Mark, Khabbaz [1]).

The total damage is assumed to be a function of time, and for large values of time T , the total damage $D(T)$ is represented as the sum of incremental damages accrued during incremental time intervals. Thus, let the interval $[0, T]$ be divided into M equal sub-intervals. Let the damage d_i be associated with the i^{th} sub-interval. Hence,

$$2.37) \left\{ \begin{array}{l} D = \sum_{i=0}^{M-1} d_i \\ E\{D\} = \sum_{i=0}^{M-1} E\{d_i\} = (M-1) E\{d_0\}. \end{array} \right.$$

The last equality holds from the assumed stationarity of the stress history and hence from the stationarity of the damage. Ordinarily, in the Palmgren-Miner hypothesis, the damage d_i is given as $\frac{1}{N_i}$, a constant factor (the inverse of the number of cycles to failure at stress level S_i) as given by the assumed S-N diagram. However, Crandall et al find it more convenient to look at half cycles and write the damage $d_i = \frac{1}{2N_i}$ as the damage per half cycle. This is simply an analytical aid and has no physical significance. For narrow band stationary Gaussian processes it is well known that the expected frequency ω_0 of the

process may be associated with the zero crossings (Rice [1]) and is given as

$$2.38) \quad \omega_0^2 = \frac{\int_{-\infty}^{\infty} \omega^2 G(\omega) d\omega}{\int_{-\infty}^{\infty} G(\omega) d\omega}$$

where $G(\omega)$ is the power spectral density function of the assumed stationary Gaussian process. Using the assumption of the narrow band Gaussian process, as represented by a sinusoidal function of frequency ω_0 and varying stress levels, the authors write the stress amplitude as

$$2.39) \quad s = \frac{|\dot{s}|}{\omega_0},$$

where $\dot{s}$ is the slope of the zero crossing.

Hence, N may be obtained from $\dot{s}$ through the assumed S-N relation. Now dividing $[0, T]$ into small intervals of length Δt , it follows that only for those sub-intervals for which there is an axis crossing, will there be damage. Thus, upon applying the now classical technique of Rice, the authors obtain the expected damage for an interval of length Δt to be given by

$$2.40) \quad \left\{ \begin{aligned} E \{d\} &= \int_{-\infty}^0 \frac{1}{2N} d\dot{s} \int_0^{\dot{s}\Delta t} p(s, \dot{s}) ds + \int_0^{\infty} \frac{1}{2N} d\dot{s} \int_{-\dot{s}\Delta t}^0 p(s, \dot{s}) ds \\ &= \Delta t \int_{-\infty}^{\infty} \frac{1}{2N} |\dot{s}| p(o, \dot{s}) d\dot{s} . \end{aligned} \right.$$

Since s is Gaussian, it follows that

$$2.41) \quad p(o, \dot{s}) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[- \frac{\dot{s}^2}{2\sigma_s^2} \right] .$$

Applying 2.39) to the S-N relation yields.

$$2.42) \quad \frac{1}{2N} = \frac{1}{2} \left(\frac{|\dot{s}|}{\omega_0 s_1} \right)^\alpha .$$

Substituting 2.41), 2.42) into 2.40) gives

$$2.43) \quad E \{d_o\} = \Delta t \frac{\omega_o}{2\pi} \left(\frac{\sqrt{2\sigma_s}}{s_1} \right)^\alpha \Gamma \left(1 + \frac{\alpha}{2} \right) .$$

Now inserting 2.43) into 2.37), letting M become large and Δt become small ($M\Delta t = T$), Miles' result

$$2.44) \quad E \{D(T)\} = \frac{\omega_o}{2\pi} T \left(\frac{\sqrt{2\sigma_s}}{s_1} \right)^\alpha \Gamma \left(1 + \frac{\alpha}{2} \right)$$

is obtained.

Upon making further assumptions, the authors proceed to study the variance of the damage and apply their results to various special cases. Physically speaking the work suffers from the same faults cited above concerning the Palmgren-Miner hypothesis and "static" material properties. However, it does point out a few of the analytical problems involved with the introduction of random stress histories. These problems are mainly concerned with the axis crossing and extreme value properties of the stress history. These problems are among the most difficult and at the same time the most important to be encountered in physical applications

of stochastic models. The use of the narrow band assumption in order to obtain an "average frequency" from which the maximum amplitude may be obtained from the slope at the axis crossing has been the only physical approach used so far. This leads mainly to random amplitude variations in the successive cycles and leaves the field wide open for the study of arbitrary stress input spectra. On this point we mention again the work carried on by Leybold and Naumann (Leybold and Naumann[1]), who have been actively engaged in the determination of the pertinent statistics of various random stress functions. These statistics are concerned with extreme value and axis crossing statistics. It is just this type of information that will certainly become of increasingly greater importance in the future analysis of fatigue properties under random loadings.

In all of examples of fatigue models cited above not one has taken into consideration the fundamental characteristic of a material specimen. That is, the fact that the physical properties are changing as the specimen undergoes stressing. Such properties cannot be fully determined by stress exceedences alone. This very fact, in our opinion, rules out any theory that is based upon stress exceedence alone. However in the work of Freudenthal and Heller, cited

previously, an attempt to incorporate material properties implicitly was made via a modified Palmgren-Miner rule. The Palmgren-Miner hypothesis states that failure occurs when,

$$2.45) \quad \sum_{i=1}^n p_i N \frac{1}{N_{s_i}} = 1 ,$$

where $s_1, \dots, s_n$ are the stress amplitudes present, N_{s_i} is the required number of cycles for failure at stress amplitude s_i , N is the total number of cycles, and p_i is the proportion of cycles at amplitude s_i .

Freudenthal and Heller replace 2.45) by

$$2.46) \quad \sum_{i=1}^n p_i N \left(\frac{\omega_i}{N_{s_i}} \right) = 1 ,$$

where $\{\omega_i\}$ is a sequence of "stress interaction" factors that depend upon all the other amplitudes and number of cycles present at each amplitude, as well as the order of appearance. At present there appears to be no physical basis for determining such factors other than

laboratory experiments for a great variety of input stress functions. Although, Freudenthal and Heller, correctly, try to introduce the fact that there is interaction by attempting to account for it, we feel that their approach still misses the heart of the matter. In our opinion, their approach is in one sense too naive by merely trying to modify the Palmgren-Miner hypothesis and in another sense too complex since their interaction factors have no physical basis and thus a new set of factors would have to be determined for every conceivable situation. This results in a lack of physical and engineering interest since few general statements can be made. But we must not overlook the fact that their approach is among the first attempts to construct a more realistic fatigue failure theory.

The approach that is, perhaps, more in line with what is occurring in the cycling of a material is due to Gatts (Gatts [1]). His approach is not only one of the first attempts to base fatigue failure on the more realistic stress-strain relations, accounting for changing material properties, but it is the first to use strength as the fundamental quantity and the stress-strain hysteresis loop as

the phenomenological basis.

Gatts postulates that the damage due to a given stress history should be measured by the reduction of the two strengths, namely the endurance limit and the failure stress. Furthermore, no damage occurs until the stress amplitude exceeds the endurance limit and failure will not occur until the applied stress at least equals the failure stress. He writes the functional relationship describing his postulates as

$$2.47) \quad \frac{dS_f(n)}{dn} = -k D(S(n), S_e(n)) ,$$

where n is the number of cycles, $S(n)$ is the applied stress, $S_e(n)$ is the endurance limit, $S_f(n)$ is the failure stress, $k > 0$, and

$$2.48) \quad D(S, S_e) = 0 \text{ for } S \leq S_e.$$

The equation 2.47) must be solved according to the boundary conditions

$$2.49) \quad \begin{cases} n = 0, & S_e = (S_e)_0 \\ n = N, & S_f(N) = S(N). \end{cases}$$

With this basis, Gatts has put forth a very general approach. In fact, his approach is so general that a number of further assumptions are required in order that specific results may be obtained. The important question is the nature of the essential and, always, unknown function D . For it is this function that analytically expresses the mechanism of failure. Gatts makes a phenomenological hypothesis upon the irreversible work put into the material as given by the shaded portion of the stress-strain hysteresis loop shown in the Figure 1.

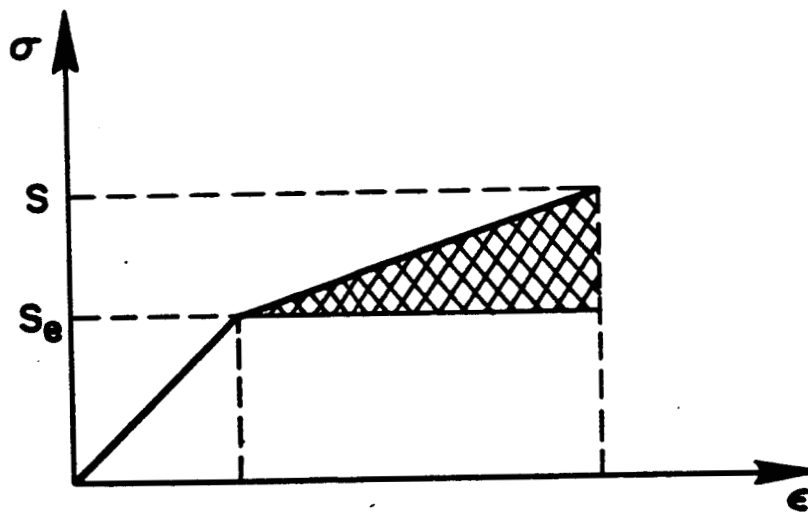


FIGURE 1
STRESS-STRAIN HYSTERESIS LOOP

Assuming the hysteresis loop to be made up of straight lines, the area of the shaded portion is proportional to $(S-S_e)^2$. Hence, 2.47) becomes

$$2.50) \quad \frac{dS_f(n)}{dn} = \begin{cases} -k' [S(n) - S_e(n)]^2 & S > S_e \\ 0 & \text{otherwise.} \end{cases}$$

Further hypothesis must be made since 2.50) includes two unknown functions S_e , S_f . Gatts then proposes the doubtful hypothesis that $S_e:S_f$ is constant, obtaining

$$2.51) \quad \frac{dS_e(n)}{dn} = \begin{cases} -K [S(n) - S_e(n)]^2 & , S > S_e \\ 0 & \text{otherwise.} \end{cases}$$

After analyzing the equation 2.51), he then assumes a random function for $S(n)$ (Gatts [2]). However, in this case it is not too clear what his results will be. Gatts formally takes averages and then integrates the equation 2.50). However, with his apparent assumptions of the independence of $S(n)$ and $S_e(n)$ at each instant the differential equation cannot be integrated formally. Indeed, in this

case it is best to revert back to the variable n as an integral variable (the number of cycles) and use a difference equation. Therefore, let us assume that at each cycle the amplitudes are identically distributed between zero and some maximum possible amplitude A , with distribution $P(s)$. Furthermore, we assume that the amplitudes are independently distributed for distinct cycles. We desire to obtain the expected change in the endurance limit during the $n + 1^{\text{rst}}$ cycle. The correct results, up to a point, are presented below. To proceed further one must make specific assumptions on $P(s)$, as we shall see.

