1. Introduction

Linear section of grains in polyhedral material microstructure is a system of chords on a test line. The mean chord length is the linear grain size of the material. In the prior austenite microstructure of low alloy structural steels, the chord length is a random variable which has the gamma distribution. The distribution parameters are closely related to the austenite linear grain size. Because the sequent chord lengths are statistically independent, the grain size interval estimation (confidence interval) may be performed by the chi-square distribution.

Keywords: Modeling, austenite grain size, linear section method

Przekrój liniowy ziarn mikrostruktury materialu polikrystalicznego jest układem cięciw linii testowej. Średnia długość cięciw jest liniową wielkością ziarna materiału. W mikrostrukturze pierwotnego austenitu niskostopowych stali konstrukcyjnych długość cięciwy jest zmienną losową, która ma rozkład gamma. Parametry rozkładu są ściśle związane z liniową wielkością ziarna austenitu. Ponieważ kolejne długości cięciw są statystycznie niezależne, przedziałową ocenę wielkości ziarna (przedział ufności) można wykonać za pomocą rozkładu chi-kwadrat.

AUSTENITE GRAIN SIZE ESTIMATION IN STRUCTURAL STEELS BY LINEAR SECTION METHOD

OCENA WIELKOŚCI ZIARNA AUSTENITU METODĄ PRZEKROJÓW LINIOWYCH

Linear section of grains in polyhedral material microstructure is a system of chords with random length along a test line, Fig.1, [1,2].

Fig.1. Polyhedral metal microstructure with test line and the length of grain chords, $x_1$, $x_2$, ..., $x_i$

The sequent chord lengths: $x_1$, $x_2$, ... are values (realizations) of random variables $X_1$, $X_2$, ... For isometric microstructure (homogeneous and isotropic [2]) the $X_1$, $X_2$, ... are independent and have the same distribution, so it may be represented by a simple random variable $X$ [3]. The so-called chord length distribution of $X$ [2] has the probability density function (PDF) $f(x)$ with the statistical parameters: the mean (expected value) $\langle X \rangle = m$ and variance $\sigma_X^2 = \sigma^2$. The mean chord length ($m$) may be used as a measure of the linear grain size of the microstructure. The linear grain size estimation is subject of the quantitative metallography [1,4,5]. For $n$ values (realizations) of $X$: $x_i$ ($i=1$, ..., $n$) the arithmetic mean

$$\overline{x_n} = \frac{1}{n} \sum_{i=1}^{n} x_i$$  \hspace{1cm} (1)

i.e., the empiric grain size, is a value of the random variable $\overline{X_n}$ - the random mean. The $\overline{X_n}$ is characterised by the PDF $f(\overline{x_n})$ and statistical parameters, the mean $\langle \overline{X_n} \rangle$ and the variance $\sigma_{\overline{X_n}}^2$ which may be expressed by parameters of the $X$,

$$\langle \overline{X_n} \rangle = m$$  \hspace{1cm} (2)

and

$$\sigma_{\overline{X_n}}^2 = \frac{\sigma^2}{n}$$  \hspace{1cm} (3)

Depending on the PDF $f(x)$ and the number, $n$, of chords, the $\overline{X_n}$ may be used for statistical grain size estimation.

---

* A GH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF METALS ENGINEERING AND INDUSTRIAL COMPUTER SCIENCE, 30-059 KRAKÓW, 30 MICKIEWICZA STR., POLAND

* Corresponding author: wiencek@agh.edu.pl
well known point estimation is rather simple. The interval estimation is the main subject of the present work.

Statistical studies of the austenite grain chord lengths in structural steels of isometric microstructure show that the random variable $X$ has the gamma distribution [6], its PDF is:

$$f(x) = \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx}, \quad x \in [0, \infty);$$

where $b$ and $p$ are non-negative parameters ($b$ is the scale, and $p$ is the shape parameter of the PDF), and $\Gamma$ is the Euler gamma function, [7]. (Henceforth, the symbol $G(b,p)$ denotes the gamma distribution of $b$- and $p$-parameters; the symbol $X\sim G(b,p)$ denotes the random variable $X$ which has the $G(b,p)$.)

