PERGAMON

International Journal of Solids and Structures 36 (1999) 2585–2612

INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES

Similarity and fractality in the modelling of roughness by a multilevel profile with hierarchical structure

F. M. Borodich*, D. A. Onishchenko†

Department of Mathematics, Glasgow Caledonian University, Glasgow G4 0BA, U.K.

Received 12 December 1997; in revised form 5 March 1998

Abstract

Some modern scaling and fractal approaches to the characterization of surface roughness are under consideration. A new mathematical model of a multilevel profile suitable for flexible simulation of roughness is presented as a development from these approaches. The iterative construction of the model is described. It is shown that the profile is a hierarchical structure with self-affine and fractal features. Its geometrical properties are studied both analytically and numerically, and some of them are proved to be of the renormalization type. The profile contour length can be finite or infinite depending on the chosen values of the structural parameters. Some characteristics including the bearing curve and the upper volume function of the profile and the corresponding prestructures are calculated.

1. Introduction

Surface topography plays a significant role in tribology, i.e., in problems of friction, wear, lubrication and contact. This is the reason why the problem of analysis of rough surfaces attracts the attention of engineers and applied mathematicians. Historically, the following engineering parameters, statistical in nature, were used for the characterization of surface roughness: (i) the root mean square (rms) height σ; (ii) the rms slope σ_m^2, and (iii) the rms curvature σ_k^2 of the surface. These parameters, which are included in national standards of a number of countries, can be obtained for real surfaces using contact profilometry, scanning electron microscopy and other experimental methods.

However, it was realized later that the topography of engineered surfaces is too complex to be described completely by a few parameters. Thus, it was found that roughness has a multiscale nature and requires sophisticated mathematical techniques for its description.

* Corresponding author.

† EPSRC Visiting Research Fellow on leave from Institute for Problems in Mechanics (Russian Academy of Sciences), 101 Vernadskogo pr., Moscow 117526, Russia. E-mail: onish@ipmnet.ru

0020–7683/99/$ - see front matter
PII: S0020–7683(98)00116–4

Statistical modelling based on random field theory is the standard approach to surface geometry characterization. One of the first attempts to model the distribution of heights of surface asperities was presented by Longuet-Higgins (1975a, b); further development was provided by Nayak (1971). These authors and most other users of classical random field theory assumed that the functions of surface model are differentiable. In particular, this implies that limiting values for σ_m^2 and σ_k^2 should exist as the sample interval tends to 0. However, it turned out that such limiting behaviour is in contradiction with the results of advanced investigations of surfaces. For example, the exponential behaviour of the auto-correlation function implies that the engineering parameters should tend to infinity rather than to constant values when the sampling interval is infinitely reduced (see, e.g., Greenwood, 1992). Sayles and Thomas (1978) showed that the profiles of a large number of both natural and artificial surfaces have the following form of the spectral density $G(\omega) = C/\omega^{\alpha}$, where $\alpha \approx 2$. It follows from this that all wavelengths are equally represented in the profile and that there exists no characteristic scale; in other words, after arbitrary magnification roughness looks like before. Moreover, it was found that the values of engineering parameters depend on the measurement scale, i.e., these parameters are scale-dependent (Thomas, 1982; Madjumdar and Bhushan, 1990; Greenwood, 1992).

The fractal approach was introduced as an attempt to give a scale-invariant characterization of surface topography. The idea of fractality of roughness was experimentally verified on real surfaces as well as when applied to mathematically simulated profiles (see, e.g., Mandelbrot et al., 1984).

The term fractal was coined by Mandelbrot (1977, 1983). Roughly speaking, fractals can be defined as sets with non-integer fractal dimension. Naturally, the term fractal dimension requires a precise definition. Mandelbrot (1977, 1983) argues that the language of fractal geometry often describe physical phenomena better than the language of smooth classical objects. Now fractal geometry, which is a branch of mathematics dealing with highly irregular sets, is one of those modern topics, whose results have had a great impact on researchers in various areas of science.

Evidently, roughness of the surface of a body has a great influence on stress fields that arise when two deformable bodies are pressed together. Analysis of the effect of roughness on the contact interaction of solids has attracted wide attention (see, e.g., Johnson, 1985). One of the most popular models for studying contact of rough bodies is the Greenwood and Williamson (1966) model based on the use of the Hertz theory. However, Majumdar and Bhushan (1991) criticised the Greenwood and Williamson (GW) model as a non-scale-invariant. Currently, the development of models of contact between nominally flat fractal rough surfaces presented by Borodich and Mosolov (1991, 1992) for the Cantor profile is an active area of research. Various contact problems for nominally flat fractal punches were considered by Borodich and Onishchenko (1993), Warren and Krajcinovic (1995, 1996a, b) and Warren et al. (1996). All these models consider the one-level Cantor profile. It was noted by Borodich and Onishchenko (1993) that models incorporating multilevel self-affine profiles, when the asperities of the next generation lie on the tops of those of the previous level or when the even and odd generations are directed upwards and downwards correspondingly, could give much more natural rough profiles. Nevertheless, there has been no progress in study of such models yet. It is necessary to note that a multilevel self-affine structure were used by Liu (1985) and Kaplan et al. (1987) for surface characterization in application to metal–electrolyte interface. However, their construction procedure led to unbounded profile height in most of the considered examples. The present paper is to rectify this situation.

In this paper, the authors propose to study the bounded similarity of surface roughness on

iteratively generated rough profiles. The new multilevel profile model, described in the paper, is suitable for the flexible modelling of roughness. The profile is easily simulated and drawn, which is important for visual analysis.

It is known that fractal dimension is not a compressive geometric parameter that could characterize alone the behaviour of contacting rough bodies (see, e.g., Borodich and Onishchenko, 1993). Moreover, as we will see below the employment of the fractal approach in the study of surfaces has several drawbacks. The proposed model can be both fractal and non-fractal depending on values of the structural parameters. Regardless of this, the model profile remains rough and possesses certain self-affine properties. The iterative regular construction of the profile allows us to analyse its prestructures (prefractals) of arbitrary generation.

The paper is organised as follows: in Section 2 we recall some relevant definitions and methods attributed to fractal geometry in application to the modelling of rough surfaces and underline the difference between mathematical and physical fractals. We consider the properties of self-similarity and self-affinity of surface profiles and discuss some approaches for the fractal description of roughness involving Brownian surfaces and Weierstrass type functions. In Section 3 a new mathematical model suitable for the flexible simulation of rough profiles is presented. The iterative construction of the model profile is described. It is shown that the profile is a hierarchical structure with certain self-affine and fractal features. Then, geometrical properties of the profile are studied both analytically and numerically, and some of them are proved to be of the renormalization type. Some characteristics of the profile and corresponding prestructures are calculated, in particular the bearing curve and the upper volume function.

2. Fractal description of rough surfaces

During the last two decades methods of fractal geometry were of great interest as a possible tool for describing surface roughness. Mandelbrot et al. (1984) presented the results of experimental studies of fracture surfaces of a selection of steel samples. They argued that the irregularity of these fracture surfaces exhibits fractal characteristics and that, at least on the mesoscopic scale, these surfaces have irregularities of all scales. Afterwards, there were numerous papers published in the field. Below we consider various approaches to the fractal analysis of surface topography.

2.1. Mathematical definition of fractal

What is a fractal? Greenwood (1992) noted that Mandelbrot is somewhat reluctant to define 'fractals' or 'fractal dimension' preferring to offer examples. This indeed is the case (see, e.g., Mandelbrot, 1977, 1980, 1983). Nevertheless, we use the definition (Mandelbrot, 1977, 1983), according to which a set in a metric space is called a fractal set if the Hausdorff–Besicovitch dimension of the set is greater than its topological dimension. Note that this is the definition of mathematical fractals.

Let us recall some definitions which are used in fractal modelling.

Let X be a compact metric space and O be the totality of open balls in X. The Hausdorff s-measure of a subset $S \subset X$ is defined for $s \geqslant 0$ as the following limit

$$m_{\mathrm{H}}(S,s) = \lim_{\sigma\to 0+} \inf_{G\in O} \left\{ \sum_{V\in G} (\operatorname{diam} V)^s : S \subseteq \bigcup_{V\in G} V, \operatorname{diam} V \leqslant \delta \right\}.$$

Here G's are finite or denumerable subsets of O. It was proved that there exists a value s_0 such that

$$m_{\mathrm{H}}(S,s) = \begin{cases} \infty, & \text{for } s < s_0, \\ 0, & \text{for } s > s_0. \end{cases} \tag{1}$$

The Hausdorff dimension of the set S, denoted by $\dim_{\mathrm{H}} S$, is the number s_0 such that (1) holds.

Unfortunately, the calculation of the Hausdorff dimension of mathematical objects often demands a lot of effort. Even to find some estimations of the dimension, it is necessary to overcome a number of rather complex mathematical difficulties (see, e.g., Mauldin and Williams, 1986). This called for the use of other definitions of dimension which are useful in applied mathematics for the characterization of fractal objects.

One such alternative is the box dimension, whose definition is usually attributed to Pontrjagin–Schnirelman and Kolmogorov. The analytical calculation of the box dimension is usually easier since the corresponding definition of this dimension involves coverings by spheres of equal radii.

Let E be the Euclidean dimension of the space in which a set S is embedded. For $\delta > 0$, let $N(\delta)$ be the smallest number of E-dimensional balls or cubes of diameter δ needed to cover the set S. The box counting dimension or box dimension, denoted by $\dim_{\mathrm{B}} S$, can be defined if the following limit exists

$$\dim_{\mathrm{B}} S = \lim_{\delta\to 0+} \frac{\log N(\delta)}{-\log \delta}. \tag{2}$$

It can be proved that $\dim_{\mathrm{B}} S$ does not change if one takes $N(\delta)$ as: (i) the smallest number of δ-cubes that cover S; (ii) the number of δ-mesh cubes that intersect S; (iii) the smallest number of sets of diameter at most δ that cover S; (iv) the largest number of disjoint δ-balls with centres in S. In general, one can define two numbers, called the upper and lower box dimensions. To calculate these numbers, one has to replace the expression lim in the above definition by lim sup and lim inf, respectively. If the upper and lower dimensions are equal to each other, then the true limit exists, and this value is defined as box dimension (see, e.g., Falconer, 1990).