The difference equation is

$$2.52) \quad S_e(n+1) = \begin{cases} S_e(n) - K [S(n+1) - S_e(n)]^2, & S(n+1) > S_e(n) \\ S_e(n) & S(n+1) \leq S_e(n) \end{cases} .$$

From elementary probability (Feller [1]) if events B, C satisfy $B \cap C = \emptyset$ and $\text{Prob } \{B\} + \text{Prob } \{C\} = 1$, it follows that for any random variable Y ,

$$2.53) \quad E \{Y\} = E \{Y|B\} \text{Prob } \{B\} + E \{Y|C\} \text{Prob } \{C\}.$$

Therefore, 2.52) yields

$$\begin{aligned} 2.54) \quad E \{S_e(n+1) | S_e(n)\} &= E \{S_e(n+1) | S_e(n), S(n+1) \\ &> S_e(n)\} \text{Prob } \{S(n+1) > S_e(n)\} + E \{S_e(n+1) | S_e(n), \\ &S(n+1) \leq S_e(n)\} \text{Prob } \{S(n+1) \leq S_e(n)\} \\ &= S_e(n) - K E \{[S(n+1) - S_e(n)]^2 | S_e(n), \\ &S_1(n+1) > S_e(n)\} \text{Prob } \{S(n+1) > S_e(n)\} \\ &= S_e(n) - K \int_{S_e(n)}^A [S - S_e(n)]^2 dP(s). \end{aligned}$$

Hence, finally

$$2.55) \quad E \{S_e(n+1)\} = E \{S_e(n)\} - K E \left\{ \int_{S_e(n)}^A [S - S_e(n)]^2 dP(s) \right\},$$

which is quite different from the results obtained by Gatts. As is easily seen, we cannot proceed further with 2.55) without a specific function $P(s)$. However, this in no way discredits Gatts basic fatigue theory. Indeed his basic approach is very interesting and certainly has merit. His attempt to base a cumulative damage theory upon the changing material properties is truly a step in the right direction. His theory is quite general and many hypotheses must be made. This is somewhat of a disadvantage since one would like to make as few assumptions on the nature of the damage mechanism as is possible. The major lack in his theory is that the inherent variability of fatigue damage has not been accounted for, as he states in his discussion (Gatts [2]). Furthermore, his assumption of constant $S_e:S_f$ is seriously attacked in the discussions of his paper.

We now conclude this brief summary of pertinent models that have been proposed. The various approaches, as we have seen, have possessed good points and bad points. This is probably true of most theories that one can construct, and is no less true for the theory to be presented in the following section. However, in the development to follow we feel we have the most realistic theory yet proposed. In addition, the theory is one for which relatively straightforward laboratory experiments can provide the required quantities incorporated in the analytical development.

CHAPTER III

A STOCHASTIC THEORY OF FAILURE

BASED UPON THE STRESS-STRAIN HYSTERESIS LOOP

A.) The General Theory of Failure

The general theory of failure of any system, whether it is a material specimen, a complex missile or communications system, an economic system, etc. may be given a rather straight forward mathematical formulation. We may let $X(t) \equiv [x_1(t), \dots, x_n(t)]$ represent the general state vector of the system as a function of the generalized "age" parameter t . The age parameter may represent the time or it may represent other measures such as the number of on-off operations, the amount of time above some predetermined value or the number of sinusoidal cycles at a given stress level, etc. We shall further let $\mathcal{J}_t[X]$ denote the failure functional. That is $\mathcal{J}_t[X]$ denotes the failure state of the system at age t , depending upon the entire history of the general state vector X over the entire age

interval, $[0, t]$. In the most general situation the functional $\mathcal{G}_t[X]$, as a function of the age parameter t , will perform a one-dimensional random walk in the $(\mathcal{G}_t[\cdot], t)$ - plane. We assume furthermore that there is a predetermined positive function $F(t)$ that defines the failure criterion. It is then said that the system, with general state vector X , fails at the first age for which the failure function $\mathcal{G}_t[x]$ passes outside the region bounded by $\pm F(t)$ having initially started within the region. This situation is demonstrated in Figure 2.

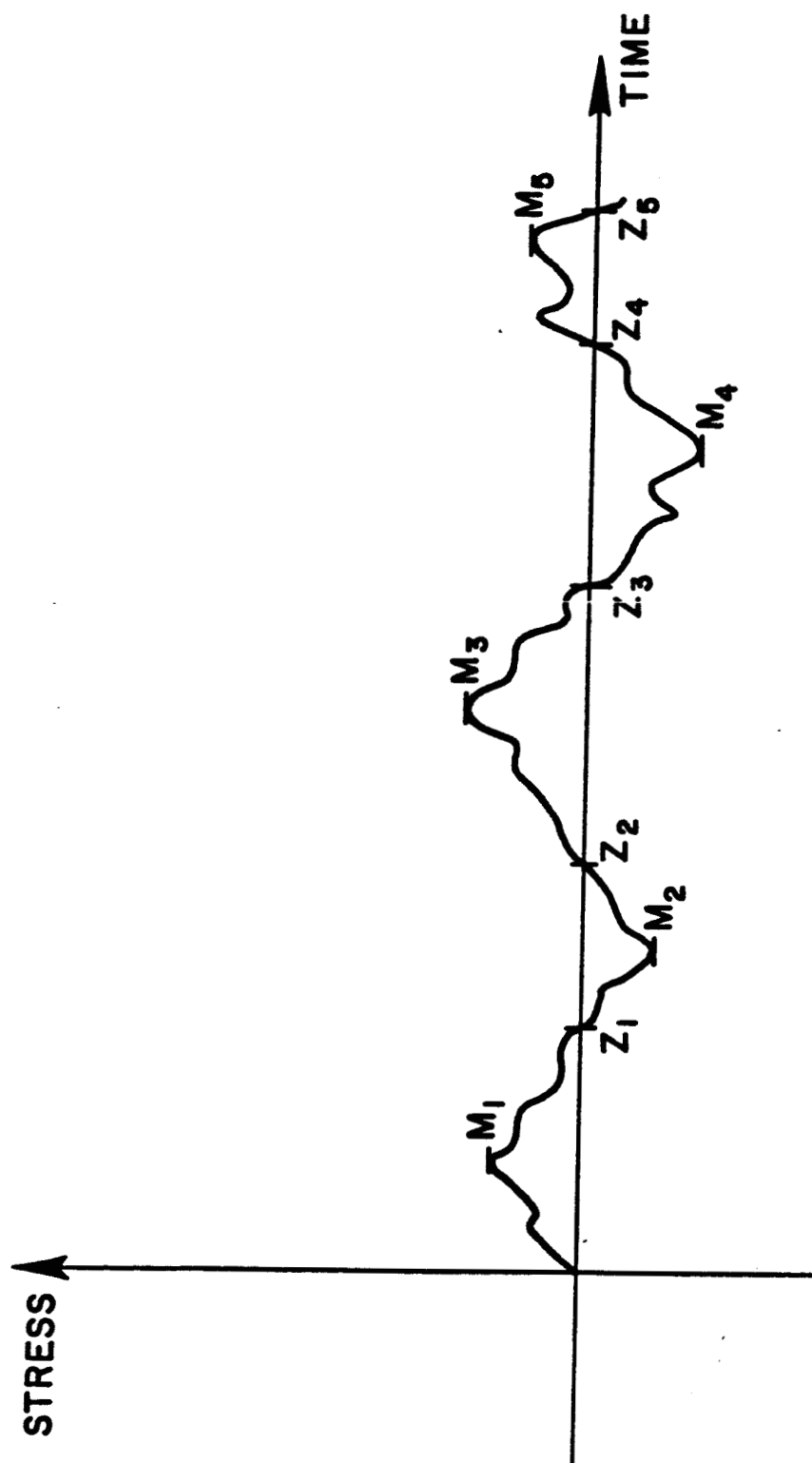


FIGURE 2
Failure Functional

As shown in the figure, the system fails at age t' . We may mention here that since $F(t)$ is a known curve, never zero, we may divide $\mathcal{F}_t[\cdot]$ by $F(t)$, and replace the curves $\pm F(t)$ in Figure I, by horizontal lines at ± 1 . In the following notation we shall assume this division has taken place.

In the most general practical setting, the system and its environment are not explicitly known and can at best be given as stochastic processes. It follows that the problem of greatest interest, whose solution would yield most of the desired failure information, is the determination of

$$3.1) \quad \text{Prob} \left\{ \sup_{0 \leq \tau \leq t} |\mathcal{F}_\tau[X]| < 1 \right\}.$$

The expression 3.1) represents the probability that the system is still operational at age t , never having failed prior to t . This problem is related to the first passage problem of stochastic processes, and is basic to any failure formulation, (Bogdanoff and Kozin [1]).

We wish to specialize the general formulation somewhat. In particular, let us consider the case that the failure of a system is due to an aging process. That is, instead of assuming that $\mathcal{F}_t[X]$ undergoes a general one dimensional random walk with increasing age, we assume that it is a monotone non-decreasing random function of the age t . Physically such an assumption represents an accumulation of non-negative random quantities with increasing age. In view of this situation, it follows that

$$3.2) \quad \text{Prob} \left\{ \sup_{0 \leq \tau \leq t} |\mathcal{F}_\tau[X]| < 1 \right\} = \text{Prob} \{ \mathcal{F}_t[X] < 1 \}.$$

simply because of the non-decreasing property of $\mathcal{F}_t[X]$. The complicated first passage problem has its simplest form in this case. Fortunately, the theory of cumulative damage just fits this situation.

The mathematical formulations given by 3.1) or 3.2) are well defined. However, an even more difficult physical problem remains. The explicit form of the failure functional $\mathcal{F}_t[X]$ depending upon the properties of the system remains unknown, especially in the case of material properties. Knowledge of this functional is tantamount to knowing the

underlying physical failure mechanism. If, as in the case of a large scale missile system, one requires only that the generalized positions and momentum variables be within some required range, then the failure functional can be derived from straight forward input-output relations for dynamical systems. Thus, the failure functional, although quite complex, can easily be defined. Of course, determining 3.1) even for this case is a very difficult problem that in general will require many hours of computer studies. On the other hand, failure due to the aging of physical properties is a problem of a much greater magnitude of complexity. Indeed there appears to be no means at present by which we can analytically express the physical aging due to the accumulation of damage quantities, using the observable states of a system. This happens to be true whether the system is mechanical, electrical or biological. In lieu of having an explicit failure functional we must seek a phenomenological means by which we may measure an aging effect through appropriately chosen, experimentally observable, physical quantities. At the same time we must account for our basic ignorance of the failure functional. But even this is quite difficult and has not been accomplished to any satisfactory degree heretofore.

The extreme value, the Palmgren-Miner and the Parzen approaches all neglect the fact that the material is undergoing internal changes, and the Gatts approach assumes a distinct failure functional. In our opinion, all of these approaches miss the heart of the problem in one way or another. Indeed, it is because the problem is so difficult that such an enormous amount of literature has been written on the subject.

The problem as we see it is two-fold. We must assume a damage variable that is a function of the age through a readily observed statistic and then have a method of deriving statistical quantities necessary to make predictions of the fatigue state of the material specimen. Heretofore, the observations have been concerned mainly with the numbers of cycles to failure at given stress levels, neglecting the strains that the material is undergoing. Only relatively recently has the idea of the strain that the material undergoes become seriously considered, especially due to the success of Wood's theory (W. A. Wood [1]). Even more recent is the consideration of the stress-strain hysteresis loop in fatigue studies (Feltner and Morrow [1], Gatts [1], Kawamoto and Koibuchi [1]). This latter concept we discuss further in the next paragraph.

B.) The Stress-Strain Hysteresis Loop

In the early work of Gough and Haigh performed in the 1920's (Gough [1], Haigh [1]), studies were made of the changes of mechanical properties occurring during fatigue tests in metals. The changes of the mechanical properties were observed through the energy dissipated in the specimen during the test. One method of observing the energy dissipated in the specimen is through studies of the stress-strain hysteresis loop as given in Figure 3.