The statistical parameters of the $G(b,p)$ may be expressed by the $b$- and $p$- parameters, [7],

$$m = \frac{p}{b}$$

and

$$\sigma^2 = \frac{p}{b^2}.$$  

The variation coefficient, $v=\sigma/m$ is:

$$v = \frac{1}{2}.$$  

If the sequent chord lengths along a test line are statistically independent values (realizations) of $X\sim G(b,p)$, then the $\bar{X}_n$ has a gamma distribution too, [7], what make possible the grain size interval estimation.

The particular subjects of this article are:
(i) general properties of the austenite chord length distributions given in [6];
(ii) statistical independence of sequent chord lengths along test line (the independence is a condition for the assumed isometric microstructure);
(iii) construction of the grain size confidence interval.

For more advanced statistical analysis of the austenite chord lengths in structural steels some linear transformations of the $X\sim G(b,p)$ are required. They are presented in the Appendix.

The studies are also supported by computer simulation with one dimensional (1D) models of random segments on a line, its lengths are independent values of the $X\sim G(b,p)$.

2. Experimental

A set of 20 isometric prior austenite microstructures of low alloy structural steel 40Cr8 (with different contents of microalloying elements, V, Nb and Ti, heat treated with austenitizing for 30 min in argon atmosphere at temperature in the range of 840-1200°C and quenching in water) were investigated [6]. Fig.2 shows a typical microstructure of a steel specimen polished surface, after etching in saturated aqueous picric acid solution. The quantitative metallography was carried out by linear sections of the microstructure [1]. For each specimen, in the microscope field of view, on 5 random test lines (having minimum 10 chords and which end points are on the austenite grain boundaries), the length of chords were measured automatically using the computer image analysis program SigmaScan-Pro [6]. For each steel in several fields of view the length of $N>450$ chords was measured. In a measurement results data set the length of chords: $x_1, x_2, \ldots, x_N$ are written for the particular test lines, which for the isometric microstructure are chosen in an arbitrary (random) sequence.

The statistical analysis shows (chi-square test), that most of the austenite grain chord length distributions correspond to the gamma distribution [6]. It is necessary to emphasize that part of investigated austenite grain chord length distributions correspond to log-normal size distributions, whilst size distributions of austenite grains of specimen austenitised at temperature 1200°C had irregular shape different from gamma and log-normal distributions. As an example, Fig.3 shows some empirical PDFs in comparison with the appropriate $G(b,p)$ PDFs, given by Eq.(4). In Table 1. are given the austenite chord lengths characteristics: the statistical parameters (the arithmetic mean, $\bar{x}$, and standard deviation, $s$, in μm, and variation coefficient, $v=s/\bar{x}$, μm$^2$), the empirical parameters of the $G(b,p)$, ($b$- in μm$^{-1}$ and $p$-dimensionless) - calculated by Eq.(5) to Eq.(7) for suitable $x$ and $s$ values, and the number of chords, $N$, as well. In Table 1. the data sets (denoted: S1, ..., S20) are ordered relative to increasing of the $\bar{x}$ value.

From Table 1 it results that the range of values of the $\bar{x}$, $s$ and $b$-parameters is relatively large (max/min>15) in contradiction to the $v$ and $p$-parameters (max/min<2).
According to the equation (7), the $p$ parameter is determined by the variation coefficient, $v$.

Fig.4 shows the empirical $p$ values as a function of $x$. Because $p$ only changes a little (linearly) with $x$, therefore in practice the independence of $p$ from $x$ may be assumed. The arithmetic mean $<p>$ - calculated for the $p$-values given in Table.1 - is, $<p>=3$.

A more detailed approach uses the relative chord length distribution. For $X:G(b,p)$, the random variable

$$U = \frac{X}{m}$$

has the one parameter $G(p,p)$ (Appendix A2 for $a=m^{-1}$ and Eq.(5)). The statistical parameters of the $U:G(p,p)$ are $\{U\} = 1$ and $\sigma_u = p^2$. For a given data-set,

$$u = \frac{x}{x}$$

is the relative chord length, which may be regarded as empirical value (realization) of the random variable $U$. Fig.5 shows the empirical PDFs $f(u)$ for the particular data sets (S1,...,S20) in comparison with the PDF $f(u)$ of the $G(<p>,<p>)$ for $<p>=3$.