Unfortunately, the box dimension is not always equal to the Hausdorff dimension. For example, the set $S = \{0, 1, 1/2, \ldots, 1/n \ldots\}$ has unequal values for the Hausdorff and box dimensions:

$$\dim_{\mathrm{H}} S = 0 \neq \dim_{\mathrm{B}} S = 1/2.$$

However, it can be proved that $\dim_{\mathrm{H}} S \leqslant \dim_{\mathrm{B}} S$.

There are also many other names which are often used for the fractal dimensions. Several examples of these are the Bouligand, Minkowski, self-similarity, Frostman and packing dimensions. Various methods can be employed to calculate these dimensions for a mathematical fractal (see, e.g., Falconer, 1990; Tricot, 1995). Sometimes, when using different definitions, different values of the fractal dimension can be obtained. At the same time, it was shown that some of the above mentioned definitions are mathematically equivalent to each other.

As a simple alternative to the Hausdorff measure, we can introduce the s-measure m_s of a set as the following limit

$$m_s(S) = \varlimsup_{\delta \to 0+} N(\delta)\delta^s. \tag{3}$$

and define the box dimension as the value $s = D$ such that $m_s(S)$ has a jump from 0 to ∞ similar to the behaviour of $m_{\mathrm{H}}(S, s)$ in (1). However, m_s is not a σ-additive measure. One consequence of this, however, is that box dimensions have a number of unfortunate properties, and can be awkward to handle mathematically (Falconer, 1990). On the other hand, the difficulties involved with calculating the Hausdorff dimension are the reason for the opinion that the Hausdorff dimension has no practical application in the study of curves originated in other sciences: physics, biology or engineering (Tricot, 1995).

2.2. Physical concept of fractals

Evidently, it is impossible to carry out the scaling procedure for any real physical object down to infinitely small scales. Hence, the mathematical concept of the Hausdorff measure is applicable only to mathematical models of objects rather than to the objects themselves and, of course, the Hausdorff dimension cannot be obtained by experimental procedures. In this sense there are no actual fractal objects in nature.

For physical objects the box dimension cannot be calculated analytically but it is estimated by experimental or numerical calculations. However, various errors can arise during such numerical calculations.

There is no canonical definition of physical fractals and there are numerous methods for the practical estimation of the fractal dimension of an object. The cluster fractal dimension is taken as the first example of a physical fractal dimension definition.

Let a whole cluster be imagined as consisting of elementary parts of the size δ_*. An object can be modelled as a fractal cluster with dimension D when the model considers scales $\mathscr{R}$ such that $\delta_* < \mathscr{R} < \Delta_*$, where δ_* and Δ_* are the upper and lower cutoffs for the fractal representation.

To get the value D of the dimension, the considered region is discretized into cubes with side length δ_*. Then the smallest number of E-dimensional cubes needed to cover the cluster ($N(\delta_*)$) is counted. One says that the cluster is fractal if the numbers $N(\delta_*)$ satisfy the so-called number-radius relation for different sizes of the considered region of the cluster $\mathscr{R}$

$$N(\delta_*) \approx (\mathscr{R}/\delta_*)^D, \quad \delta_* < \mathscr{R} < \Delta_*. \tag{4}$$

The value of D is estimated as the slope of linear growth of $\ln(N(\delta_*))$ plotted against $\ln(\mathscr{R})$. The power D is usually called the cluster dimension or mass dimension.

The name of the latter term can be explained in the following way. Let some 'mass' $M(S_*)$ be assigned to the elementary particle S_* of the size δ_*. Then instead of the s-measure m_s of the cluster S used in the definition of box dimension one has the 'mass' of the whole cluster $M(S) = N(\delta_*)M(S_*)$.

As an example of the use of the definition, let us consider a profile which is imaged on a computer screen as a union of points (pixels) of the size δ_*. Then we can obtain a computerised estimation of the number of pixels $N(\delta_*)$ forming the line and lying inside a circle or a square box of radius $\mathscr{R}$ centred at a point $\mathbf{x}$. If the profile has fractal properties then repeating the procedure of estimation for different values of $\mathscr{R}$ will always give relation (4).

Another definition of the physical fractal dimension is based on the Richardson method. This method uses dividers which are set to a prescribed opening δ (Mandelbrot, 1983). Moving with these dividers along the contour so that each new step starts where the previous step leaves off, one obtains the numbers of steps $N(\delta)$. The contour is said to be of fractal nature if by repeating this procedure for different values of δ the relation

$$N(\delta) \sim (\delta)^{-D} \tag{5}$$

is obtained in some interval $\delta_* < \delta < \Delta_*$ of sizes δ. The power D is usually called the Richardson dimension D_R.

The last method which we would like to mention is the following. We cover the contour S by a grid of squares of size δ_i. Then successively dividing each initial square of the grid into four subsquares of size $\delta_{i+1} = \delta_i/2$ and calculating the number of subsquares which contain points of S, we may obtain the relation (5) which holds in some interval of sizes δ. In this case, the power D is usually called the physical box dimension D_B.

We see that various methods are utilized to estimate the fractal dimension of a physical object. However, the notion fractal dimension is not well-defined in that the relative value does depend on the approach used. Indeed, only for the mathematical box dimension of a fractal set S it is proved that $\dim_B S$ is the same when using various specific schemes of covering (see, e.g. Falconer,1990), while for physical fractals the estimations of the fractal dimension inevitably involve various techniques, distinct scale ranges, and various computation rules. Therefore, the obtained values can differ strongly and it is unlikely that they could be fruitfully compared for distinct objects. Thus, even in the case of physical objects of a similar nature, it would be wrong to consider fractal dimension of these objects as their specific property without referring to the estimation technique involved.

2.3. Self-similarity and self-affinity of surfaces

Let us recall that a one-to-one mapping M of a plane π onto a plane π' is called a similarity mapping with coefficient $\lambda > 0$, or simply a similarity, when the following property holds: if A and B are any two points of π, and A', B' are their images under M, then $|A'B'| = \lambda|AB|$ (see, e.g., Modenov and Parkhomenko, 1965).

It is known that any similarity transformation of a plane is a homogeneous (isotropic) dilation of coordinates $x' = \lambda x$, $z' = \lambda z$ up to a rotation and translation.

A set S is called statistically self-similar if under homogeneous scaling with the coefficient λ, $1 > \lambda > 0$, it is identical from the statistical point of view to the set $S' = \lambda S$.

In practice, it is impossible to verify that all statistical moments of the two distributions are identical. Frequently, a set S is said to be self-similar if only a few moments do not change under scaling (see, e.g., Voss, 1985).

A one-to-one mapping M of a plane π onto a plane π' is called an affine mapping, if the images of any three collinear points are collinear in turn (see, e.g., Modenov and Parkhomenko, 1965). In general, an affine transformation of a plane may be given in any coordinate system as a non-degenerative linear transformation. In practical studies of rough surfaces, one often considers a particular affine mapping, with anisotropic scaling, that is given coordinatewise by

$$x' = \lambda x, \quad z' = \lambda^H z.$$

Here z is a graph of a surface profile and H is some scaling exponent.

One says that a fractal is self-affine if it is invariant from the statistical point of view under quasi-homogeneous (anisotropic) scaling.

It is possible to show that usually a quasi-homogeneous transformation is a particular case of Lipschitz homeomorphism (Borodich, 1994, 1995). The Hausdorff dimension of a set S does not change under the action of the Lipschitz homeomorphism L, i.e.,

$$\dim_{\mathrm{H}} S = \dim_{\mathrm{H}} L(S). \tag{6}$$

The ideas of self-similarity and self-affinity are very popular in studying surface roughness because experimental investigations show that usually profiles of vertical sections of real surfaces are statistically similar to themselves under repeatedly magnifications; however, the profiles should be scaled differently in the direction of nominal surface plane and in the vertical direction. The self-affine fractals were used in a number of papers as a tool for description of rough surfaces (Majumdar and Tien, 1990; Majumdar and Bhushan, 1990; Moreira et al., 1994; Tricot et al., 1994; Schmittbuhl et al., 1994, 1995; Hansen et al., 1995; Plouraboué et al., 1995; Lopez et al., 1995; Blackmore and Zhou, 1996; Vandembroucq and Roux, 1997).

Two standard examples of self-affine fractals are the trace of the fractional Brownian motion and the Weierstrass function. The former is a statistical fractal, while the latter is a deterministic fractal. We will consider them below.

