4. *P—S* Curves

If we examine the general properties of the distribution function $F(S)$ at a constant N, illustrated by the diagrams in Fig. 4, field 2, we shall immediately find that this function is a positive, non-decreasing function where

$$F(U) = 0 \text{ for } U \geq 0 \text{ and } F(\infty) = 1.$$

Without loss of generality $F(S)$ may be written in the form

$$F(S) = 1 - e^{-\varphi(S)} \dots \dots \dots \dots \dots \dots (19)$$

as for every given distribution function F, the corresponding function can be uniquely determined by

$$\varphi = -\log(1-F) \dots \dots \dots \dots \dots \dots (20)$$

Evidently $\varphi(S)$ is a positive, non-decreasing function of S, satisfying the conditions

$$\varphi(S) = 0 \text{ for } S \leq U \geq 0 \dots \dots \dots \dots \dots (21)$$

and

$$\varphi(S) \to \infty \text{ for } S \to \infty \dots \dots \dots \dots \dots (22)$$

The condition (21) is a very important one, as it states that for values less than U the probability of failure is not only very small but exactly equal to zero. There is no doubt that this statement corresponds to a real physical property of all existing engineering materials.

The advantage of writing the expression of the distribution function in the form (19) is to be found in the fact that

$$[1 - F(S)]^n = e^{-n\varphi(S)} \dots \dots \dots \dots \dots (23)$$

as will be shown in the next chapter.

Experience has shown that the expression

$$\varphi(S) = k(S - U)^m \dots \dots \dots (24)$$

which satisfies (21) and (22), in many cases gives an excellent reproduction of the observations.

It contains three parameters, which have to be fitted to given data.

Introducing (24) in (19) and considering (1) we have

$$P = 1 - e^{-k(S-U)^m} \dots \dots \dots (25)$$

which by (14) gives

$$\log \log \frac{1}{1-P} = \log \log \frac{n+1}{n+1-\nu} = \log k + m \log (S - U) \ . \ (26)$$

If we put

$$y = \log \log \frac{1}{1-P}; \ a = \log k; \ x = \log (S - U) \ \dots \ (27)$$

we have, in fact, a linear relation

$$y = a + mx \dots \dots \dots (28)$$

which is formally the same expression as (18) though, of course, with a quite different sense.

The conformity of (18) and (28) may be extended to the computation of their parameters.

In this case, too, it is advisable to start with a graphical examination of the observations, in order to check the statistical homogenity of the material. After this check-up, the determination of U and the two other parameters is carried through in the same way as has already been demonstrated.

It should be observed that the values x and y in (28) are obtained by an interpolation, as they are the intersections between the smoothed $S - N$ curves and a vertical line $N = $ constant.

Many authors have been strongly in favour of the idea that the distribution of failures in solids must necessarily be a normal one, and this because the final fracture of a specimen is the total effect of a large number of independent causes. I do not find this to

be a very convincing argument. Who knows, a priori, if the decisive causes are independent or even if they are plural?

The arguments are sometimes of a more sentimental character, frankly stating that the normal distribution function is »nature's own distribution». This standpoint seems, in any cases, to have been abandoned by the statisticians. Years ago, the famous scientist Pearson asked: »Is it not possible to find material which obeys the normal law?» He answered himself: »Yes, but this law is not a universal law of nature. We must hunt for cases.» But, nevertheless, in the discussions on the strength of material, the obsolete idea that the normal distribution is the only correct one, turns up from time to time, e. g. FRENKEL and KONTOROWA (8) et. al.

As it is the opinion of the present Author that it is not of any use even to hunt for normal distributions of failures in solids, because they do not exist at all, it might be necessary to discuss the question at some length.

The normal distribution function is defined by the relation

$$\Phi(S) = \frac{1}{a\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(S-S_m)^2}{a^2}} \cdot dS \dots \dots (29)$$

It is easy to see that this expression does not satisfy the condition (21) as $\Phi(0) \neq 0$, which means that there is still a finite probability of fracture at the load 0. I think that, if you believe in this statement, it is to exaggerate your pessimism.