Consequently, for the analysed steels, the austenite grain chord length distributions may be approximated by the gamma distribution $G(<p>/m,<p>)$ for $<p>=3$ and $m=x$.

### 3. Independence

If the austenite sequent chord lengths along the test line: $x_1, x_2, ...$ are independent values (realizations) of the $X:G(b,p)$ (exactly, they are the values of independent random variables $X_1, X_2, ...$ which have the same $G(b,p)$), then for a given $n$, the arithmetic mean $\bar{x}_n$ is a value of the $\bar{X}_n:G(nb,np)$ (Appendix A4). Consequently, the statistical independence of the chord lengths may be analysed by the PDF $f(\bar{x}_n)$ of the random mean $\bar{X}_n$.

In a data-set for given number $n$ ($n=2,3,4,5$), the are at most $r_n$ disjoint subsets of $n$ elements. The arithmetic mean distribution of the $r_n$ subsets are characterised by the empirical PDF $f(\bar{x}_n)$.

The chi-square test shows that most of the 80 empirical PDFs $f(\bar{x}_n)$ (the 20 data sets and the $n=2,3,4,5$ give 80 subsets) are consistent with the corresponding PDF of the $G(nb,np)$.

As an example Fig.6 shows the comparison of empirical $f(\bar{x}_n)$ functions with the corresponding gamma distribution functions for $S6 (N=664, n=2$ and $n=5$).
From the analysis of the empirical PDFs \( f(x) \) it results, that the sequent chord lengths \( x_i, (i=1, \ldots, N) \) in the data-sets (S1, \ldots, S20) can be interpreted as independent values (realizations) of the corresponding \( X \sim G(b,p) \). This result is the basis for statistical grain size interval estimation.

4. Estimation

If the austenite sequent chord lengths along the test line are independent values (realizations) of \( X \sim G(b,p) \), then

\[
Z = 2nbX_n
\]

has the \( G(1/2,np) \), i.e., \( Z \sim G(1/2,np) \). The \( G(1/2,np) \) is a chi-square distribution with \( 2np \) degrees of freedom (Appendix A5). For given \( n \) and \( p \), the \( Z \sim G(1/2,np) \) taking into consideration of Eq.(5), leads to the grain size confidence interval:

\[
\binom{2np X_n \chi^2_{\alpha/2,2np}}{2np X_n \chi^2_{1-\alpha/2,2np}}
\]

where the so-called critical values \( \chi^2_{\alpha/2,2np} \) and \( \chi^2_{1-\alpha/2,2np} \) are determined by the chi-square distribution for \( 2np \) degrees of freedom and the given \( \alpha \) probability (\( \alpha<0.1 \)).

### TABLE 1.

| Data set | \( \bar{x} \) | \( s \) | \( v \) | \( b \) | \( p \) | \( N \) |
|----------|------|------|------|------|------|------|
| S1       | 4.70 | 2.44 | 0.52 | 0.79 | 3.71 | 654  |
| S2       | 5.42 | 2.73 | 0.50 | 0.73 | 3.94 | 591  |
| S3       | 5.72 | 2.93 | 0.51 | 0.67 | 3.81 | 632  |
| S4       | 5.74 | 3.42 | 0.60 | 0.48 | 2.82 | 1265 |
| S5       | 5.90 | 3.22 | 0.57 | 0.56 | 3.36 | 641  |
| S6       | 6.68 | 3.65 | 0.55 | 0.50 | 3.35 | 664  |
| S7       | 6.72 | 3.81 | 0.57 | 0.46 | 3.11 | 591  |
| S8       | 6.90 | 4.24 | 0.61 | 0.38 | 2.65 | 622  |
| S9       | 7.79 | 4.17 | 0.54 | 0.45 | 3.49 | 560  |
| S10      | 13.32| 8.19 | 0.61 | 0.20 | 2.65 | 679  |
| S11      | 17.63| 10.54| 0.60 | 0.16 | 2.80 | 604  |
| S12      | 23.65| 14.56| 0.62 | 0.11 | 2.64 | 557  |
| S13      | 25.36| 15.22| 0.60 | 0.11 | 2.78 | 547  |
| S14      | 28.76| 17.03| 0.59 | 0.10 | 2.85 | 548  |
| S15      | 32.77| 17.20| 0.52 | 0.11 | 3.63 | 462  |
| S16      | 34.48| 19.84| 0.58 | 0.09 | 3.02 | 472  |
| S17      | 38.74| 22.98| 0.59 | 0.07 | 2.84 | 806  |
| S18      | 48.92| 33.79| 0.69 | 0.04 | 2.10 | 852  |
| S19      | 56.53| 35.26| 0.62 | 0.05 | 2.57 | 624  |
| S20      | 64.19| 39.75| 0.62 | 0.04 | 2.61 | 561  |