2.4. Brownian surfaces and random fractals

Fractional Brownian motion (FBM) is a generalisation of the ordinary Brownian process. Fractional Brownian processes are widely used in creating computer-generated surfaces, in particular landscapes. For example, a profile can be constructed as a graph of 1-D FBM $V_H(x)$ of index H, i.e., we have

$$z = z(x), \quad z(x) = V_H(x)$$

where x is taken as the time and z is the random variable of the single valued function $V_H(x)$ with

$$\langle [V_H(x+\delta) - V_H(x)]^2 \rangle \sim \delta^{2H}, \quad 0 < H < 1 \tag{7}$$

or, in nonstrict notation,

$$|V_H(x+\delta) - V_H(x)| \sim \delta^H. \tag{8}$$

It is known (see, e.g., Falconer, 1990) that with probability equal to 1

$$\dim_{\mathrm{H}} V_H(x) = \dim_{\mathrm{B}} V_H(x) = 2 - H.$$

The auto-correlation function is one of the main tools for studying statistical models of rough surfaces. The auto-correlation function $R(\delta)$ of the profile is

$$R(\delta) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} [z(x+\delta) - \bar{z}]\,[z(x) - \bar{z}]\,\mathrm{d}x = \langle [z(x+\delta) - \bar{z}]\,[z(x) - \bar{z}] \rangle$$

or

$$R(\delta) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} z(x+\delta) z(x)\,\mathrm{d}x - (\bar{z})^2$$

where $\bar{z}$ is the average value of the profile function $z(x)$

$$\bar{z} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} z(x)\,\mathrm{d}x.$$

Another tool for the characterization of surfaces is the spectral density function $G(\omega)$, which is the Fourier transform of $R(\delta)$, i.e.,

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} R(\delta) \cos\omega\delta\,\mathrm{d}\delta \quad \text{and} \quad R(\delta) = \int_0^{\infty} G(\omega) \cos\omega\delta\,\mathrm{d}\omega.$$

It is believed (see, e.g., Falconer, 1990) that

(i) if the auto-correlation function $R(\delta)$ of the profile $z(x)$ satisfies

$$R(0) - R(\delta) \sim \delta^{2(2-s)}$$

then it is reasonable to expect that the box dimension of the graph $z(x)$ is equal to s [note that one can find $R(0) - R(\delta) \sim \delta^{2H}$ for the FBM defined by (7)];
(ii) if the profile $z(x)$ has spectral density

$$G(\omega) \sim 1/\omega^{\alpha} \tag{9}$$

then it is reasonable to expect that the box dimension of the graph $z(x)$ is equal to $(5-\alpha)/2$.

The above conclusions are valid for mathematical models of the profile, for which the relation $2(2-s) = \alpha - 1$ or $\alpha = 5 - 2s$ holds. Usually, it is assumed that the same conclusions concerning the box dimension are valid for physical fractals as well. An effective method for determining the physical fractal dimension of a random structure is based on this idea. It is shown that real surfaces approximately satisfy the property (9) in wide range of scales (see, e.g., Sayles and Thomas, 1978; Brown, 1995). The exponent α varies typically between 1 and 3.

The moments m_n of the spectral density $G(\omega)$ provide a useful description of the surface roughness

$$m_n = \int_{\omega_0}^{\infty} \omega^n G(\omega)\,\mathrm{d}\omega$$

where $\omega_0 = 2\pi/\lambda_0$ is the wavenumber corresponding to the profile length λ_0. It is possible to show (see, e.g., Brown, 1995) that m_0 is the variance of heights (rms height) of the surface, m_2 is the variance of slopes (rms slope) and m_4 is the variance of curvatures (rms curvature). This gives us a connection with the standard engineering parameters.

2.5. Weierstrass type functions and modelling of rough surfaces

A number of researchers have used the Weierstrass type functions for fractal modelling of surface roughness (see, e.g., Roques-Carmes et al., 1986; Sun and Jaggard, 1990; Majumdar and Tien, 1990; Majumdar and Bhushan, 1990; Borodich, 1993; Moreira et al., 1994; Tricot et al., 1994; Lopez et al., 1995; Blackmore and Zhou, 1996).

The real Weierstrass type function can be defined as

$$W(x;p) = \sum_{n=0}^{\infty} p^{-\gamma n} h(p^n x), \quad p > 1, \quad 0 < \gamma < 1$$

where h is a bounded Hölder function of order greater than β. The following complex generalisation of the $W(x;p)$ was considered by Mandelbrot (1977)

$$\tilde{W}(x;p) = \sum_{n=-\infty}^{\infty} p^{-\gamma n}[(1-\mathrm{e}^{ip^n t})\,\mathrm{e}^{i\Phi_n}], \quad p > 1, \quad 0 < \gamma < 1$$

where Φ_n are arbitrary phases.

The Weierstrass type functions are continuous everywhere and differentiable nowhere. In addition, their graphs are curves whose fractal dimension exceeds one. Fractal properties of these functions including the Weierstrass–Mandelbrot (WM) function C (Fig. 1) and the Takagi–Hopson function T

$$C(x;p) = \sum_{n=-\infty}^{\infty} p^{-\gamma n}(1-\cos p^n x), \quad T(x;p) = \sum_{n=-\infty}^{\infty} p^{-\gamma n}\left|p^n x - \left[p^n x + \frac{1}{2}\right]\right|, \quad p > 1, \quad 0 < \gamma < 1. \tag{10}$$

have been studied in numerous papers (see, e.g., Mandelbrot, 1977; Berry and Lewis, 1980; Mauldin and Williams, 1986; Falconer, 1990; Hu and Lau, 1993). By direct calculations, one may obtain

$$|\tilde{W}(x+\delta;p) - \tilde{W}(x;p)| \sim \delta^{\gamma}$$

which is similar to the behaviour (8) of fractional Brownian motion. The box dimension of the Weierstrass function graphs is $D = 2-\gamma$ and it is believed (see, e.g., Berry and Lewis, 1980; Falconer, 1990) that their Hausdorff dimension is the same. Currently, the only known bounds for the Hausdorff dimensions are obtained by Mauldin and Williams (1986), i.e.,

$$D - (c/\log p) \leqslant \dim_{\mathrm{H}} \text{graph}\, C \leqslant D$$

provided that p is large and constant c is large enough.

It is possible to calculate the spectral density of the WM function $\tilde{W}(x;p)$

$$G(\omega) = \sum_{n=-\infty}^{\infty} \frac{\delta(\omega - p^n)}{p^{2(2-D)n}}$$

where δ is the Dirac delta. Berry and Lewis (1980) suggested some arguments for approximating this discrete spectral density by a continuous spectral density

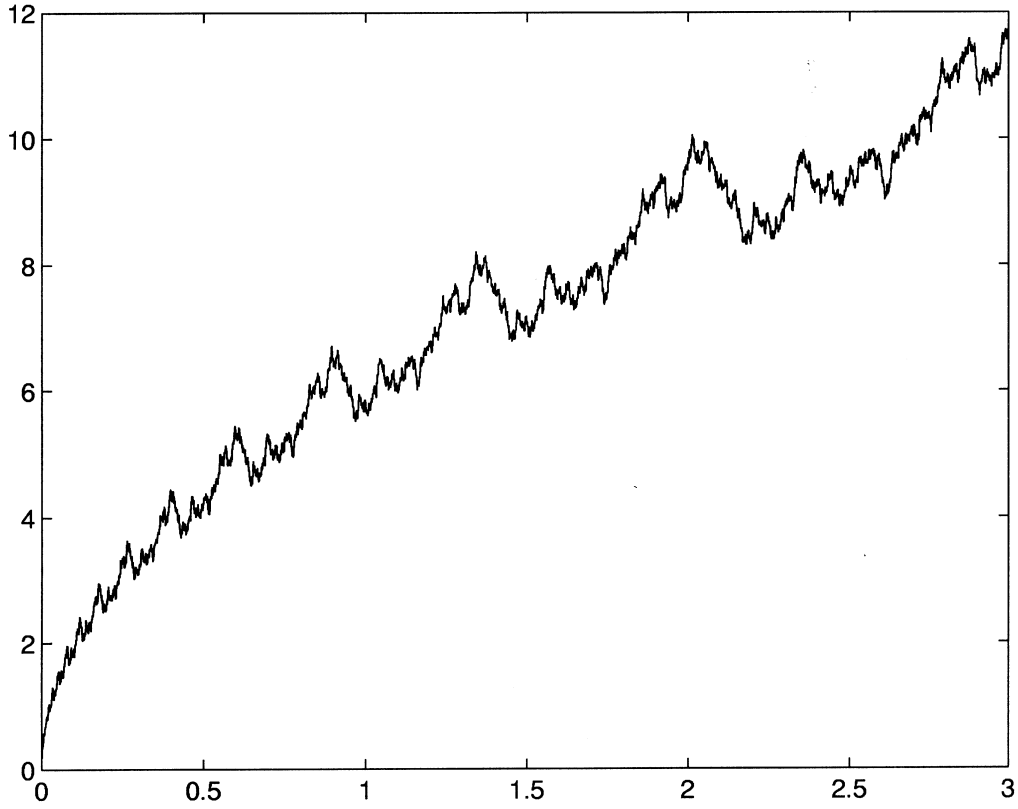

Fig. 1. Graph of the Weierstrass–Mandelbrot function C in the range $0 \leqslant x \leqslant 3$; $p = 1.5$, $\gamma = 0.5$. The trend of the function is $\sim x^{\gamma}$.

$$\bar{G}(\omega) \sim \frac{1}{\omega^{5-2D}}$$

whose exponent $5-2D$ is in agreement with (9) with respect to the box dimension.

The following truncated WM function

$$\tilde{W}_1(x;p) = A^{(D-1)} \sum_{n=n_1}^{\infty} p^{(D-2)n} \cos 2\pi p^n x, \tag{11}$$

is often used for fractal characterization of the surface topography (see, e.g., Majumdar and Tien, 1990; Majumdar and Bhushan, 1990; Lopez et al., 1995; Wang and Komvopoulos, 1995). Here n_1 is an integer number, which corresponds to the low cut-off frequency of the profile,and A is the so-called characteristic length scale of the profile. The number n_1 depends on the length L of the sample and is given by $p^{n_1} = 1/L$ and the parameter A determines the position of the spectral density along the $\log G$ axis. It was stated that both parameters A and D of the function W_1 are scale-invariant characteristics of the roughness. However, the extensive experimental studies of

this fractal characterization model showed that the values of parameters A and D are not unique and depend on instruments or resolution of a given instrument (Bhushan, 1995).