Then you will be told that the probability $\Phi(0)$ is certainly not zero but its value is insignificantly small. The objection to this view is that this difference, even if small, is of fundamental consequences as to the factor of safety.

If $U > 0$ — and without doubt such materials do exist — it is possible to define the factor of safety in an objective way. If, on the contrary, $U = 0$ or $U = -\infty$ as in (29), the only way out of this dilemma is to select subjectively a certain probability of fracture, which has to be tolerated. In the previously cited book by the Staff of the Battelle Memorial Institute (5) you will find this: »Engineering design of roller bearings is based on the stress which would permit the endurance of a given number of cycles with only 10 percent failed bearings, as shown in laboratory tests af actual bearings.»

This is the situation illustrated by Fig. 4, if the line $P = 0$ had

been omitted in field 1, which makes a considerable reduction of the information available through the complete diagram.

Now, the expression (29) may, of course, easily be modified in order to obtain a U-value, for instance, by substituting for S the function $\log (S - U)$ or introducing a term $- (S - U)^{-m}$ or the like. But there is an unlimited number of substitutions possible and which of them is »the nature's own»?

As a final objection against the normal distribution it might be mentioned that if the distribution were exactly normal for a specimen of a certain length, then the distribution cannot be normal for any other length. This statement will be proved in the next chapter. At the same time it will be seen that the normal distribution is very badly suited for the manipulations that arise in connection with a change of the dimensions of the specimen.

5. Size Effect

The original source of the statistical theory of strength is to be found in the Author's effort to explain the fact, known for a long time, that the ultimate strength of a specimen increases with reduced dimensions and that it is greater in bending than in tension. As far as he knows it was previously unknown that these two phenomena were intimately connected and could be explained by the same principle based on statistical considerations. This idea will now be briefly recapitulated.

Suppose that a specimen of the length L_1 and its distributions function F_1 are given for an arbitrary life N lying anywhere between 1 and ∞. Suppose further, that two such, nominally equal, specimens are coupled end to end, thus forming a specimen with the length $2 L_1$ but with an unchanged cross-section. The probability that failure will not occur in one of the two halves is obviously equal to $1 - F_1 (S)$ and the probability that failure will not occur either in one or in the other half-part is equal to the product of the two independent events. If $F_2 (S)$ denotes the distribution function of a failure in the specimen with the length $2 L_1$ it is easy to see that

$$1 - F_2 (S) = [1 - F_1 (S)]^2 \ldots \ldots \ldots \ldots \ldots (30)$$

or

$$F_2 (S) = 1 - [1 - F_1 (S)]^2 \ldots \ldots \ldots \ldots \ldots (31)$$
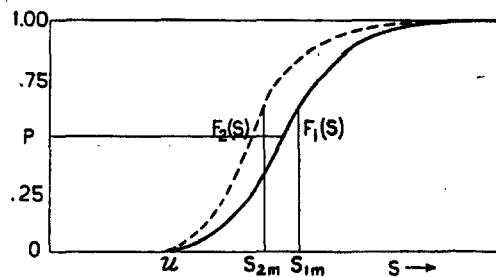
Fig. 18. $P - S$ curves for two different lengths.

For a specimen with the arbitrary length L its distribution function F_L is determined by

$$F_L(S) = 1 - [1 - F_1(S)]^{\frac{L}{L_1}} \dots \dots \dots \dots (32)$$

This is evidently an application of the weakest link principle and it may be found in a previous work by the Author (1).

If F_1 is given, graphically or numerically, it is extremely easy to construct F_2 as shown in Fig. 18.

It should be well observed that the two functions in (32) vanish for the same value of S, independently of the form of the distribution function, i. e. the change of the length does not change the value of U.

Furthermore, the value of the median as well as that of the arithmethic mean will be decreased when the length is increased. The experimentally determined ultimate strength is obtained by taking the arithmetic mean $S_m = \dfrac{\Sigma S_\nu}{n}$ of the n values, and will accordingly have a lower value if the length is increased. The only exception to this rule occurs if there is no dispersion of S, i. e. if the distribution function is a step function with a step of the height 1 in the point $S = S_m$. This is exactly a tacit assumption of the classical theory of strength.