For given \( n \), because of the random variable \( X_n \) :\( G(np,np) \), the end points of the confidence interval, i.e.,

\[
\left[ \frac{2np x_n \chi^2_{\alpha/2,2np}}{2np x_n \chi^2_{1-\alpha/2,2np}} \right]
\]

where the so-called critical values \( \chi^2_{\alpha/2,2np} \) and \( \chi^2_{1-\alpha/2,2np} \) are random variables with the means,

\[
a_1 = \frac{2npm}{\chi^2_{\alpha/2,2np}} \quad a_2 = \frac{2npm}{\chi^2_{1-\alpha/2,2np}}
\]

respectively. The algebraical variables \( a_1, a_2 \) as functions of \( n \) characterise the confidence intervals given by Eq.(11). These functions for the \( G(0.49,2.82) \) are shown in Fig.7. For given \( n \), the \( a_2-a_1 \) length of the \( [a_1,a_2] \) interval characterises the length of the confidence interval. Fig.7 shows that the \( a_2-a_1 \) length is a decreasing function of \( n \) (a part magnified 2 times for \( n<200 \) is shown only). With the \( a_2-a_1 \) length may be connected the precision of the grain size interval estimation. In this way, if the number of measured chords, \( n \), increase, the \( a_2-a_1 \) length decreases and the precision of the estimation increases.

Fig.7. The \( a_1,a_2 \) variables and the length \( a_2-a_1 \) as functions of the \( n \) for the \( G(0.49,2.82) \) of S4

Because the confidence interval, Eq.(11), depends on the \( p \)-parametr, in the subsequent analysis two cases will be distinguished: (i) the \( p \)-parametr is known; and (ii) the \( p \)-parametr is unknown.

4.1. The \( p \)-parametr is known

For given \( p \), Eq.(11) represents the exact grain size, \( m \), confidence interval. As an example, the confidence intervals for the data set S4 (\( p=2.82 \) and the largest data number, \( N=1265 \)) were analysed. For given \( n \) and \( X_n \), the grain size confidence interval value (realization) is

\[
\left[ \frac{2np x_n \chi^2_{\alpha/2,2np}}{2np x_n \chi^2_{1-\alpha/2,2np}} \right]
\]

The \( w_1, w_2 \) are the end points of Eq.(12) interval; \( w_1=2np X_n \chi^2_{\alpha/2,2np} \) and \( w_2=2np X_n \chi^2_{1-\alpha/2,2np} \). Fig.8 shows the variables \( w_1, w_2 \) and \( a_1, a_2 \) as functions of the \( n \) (for \( 2np=1132 \) and \( 1-\alpha=0.95 \)).
Fig. 8. Comparison of the S4 grain size confidence interval value \( w_1, w_2 \) variables with the \( a_1, a_2 \) variables as functions of \( n \)

From Fig. 8 it results, that for small \( n \) the scatter of the \( w_1, w_2 \) is quite large and with increasing of \( n \) it decreases (in accordance with Eq. (3)). For greater \( n (n>600) \) the \( w_1, w_2 \) slowly converge to the \( a_1, a_2 \). The precision measure of the grain size interval estimation may be connected with the half length, \( g \), of a confidence interval value, i.e.,

\[
g = \frac{w_2 - w_1}{2}
\]

(13)

For the S4 (Fig. 8), for \( n>600 \), the \( g<0.10 \) m.