Evidently, the function $C(x;p)$ is not homogeneous. Nevertheless, it exhibits the following property

$$C(p^k x;p) = p^{k\gamma} C(x;p), \quad k \in \mathbb{Z},$$

where $\mathbb{Z}$ is the set of all integers, which looks similar to the definition of a homogeneous function h_d of degree d

$$h_d(\lambda x) = \lambda^d h_d(x), \quad \forall \lambda > 0.$$

Thus, the graph of the function $C(x;p)$ near any point x_0 is repeated in scaling form near all points $p^k x_0$, $k \in \mathbb{Z}$. This scaling (self-affine) property was often attributed to fractal features of the graph. However, this discrete scaling property is the main property of the so-called parametric-homogeneous (PH) functions introduced by Borodich (1992, 1994) which strictly satisfy the following equation

$$b_d(p^k \mathbf{x};p) = p^{kd} b_d(\mathbf{x};p), \quad k \in \mathbb{Z}$$

where d is degree of homogeneity. The graphs of these functions can be both continuous and discontinuous, they can also be smooth, piecewise smooth, with singular points of growth, fractal, non-fractal nowhere differentiable (Borodich, 1995, 1997). As examples of 1-dimensional fractal PH-curves we can consider the graphs of functions b_1 and b_2 with degrees $d = 1$ and $d = 2$, respectively (Figs 2 and 3)

$$b_1(x;p) = x b_0(x;p), \quad b_2(x;p) = x^2 b_0(x;p), \quad b_0(x;p) = x^{-\gamma} C(x;p). \tag{12}$$

Because of (6) these functions have the same Hausdorff dimension as the WM function $C(x;p)$, whose box-dimension is D.

The above examples (12) of PH-functions show that the trend of a function is not necessarily connected to fractal dimension. Indeed, both function b_1 and b_2 have the same Hausdorff dimension. However, if we consider the trend of the functions then we get that b_2 is a self-affine function with the exponent $H = 2$ and b_1 is a self-similar function. Thus, the common statement, that the scaling exponent H of self-affine fractals is always closely related to the fractal dimension, is wrong.

Another consequence is that the WM function $C(x;p)$, which has trend $C(x;p) \sim x^{2-D}$, can be used only as an example of fractal profile and it cannot be considered as the general fractal functional model for simulations of the rough surface profiles. The assumption that the WM function represents the general fractal properties of rough profiles can lead to wrong conclusions concerning surface roughness parameters and their distributions.

3. Multilevel hierarchical profile

In this section we present a new model of self-affine profile which can be used for roughness description. It has a hierarchical structure, and we call it a multilevel profile. The profile is constructed using a certain recursion procedure. Note that a number of classical fractals are created in such a way. Examples are the Cantor dust (the Cantor discontinuum), the von Koch curve, the

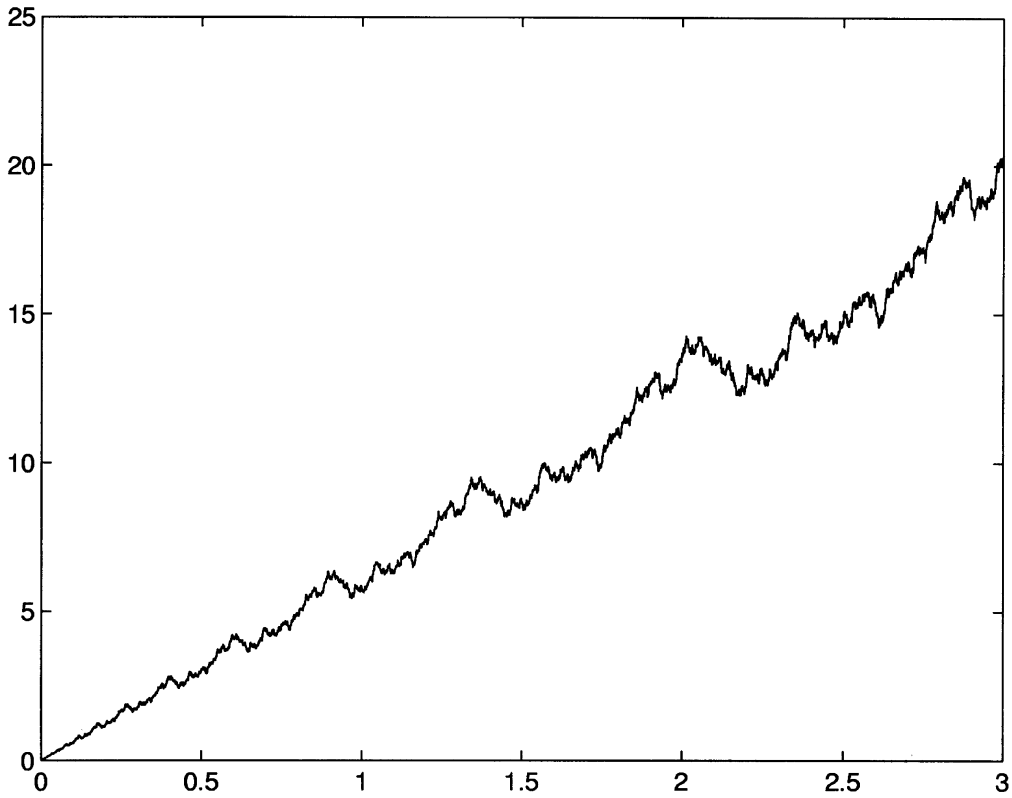

Fig. 2. Graph of a PH-function b_1 of degree $d = 1$ in the range $0 \leqslant x \leqslant 3$; $p = 1.5$, $\gamma = 0.5$. The trend of the function is $\sim x$.

Sierpinski carpet, and the techniques of iterated function systems which were recently introduced as a unified way of generating a broad class of fractals (Barnsley and Demko, 1985; Barnsley, 1988).

3.1. Construction of multilevel profile

We denote the presented multilevel profile by $\mathscr{F}$. We construct it in a step-by-step manner (see Fig. 4). At every step i, $i = 1, 2, \ldots$, we add several rectangles ('blocks' B_i for further reference) to the structure obtained at the previous step. We will refer to these intermediate structures as i-th generation prestructures and denote them by $S_1^{(i)}$. The contours of the prestructures are in some sense i-th iterations for the required multilevel profile $\mathscr{F}$. Thus, the profile itself is the 'contour' of the final structure, which includes the contributions given by all subsequent construction steps beginning from the first one and continuing to ad infinitum. Note, that under this iterative construction some points can be temporarily included in the profile at one step and then be removed from it at another step.

We will use the notation $\mathscr{F}$ in two senses, namely (i) as a 2-D-region that we obtain in normal

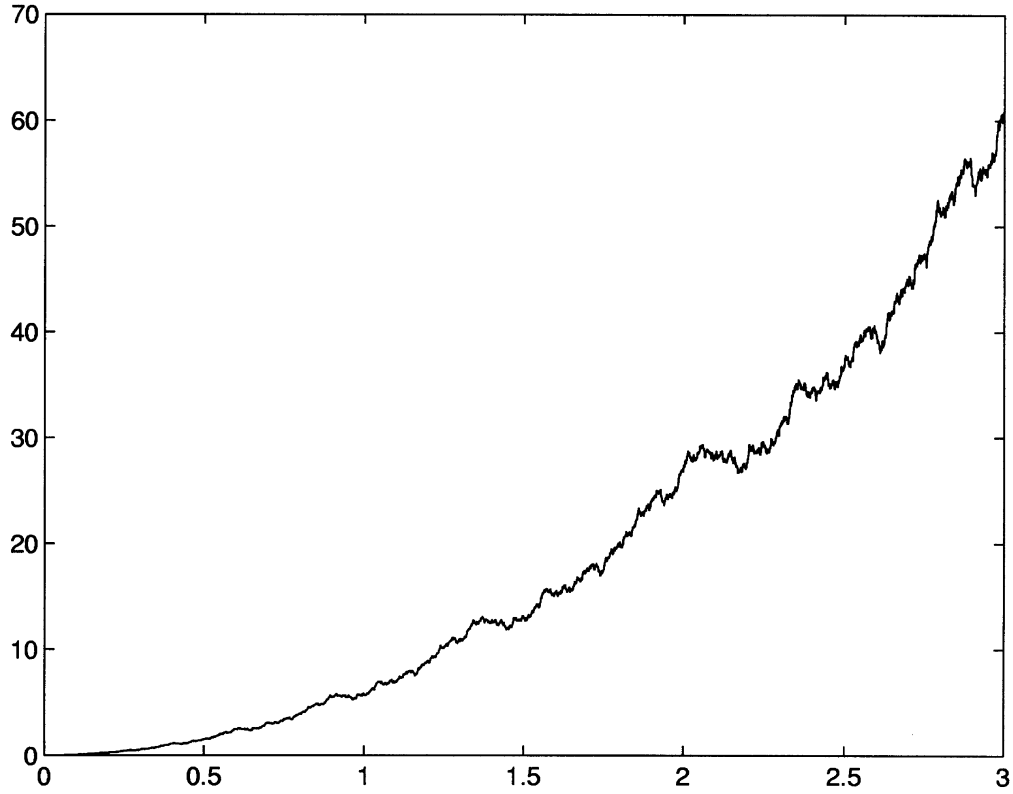

Fig. 3. Graph of the Weierstrass–Mandelbrot parabola b_2 (after Borodich, 1995) in the range $0 \leqslant x \leqslant 3$; $p = 1.5, \gamma = 0.5$. The trend of the function is $\sim x^2$.

section of rough solid, and (ii) as a 1-D-contour that is the boundary of the 2-D-region. We will analyse properties of both of them and will use the same name 'profile $\mathscr{F}$'. We believe this cannot result in any misunderstanding, because the object will be uniquely defined by context.