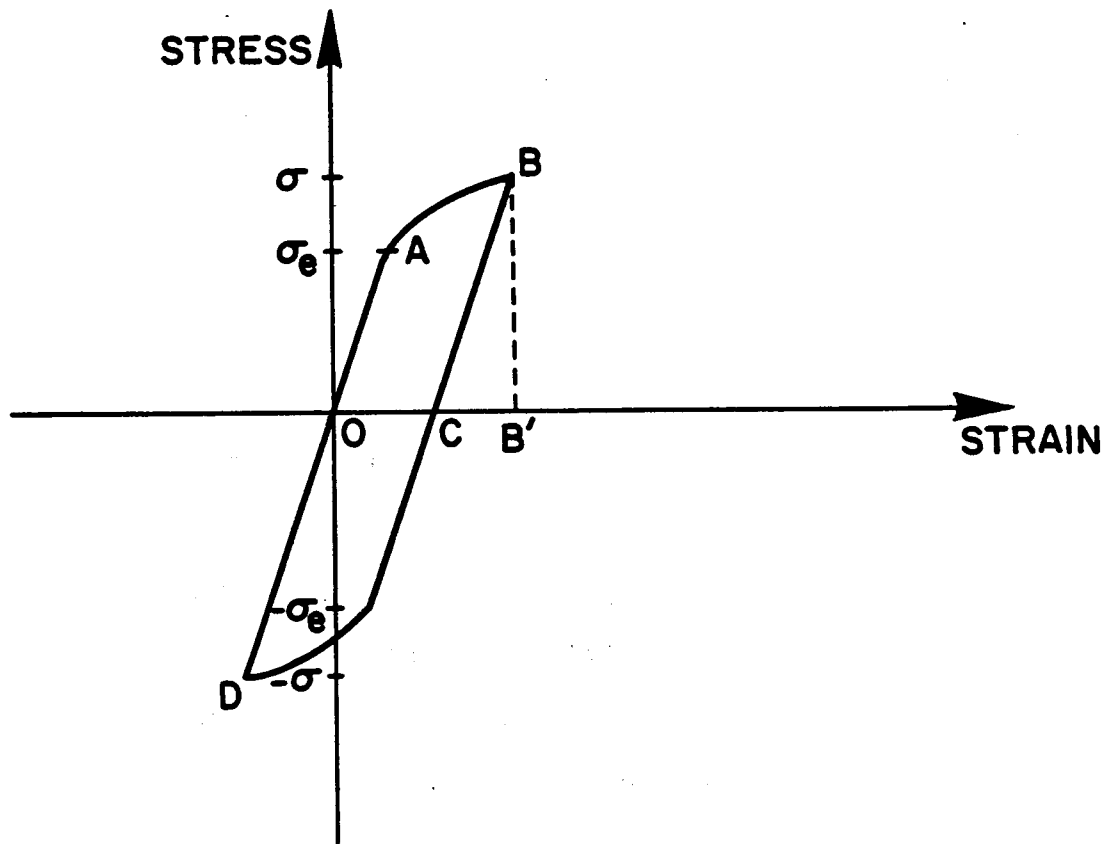


FIGURE 3
STRESS-STRAIN HYSTERESIS LOOP .

The kinematical process producing this loop is as follows: As the stress is increased from zero to the elastic limit σ_e , the material will traverse the curve along the linear portion from o to A. If the stress is reduced to zero before reaching σ_e , then the material will retrace its path along the line segment back to the origin.

However, if the stress is increased past σ_e to a level σ , then the material will continue along the plastic portion of the stress-strain curve AB. We notice that the total strain involved is the projection OB' on the strain axis. Upon decreasing the stress to zero, the material does not retrace the curve BAO, but travels back along the line BC which is parallel to OA. The segment OC is referred to as the plastic strain. If the stress is increased again from zero to σ after the material has reached state C, then the path of the material will be along the elastic line segment CB. If the stress is decreased from zero to $-\sigma$ and released back to zero, then the lower half of the loop CDO of Figure 3 will be traced to form the closed stress strain hysteresis loop whose area represents the irreversible work put into the material for the entire symmetric stress of range $(-\sigma, \sigma)$. The area of the half loop OABC represents the irreversible work put into the material for the stress

range $(0, \sigma)$. Assuming that the strain rates are not too high the half loop OABC shown in Figure 3 will be generated by each of the stress functions in Figure 4.

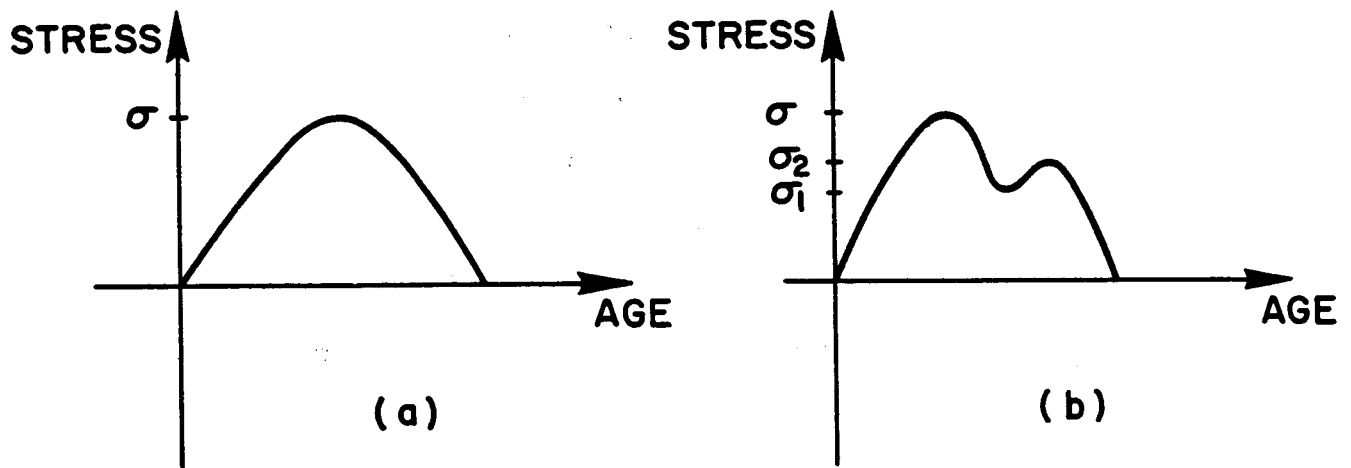


FIGURE 4
STRESS FUNCTIONS

We mention at this point that the property just described makes the hysteresis loop the most significant observable feature of the material when it is subjected to a random stress function. We also mention that since real materials do not possess an exact straight line OA as shown in Figure 2, the actual hysteresis loop generated by the stress function in Figure 4(b) is shown in Figure 5 (Kawamoto and Koibuchi [1]).

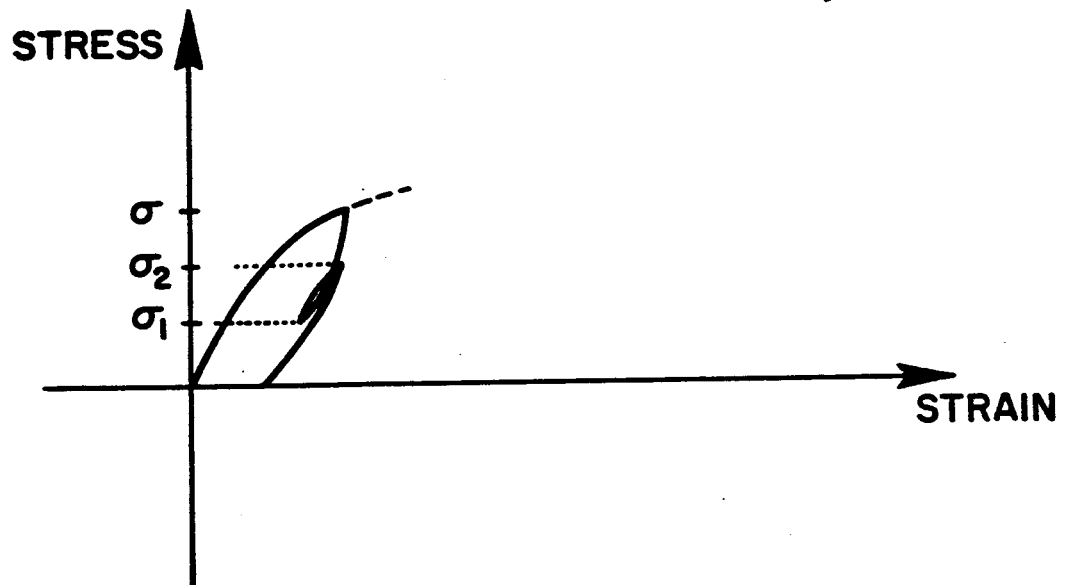


FIGURE 5
ACTUAL HYSTERESIS LOOP

However, the more nearly linear the elastic portion of the stress-strain curve is, the smaller is the area of the small interior loop, and will in general be neglected relative to the entire area of the half loop.

In studies of the aging process of a material during cyclic constant amplitude fatigue testing, beginning with the fundamental work of the 1920's, it has been well established that the stress-strain hysteresis loop undergoes three periods of evolution until final failure takes place. These periods are referred to as Stage I, Stage II and Stage III. They are characterized by the nature of the changes in the areas of the hysteresis loops. Stage I is mainly prevalent in softer materials and is characterized by a decrease in the loop areas for successive cycles. This is a hardening period and lasts for a relatively few number of cycles as compared to the entire fatigue life of the material. In the case of hardened materials this period may only last a few hundred cycles, and in general is less than 1000 cycles.

Stage II, on the other hand, occupies the major portion of the fatigue test and it is this stage that we are mainly concerned with in the present work. It is this stage that appears to be related to the formation of slip bands associated

with the damage process, and hence the cumulative damage concept as well. Stage II is characterized by a gradual increase in the hysteresis loop areas from the minimum area attained in Stage I. The rate of increase of the loop areas depend upon the magnitude of the stress level, and it is a monotone increasing function of the stress level. Furthermore, the areas remain constant for stresses near the endurance limit. (Kawamoto and Koibuchi [1], Feltner and Morrow [1], Thompson and Wadsworth [1]).

Stage III is of a much shorter duration, perhaps even shorter than Stage I. It is characterized by an extremely high growth rate of the hysteresis loop areas just prior to actual fracture of the specimen. Stage III is physically associated with the final stages of the propagation of surface fatigue cracks and is of no interest to us, since when cracks appear the specimen is usually considered to have failed.

It is an accepted fact that the work done on a specimen is converted into an energy that finally fractures the specimen. In fact, a number of researchers (e.g. Enomoto [1], Hanstock [1]) have advanced the postulate that the total amount of irreversible work per unit volume

required to cause fatigue failure is constant.

This is based upon the fact that not all of the irreversible work in a given cycle is converted into fatigue energy, but some unknown portion of it that will cause slip bands to form and internal bands to weaken. The other portion may be lost as heat to the surrounding medium. The fact that work is the essence of the problem was even stated by Miner in his now famous paper (Miner [1]). His basic error is that he considered the total work done on the material, rather than the work absorbed by the material. It is this point that led Miner to consider the stress amplitude as basic in the formation of his damage accumulation hypothesis.