The deduction of (32) postulates no knowledge of the properties of the material or its homogeneity within the cross section; only that the material is statistically homogeneous as to the length direction. For this reason (32) may be considered as having a mathematical rather than a physical character.

TABLE 3. *Rotating-beam endurance tests at two different effective lengths of the specimens* WEIBULL (9).

ν	Number of millions of cycles N at an effective length of		ν	Number of millions of cycles N at an effective length of	
	50 mm.	25 mm.		50 mm.	25 mm.
1	.50	.61	14	2.62	> 10.00
2	.68	.80	15	2.76	> 10.00
3	1.07	1.31	16	3.62	> 10.00
4	1.13	1.41	17	8.63	> 10.00
5	1.14	1.66	18	9.38	> 10.00
6	1.72	2.07	19	9.99	> 10.00
7	1.76	2.65	20	> 10.00	> 10.00
8	1.98	2.88	21	> 10.00	> 10.00
9	1.99	4.32	22	> 10.00	> 10.00
10	2.11	4.33	23	> 10.00	> 10.00
11	2.21	> 10.00	24	> 10.00	> 10.00
12	2.32	> 10.00	25	> 10.00	> 10.00
13	2.57	> 10.00	26	> 10.00	> 10.00

We shall apply the result to an investigation performed at AB Svenska Kullagerfabriken, Gothenburg, and published by the AUTHOR (9) some time ago. In a rotating-beam endurance test, 52 specimens were tested, half of them with an effective length of 25 mm and the other half with a length of 50 mm. The test pieces were made from a first-class ball bearing steel; the heat treatment and machining were very carefully done. Thus, the specimens may be regarded as unusually homogeneous ones. The reversed bending stress was $\pm$ 31.0 kg/mm² for all specimens. Observed lives N are given in Table 3, from which the influence of the length of the specimen is easily seen.

When the results are plotted as probability against number of load cycles — corresponding to field 3 of Fig. 4 — the curves look like those in Fig. 19. If there had been no influence of the length, as postulated by the classical theory, the observations would have fallen on approximately the same curve instead of on two separate ones.

The most perspicuous way to verify (30) seems to be the following: We have by (1), (4), and (30)
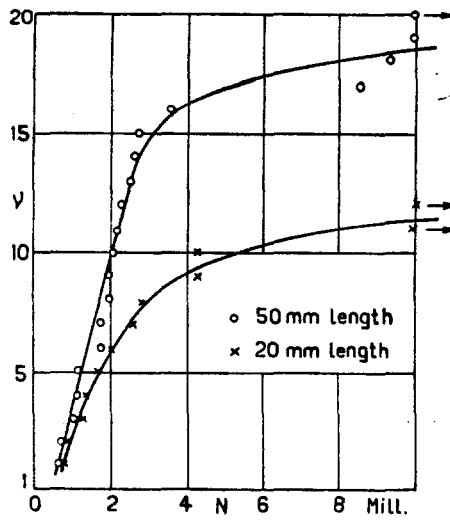
3

Fig. 19. Rotating-beam test on steel specimens of two different lengths.

$$1 - \frac{\nu_2}{n+1} = \left(1 - \frac{\nu_1}{n+1}\right)^2 \dots\dots\dots\dots (33)$$

or

$$\nu_2 = 2\,\nu_1 - \frac{\nu_1^2}{n+1} \dots\dots\dots\dots (34)$$

As (30) is deduced on the assumption of the same life N for the two cases, (34) gives the corresponding values of ν for the same arbitrary N. Accordingly, if ν_2 and $2\nu_1 - \dfrac{\nu_1^2}{n+1}$ are plotted against N, all the points will fall approximately, on the same curve. Fig. 20 verifies this in a satisfactory way, considering the small number of specimens.