Let us describe the construction procedure in detail. With l_0 as a base length of the profile, consider a segment of length l_0 located along an axis Ox with the point O in the middle (see Fig. 4). Let us take some value h_0, which we will connect later with the total height H^* of the profile, and consider a block B_0 with length l_0 and height h_0. This is an initial element for constructing our structure, but it itself is not included in it. We will call this object and its values an initiator in further references. Then, we fix two structural parameters of the model, α and β such that

$$0 < \alpha < 1/2, \quad 0 < \beta < 1. \tag{13}$$

At the first step of the construction, we divide the segment l_0 into three parts with lengths αl_0, $(1-2\alpha)l_0$, and αl_0, correspondingly, and set a block B_1 of height $h_1 = \beta h_0$ and length $(1-2\alpha)l_0$ on the middle part of the initial segment l_0. This primitive structure, which contains only the block B_1, is the prestructure of the order one or the first-generation prestructure. Its contour is a continuous broken line consisting of two vertical segments of length h_1 and $3^1 = 3$ horizontal

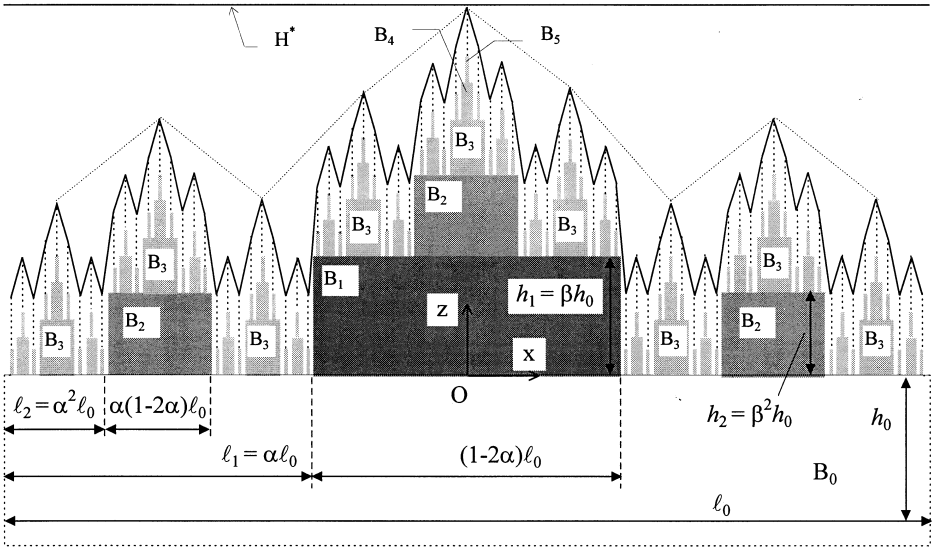

Fig. 4. Step-by-step construction of the multilevel profile and its contour passing through true peaks after third (- - -) and fifth (—) iterations of the construction.

segments, two of which having length αl_0 are located at the level $z = 0$, while the third is located at the top of the block B_1 and, hence, has length equal to $(1-2\alpha)l_0$. Note that the vertical segments belong to the final contour.

Next we apply a similar operation to each of the three horizontal segments. We set three blocks B_2 in the middle of these horizontal segments. The heights of the blocks B_2 are identical and are equal to $h_2 = \beta h_1$. However, their lengths are not identical. By construction, we have the value $\alpha(1-2\alpha)l_0$ as length for the new left and right blocks and $(1-2\alpha)^2 l_0$ as the length for the lifted middle block. Thus, the second-generation prestructure contains one block B_1 of the first-order and three second-order blocks B_2. The contour of this prestructure is a broken line consisting of $3^2 =$ nine horizontal segments, as well as eight vertical segments; two are of height h_1 and the other six are of height h_2. Note again that all vertical segments are part of the final profile.

Repeating the procedure in the same way, we add at every step $n_i = 3^{i-1}$ blocks B_i of height

$$h_i = \beta h_{i-1} = \beta^i h_0,$$

where i is the number of the corresponding generation. Thus, we obtain for any m, $m = 1, 2, \ldots$, the m-th prestructure $S_1^{(m)}$ that consists of blocks B_i, $i = 1, \ldots, m$, of which there are exactly n_i copies of each of them. It is then evident that $\mathscr{F} \equiv S_1^{(\infty)}$.

The lengths of the blocks may differ as we pointed out above. However, it is easy to prove that in any case these lengths lie within the range $[\alpha^i l_0, (1-2\alpha)^i l_0]$. Moreover, we see that if $\alpha = 1/3$, all the lengths coincide and are equal to

$$l_i = \alpha l_{i-1} = \cdots = \alpha^i l_0.$$

Figure 4 shows an example of the profile construction when $\alpha = 1/3$ and $\beta = 0.682$. All blocks

up to order five, which form the prestructure $S_1^{(5)}$, as shown in grey. The vertical dashed segments designate the rest of the final structure. The heights of these segments are calculated as limits in the infinite construction procedure. Thus, their tops are points of the final contour, i.e., they are 'true' peaks.

It follows from this observation that we have at least two opportunities for obtaining approximations to the final contour after ith iteration. With such an approximation we may take either (i) the contour of i-th prestructure, or (ii) the line resulted by the consecutive connection of neighbouring true peaks. With this scheme (Fig. 4), the latter is shown with dotted lines after the third iteration and with solid lines after the fifth iteration. By construction, the two variants lead to the same limit, namely $\mathscr{F}$. Therefore, both approximations can be used in the analysis of profile properties. For example, the former is very convenient for the calculation of geometric parameters, while the latter looks quite attractive from the viewpoint of a possible graphic roughness simulation.

3.2. Geometric characteristics and properties of the profile

Let us study now some geometric properties of the profile. In the calculation we will use the sign '$\equiv$' as identity by definition.

3.2.1. Calculation of the heights of structural parts of the profile

Here we calculate the heights of some structural parts involved in the construction of the profile.

First, we determine the height $H_1^{(m)}$ of an m-th order prestructure, denoting by $\mathscr{H}\{\cdot\}$ the height of the object in the brackets:

$$H_1^{(m)} \equiv \mathscr{H}\{S_1^{(m)}\} = \sum_{j=1}^{m} h_j = \sum_{j=1}^{m} \mathscr{H}\{B_j\} = \beta \frac{1-\beta^m}{1-\beta} h_0, \quad m = 1, 2, \ldots. \tag{14}$$

Hence, we obtain the maximal height (H^*) of the profile as

$$H^* \equiv H_1^{(\infty)} \equiv \mathscr{H}\{\mathscr{F}\} = \sum_{j=1}^{\infty} h_j = \frac{\beta}{1-\beta} h_0. \tag{15}$$

This gives us the relation between the height H^* of the profile and the height h_0 of the initiator.

Further, let us denote by F_j, $j = 1, 2, \ldots$, the substructures of $\mathscr{F}$, where every F_j rests on the block B_{j-1} located along the central axis Ox. Each F_j can be obtained by applying exactly the same construction procedure as above. Evidently, different initiators should be used. For convenience, we will use the notation F_1 for $\mathscr{F}$ as well. It will be shown below that F_j can be treated as scaled copies of $\mathscr{F}$.

To find the height of F_i we need first to find the height $H_i^{(m)}$ of the prestructure resting on the block B_{i-1}

$$H_i^{(m)} \equiv \mathscr{H}\{S_i^{(m)}\} = \sum_{j=i}^{m} h_j = \beta^i \frac{1-\beta^m}{1-\beta} h_0, \quad i = 1, 2, \ldots; \quad m = 1, 2, \ldots. \tag{16}$$

Since F_i is the final structure obtained from $S_i^{(m)}$ as $m \to \infty$, its height $H_i^{(\infty)}$ may be found from (16) as

$$H_i^{\infty} \equiv \mathscr{H}\{F_i\} = \frac{\beta^i}{1-\beta} h_0, \quad i = 1, 2, \ldots \tag{17}$$

which may be rewritten in the following form

$$H_{i+1}^{(\infty)} = \beta H_i^{(\infty)}, \quad i = 1, 2, \ldots. \tag{18}$$

Obviously, we also have

$$H_i^{(\infty)} = \beta^{i-1} H_1^{(\infty)}, \quad i = 1, 2, \ldots. \tag{19}$$

3.2.2. Self-similar and self-affine properties of the multilevel profile

In this subsection we reveal certain features of self-similarity and self-affinity of the profile. First, it can be easily seen that for any integer m, $m > 0$ the structure $\mathscr{F}$ if treated as a 'jigsaw puzzle' can be formed as a union of the prestructure $S_1^{(m)}$ and a set of scaled copies F_{m+1} of the entire structure $\mathscr{F}$, where the substructures F_{m+1} are 'installed' on the tops of m-th order blocks, as well as just near them, symmetrically on the left and on the right. In fact, the values of the initiators for $\mathscr{F}$ are multipliers with respect to the lengths and heights of all blocks that form the substructures. Hence, using certain factors, we can equalise the initiators of any substructure and the whole structure $\mathscr{F}$. Doing so, we automatically obtain the same factors for all corresponding blocks. This means that the substructure is similar to the $\mathscr{F}$ from the viewpoint of affinity.

To be more precise, we may assert that every copy F_{m+1}, $m = 1, 2, \ldots$, may be generated from the whole structure $\mathscr{F} \equiv F_1$ by applying a certain scale factor s_v in the vertical direction, which will be called the v-factor, and another scale factor s_h in the horizontal direction, which will be called the h-factor. In such a case we write, by convention

$$F_{m+1} \sim (s_h, s_v) F_1. \tag{20}$$

By construction, we have the common v-factor for all members of the set $\{F_{m+1}\}$

$$s_v = \beta^m, \tag{21}$$

while the h-factors are different in the general case, i.e., when $\alpha \neq 1/3$, and are given by the following relations

$$s_h = \alpha^{m-j}(1-2\alpha)^j, \quad j = 0, 1, \ldots, m. \tag{22}$$

Note, that all members of the set $\{F_{m+1}\}$ are of the same shape and size only when

$$\alpha = 1/3. \tag{23}$$

In this case, the h-factors (22) are

$$s_h = (1/3)^m, \tag{24}$$

for F_{m+1}, $m = 1, 2, \ldots$.