If we postulate that the irreversible loop energy is the significant phenomenological observation, then the question becomes simply: "How do we incorporate it into a cumulative damage theory?" Gatts incorporated the irreversible work into his theory by assuming a specific damage relation in terms of the weakening of the ultimate failure strength and the endurance limit strength of the material. Hence, Gatts essentially assumed a specific failure functional D_t . As we have stated in the previous

section, $\mathcal{J}_t$ can never be explicitly given. Therefore, we shall postulate certain average properties to hold and let actual values be derived only through experimentally observed statistics.

The fundamental idea upon which we will base the theory to be developed in the following paragraphs is that the total area enclosed within the hysteresis loop, or half loop (Crandall, Mark and Khabbaz [1]), in the complex stress function case, is a measure of the damage done to the material. Furthermore, this measure is a random quantity whose distribution must remain unknown, but whose mean value may be determined through experiment. Larger areas will be presumed to be associated with larger damage quantities, as a result of the fact that failure appears to occur at a constant amount of irreversible work input associated with the plastic strain energy.

On the basis of relating damage to the area of the hysteresis loop, a reversible physical picture presents itself. Under pure sinusoidal cycling, the increase in area of the hysteresis loop during Stage II may be viewed as a weakening of the material to the same input as the damage is accumulated. In other words, successive cycles

of a fixed stress amplitude have greater and greater damaging effects. Furthermore, the fact that larger stress amplitudes should produce a greater damaging effect is reflected in the fact that the increase in the loop area occurs at a higher rate. Also, stress near the endurance limit possessing almost zero rates of change of hysteresis loop areas imply their much smaller damaging effect. One would expect that the Palmgren-Miner hypothesis would be most accurate in the endurance limit range of stresses. We shall see in Section E that the change of hysteresis loop areas explains the inaccuracies observed in applying the Palmgren-Miner hypothesis when subjecting a material to two different stress levels.

In the next section, we shall make the above statements precise.

C.) Postulates For A Cumulative Damage Theory

The entire theory to follow is set forth in the six postulates below concerning the damage variable D and the hysteresis loop area A .

Postulate I - Randomness

Damage, denoted by D , is a non-negative random variable.

Postulate II - Damage Function

The damage random variable, D , associated with an hysteresis loop (or half loop in the complex stress function case) of area A , is assumed to be a random function of A .

That is

$$D = D(A) \geq 0.$$

Postulate III - Independence

The random variables $D(A_1)$, $D(A_2)$, --- associated with successive loops (or half loops in the complex stress function case) of areas A_1 , A_2 , --- are assumed to be independent random variables.

Postulate IV - Failure Criterion

We assume that there is a value $D > 0$, depending only upon the material, such that when the total accumulated damage reaches this value, the material is said to have failed.

Postulate V - Proportionality Factors

The expected values $E \{D(A_1)\}$, $E \{D(A_2)\}$ of the damage for hysteresis loops of areas A_1 , A_2 shall satisfy the relation

$$3.3) \quad \frac{E \{D(A_1(\sigma_1))\}}{E \{D(A_2(\sigma_2))\}} = \left(\frac{A_1(\sigma_1)}{A_2(\sigma_2)} \right)^\alpha$$

where $\alpha > 0$ is the exponent in the ϵ_p -N diagram for the given material. That is, α is defined by

$$3.4) \quad N(\epsilon_p) = K \epsilon_p^{-\alpha},$$

where N is the average number of cycles to failure at constant plastic strain ϵ_p .

Postulate VI - Load Interaction

The effect of a load of stress level σ on any given cycle is to increase the hysteresis loop areas on all future load applications by an (experimentally determined) factor associated with σ and independent of all past or future stress levels. Furthermore, the effect of two loads of level σ on subsequent hysteresis loops is independent of whether the loads are on contiguous cycles or separated by cycles of other loads.*

We shall now present a short discussion describing the postulates set forth above.

Postulate I sets forth our intention to develop a stochastic theory of cumulative damage. This assumption reflects the universally accepted fact that there is an inherent variability in the fatigue phenomenon or, equivalently, it acknowledges our ignorance of the basic damaging mechanism and the fundamental failure functional $\mathcal{F}_t$.

$$A_n(\sigma_n) = A_1(\sigma_n) \prod_{i=1}^{n-1} \gamma(\sigma_i), \text{ where } \gamma(\sigma_i) \text{ is defined in}$$

3.17).

Postulate II is the basic departure from the majority of cumulative damage approaches, for it immediately assumes damage to be related to the fundamental phenomenon of irreversible work input into the material rather than the usual stress exceedence. Thus, we bring in the non-stationarity or aging process that the material actually undergoes during the evolution of fatigue, particularly regarding the Stage II development of the hysteresis loops.

Postulate III is difficult to condemn or defend on any observational basis. It essentially states that the amount of irreversible work that goes into damage on any given cycle is independent of the amount of irreversible work that goes into damage on any other cycle. This independence has been a part of all of the stochastic theories so far. Furthermore, it is required to make the renewal theory ideas go through in the analysis. Our feeling on this matter is that the age-dependence or non-stationarity is the most significant aspect here. Since we have taken age into account in our theory, we do not feel that independence is a serious assumption.

Postulate IV is our failure criterion and corresponds to the function $F(t)$ discussed in Section III A. This is a postulate used by many researchers. Furthermore, experimental evidence appears to lend credence to a fixed upper limit of work input or total damage as we have discussed above in Section III B.

Postulate V may at first appear quite specific in its form. However, there is a reasonable direct argument that leads us to this choice. The fact that larger loop areas should imply larger damage variables on the average, as we have seen, is the result of actual experimental investigations. However, the question is how to incorporate this fact into a theory. Our approach to answering this question is via the ϵ_p -N curve that is rapidly becoming as important as the S-N diagram in cumulative damage literature. Let us consider the case in which the stress amplitude is near the endurance limit so that the Stage II increase in areas of the hysteresis loops is so small that we may be allowed to neglect the change in area without significant error as has been observed (Kawamoto and Tanaka [1], Feltner and Morrow [1]). In that case since the hysteresis loop areas are constant, then the expected values of the damage are constant, so well as the plastic

strain ϵ_p , for successive cycles. It follows, for example, from Parzen's theory described in Chapter II that for a stress level σ near the endurance limit, we have

$$3.5) \quad \frac{1}{E \{N_D(\epsilon_p)\}} = \frac{1}{E \{N_D(\sigma)\}} \\ = E \left\{ \frac{D(A(\sigma))}{N} \right\} ,$$

where $A(\sigma)$ is the constant area of the N_D hysteresis loops required for failure at constant stress σ , and by our assumption, constant plastic strain ϵ_p .

Now we must reflect upon the fact that any empirical diagram such as the S-N diagram or the ϵ_p -N diagram is an average relation. That is, the curve represents the average number of cycles to failure for a given stress σ or a given plastic strain ϵ_p . Hence, physically, we must identify $N(\epsilon_p)$ in the ϵ_p -N diagram with $E \{N(\epsilon_p)\}$ in formula 3.5). Therefore, let us consider two stress levels σ_1, σ_2 that are close to the endurance limit so that the hysteresis loops do not vary. Therefore, $\epsilon_{p1}, \epsilon_{p2}$ also are assumed to remain constant. It follows from 3.4) and 3.5) that

$$3.6) \quad \frac{E \{D(A_1)\}}{E \{D(A_2)\}} = \frac{E \{N(\epsilon_{p2})\}}{E \{N(\epsilon_{p1})\}} = \frac{N(\epsilon_{p2})}{N(\epsilon_{p1})} = \left(\frac{\epsilon_{p1}}{\epsilon_{p2}} \right)^\alpha.$$

Consider a simplified hysteresis loop, in the shape of a parallelogram, as shown in the Figure 6.

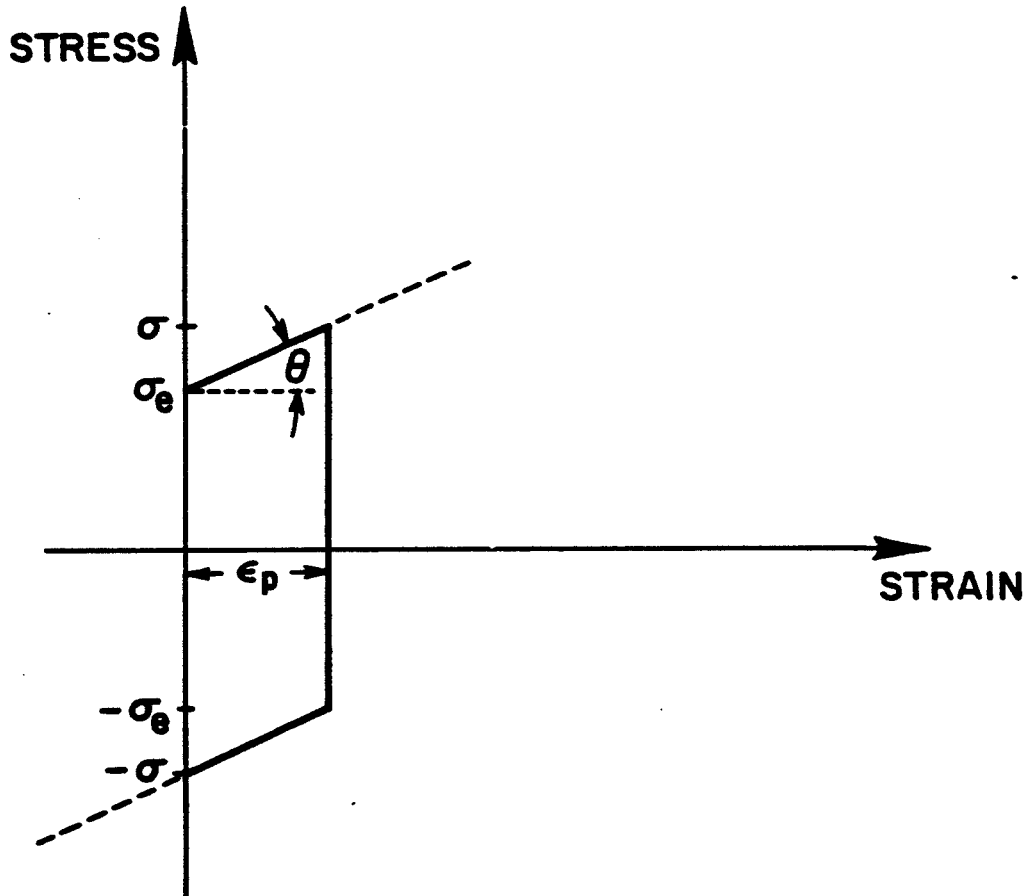


FIGURE 6
SIMPLIFIED HYSTERESIS LOOP

The area of the loop in Figure 6 is simply given by $\epsilon_p(\sigma + \sigma_e)$, where $\epsilon_p = (\sigma - \sigma_e) \cot \theta$.

Now in the case that σ_1, σ_2 are close to σ_e , it follows that the areas of the hysteresis loops satisfy

$$3.7) \quad \frac{A_1}{A_2} = \frac{\epsilon_{p1} (\sigma_1 + \sigma_e)}{\epsilon_{p2} (\sigma_2 + \sigma_e)} = \frac{(\sigma_1 - \sigma_e) (\sigma_1 + \sigma_e)}{(\sigma_2 - \sigma_e) (\sigma_2 + \sigma_e)} = \frac{\sigma_1 - \sigma_e}{\sigma_2 - \sigma_e} = \frac{\epsilon_{p1}}{\epsilon_{p2}},$$

to the first order in small quantities $\sigma_1 - \sigma_e, \sigma_2 - \sigma_e$.

Hence, we have the approximate equality

$$3.8) \quad \frac{E \{D(A_1)\}}{E \{D(A_2)\}} = \left(\frac{A_1}{A_2} \right)^\alpha,$$

which we postulate as an equality throughout the range of loading of the material.