By (32) we are now in a position to prove that there exists at the most one single length of the specimen that may have a normal distribution. If this be L_1, the distribution function for a length L is by (29) and (32)

$$\Phi_L = 1 - \left[1 - \frac{1}{a\sqrt{2n}} \int\limits_{-\infty}^{+\infty} -\frac{(S - S_m)^2}{a^2} \cdot d\,S\right]^{\frac{L}{L_1}} \dots\dots (35)$$
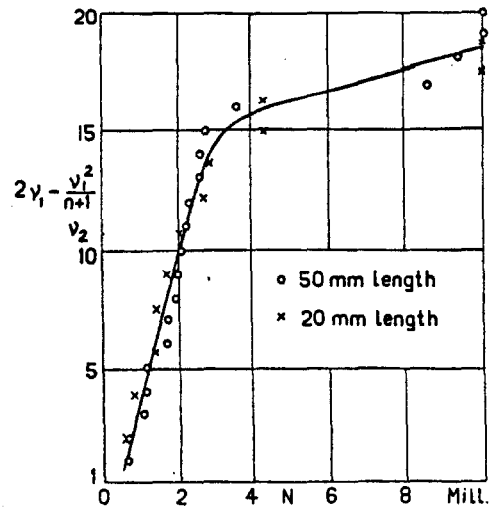
Fig. 20. Rotating-beam test on steel specimens of two different lengths.

This is not a normal distribution and, besides, not a very convenient expression for numerical computations. If one wants, nevertheless, to maintain the fiction that the distribution is approximately normal, it is necessary to compute new values of the parameters for each length of the specimen.

As a contrast, we now take a distribution according to (25) Then

$$F_L(S) = 1 - e^{\frac{L}{L_1} k (S - U)^m} \quad \dots\dots\dots\dots (36)$$

and there is no change in the values of the parameters.

This rule holds good even in the general expression (19), which gives

$$F_L(S) = 1 - e^{-\frac{L}{L_1} \varphi(S)} \quad \dots\dots\dots\dots (37)$$

or

$$\log \log \frac{1}{1 - P} = \log \varphi + \log L - \log L_1 \quad \dots\dots\dots (38)$$

If we then take $\log \log \dfrac{1}{1 - P}$ as ordinate and S as abscissa, the change of the length L means a displacement of the curve without any change of the form, as shown in Fig. 21.
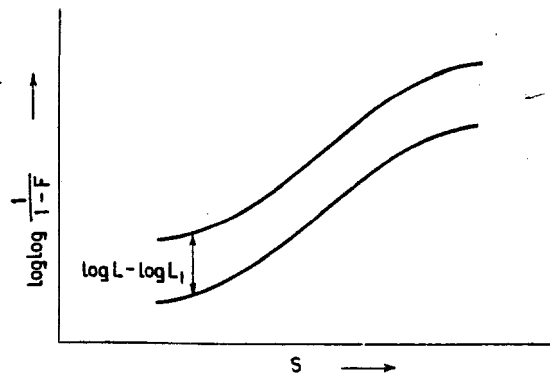
Fig. 21. $P - S$ curves for different lengths of the specimen.

In the case just discussed, where the length only, and not the cross-sectional area, was supposed to be changed, it is quite correct to put in the volume V instead of the length L in (36) thus obtaining

$$F_V(S) = 1 - e^{\frac{V}{V_1} k (S-U)^m} \quad \dots \dots \dots \dots (39)$$

The question now arises, if (39) has the same general validity as (32).

In order to elucidate this problem, we suppose that we couple the two specimens side by side, thus having a specimen with the original length L_1 but an area twice as great. Even in this case (39) is valid, because, if failure occurs in one of the specimens, the load on the other one will be doubled. The probability that it will endure more than a few cycles is extremely small.

If the two bars are allowed to melt together, so to speak, i. e. if we have a single specimen with a doubled cross-section, the validity of (39) is not at all self-evident. It depends on the manner in which the fracture is spread over the area, but it depends also on the homogeneity of the material. Sometimes the surface layers of the specimen may have other properties than the inner part, e. g. cast iron, ceramic material, etc. and then (39) certainly will be invalid. The same effect may be produced by machining, corrosion and similar processes, deteriorating the surface layers. It may also happen that the material is homogeneous.