Thus, if (23) does not hold, the set $\{F_{m+1}\}$ contains scaled copies of $\mathscr{F}$ with various scale factors. For example, all substructures F_j, $j = 2, 3, \ldots$, which are located along the central axis Oz, have the following h-factors:

$$s_h = (1-2\alpha)^{j-1}$$

and those, which are located symmetrically at the left and right ends of the initial segment l_0, have the h-factors

$$s_h = \alpha^{j-1}.$$

Recall now Mandelbrot's definition (1983) that a bounded set S on a plane is said to be self-affine with respect to a ratio vector (s_h, s_v) is S is a union of N non-overlapping subsets $S_1, \ldots, S_N$ such that they all are scaled copies of the S.

Obviously, the structure $\mathscr{F}$ is not self-affine in the sense of this definition, and furthermore it is not even self-similar in the 'very regular' case $\alpha = \beta = 1/3$, because to collect the whole structure, except scaled copies F_k we need additional blocks B_i. However, we see that 'principal' parts of the $\mathscr{F}$, namely, substructures F_m are exactly scaled copies of $\mathscr{F}$. Moreover, it is easy to show that if we slightly modify the structure, building it downward ad infinitum by an extrapolation procedure, then this new unbounded structure $\hat{\mathscr{F}}$ can be considered to be self-affine with respect to a scale vector $(1-2\alpha, \beta)$ according to the usual geometric notion of affine transformation

$$\hat{\mathscr{F}} \sim (1-2\alpha, \beta)\hat{\mathscr{F}}. \tag{25}$$

Note that $\hat{\mathscr{F}}$ includes as its parts F_i, $i = \ldots, 2, 1, 0, -1, \ldots$.

Evidently, the similarity features can be useful for the analysis of geometric properties of the structure, and below we give some examples in which we employ this property. The first immediate corollary of self-affinity is the following relation

$$\mathscr{V}\{F_{i+1}\} = \alpha\beta\mathscr{V}\{F_i\}, \quad i = 1, \ldots \tag{26}$$

where for a given j the most left (the most right) among all $\{F_j\}$ should be substituted into (26). Here we have denoted by $\mathscr{V}\{\cdot\}$ the area bounded by the contour of a structure in brackets.

Note that relation (26) can be called a renormalization group (RG) relation because it connects values of a certain characteristic (an area here) of the profile for adjacent scale levels of consideration and is invariant with regard to scaling (for details see Borodich and Onishchenko, 1997).

3.2.3. *Behaviour of the heights of the substructures* F_i *for different* i

In problems of contact between rough surfaces, it is important to characterize the tops of highest asperities (see, e.g., Borodich and Onishchenko, 1993). Therefore, the question arises whether only the central 'column' of our profile is involved with the contact or will other substructures also contribute to it?

To answer this question, consider the relation between the heights of the structures F_i. It follows from (19) that for arbitrarily large integer k, we always can choose β near one such that the values $H_2^{(\infty)}$, $H_3^{(\infty)}, \ldots, H_k^{(\infty)}$ are near the level H^*. This means that by 'tuning' β we may obtain an arbitrary number of asperities of the profile with approximately equal heights.

It is useful to compare the heights of scaled copies of the structure with the heights of its m-th generations. It follows from (14) and (16) that if the following inequalities hold

$$\beta\frac{1-\beta^{k-1}}{1-\beta}h_0 < \frac{\beta^2}{1-\beta}h_0 \leqslant \beta\frac{1-\beta^{k}}{1-\beta}h_0 \tag{27}$$

or, after division by a common factor

$$1-\beta^{k-1} < \beta \leqslant 1-\beta^{k} \tag{28}$$

then the peak (top point) of F_2 lies in the gap between the tops of the prestructures of generations $k-1$ and k. We can write this in such a form

$$\mathscr{H}\{S_1^{(k-1)}\} < \mathscr{H}\{F_2\} \leqslant \mathscr{H}\{S_1^{(k)}\}. \tag{29}$$

Consider now the case of equality in (28), i.e., $\beta = 1-\beta^k$ and rewrite this as

$$\beta^k = 1-\beta. \tag{30}$$

It follows then from (29) that

$$\mathscr{H}\{F_2\} = \mathscr{H}\{S_1^{(k)}\}. \tag{31}$$

In other words, the height of the second-order structure F_2 is equal to the height of the k-th prestructure $S_1^{(k)}$.

3.2.4. Volumes of the structures $\mathrm{F_i}$

Let us now calculate the volumes per unit depth, or areas of the structures $F_i, i = 1, 2, \ldots$, for the case $\alpha = 1/3$. First, due to the scaling property (20) and using (21) and (24), we obtain

$$\mathscr{V}\{F_i\} = (\alpha\beta)^{i-1}\mathscr{V}\{F_1\} \tag{32}$$

and, in particular, for $i = 2$

$$\mathscr{V}\{F_2\} = \alpha\beta\mathscr{V}\{F_1\}.$$

On the other hand, by construction (see Fig. 5)

$$\mathscr{V}\{F_1\} = \mathscr{V}\{B_1\} + 3\mathscr{V}\{F_2\}.$$

Whence, using notation

$$V_i \equiv \mathscr{V}\{F_i\}; \quad v_0 \equiv \mathscr{V}\{B_0\} = l_0h_0$$

we find that

$$V_2 = \frac{(\alpha\beta)^2}{1-3\alpha\beta}v_0$$

and, consequently, that

$$V_1 = \frac{\alpha\beta}{1-3\alpha\beta}v_0; \quad i = 1, 2, \ldots. \tag{33}$$

It follows from (32) that

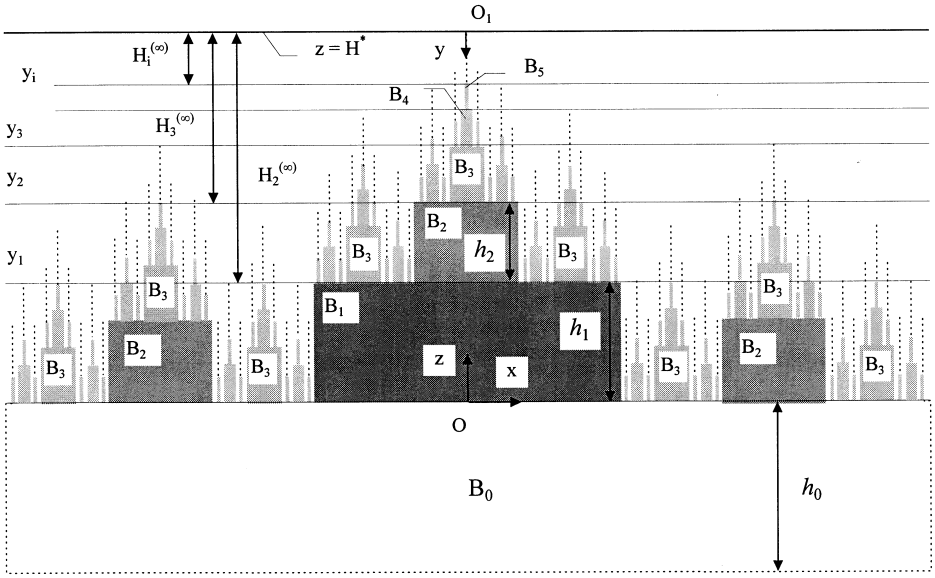

Fig. 5. Geometrical scaling relations between various structural parts of the multilevel profile; $\alpha = 1/3$, $\beta = 0.682$.

$$V_i = (\alpha\beta)^{i-1} V_1.$$

Taking $\alpha = 1/3$, we finally obtain

$$V_i \equiv \mathscr{V}\{F_i\} = (\beta/3)^{i-1} V_1 \tag{34}$$

where V_1 is given by (33).

3.3. Fractal dimension of the profile

Throughout what follows, we consider only the case $\alpha = 1/3$ unless otherwise is stipulated, using the notation α for the sake of convenience only. Keeping in mind that the 'length' of a fractal line has to be infinite, let us calculate the contour length $L^{(m)}$ of m-prestructure $S_1^{(m)}$. As we know from the above, $S_1^{(m)}$ consists of blocks B_i, $i = 1, \ldots, m$, and there are exactly $n_i = 3^{i-1}$ copies of each of them. The contour $L^{(m)}$ is the union of horizontal and vertical segments. The sum of the lengths of all horizontal segments is equal to l_0. The vertical segments forming the contour are the vertical sides of the blocks B_i. Therefore,

$$L^{(m)} = l_0 + \sum_{i=1}^{m} 2n_i h_i.$$

By simple calculations, we obtain

$$L^{(m)} = \begin{cases} l_0 + 2\beta h_0[1-(3\beta)^m]/(1-3\beta), & \beta \neq 1/3 \\ l_0 + 2\beta m h_0, & \beta = 1/3. \end{cases} \tag{35}$$

We see that the contour length of prestructures is limited when $\beta < 1/3$ and is unlimited (diverges) in the opposite case. Moreover, when

$$\beta > 1/3 \tag{36}$$

the contour length tends to infinity in accordance with the power law. Thus, only in case (36) we have some reasons to assume the multilevel self-affine profile to be fractal.