Postulate VI is somewhat of a different nature than the other postulates, in that we are here assuming a particular type of aging to take place. We are motivated by the observed fact that the presence of interspersed stresses of levels greater than the remaining stress levels cause a large increase on the subsequent accumulation of damage even

at lower stress levels. Their effect upon the future must be accounted for by assuming a proportionately greater amount of expected damage for all future cycles at any stress level. Furthermore, since stresses of different amplitudes cause distinctly different internal structural changes to a material, the change in areas of the hysteresis loop must be dependent upon the past. This last point also motivates our assumption that two cycles of the same stress have the same effect on subsequent cycles independent by other cycles. We must also point out that whereas Postulate III assumes independence of the damage variables themselves, Postulate VI assumes dependence of their average values, through the sequence of stresses that occur, as we shall see in the next section.

Upon applying the six postulates above to a combination of physical and statistical arguments we shall develop a stochastic theory of cumulative damage. We shall at first consider the simple sinusoidal input of constant amplitude, then the input consisting of two blocks of sinusoidal cycles at distinct constant amplitudes. We will finally study the random situation. In every case we shall be interested in the average damage at any given age. The age parameter

will always be generated by the number of cycles of sinusoidal stress inputs or by the number of axis crossings for complex random loading functions. Time will be brought into the picture only through the above age measures.

D.) Pure Sinusoidal Input

Let a material specimen be subjected to a pure sinusoidal stress function of load amplitude σ . The areas of the successive hysteresis loops generated by this stress function will be denoted by $A_1(\sigma)$, $A_2(\sigma)$, ..., and the associated random damage quantities will be denoted by $D(A_1(\sigma))$, $D(A_2(\sigma))$, Hence, the damage after N cycles is given by the sum of the independent random variables

$$3.9) \quad \sum_{i=1}^N D(A_i(\sigma)).$$

The number of cycles to failure at stress level σ is a random variable which we denote by $N(\sigma)$, and by our failure criterion, Postulate IV, it is defined as the smallest integer M for which

$$3.10) \quad \sum_{i=1}^M D(A_i(\sigma)) > \mathcal{D}.$$

It therefore follows that, at failure,

$$3.11) \quad \sum_{i=1}^{N_{\mathcal{D}}(\sigma)} D(A_i(\sigma)) = \mathcal{D} + \delta,$$

where δ is a non-negative random residual.

We may take the expected values of each side of 3.11) to yield, recalling $\mathcal{D}$ is a constant depending upon the material,

$$3.12) \quad E \left\{ \sum_{i=1}^{N_{\mathcal{D}}(\sigma)} D(A_i(\sigma)) \right\} = \mathcal{D} + E\{\delta\}.$$

Recalling that $N_{\mathcal{D}}(\sigma)$ is a random variable, it follows from Postulate III and Postulate V that

$$3.13) \quad E \left\{ \sum_{i=1}^{N_{\mathcal{D}}(\sigma)} D(A_i(\sigma)) \right\} = E \{ D(A_1(\sigma)) \}$$

$$E \left\{ \sum_{i=1}^{N_{\mathcal{D}}(\sigma)} \left(\frac{A_i(\sigma)}{A_1(\sigma)} \right)^\alpha \right\}.$$

However, the proof of 3.13) required the use of Martingale Theory of stochastic processes. We shall present the proof in the appendix.

Upon applying the result of 3.13) to 3.12) and dividing by the non-negative quantity $\mathcal{D}$, we obtain

$$3.14) \quad E \left\{ \frac{D(A_1(\sigma))}{\mathcal{D}} \right\} = E \left\{ \sum_{i=1}^{N_{\mathcal{D}}(\sigma)} \left(\frac{A_i(\sigma)}{A_1(\sigma)} \right)^{\alpha} \right\} = 1 + \frac{E\{\delta\}}{\mathcal{D}} .$$

We must now study the quantity $\frac{E\{\delta\}}{\mathcal{D}}$ on the right hand side of equation 3.14), as $\mathcal{D}$ becomes large. As a general mathematical question, this is beyond the scope of the present work. It is theoretically difficult and has been answered in only a relatively few number of cases in renewal theory (Smith [1]). Fortunately for us we have a rather well defined physical situation at hand that we may fall back upon. The fundamental physical feature of the effect of cyclic loading in producing fatigue failure is that the total damage accumulated consists of a large number of small contributions, one contribution from each cycle. Hence, no one cycle is the predominant cause of failure, and the damage accumulated on each cycle is a very small percentage of the overall damage produced at fatigue failure. Therefore, although we

cannot state rigorously that $\frac{E\{\delta\}}{D}$ approaches zero almost surely as D approaches infinity, we certainly can say that $\frac{E\{\delta\}}{D} \ll 1$, being of the order of tenths or even hundredths of a per cent. This is unquestionably outside the range of engineering significance, hence we may neglect this quantity to yield the final formula,

$$3.15) \quad E \left\{ \frac{D(A_1(\sigma))}{D} \right\} = E \left\{ \frac{1}{\sum_{i=1}^N \left(\frac{A_i(\sigma)}{A_1(\sigma)} \right)^\alpha} \right\}.$$

Equation 3.15) is the basic equation of the entire theory. It is through this equation that we may empirically determine the average damage per cycle in the laboratory by observing the areas of the hysteresis loops, and recording the number of cycles to failure for a collection of specimens. Then statistical averaging may be applied to produce the quantity on the left hand side of the equation 3.15). We wish to stress the fact that we have not assumed any functional forms for the actual damage quantities themselves or their distributions other than Postulate V, which relates average damage to hysteresis loop areas. But as we have already shown Postulate V itself is the result of actual empirical observations.

It is important to mention at this time that many investigators (Feltner and Morrow [1], Thompson and Wadsworth [1], Duce [1], Enomoto [1], Kawamoto and Koibuchi [1]) have found that simple exponential formulas fit the empirical data concerning the nature of the change in areas of the hysteresis loop for successive cycles at a constant stress amplitude. Indeed for Stage II, the formula

$$3.16) \quad A_{i+1}(\sigma) = \gamma(\sigma)A_i(\sigma),$$

where $\gamma(\sigma) > 1$ and is an increasing function of σ appears to fit the data quite well prior to Stage III, when a sharp upward turn in the rate of change of areas takes place for the few cycles before actual fracture. Furthermore, for $\sigma = \sigma_e$, it follows that $\gamma(\sigma) = 1$ for practical purposes.

On the basis of formula 3.16), the formula 3.15) becomes

$$3.17) \quad E \left\{ \frac{D(A_1(\sigma))}{D} \right\} E \left\{ \frac{1 - [\gamma^a(\sigma)]^{N_D(\sigma)}}{1 - \gamma^a(\sigma)} \right\} = 1,$$

where again $\gamma(\sigma)$ and $N_D(\sigma)$ are experimentally observable quantities.

On the basis of our theory and formula 3.17), it follows that after N cycles at stress level σ , the expected value of the accumulated damage $E \left\{ \frac{D_N(\sigma)}{\mathcal{D}} \right\}$ is given as

$$3.18) \quad E \left\{ \frac{D_N(\sigma)}{\mathcal{D}} \right\} = E \left\{ \frac{D(A_1(\sigma))}{\mathcal{D}} \right\} \left[\frac{1 - [\gamma^a(\sigma)]^N}{1 - \gamma^a(\sigma)} \right].$$

As we have already discussed in section A, the failure criterion function, in our case the constant $\mathcal{D}$, can always be absorbed so that the failure criterion becomes in our case the first exceedence of unity. Hence in the following sections we shall drop the explicit use of the limit damage $\mathcal{D}$, and let D represent the per cent of damage to failure.

E.) Two Amplitude Level Sinusoidal Input

The major motivation for our approach to the problem of cumulative damage has been to try to develop a rational means of accounting for the damage, in the presence of complex load histories, on the material, as well as take into account the aging of the material. It is a well established fact that the commonly used Palmgren-Miner hypothesis does not fit a rather consistent set of observations. An exper-

iment that shows the hypothesis to be inaccurate is to run a sample of specimens for n_1 cycles at stress level σ_1 for which the average life to failure would be N_1 cycles, and then run the specimens to failure for n_2 cycles, at stress level σ_2 for which the average life to failure is N_2 cycles. The Palmgren-Miner hypothesis should yield $\frac{n_1}{N_1} + \frac{n_2}{N_2}$ approximately equal to unity. However, it is invariably found that the value this quantity attains depends upon the relative magnitudes of σ_1 and σ_2 .

In particular, for un-notched specimens,

$$3.19 \quad \left\{ \begin{array}{ll} \text{a) } \frac{n_1}{N_1} + \frac{n_2}{N_2} > 1 & \text{for } \sigma_1 < \sigma_2 \\ \text{b) } \frac{n_1}{N_1} + \frac{n_2}{N_2} < 1 & \text{for } \sigma_1 > \sigma_2 . \end{array} \right.$$

We shall show in this section that the theory presented in this report, in fact, predicts this non-commutativity of the application of cycles of two different stress levels. We shall demonstrate this by example. In general, actual numerical values would have to be obtained.

In order to consider the problem we must first determine the so-called interaction effects of the presence of different stress levels. Therefore, let us consider the expected value of damage accumulated after N cycles of various stress amplitudes $\sigma_1, \sigma_2, \dots, \sigma_n$. We shall denote this damage (in fact, per cent of failure) by $E \{D(N)\}$. In order to determine this we must determine the expected value of damage which occurs on the $n + 1^{\text{rst}}$ cycle of stress level σ_{n+1} . By Postulate VI, using the empirically determined area factors defined in 3.16), we have the damage on the $n + 1^{\text{rst}}$ cycle given as,

$$3.20) \quad \gamma^a(\sigma_1) \dots \gamma^a(\sigma_n) E \{D(A_1(\sigma_{n+1}))\},$$

where $\sigma_1, \dots, \sigma_n$ are the first n stress levels to have been applied. They may or may not be distinct. The quantity $E \{\gamma(A_1(\sigma_{n+1}))\}$ is the empirically determined quantity defined by our basic equation 3.15) for the stress level σ_{n+1} .

Thus, we must have from our theory and 3.20)

$$3.21) \quad E \{D_N\} = \sum_{i=1}^N \left([\gamma^a(\sigma_1) \dots \gamma^a(\sigma_{i-1})] E \{D(A_1(\sigma_i))\} \right).$$

Let us consider now two stress levels σ , $\bar{\sigma}$, where $\bar{\sigma} = \sigma_e$ and $\sigma > \bar{\sigma}$. Thus, we have $\gamma(\bar{\sigma}) = 1$, $\gamma(\sigma) > 1$.

We shall first apply $\bar{n}$ cycles of stress level $\bar{\sigma}$ to the specimen, and then apply cycles of stress level σ until failure. We shall denote the number of cycles to failure as predicted by our theory by n and the number of cycles as predicted by the Palmgren-Miner theory by n' .

By the Palmgren-Miner Theory we must have

$$3.22) \quad \frac{\bar{n}}{\bar{N}} + \frac{n'}{N} = 1,$$

where $\bar{N}$, N are the average number of cycles to failure for stress levels $\bar{\sigma}$, σ respectively, that is

$$3.23) \quad \bar{N} = E\{N_D(\bar{\sigma})\}, \quad N = E\{N_D(\sigma)\}.$$

However, by our theory, the expected damage accumulated on the first $\bar{n}$ cycles is given by 3.21) as

$$3.24) \quad E\{D(A_1(\bar{\sigma}))\} \left\{ 1 + \gamma^a(\bar{\sigma}) + \frac{\bar{n}}{N} + [\gamma^a(\bar{\sigma})]^{\bar{n}-1} \right\} \\ = \bar{n} E\{D(A_1(\bar{\sigma}))\} = \frac{\bar{n}}{\bar{N}}$$

since the hysteresis loop areas remain constant and therefore yield the same expected value of damage on each cycle.