The precision of analysis may be supported by computer simulation with the one dimensional (1D) model of S4 in form of random segments on a line. The segment lengths are independent values of the \( X \sim G(0.49,2.82) \). For given \( n \), the model confidence interval value is characterised by the \( w_1, w_2 \) points. Fig. 9 shows the \( w_1, w_2 \) variables for three independent simulations and the model \( a_1, a_2 \) variables as functions of \( n \). Fig. 9 is in accordance with Fig. 8 and confirms the empirical conclusion above: for \( n>600 \), the \( g<0.10 \) m.

Fig. 9. Comparison of the simulated grain size confidence interval value \( w_1, w_2 \) variables with the \( a_1, a_2 \) variables as functions of \( n \)

4.2. The \( p \)-parameter is unknown

If the \( p \)-parameter is unknown, assuming \( p=\bar{p}=3 \) the chord lengths distribution has the form \( G(3/m, 3) \). From Eq. (11) results the approximate grain size, \( m \), confidence interval,

\[
\left( \frac{6n\bar{X}_n}{\chi^2_{\alpha/2}}, \frac{6n\bar{X}_n}{\chi^2_{1-\alpha/2}} \right)
\]

(14)

In order to gain an idea relative to the approximation of an unknown \( p \) by \( <p> \) an example is given. For the data sets S2 and S18 (with the extreme \( p \)-parameter values, i.e, for S18: \( p=2.10 \) and for S2: \( p=3.94 \)) the confidence intervals were compared for two cases: (i) the \( p \)-parameter is known (an exact interval); and (ii) the \( p \)-parameter is unknown and is assumed \( p=\bar{p}=3 \) (an approximate interval).

For the data sets S2 and S18, the \( a_1, a_2 \) variables as functions of \( n \) were analysed (for \( m=x, 2np \) and \( 6n \) degrees of freedom, and \( 1-\alpha=0.95 \)). In Fig. 10 the functions for the exact and the approximate cases are shown. For S2 (Fig. 10a) and S18 (Fig. 10b) the difference between the \( a_1, a_2 \) variables as functions of \( n \) for \( p \) and \( <p> \), respectively, is not large. So, the difference between the approximate confidence interval and the exact one is not large too. Consequently, when the \( p \)-parameter is not known, by using \( p=\bar{p}=3 \) it is possible to determine an approximate grain size confidence interval.

Fig. 10. Comparison of the \( a_1, a_2 \) variables as functions of \( n \) for the exact and approximative grain size confidence intervals (data sets: S2 (Fig. 10a), S18 (Fig. 10b))

4.3. The chord counting method

In metallography, the chord lengths measurement may be reduced to simple counting measurements of \( n \) chords (or it end points) along a test line of length \( L \) - the so-called chord counting method [1,4,5]. Two cases will be distinguished:
(i) the test line length $L$ is equal the total length of chords, \( L = x_1 + x_2 + \ldots + x_n \), the $n$ and the $L$ should be measured; (ii) for given random test line length, $L$, the $n$ are to be determined be counting measurements ($n = n + 1$) where $n$, are chords in the test line). The chosen test line length, $L$, should ensure a sufficient number of chords, $n$, to secure the precision of the grain size interval estimation (e.g., $g < 0.10 m$).

Now, the arithmetic mean length $\overline{x_n}$ may be written,

$$\overline{x_n} = \frac{L}{n} \quad (15)$$

The Eq.(15) may be used for the determination of the grain size, $m$, confidence interval values if the $p$- or $<p>$- parameter are given.