In any case, it is no trivial mathematical affair to determine the influence of changing the diameter of the specimen.

The relation (39) is based on the assumption that the stress is uniformly distributed over the entire volume. This condition is not satisfied, for instance, at bending and torsional loads. As has already been demonstrated in another place (1, 2), the actual volume of the specimen then has to be reduced.

This is a rule which has some bearing on fatigue tests at combined loads. By a very ingenious method, GOUGH (10) has investigated the influence of combined stresses on fatigue. His testing machines had the capacity of changing the mode of loading in a continuous way, from bending to torsion.

Now, the effectively tested volume is much greater in torsion than in reversed bending. For this reason it is not quite correct to use the medians (or still worse the means) when evaluating the results without introducing corrections for the volume.

As a better method, I should propose the use of the U-values, which are, as previously mentioned, independent of the tested volume.

6. Computation of the Parameters

If one has obtained the lives N_ν, where $1 \leq \nu \leq n$, at a number of loads S_μ, where $1 \leq \mu \leq i$ the first procedure is to arrange the values according to increasing magnitudes of N and tabulate them as shown in Table 1.

Every column ($\nu =$ const.) may be considered as containing specimens with approximately the same value of E. The relation (15) is, accordingly, valid for each column, and the parameters k, m, and E may be computed.

For this reason, we start by plotting the S_ν and N_ν of each column on logaritmic scales. If the points fall approximately on a straight line, this means that the endurance limit $E = 0$.

In most of the cases, the points fall on a curve, which is concave upwards. Then, by way of trial, different values of E are tried. If a chosen value is greater than the actual one, the curve will be concave downwards as shown in Fig. 10. In this way the correct value will be found. After some experience, one or two trials will be sufficient.

It is difficult to judge, in this way, which value is the better, if the scatter is great, and the result will be, to some extent, subjective but, as a rule, quite sufficiently exact to decide whether the material is statistically homogeneous or not.

As previously mentioned, the observations sometimes have to be divided into two or more homogeneous parts, inside of which the parameters are constants and may be computed by a numerical method, giving objective and more correct values. This method, which originally was proposed and used by HANS WEIBULL (11), is based on the idea of determining the correlation between the S_μ and the N_μ values of each column (v = constant). Obviously, if there had been no scatter and (15) is valid, the points (x, y) of (17) should fall exactly on a straight line and the correlation should be complete, i. e. the correlation coefficient $r = 1$. Every deviation from the straight line results in a decrease of r. Then, it follows that every deviation of E from the correct value as a rule will be followed by a decrease of r.

The method now consists of chosing appropriate values of E, calculating corresponding values of

$$ r = \frac{\Sigma\, x\, y - \dfrac{\Sigma\, x \cdot \Sigma\, y}{i}}{\sqrt{\left[\Sigma\, x^2 - \dfrac{(\Sigma\, x)^2}{i}\right]\left[\Sigma\, y^2 - \dfrac{(\Sigma\, y)^2}{i}\right]}} \quad\dotso (39) $$

and selecting that value of E, which gives the greatest r.

From (17) it follows that only x, but not y, is influenced by a change of E.

When E has been determined in this manner, the a and m may be computed by the method of least squares through the formulae

$$ \log k = \frac{\Sigma\, y \cdot \Sigma\, x^2 - \Sigma\, x\, y\, \Sigma\, x}{i\, \Sigma\, x^2 - (\Sigma\, x)^2}; \quad m = \frac{i\, \Sigma\, x\, y - \Sigma\, y \cdot \Sigma\, x}{i\, \Sigma\, x^2 - (\Sigma\, x)^2} \quad (40) $$

These computations of E, k, and m are rather tedious, as they have to be carried through for every column of the table. For this reason, it is better to use a punched-card method, which is drawn up by BENGT W. WEIBULL, who has given the description and the schemes in Appendix I.