It follows from 3.21) and 3.23) that the number of cycles, n , to failure at stress level σ is determined from

$$3.25) \quad \frac{\bar{n}}{\bar{N}} + [\gamma^a(\sigma)]^{\bar{n}} \{1 + \gamma^a(\sigma) + \dots + [\gamma^a(\sigma)]^{n-1}\} E\{D(A_1(\sigma))\} = 1.$$

But, since $\gamma(\sigma)$ is unity and $E\{D(A_1(\sigma))\}$ is given by our basic formula 3.15), then 3.24 becomes

$$3.26) \quad \frac{\bar{n}}{\bar{N}} + \frac{1 + \gamma^a(\sigma) + \dots + [\gamma^a(\sigma)]^{n-1}}{E \left\{ \sum_{i=1}^{N_D(\sigma)} [\gamma^a(\sigma)]^{i-1} \right\}} =$$

$$\frac{\bar{n}}{\bar{N}} + \frac{[\gamma^a(\sigma)]^n - 1}{E \{ [\gamma^a(\sigma)]^{N_D(\sigma)} \} - 1} = 1.$$

But from 3.22) and 3.25) it follows that the predicted n' from the Palmgren-Miner hypothesis satisfies

$$3.27) \quad \frac{n'}{N} = \frac{[\gamma^a(\sigma)]^n - 1}{E\{[\gamma^a(\sigma)]^{N_{\mathcal{D}}(\sigma)}\} - 1}.$$

We may now apply Jensen's inequality that states that if $f(y)$ is a continuous convex function, then

$$3.28) \quad f(E\{y\}) \leq E\{f(y)\},$$

for any random variable Y .

Therefore, we have from 3.28), since γ^{aN} is convex for $\gamma > 1$,

$$3.29) \quad \frac{1}{E\{[\gamma^a(\sigma)]^{N_{\mathcal{D}}(\sigma)}\} - 1} \leq \frac{1}{[\gamma^a(\sigma)]^{E\{N_{\mathcal{D}}(\sigma)\}} - 1} = \frac{1}{[\gamma^a(\sigma)]^N - 1}.$$

Hence, from equations 3.27) and 3.29) we have

$$3.30) \quad \frac{n'}{N} \leq \frac{[\gamma^a(\sigma)]^n - 1}{[\gamma^a(\sigma)]^N - 1} = \frac{1 + \gamma^a(\sigma) + \dots + [\gamma^a(\sigma)]^{n-1}}{1 + \gamma^a(\sigma) + \dots + [\gamma^a(\sigma)]^{N-1}} \leq \frac{n}{N}$$

for $\gamma^a(\sigma) > 1$.

The last inequality obtains since setting $\gamma^a(\sigma) = 1 + \epsilon$,
 $\epsilon > 0$,

$$3.31) \quad \frac{1 - (1 + \epsilon)^n}{1 - (1 + \epsilon)^N} = \frac{n\epsilon + \frac{n(n-1)}{2!} \epsilon^2 + \dots}{N\epsilon + \frac{N(N-1)}{2!} \epsilon^2 + \dots} =$$

$$\frac{n}{N} \cdot \frac{1 + \frac{n-1}{2!} \epsilon + \frac{(n-1)(n-2)}{3!} \epsilon^2 + \dots}{1 + \frac{N-1}{2!} \epsilon + \frac{(N-1)(N-2)}{3!} \epsilon^2 + \dots} < \frac{n}{N},$$

for $n < N$.

But, the inequality 3.30) along with 3.22) imply

$$3.32) \quad \frac{\bar{n}}{\bar{N}} + \frac{n}{N} > 1,$$

which agrees with the observed physical phenomenon.

If we reverse the order to first apply $\bar{n}$ cycles of stress σ , then cycle the material specimen to failure at the stress level $\bar{\sigma}$, we are led, by the same considerations as above, to the inequality

$$3.33) \quad 1 < \frac{\bar{n}}{N} + [\gamma^{\alpha}(\sigma)]^{\bar{n}} \frac{n-1}{\bar{N}} .$$

We cannot go any further with the inequality 3.33) without specific numbers, especially for $\gamma(\alpha)$. However it is certainly possible that 3.33) holds and yet we have

$$3.34) \quad \frac{\bar{n}}{N} + \frac{n}{\bar{N}} < 1.$$

We shall now look at the random input case.

F.) Random Stress Functions

We now have sufficient material with which to present the case of the random stress input. In order to illustrate the ideas involved, we shall first treat the situation of pure sinusoidal cycles of random amplitudes, without considering the time factor involved. Therefore, we shall assume that the material is subjected to a random sequence of sinusoidal cycles of amplitudes $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ which are independent identically distributed continuous random variables with common distribution $P(\sigma)$.

As given by our postulates and through equation 3.20) we find that the expected damage accumulated on the n^{th} cycles is given, recalling that the σ_i 's are random variables which are independently distributed, by

$$\begin{aligned}
 3.35) \quad & E\left\{\gamma^{\alpha}(\sigma_1) \cdots \gamma^{\alpha}(\sigma_{n-1}) E(D(A_1(\sigma_n)) | \sigma_1, \dots, \sigma_n)\right\} \\
 & = E(D(A_1(\sigma_n))) \prod_{i=1}^{n-1} E(\gamma^{\alpha}(\sigma_i)).
 \end{aligned}$$

Because the random variables involved are assumed to be identically distributed, it follows that

$$3.36) \quad \begin{cases} \text{a) } E(D(A_1(\sigma))) = \int E(D(A_1(\sigma)) | \sigma) dP(\sigma) \equiv \bar{d} \\ \text{b) } E(\gamma^{\alpha}(\sigma)) = \int \gamma^{\alpha}(\sigma) dP(\sigma) \equiv \gamma \end{cases}$$

Hence it follows that the expected value of the damage after N cycles is simply given by

$$3.37) \quad d + \gamma d + \gamma^2 d + \cdots + \gamma^{N-1} d = d \frac{1 - \gamma^N}{1 - \gamma}$$

It is most interesting to note that the independence of the stresses allows the expected damage to be independent of the order in which the stresses have occurred. Indeed the formula reflects that the average damage accumulates as though it is due to some equivalent stress possessing an initial damage d , with a rate of increase given by γ . What occurs in the general random stress input as a function of time is much more complicated than the example above as we shall now see.

Let us consider a random process with a sample function as shown in Figure 7.

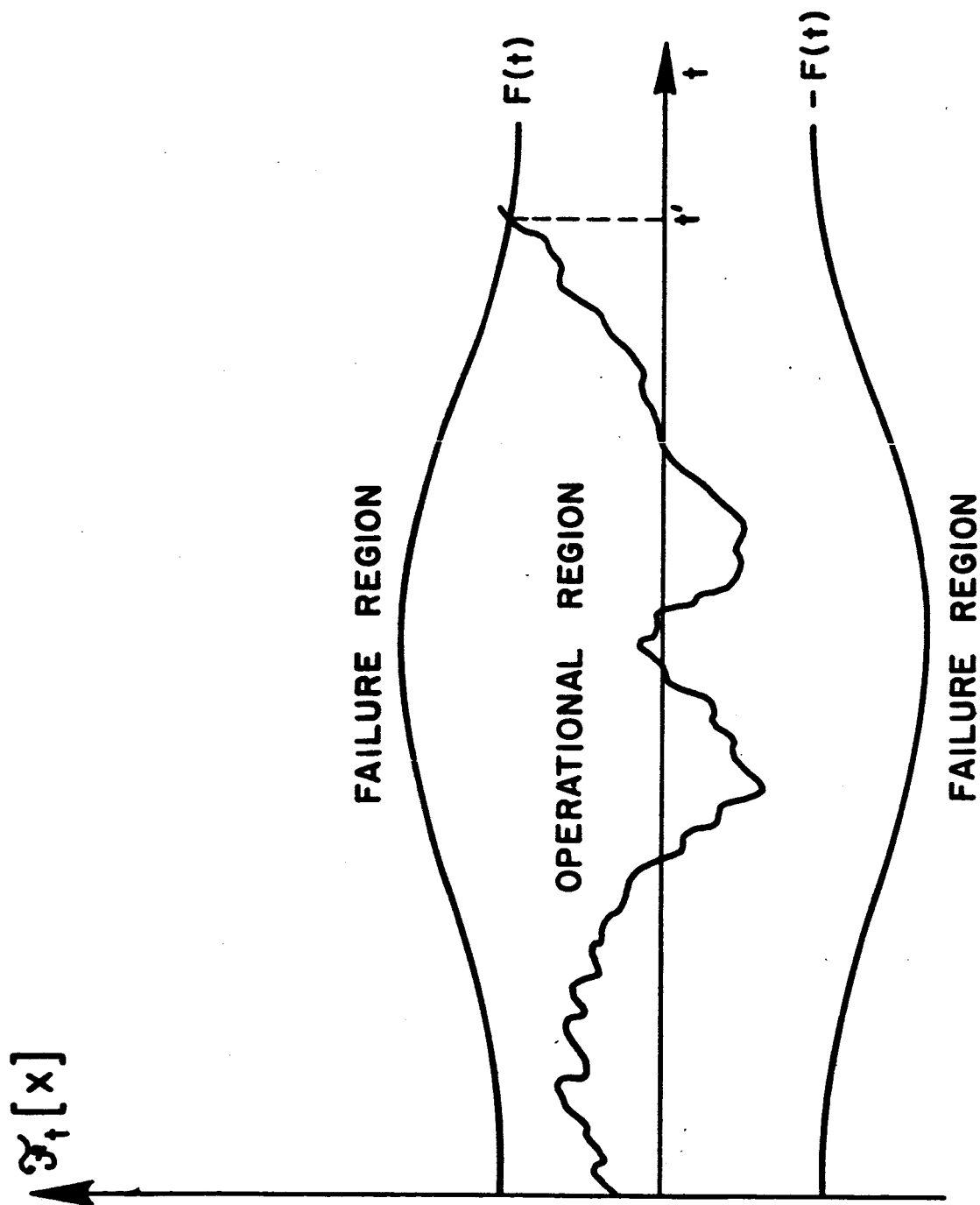


FIGURE 7
A SAMPLE STRESS FUNCTION

As we have discussed in Section B, if the strain rates are not too great, the only significant statistics of the sample stress function in Figure 7 are the zero crossings $z_1, z_2, \dots$, and the absolute extremes between zero crossings $M_1, M_2, \dots$. These are the points that will determine the hysteresis half loops, for which we shall associate half the average damage as accumulated for the full loop at the same stress level.

Hence, it follows that at any fixed time t , there will be a random number $N(t)$ of axis crossings or half loops which have formed. Furthermore, the areas of these half loops will depend upon the sequence of absolute extremes that have occurred up to time t . Let us make these statements precise.