5. Discussion and conclusions

For the analysed steels, the austenite grain chord lengths along a test line are the values of $G(\lambda, b, p)$. In the first approximation the $p$-parameter is independent of the particular austenite microstructure; it may be assumed to be constant, i.e., $p < < p > = 3$. The chord lengths distribution has the form: $G(\lambda, b, p)$, which is the main statistical property of austenite grain structure. From the metallurgical point of view, it seems to be determined by special interaction between the disperse phase (carbide, nitride, ...) and the austenite grain boundaries at elevated temperature (during the austenitizing heat treatment), [9]. The sequential chord lengths along a test line may be interpreted as statistically independent values (realizations) of the $X:G(\lambda, b, p)$. The expected value of $X_i$, i.e., the mean $m$, is a measure of the linear grain size. For a given number $n$, the arithmetic mean of the chord lengths, $\overline{x_n}$, is a value of the random mean $\overline{X_n} = G(n, b, np)$ - an estimator which may be used for the grain size, $m$, interval estimation. The grain size interval estimation takes into account independent values (realizations) of the $X:G(b, p)$. The expected value of $X_i$, i.e., the mean $m$, is a measure of the linear grain size. For a given number $n$, the arithmetic mean of the chord lengths, $\overline{x_n}$, is a value of the random mean $\overline{X_n} = G(nb, np)$ - an estimator which may be used for the grain size, $m$, interval estimation. The grain size interval estimation takes into account independent values (realizations) of the $X:G(b, p)$. The expected value of $X_i$, i.e., the mean $m$, is a measure of the linear grain size.

The distribution of a random variable may be characterized by the characteristic function (CHF). On the basis of a given CHF some conclusions according to the distribution may be made, [7]. In the Appendix are presented (the parts A1, .., A5) the CHFs of gamma distributions for random variables which are linear transformations of the $X:G(b, p)$.

A1. The $X:G(b, p)$ has the CHF

$$\varphi_x(t) = \left(1 + \frac{jt}{b}\right)^{-p} ; \quad t \in (-\infty, +\infty), \quad j^2 = -1 \quad (16)$$

A2. The random variable

$$Y = aX, a > 0$$

has the CHF

$$\varphi_Y(t) = \left(1 + \frac{ja}{b}\right)^{-p} \quad (18)$$

it is the CHF of the $Y:G(b/a, p)$.

A3. For $n$ independent random variables $X_i:G(b, p)$ and numbers $a_i > 0, i = 1, \ldots, n$; the random variable

$$Y = \sum_{i=1}^{n} a_i X_i \quad (19)$$

has the CHF

$$\varphi_Y(t) = \prod_{i=1}^{n} \left(1 + \frac{j a_i t}{b_i} \right)^{-p_i} \quad (20)$$

A4. A special case of A3. If the $X_i$ have the same distribution, i.e., $X_i:G(b, p)$ and $a_i = 1/n$, the random arithmetic mean

$$\overline{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad (21)$$

5.1. Conclusions

- In the analysed low alloy structural steels, the sequential austenite grain chord lengths $x_1, x_2, \ldots$ along a test line are values of the $X:G(b, p)$.
has the CHF
\[ \varphi_{X_n}(t) = \left(1 + \frac{jt}{nb}\right)^{-np} \]  \hspace{2cm} (22)
it is the CHF of the G(nb, np).

A5. The random variable is:
\[ Z = 2nb \overline{X_n}, \]  \hspace{2cm} (23)
If in Eq.(17), one substitutes the \( a \) by \( 2nb \) and the \( X \) by \( \overline{X_n} \),
then the \( Y \) is equal to \( Z \) which CHF is
\[ \varphi_{Y}(t) = \left(1 + j2t\right)^{-np}; \]  \hspace{2cm} (24)
it is the CHF of the G(1/2, np). The G(1/2, np)) may be interpreted
as chi-square distribution with \( 2np \) degrees of freedom.

Acknowledgement

The main parts of this paper were presented at the
German-Polish Joint Conference on Probability and
Mathematical Statistics, 6-9 June 2013 in Toruń (Poland). The
financial support work within the AGH project 11.11.110.299
is gratefully acknowledged.

REFERENCES

[1] J. Ryś, Stereology of materials (in Polish). Fotobit, Kraków
1995.

Received: 20 January 2015.