Now, when the parameters are computed, it is easy to calculate the relations between P and S for given values of N, i. e. a vertical section through the $S - N$ curves of field 1 in Fig. 4. It should be recommended to take three N values: one of them $N = \infty$, corresponding to the computed E, the second of them at the smallest N

inside the homogeneous part, and the third, an appropriate value between the two others.

For each of these N, corresponding values of S and P have to be calculated, being altogether n pairs of values, which are connected by (25).

The parameters U, k, and m of these $P - S$ curves, as illustrated by Fig. 4, field 2, will be computed in exactly the same manner as demonstrated concerning the $S - N$ curves. First, the homogeneity is graphically examined, then, the U value making r a maximum is determined. Here, of course, n has to be substituted for i, and x and y are defined by (27) instead of by (17).

After the parameters have been computed, corrected S values are easily obtained for appropriate P, including $P = 0$, i. e. the computed U values. The P values may be arbitrarily selected. There is no necessity of taking the same values as the ones obtained experimentally.

Finally, for each such P, corrected values of the parameters will be calculated using (14), after which the curves of fields 1 and 2 in Fig. 4 may be computed and drawn.

7. Numerical Examples

As the first example, we take the rotating-beam endurance test on aluminum by JOHNSON and OBERG (4). From Fig. 17, the probable value of E seems to be $E = 0$. In order to check this assumption, the correlation coefficient r has been calculated according to (39) for some values of E in the neighbourhood of $E = 0$. The result is illustrated in Fig. 22, giving for alloy A a maximum of $-r = 0.99\,960$ at $E = -0.4$ and for alloy C of $-r = 0.99\,965$ at $E = +2.6$.

It should be well observed that the values of log N are in both cases taken from Fig. 9 by measurement after some photographic magnification. This procedure may certainly introduce some additional errors, and for this reason the deviation of E from zero cannot be considered as significant.

By (40) the values of log k and m are computed giving for alloy A

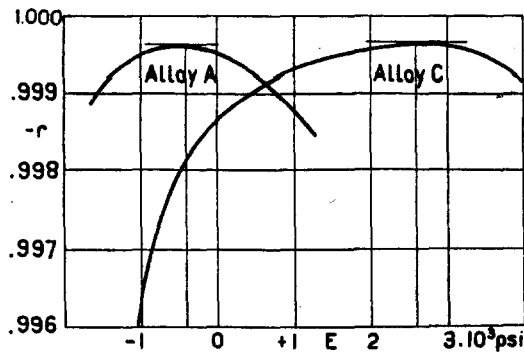$$\log k = 17.8346 \text{ and } m = -7.973$$

Fig. 22. Correlation coefficient vs. endurance limit for two aluminum alloys.

from which it follows that

$$N = 6.833 \cdot 10^{17} \cdot S^{-7.973}$$

where S is given in thousands of psi.

As the second example, we take the reversed-torsion endurance test on copper, by RAVILLY (6). The two hundred observed values of N are given in Table 1.

The first step now consists of computing the E-values for each one of the 20 columns of the table by chosing different E values between 5.5 and 6.8, and calculating corresponding values of the correlation coefficient r. The easiest way to do this is by means of punched-card machines, as described in the Appendix. As an example, the calculated values for $\nu = 14$ are shown in Fig. 23, from which it may be seen that $E = 6.4$ gives the maximum of correlation, indicating the closest fit to a straight line of the x and y values.

In this way E has been determined for each ν. The values thus obtained are denoted E_1 in Table 4 and plotted in Fig. 24.