We shall assume the areas of the successive hysteresis half loops for a given time t , to be denoted by $A_1, A_2, \dots, A_{N(t)}$, with the expected total damage given as

$$3.38) \quad E \left\{ E\{D(A_1) | N(t)\} \right\} + \dots + E \left\{ E\{D(A_{N(t)}) | N(t)\} \right\} .$$

However, the areas involved depend upon the previous amplitude values $M_1, M_2, \dots$, that are present. We can represent this dependence explicitly as

$$3.39) \quad E \left\{ E(D(A_1) | M_1, N(t)) \right\} + E \left\{ E(D(A_2) | M_1, M_2, N(t)) \right\} \\ + \dots + E \left\{ E(D(A_{N(t)}) | M_1, M_2, \dots, M_{N(t)}, N(t)) \right\} .$$

From our above analysis, it follows that 3.39) may be expressed as

$$3.40) \quad E \left\{ E(D(A_1(M_1)) | M_1, N(t)) \right\} + E \left\{ \gamma^\alpha(M_1) E(D(A_1(M_2)) | M_1, M_2, N(t)) \right\} \\ + \dots + E \left\{ \gamma^\alpha(M_1) \dots \gamma^\alpha(M_{N(t)-1}) E(D(A_1(M_{N(t)})) | \right. \\ \left. M_1, \dots, M_{N(t)}, N(t)) \right\} .$$

If we represent the joint density function for the extreme values $M_1, \dots, M_i$ and the number of axis crossings $N(t)$ at time t , by $P_t(\sigma_1, \dots, \sigma_i; N)$, it follows that the i^{th} average may be written as

$$3.41) \quad E \left\{ \gamma^\alpha(M_1) \cdots \gamma^\alpha(M_{i-1}) E\{D(A_1(M_i)) | M_1, \dots, M_i, N(t)\} \right.$$

$$= \int \cdots \int_{i\text{-fold}} \gamma^\alpha(\sigma_1) \cdots \gamma^\alpha(\sigma_{i-1}) E\{D(A_1(\sigma_i)) |$$

$$\sigma_1, \dots, \sigma_i, N(t)\} dP_t(\sigma_1, \dots, \sigma_i; N).$$

We add that the terms

$$3.42) \quad E\{D(A_1(\sigma_i)) | \sigma_1, \dots, \sigma_i, N(t)\} \equiv E\{D(A_1(\sigma_i)) | \sigma_i\}$$

are the quantities determined from actual experimental observations through formula 3.15). The new analytical complications have been brought forth through the introduction of the joint density function P_t . We cannot

continue any further at this level of generality because of this complexity. Indeed, we begin to lose insight to the original problem. We may say, however, that it is the type of information required to evaluate the expectations 3.41) that we need to carry on further for arbitrary random inputs, especially for actual environmental situations.

It is of interest to note that in the work of Nauman and Leybold this information is being studied.

We shall carry on our discussion in this section with an example of a type of random stress function that can be constructed in the laboratory and for which the analysis may be completely carried out, but for which we still have a good deal of generality in the distributions involved. We assume that the absolute extremes between axis crossings are independent, identically distributed continuous random variables, and also that the intervals between axis crossings are independent and identically distributed with distribution function $F(t)$. However, we should not underestimate the generality of this example, for if the noise is Gaussian with a short correlation time, or if the process is strongly Markov, so that at a zero crossing the future is independent of the past, the above independence assumption is not far from the truth.

We denote the total damage accumulated at time T by $D(T)$. Then our problem is to determine $E\{D(T)\}$, which we may write from the assumption of independence and from our above considerations as,

$$\begin{aligned}
 3.43) \quad E\{D(T)\} &= E \left\{ E \left\{ \sum_{i=1}^{N(T)} \gamma^{\alpha}(M_1) \cdots \gamma^{\alpha}(M_{i-1}) E\{D(A_1(M_1)) | M_1\} \mid N(T) \right\} \right\} \\
 &= E \left\{ \sum_{i=1}^{N(T)} (E\{\gamma^{\alpha}(M)\})^{i-1} \right\} E\{D(A_1(M))\}.
 \end{aligned}$$

Using the same notation as given in 3.36), we finally obtain from 3.43),

$$3.44) \quad E\{D(T)\} = E \left\{ \frac{1 - \gamma^{N(T)}}{1 - \gamma} \right\} d.$$

Again, because of the fact that the stress strain hysteresis loop depends only upon the absolute extremes obtained, and not on the other features of the curve as long as the strain rate is not too high, we obtain the same formula as for the sinusoidal case except for the presence of the expectation operation. This expectation has been brought in due to the fact that the time has now entered into the picture. In

order to complete the explicit analysis of the problem, we must compute the expectation $E\{\gamma^{N(T)}\}$, where $N(T)$ is the number of axis crossing or intervals to time T . But, this is a straight forward renewal equation problem, as we now show. The lengths of the intervals between axis crossings are independent identically distributed non-negative random variables, and thus we are interested in the random number $N(T)$ required to just surpass T . We analyse this in the following fashion.

Let τ_1 be the length of the first interval between the first two successive zeros, then

$$\begin{aligned} 3.45) \quad G(T) &= E\{\gamma^{N(T)}\} = E\left\{E\{\gamma^{N(T)} \mid \tau_1 = \tau\}\right\} \\ &= \int_0^\infty E\{\gamma^{N(T)} \mid \tau_1 = \tau\} dF(\tau). \end{aligned}$$

However, by the usual argument (Parzen [2]), we obtain

$$3.46) \quad E\{\gamma^{N(T)} \mid \tau_1 = \tau\} = \begin{cases} 1 & \text{if } \tau > T \\ \gamma E\{\gamma^{N(T-\tau)}\} & \text{if } \tau < T. \end{cases}$$

Thus, using 3.46) in equation 3.45) yields the convolution integral equation

$$\begin{aligned} 3.47) \quad G(T) &= \int_T^{\infty} dF(\tau) + \gamma \int_0^T G(T-\tau) dF(\tau), \\ &= 1 - F(T) + \gamma \int_0^T G(T-\tau) dF(\tau). \end{aligned}$$

The integral equation 3.47) may be easily solved by LaPlace transform techniques, since $F(\tau)$ is presumed known. Explicit cases may easily be solved.

We shall terminate the discussion of the random input at this time. Other explicit cases of practical interest await further computer studies in order to obtain the necessary statistics. However, once this is accomplished the random environment problem can be attacked through the approach put forth in this report.

APPENDIX TO CHAPTER III

Martingales and The Proof of Equality 3.13)

The proof of equality 3.13) does not follow from elementary considerations because of the fact that $N_D(\sigma)$ is a random variable conditioned by the previous damage magnitudes. If $N_D(\sigma)$ was independent of these magnitudes, the equality would follow from the most elementary considerations of probability theory. As this is not the case, we require other means to establish the formula. Fortunately for us, the tools required can be found in the theory of martingales (Doob [1]). The treatise of Doob will be our reference, particularly Chapter VII, theorem 2.1 and Section 10.

A discrete martingale is defined to be a stochastic process $\{x_n, n \geq 1\}$ for which $E\{|x_n|\} < \infty, n \geq 1$ and

$$\begin{aligned} \text{A.1)} \quad x_n &= E\{x_{n+1} | x_1, \dots, x_n\} \text{ with probability one for} \\ &\text{any } n \geq 1. \end{aligned}$$

We note that from A.1) it follows that

$$A.2) \quad E\{x_{n+1}\} = E\{x_n\} = E\{x_1\}.$$

If the equality sign in A.1 is replaced by " $\leq$ ", then the process is referred to as a semi-martingale.

Let us consider the case that the process, martingale or semi-martingale, is stopped at some optional stopping time depending upon observations of the previous values of the process. What is the nature of the new process? To be precise, suppose $m(\omega)$ represents the random stopping time, let us consider the new process $\{x_n^u, n \geq 1\}$, defined by

$$A.3) \quad x_j^u(\omega) = \begin{cases} x_j(\omega) & , \quad j \leq m(\omega) \\ \frac{x(\omega)}{m(\omega)} & , \quad j > m(\omega) \end{cases}$$

where $m(\omega) = \mu$ is a condition only on the past values of the variables up to the μ^{th} stage.

This is exactly the situation we find ourselves with in cumulative damage. For, since total accumulated damage, which is the sum of non-negative independent random quantities, is terminated when the total damage first surpasses $\mathcal{D}$, we see that we have a stopped process. This random stopping time or age, as we choose to call it, is dependent only upon the previous values of the accumulated random damage quantities.

The result we require is a part of Theorem 2.1 in Doob's treatise, given as follows

Theorem - Suppose that the martingale $\{x_n, n \geq 1\}$ is transformed into the process $\left\{ \overset{u}{x}_n, n \geq 1 \right\}$ under optional stopping. Then the $\overset{u}{x}_n$ - process is a martingale and, furthermore,

$$A.4) \quad E \left\{ \overset{u}{x}_n \right\} = E\{x_1\}, \quad n \geq 1.$$

For the proof, we refer to Doob's treatise.

Now let us consider the mutually independent non-negative random variables $D_1, D_2, \dots$ each satisfying

$$A.5) \quad E\{D_i\} = a_i, \quad E\{D_1\} > 0,$$

where a_i are known non-negative constants.

We let $N_D(\omega)$ be the first value m for which

$$A.6) \quad D' = D_1 + D_2 + \dots + D_m > D.$$

We wish to determine $E(D')$ in terms of the expected values $E\{D_i\}$, $i \leq N_D(\omega)$.

We form the new sequence of random variables $D(n)$ defined as

$$A.7) \quad D(n) = \sum_{j=1}^n (D_j - E\{D_j\}) = \sum_{j=1}^n D_j - E\{D_1\} \sum_{j=1}^n a_j.$$

It can be easily shown that if $Y_1, Y_2, \dots$ are mutually independent random variables, then the necessary and sufficient condition for the X_n - process defined by

$$A.8) \quad X_n = \sum_{i=1}^n Y_i ,$$

to be a martingale, is that $E\{Y_i\} = 0$ for all i .

But, this is the case for $D(n)$, since

$$A.9) \quad E\{D_j - E\{D_j\}\} = E\{D_j\} - E\{D_j\} = 0.$$

Therefore the $D(n)$ - process is a martingale.

Hence, by the theorem, the stopped process (stopped by the condition A.6), $\left\{ \overset{u}{D}(n), n \geq 1 \right\}$, is a martingale satisfying

$$A.10) \quad E\left\{ \overset{u}{D}(n) \right\} = E\{D(1)\} = 0 \text{ for } n \geq 1.$$

$$\text{Hence, } \left\{ E \overset{u}{D}(N_D) \right\} = 0.$$

Thus,

$$\begin{aligned} A.11) \quad E\left\{ \overset{u}{D}(N_D) \right\} &= E\left\{ \sum_{j=1}^{N_D} D_j - E\{D_1\} \sum_{j=1}^{N_D} a_j \right\} \\ &= E\left\{ \sum_{j=1}^{N_D} D_j \right\} - E\{D_1\} E\left\{ \sum_{j=1}^{N_D} a_j \right\} = 0. \end{aligned}$$

Therefore, we obtain

$$\text{A.12)} \quad E \left\{ \sum_{j=1}^{N_D} D_j \right\} = E(D_1) \cdot E \left\{ \sum_{j=1}^{N_D} a_j \right\} ,$$

as was to be proved.

CHAPTER IV

DISCUSSION AND RECOMMENDATIONS

We have presented in this Report the first step in developing a stochastic theory of cumulative damage based on the irreversible work put into the material.

The distinctions between our theory and others proposed may easily be established. In our theory, damage is a non-stationary random function of the irreversible work input for which the average damage per cycle may be empirically established. Parzen's theory is a stationary random function of the stress exceedences for which the damage per cycle may be established. The Freudenthal-Heller theory is a modified Palmgren-Miner deterministic theory in which non-stationarity is introduced empirically through observed interaction effects. We, on the other hand, introduce a mechanism for determining the interaction effects. Gatts' theory is a non-stationary deterministic theory based upon the hysteresis loops area for which the damage (or reduction in strength) is assumed to be

a given function. Note that in our theory we never assume the value of damage, but only allow its average to be determined experimentally.