Now we will try 'to measure' the length of the contour of the structure $\mathscr{F}$ using dividers with continually reduced openings. Let $\delta = l_i = \alpha^i l_0$ be the opening of the dividers. Following the engineering approach (the Richardson method), we calculate the number of dividers steps, when 'walking' along all the contour of the prestructure, with dependence on the prestructure order m and then express the contour length (35) by means of δ. It can be easily seen under restriction (36) that

$$L^{(m)} \sim C(\delta/l_0)^{-d_*}, \quad \text{where } d_* = \frac{\ln(3\beta)}{\ln(1/\alpha)}$$

and C is a certain const. So, the value

$$D_R = 1 + d_* = 1 + \frac{\ln(3\beta)}{\ln(1/\alpha)}$$

can be considered as the Richardson dimension of the structure $\mathscr{F}$. For $\alpha = 1/3$ we get

$$D_R = 2 - \frac{\ln(1/\beta)}{\ln(3)}.$$

Next we determine the fractal box dimension D_B of the profile $\mathscr{F}$. Let us recall that we consider the case $\alpha = 1/3$. The box dimension may be found from relation (5), where $N(\delta)$ denotes the number of the non-overlapping squares (boxes) $\delta \times \delta$ that jointly cover all the contour.

We consider the entire structure $\mathscr{F} \equiv S_1^{(\infty)}$ as the union of the prestructure $S_1^{(m-1)}$ and its scaled copies F_m of order m. The number of the latter copies equals n_m. Take δ as

$$\delta = l_{m+1}. \tag{37}$$

Evidently, the contour of the whole structure $\mathscr{F}$ is covered if we cover both vertical parts of the prestructure boundary and all the scaled substructures. To cover the former we need, in excess,

$$n_1 = \sum_{j=1}^{m} 2n_{j-1}\left(\frac{h_j}{\delta} + 1\right) \sim C_1(9\beta)^m \tag{38}$$

boxes. Then to cover the latter we need, also in excess,

$$n_2 = n_m\left(\frac{H_{m+1}}{\delta} + 1\right) \sim C_2(9\beta)^m$$

boxes. Finally, we obtain

$$N(\delta) = n_1 + n_2 \sim C_3(9\beta)^m.$$

Taking into account (37), we obtain that (5) holds with

$$D_{\mathrm{B}} = 2 - \frac{\ln(1/\beta)}{\ln(3)} \tag{39}$$

which coincides with the value D_{R}.

Since the number n_1 (38) is, in fact, a lower bound for the adequate number of the boxes to cover all the structure $\mathscr{F}$, we obtain that (39) is the exact value for the box dimension $\dim_{\mathrm{B}}\mathscr{F}$ for the profile $\mathscr{F}$. We see that $D_{\mathrm{B}} \to 1$ as $\beta \to 1/3$, and $D_{\mathrm{B}} \to 2$ as $\beta \to 1$. Note that it is not difficult to show that the profile box dimension D_{B} is one when $\beta \leqslant 1/3$.

3.4. Upper volume curve and bearing curve

There are a lot of engineering parameters which are used for roughness characterization (Greenwood, 1992). Therefore, it is important to calculate and analyse similar characteristics for mathematical models describing roughness.

At any finite construction iteration m, when m is large enough, the contour is a very irregular but also a very simple curve at the same time. Indeed, it is a broken line with only right angles. In this situation, both analytical and numerical calculations of the various geometric parameters should not be difficult. Some results of these calculations will be presented in further publications. In this paper, we give an example of constructing two functions characterizing roughness of the multilevel profile. The first of them is a function which has a value at a point z, $0 \leqslant z \leqslant H^*$ that is equal to the area (the volume in 3-D problem) of that part of the profile which lies above a horizontal slice at the level z. We will refer to the graph of this function as the 'upper volume' curve. Another function is called the bearing area (Abbot) curve. Its value at a point z is equal to the length (the area in 3-D problem) of the slice of the profile at the level z (Greenwood, 1992).

For convenience, we introduce another vertical axis O_1y, which is directed downwards and has its origin at the point O_1 coinciding with the top of the structure $\mathscr{F}$ (Fig. 5). Obviously, for the points of the profile, we have $0 \leqslant y \leqslant H^*$. We denote by $V(y)$ the values of the upper volume function and by $A(y)$ the values of the bearing curve.

We restrict ourselves with the case when β satisfies the relation (30) for certain $k \geqslant 1$. It is not difficult to show, using the self-affinity property (25) provided $\alpha = 1/3$, that the following scaling relation holds

$$\mathscr{V}\{\mathscr{F}; \beta y\} = (\beta/3)\mathscr{V}\{\mathscr{F}; y\} \tag{40}$$

for all y such that

$$y \leqslant y_k = H^* - (h_1 + \cdots + h_k).$$

The last condition is caused by the necessity to avoid the influence of the second-order 'asperities' F_2, which are located symmetrically on the left and on the right from the central column, on the characteristics of the first-order asperity $\mathscr{F}$. Let us recall that due to (31) the height of the F_2 is exactly the value $h_1 + \cdots + h_k$ for the case under consideration (see Fig. 5, where $k = 3$).

Denote by y_i, $i = 1, 2, \ldots$ the distances from the tops of i-th order prestructures to the top of the profile

$$y_i = H^* - H_1^{(i)} = \beta^i H^* \tag{41}$$

where (15) and (14) have been used. Further, if we take a constant C_V such that

$$C_V = \frac{\mathscr{V}\{\mathscr{F}; y_k\}}{(\beta/3)^k V_1}$$

then using (40) and (41) we obtain

$$\mathscr{V}\{\mathscr{F}; y_i\} = C_V(\beta/3)^i V_1, \quad i = k, k+1, \ldots . \tag{42}$$

It follows from this relation that

$$V(y_i) = C_V V_1 \left(\frac{y_i}{H^*}\right)^\gamma, \quad i = k, k+1, \ldots \tag{43}$$

where

$$\gamma = 1 + \frac{\ln(3)}{\ln(1/\beta)}, \quad \beta^k = 1 - \beta.$$

For example, with the cases $k = 1, 2, 3$ the exponent γ equals 2.58, 3.28 and 3.87, respectively.

Evidently, the 'spline'

$$V(y) = C_V V_1 \left(\frac{y}{H^*}\right)^\gamma \tag{44}$$

is quite appropriate in the considered case. However, it should be emphasized that we have proved (43) only for the discrete set $\{y_i, i = k, k+1, \ldots\}$ of argument values. This yields that for other argument values, relation (43) should be treated as approximate unless otherwise proved.

With regard to the constant C_V of (44), it must be noted that it is possible to find it analytically for small values of k by some direct manipulation of the profile's structural parts as if they were parts of a rather complicated jigsaw puzzle, and using in addition the self-affinity properties (see Borodich and Onishchenko, 1997). However, we could not find any general algorithms for such manipulation when k in (30) is arbitrary. It seems more difficult to find C_V analytically in the case when (30) does not hold for any integer k.

Nevertheless, C_V can be easily computed with the help of any suitable algorithm taking into account the regular structure of the profile $\mathscr{F}$. One of the possible algorithms is described below.

With regards to the bearing curve, evidently, we have

$$\delta V(y) = A(y)\delta y \tag{45}$$

where δV and δy are increments of the corresponding variables. If the function $V(y)$ were differentiable, we could find $A(y)$ as the derivative of $V(y)$. However, firstly, this point is a rather questionable a priori and, secondly, even if it were so, we do not have an analytical form for the function $V(y)$.

It should be mentioned that in principal, one could calculate analytically the values of the function $A(y)$ in a way similar to that as we employed above with respect to the upper volume function. However, keeping in mind our intention to present techniques suitable for analysing the

profile in a general case, i.e., for an arbitrary value of scaling factor β, we prefer to describe how a relevant algorithm that allows us to calculate numerically the values of both functions $A(y)$ and $V(y)$.

The main idea is to analyse prestructures of the multilevel profile rather than the final multilevel profile itself giving the necessary estimations of the accuracy. Let $S_1^{(m)}$ be a prestructure of the m-th order, provided m, $m > 1$ is the number of the final iteration for the prestructure construction. Consider the base segment l_0 and points x on it such that

$$x = i_1(1/3)^1 + i_2(1/3)^2 + \cdots + i_m(1/3)^m \tag{46}$$

where $i_j = 0$ or $i_j = 1$ for all j. Here we suppose that the point $x = 0$ corresponds to the left end of the base. Note that $N = 3^m$ is the total number of points given by (46).

As we already mentioned above, the contour of this prestructure is a broken stepwise line consisting of 3^m horizontal segments of equal length, as well as a certain amount of vertical segments of easily calculated heights. Under a given value y corresponding to the level of the slice, we use item-by-item examination of points from (46). Thus, we have for a current point x the condition of 'going down', i.e., when we reach the end of a local asperity slice, we add the easily calculated relevant amounts to the previously accumulated values for the functions $A(y)$ and $V(y)$ correspondingly.

It should be mentioned that by using relation (34) we could take into account not only the volumes of the upper parts of the sliced blocks, but also the scaled copies F_j supported by these j-th order blocks, where j, $j = 2, \ldots, m$. However, in order to obtain a simple error estimation, we deal here with the prestructures only. Therefore, using this algorithm for computation, we exclude from our consideration all blocks of order higher than m, i.e., all blocks B_i with $i > m$.

It can be easily checked that the prestructure $S_1^{(m)}$ differs from the structure $\mathscr{F}$ by the totality of scaled copies F_{m+1}. The number of these copies is 3^m. Using (34) we obtain the value

$$e^{(m)} = 3^m \mathscr{V}\{F_{m+1}\} = (\beta)^m \mathscr{V}\{F_1\}$$

as the maximal possible error for the upper volume function.

Because $e^{(m)} \to 0$ as $m \to \infty$ provided $\beta < 1$, we may assert that the algorithm described is quite suitable for numerical calculation of the function $V(z)$. Of course, due to the limited capacities of the computer we will meet specific problems when β is very close to one. Definitely, this is a special case requiring a special approach.