The next step is to determine the distribution function of E. It is obvious that the scatter of the points plotted in Fig. 24 makes it impossible to decide if the expression (25) is the correct one or not.
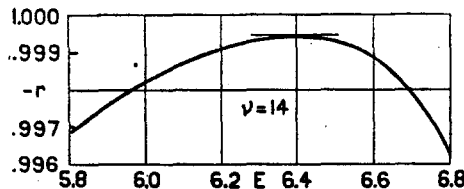

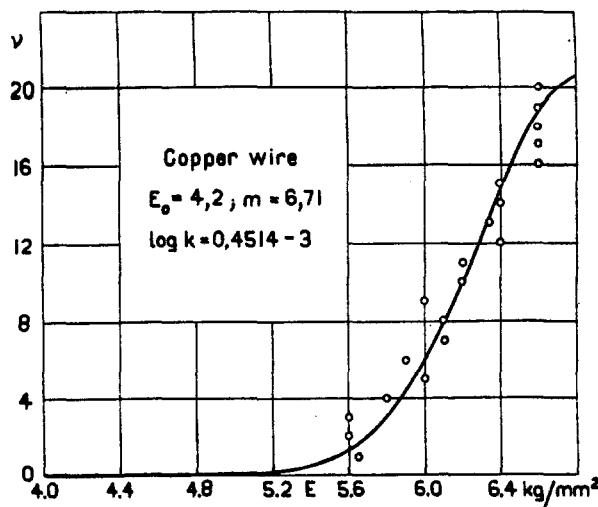Fig. 23. Correlation coefficient vs. endurance limit for copper alloys.

Fig. 24. $E - \nu$ curve for copper wire.

However, the distribution function

$$P = \frac{\nu}{n+1} = 1 - e^{-0.00283\,(E-4.2)^{6.71}} \quad \dots\dots\dots (41)$$

illustrated by the solid line in Fig. 24, reproduces in a reasonable way the values E_1, as may be seen from Fig. 24 and Table 4, where E_2 denotes the values calculated by (41).

TABLE 4. *Calculated endurance limits.*

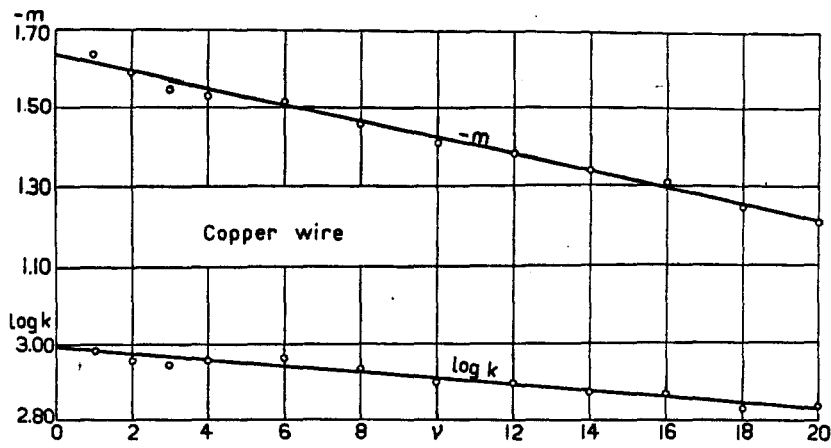|	E_1	E_2
1	5.65	5.55
2	5.6	5.70
4	5.8	5.88
6	5.9	6.00
8	6.1	6.10
10	6.2	6.19
11	6.2	6.23
12	6.4	6.27
14	6.4	6.35
16	6.6	6.44
18	6.6	6.54
20	6.6	6.70

Fig. 25. The parameters m and log k vs. ν for copper wire.

If we now calculate log k and m of (15) using E_2 and the observed values of x and y, then the relation between log k, m, and $ν$ may be put in the simple form

$$\left.\begin{array}{l} m = -1.63 + 0.021\,ν = -1.63 + 0.441\,P \\ \log k = 2.99 - 0.008\,ν = 2.99 - 0.168\,P \end{array}\right\} \quad \ldots\ldots (42)$$

These formulae may be used for $N > 10^5$ only.
As the most interesting special case we put $P = 0$ giving

$$E = 4.2, m = -1.63\,k = 977$$

Accordingly, the relation

$$N = 977\,(S - 4.2)^{-1.63} \ldots\ldots\ldots\ldots (43)$$

gives the lower boundary of the $S - N$ field.

It is obvious that (41) and (42) represent only one of the possible ways of describing the observations. Whether or not this method may be adapted to other materials must be left to future investigations.