The basic feature of the entire theory proposed here is that all quantities may be physically observed. That is, the hysteresis loop areas, the number of cycles to failure, and the rates of changes for the hysteresis loop areas at different stress levels.

One definite weakness is that we cannot predict the variance of the damage since there does not appear to be any natural method to relate the ratios of variances to the associated hysteresis loop areas in the same fashion as we have been able to for the means. This is one problem that must be looked at in the future development of this approach. This will depend upon uncovering a natural physical property associated with second moments, in the same way that the S-N, ϵ_p -N curves are associated with first moments.

In order to illustrate the utility of the above theory in predicting fatigue life, let us consider a simple deterministic example. Let there be $\bar{n}$ applications of stress amplitude $\bar{\sigma}$ followed by n applications of stress amplitude $\sigma (< \bar{\sigma})$. Equations 3.17) and 3.18) yield the result

$$\frac{\{1 - [\gamma^a(\bar{\sigma})]^{\bar{n}}\}}{E\{1 - [\gamma^a(\bar{\sigma})]^{N_{\bar{\sigma}}(\bar{\sigma})}\}} + \frac{[\gamma^a(\bar{\sigma})]^{\bar{n}} \{1 - [\gamma^a(\sigma)]^n\}}{E\{1 - [\gamma^a(\sigma)]^{N_{\sigma}(\sigma)}\}} = 1$$

Graphically this formula represents a curve of the type illustrated in Figure 8. The coordinates of a point on this curve represent

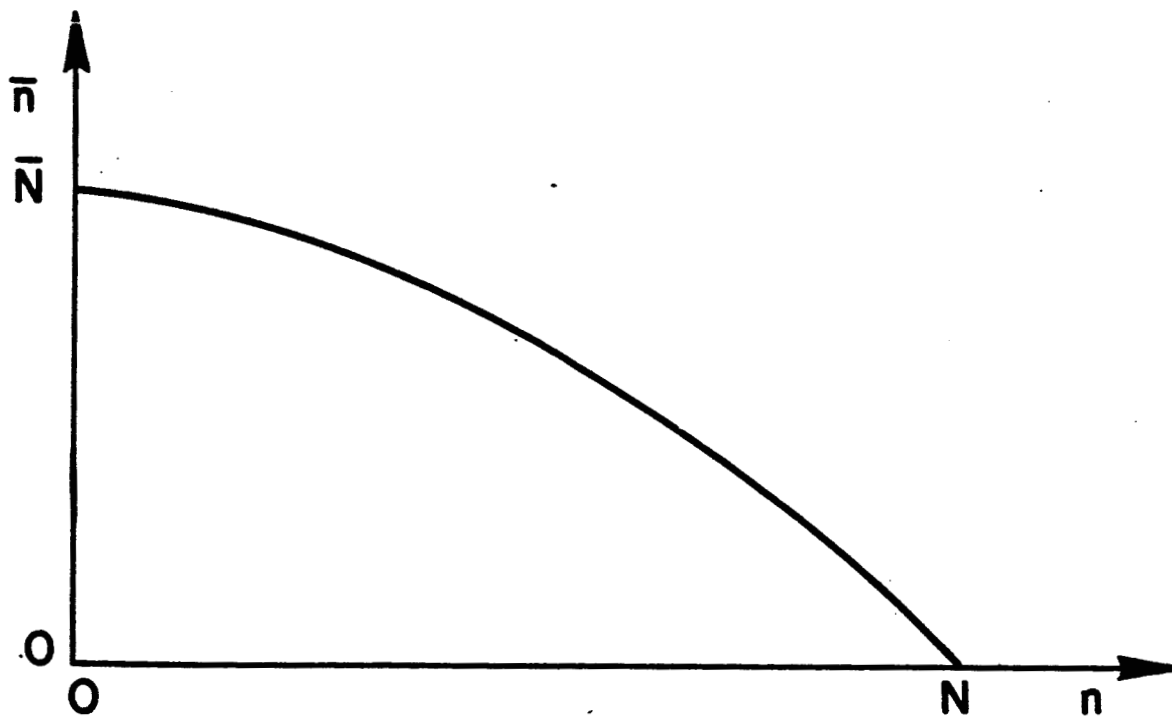


FIGURE 8
FAILURE LOCUS FOR A PARTICULAR SEQUENCE
OF TWO STRESS LEVELS

the number of applications of stress at level σ (abscissa) and the number of applications at level $\bar{\sigma}$ (ordinate) required to produce failure.

For other sequences of these two levels of stress application or with more than two levels of stress application, Equations 3.17) and 3.18) may be manipulated to yield a similar--but more complex--result.

In this first stage of the development of the present approach, we have presented the foundation and certain analytical consequences. There now remains the task of testing the theory in the laboratory to determine its effectiveness. However, since the study of fatigue in terms of hysteresis loops is relatively recent, there is not sufficient information to allow conclusions to be made immediately. The major programs of experimentation that must be accomplished as we see it are (a) The determination of the factors of increase of hysteresis loop areas as a function of the stress levels for pure sinusoidal cycling, (b) The determination of the average damage accumulated through our Equation 3.15) and (c) The study of the nature of the hysteresis loop area changes under varying inputs to determine the change in the loop geometry and loop areas due to interactions and relating these to the present theory.

We wish to conclude this section by stating that there are certain weaknesses in our theory, as one must expect in any phenomenological approach. However, it is our sincere conviction that this theory does constitute a step in the right direction. We hope that it can, in time, be developed into a practical and useful approach. It cannot be overstated that the next step is the laboratory.

BIBLIOGRAPHY

- Bogdanoff, J. L. and Kozin, F. [1], Comment on the Reliability of Aircraft Structures in Resisting Chance Failure; Ops. Res. Vol 9, No. 1, Jan. 1961, p. 123.
- Crandall, S. H.; Mark, W. D. and Khabbaz, G. R. [1], The Variance in Palmgren-Miner Damage Due to Random Loading; Proc. 4th U. S. Nat'l. Cong. of App. Mech., Vol. I, (1962), P. 119.
- Doob, J. L. [1], Stochastic Processes; J. Wiley & Sons, New York (1953).
- Duce, A. G. [1], A Study of Some Fatigue Phenomena in Pure Metals and Alloys; Ph.D. Thesis, Univ. of Cambridge, England (1950).
- Enomoto, N. [1], On Fatigue Stress Under Progressive Stress; Proceedings ASTM, Vol. 55 (1955) P. 903.
- Epstein, B. [1], Statistical Aspects of Fracture Problems; Jour. of App. Phys., Vol. 19, Feb. 1948, P. 140.
- Feller, W. [1], An Introduction to Probability Theory and Its Applications; 2nd Edition, Wiley & Sons, New York (1957).
- Feltner, C. E. and Morrow, J. D. [1], Microplastic Strain Hysteresis Energy as a Criterion for Fatigue Failure, Jour. of Basic Eng. Trans. ASME, Sec. D, Vol. 83, No. 1, March 1961, P. 15.
- Freudenthal, A. M. [1], Fatigue of Structural Metals Under Random Loading; Symp. on Acous. Fatigue, ASTM Tech. Pub. No. 284, (1960) p. 26.
- Freudenthal, A. M. [2], The Statistical Aspect of Fatigue of Materials; Proc. Roy. Soc. (London) Ser. A, Vol. 187 (1946) p. 416.
- Freudenthal, A. M. and Gumbel, E. J. [1], On the Statistical Interpretation of Fatigue Tests; Proc. Roy. Soc. (London) Ser. A, Vol. 216, (1953) p. 309.

- Freudenthal, A. M. and Heller, R. A. [1], On Stress Interaction in Fatigue and a Cumulative Damage Rule; Jour. Aero/space Science, Vol. 26, No. 7, July 1959, p. 451.
- Gatts, R. R. [1], Application of a Cumulative Damage Concept to Fatigue; Jour. of Basic Eng., Trans. ASME, Sec. D, Vol. 83, No. 4, Dec. 1961, p. 529.
- Gatts, R. R. [2], Cumulative Fatigue Damage with Random Loading, Jour. of Basic Eng., Trans. ASME, Sec. D, Vol. 84, No. 3, Sept. 1962, p. 403.
- Gnedenko, B. V. [1], Sur la distribution limite du terme maximum d'une série aléatoire; Ann. of Math., Vol. 44, 1943, p. 423.
- Gough, H. J. [1], Elastic Limits of Copper under Cyclic Stress Variation; Engineering, Vol. 114 (1922) p. 291.
- Gumbel, E. J. [1], Statistical Theory of Breaking Strength and Fatigue Failure; Bull. Int. Stat. Inst., Vol. 38, Part III (1961) p. 375.
- Gumbel, E. J. [2], Statistics of Extremes; Columbia Univ. Press, New York (1958).
- Haigh, B. P. [1], Energy Theory of Mechanical Fatigue and Hysteresis; Rep. to Aero. Res. Comm. British Assn. for Adv. of Sci. (London), (1923) p. 358.
- Hanstock, R. F. [1], Damping Capacity, Strain Hardening and Fatigue; Proceedings, Physical Society, Vol. 59, (1941) p. 275.
- Kawamoto, M. and Koibuchi, K. [1], Evaluation of the Fatigue Lines Under Actual Loads by Hereditary Method (in Japanese).
- Kawamoto, M. and Tanaka, T. [1], Elastic Hysteresis Energy and Fatigue Strength; Proc. 10th Japan Nat'l Cong. for App. Mech. (1960) p. 151.
- Leybold, H. A. and Naumann, E. C. [1], A Study of Fatigue Life Under Random Loading; Proc. of ASTM, Vol. 63, (1963) p. 717.

- Miles, J. W. [1], On Structural Fatigue Under Random Loading; Jour. Aero/space Sciences, Vol. 21, No. 11, Nov. 1954, p. 753.
- Miner, M. A. [1], Cumulative Damage in Fatigue; Jour. of App. Mech., Trans ASME, Vol. 12, No. 3 (1945) p. 159.
- Murphy R. B. [1], Reliability; Proc. First Symp. on Engineering Applications of Random Function Theory and Prob., Ed. by Bogdanoff, J. L. and Kozin, F.; J. Wiley & Sons, New York (1963).
- Palmgren, A. [1], Die Lebensdauer von Kugellagern; Zeit. des Vereines Inc., Vol. 58 (1924) p. 339.
- Parzen, E. [1], On Models for the Probability of Fatigue Failure of a Structure; App. Math. & Stat. Lab., Stanford Univ., Tech. Report No. 45, April 17 (1959).
- Parzen, E. [2], Stochastic Processes; Holden-Day, San Francisco (1962).
- Rice, S. O. [1], Mathematical Analysis of Noise; Bell Sys. Tech. Jour., Vol. 13 (1944) p. 282; Vol. 14 (1945) p. 46.
- Smith, W. L. [1], Renewal Theory and Its Ramifications; Jour. Roy. Stat. Soc., Vol. 20, No. 2 (1958) p. 243.
- Thompson, N. and Wadsworth, N. J. [1], Metal Fatigue; Advances in Physics, Vol. 7, No. 25 (1958) p. 72.
- Wood, W. A. [1], Failure of Metals Under Cyclic Strain; Proc. Int. Conf. on Fatigue of Metals, Inst. Mech. Eng. (London), (1956) p. 531.