The analytical estimation of the calculation error in the case of the bearing curve is not as straightforward as for the upper volume function. However, it may be obtained directly from the calculations. We may compare the computed values of the bearing curve obtained with increasing m and roughly estimate the error by the difference between values obtained at subsequent iterations.

As an example, we give the computational results for the case $\beta = 0.682$ [$k = 3$ in the relation (30)]. We carried out the calculations up to $m = 13$, that is for the prestructure of 13th-order. The number of calculation points z determining slice levels is equal to 500. The graphs are presented in Fig. 6. Note that any visual distinctions between successive iterations disappear as soon as $m = 9$, with the exception of the $A(z)$ function graphs for the small z range, where the theoretical value of the $A(z)$ must be equal to 0. However, this part of the bearing curve is not important for the application to contact problems.

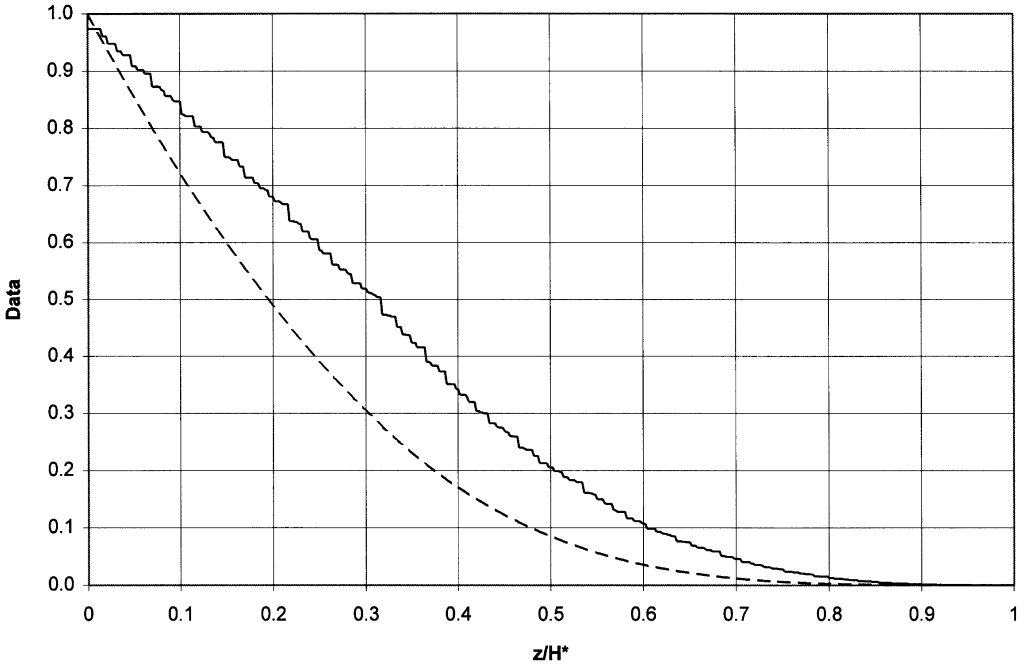

Fig. 6. The bearing curve $A(z)$ (—) and the upper volume curve $V(z)$ (- - -) for the multilevel profile; $\alpha = 1/3$, $\beta = 0.682$.

We see that the graph of the $V(z)$ is reasonably smooth, whilst the $A(z)$ graph is very 'rough', and if turned at an angle of 90°, looks similar to the so-called Cantor staircase—a curve which is almost everywhere horizontal, but has infinitely many infinitesimal vertical jumps. Of course, a more thorough analysis should be carried out, including the estimation of the possible influence of the numerical algorithm itself. Moreover, it would be very interesting to analyse relation (45) in a numerical way.

The computation time for the calculation of the coordinate of points of the graphs was about 10 min on a 486-DX2 PC. We expect a substantial reduction in the computation time for a desired accuracy, using refined algorithms.

4. Discussion and conclusion

Mandelbrot (1983) noted that natural objects do not have pure shapes of classical mathematical objects—for example, coastlines are not circles, clouds are not spheres, and mountains are not cones. We can add to this list a statement that 'roughness of real bodies is not a mathematical fractal'. All these geometrical objects: spheres, cones, circles as well as fractals are only mathematical idealizations of complex shapes of real physical bodies. We have to use different mathematical models to describe the same object on different scales. For example, to describe a nominally flat rough surface, one could use sequentially the model of a plain, the model of a

collection of spherical asperities and the model of a fractal surface depending on chosen magnification.

The first two models are customary, while the idea to use models of fractal surfaces for describing natural bodies has been developed only recently, after Mandelbrot published his papers. However, we have seen above that the employment of the fractal approach in the study of surfaces has several drawbacks. We list only several of them:

(i) It is often unclear how we can use results obtained for mathematical fractals in applications to physical fractals.
(ii) Various methods of determination of fractal dimension of real surfaces sometimes give rather different results, while mathematically these methods must be equivalent to each other. To check the accuracy of the methods, curves with analytically calculated dimensions are used (Dubuc et al., 1989; Tricot et al., 1994; Talibuddin and Runt, 1994).
(iii) It is known that the WM function C [see Fig. 1 and (10)] has trend $\sim x^{2-D}$; therefore, the use of this function as the standard model for the simulations of the rough surface profiles can lead to wrong conclusions concerning surface parameters and their distributions.

The existence of fractal behaviour in physical systems is reported so often that one can speak about 'fractals everywhere' (see, e.g., Barnsley, 1988). The question "Why are fractals so common in Nature?" was recently studied by Avnir et al. (1997). Their extended analysis of published data, in particular concerning rough surfaces, shows that the overwhelming majority of reported physical fractals span approximately 1.5 orders of magnitude. To give a possible interpretation for this observation, they considered various pure random numerical models, for example, a model in which rods of some specific length are randomly placed on the unit interval. It is shown that the characteristic relation (5) holds for the structures generated in their random models approximately over the over the above physically meaningful range. They point out that such properties of random structures can be attributed to apparent fractality. Finding that randomness obeys the dilation similarity, they conclude that 'fractals everywhere' may be caused by 'randomness everywhere'.

These results (Avnir et al., 1997) support the idea that the bounded similarity of surface profile structure can be a crucial feature for roughness. It means, in fact, that prefractals are more appropriate for roughness studies than mathematical fractals. We propose to study such a similarity on iteratively generated rough profiles which can be fractal only in the mathematical limit with respect to the number of construction iterations.

It should be emphasized that such models do not necessarily generate fractals. The multilevel profile presented in the paper is not a fractal for some values of the parameters, e.g., when $\alpha = 1/3$, $0 < \beta < 1/3$. However, it remains rough and possesses certain self-affine properties.

This new multilevel profile model is suitable for the flexible modelling of roughness. It can be easily simulated and drawn, which is important for visual analysis. The profile's 'design' points can be either determined analytically or computed using any straightforward algorithm. We have given an example of such a simulation and analysed both analytically and numerically its geometrical properties. In addition, we have presented an example of the calculation and construction of the bearing curve and the upper volume function of the profile. These characteristics can be directly used in the problem of contact between a punch and an elastic foundation (see, e.g., Johnson, 1985; Borodich and Mosolov, 1991, 1992; Borodich and Onishchenko, 1993).

With regards to the fractal dimension of the profile, it would be very interesting to apply the variation method as discussed by Dubuc et al. (1989) to the digitized data of our model and compare the result with the analytical results given in Section 3.3. We intend to fulfil this in our next paper.

It seems to us that one of the most beneficial properties of the profile is its iterative regular construction, which allows us to analyse its prestructures (prefractals) of any generation. Moreover, we believe that the standard engineering parameters and functions used for the characterization of rough surfaces can be also calculated for the profile in an efficient way. The profile $\mathscr{F}$ is defined by four structural parameters, namely the basic length l_0, the height of the profile H^*, and the scaling parameters α and β, satisfying the restrictions given in (13). Therefore, the model can be tuned for comparison with various experimental data.

Acknowledgments

One of the authors D.A.O. would like to thank the Engineering and Physical Sciences Research Council, U.K. for a Visiting Fellowship Research Grant (GR/L69176) to enable him to carry out the research at the Department of Mathematics of Glasgow Caledonian University. He is also grateful to Professor J. Gomatam for the hospitality shown there. The authors are grateful to a referee for useful comments.

References

Annir, D., Biham, O., Lidar (Hamburger), D., Malcai, O., 1997. On the abundance of fractals. In: Novak, M.M., Dewey, T.G., (Eds.), Fractal Frontiers. World Scientific, Singapore, New Jersey, London, Hong Kong, pp. 199–234.

Barnsley, M., Demko, S., 1985. Iterated function systems and the global construction of fractals. Proc. R. Soc. Lond. A 399, 243–275.

Barnsley, M., 1988. Fractals Everywhere. Academic Press, Boston.

Berry, M.V., Lewis, Z.V., 1980. On the Weierstrass–Mandelbrot fractal functions. Proc. R. Soc. Lond. A 370, 459–484.

Bhushan, B., 1995. A fractal theory of the temperature distribution at elastic contacts of fast sliding surfaces—Discussion. Trans. ASME. J. Tribology 117, 214–215.

Blackmore, D., Zhou, J.G., 1996. A general fractal distribution function for rough surface profiles. SIAM J. Appl. Math. 56, 1694–1719.

Borodich, F.M., 1992. Self-similarity properties of static and dynamic contact between a punch and anisotropic media. In: Abstracts of Reports on 18th Int. Congress of Theoretical and Appl. Mech. Haifa, pp. 26–27.

Borodich, F.M., 1993. Similarity properties of discrete contact between a fractal punch and an elastic medium. C. R. Ac. Sc. (Paris), Ser. 2, 316, 281–286.

Borodich, F.M., 1994. Some applications of the fractal parametric-homogeneous functions. Fractals 2, 311–314.

Borodich, F.M., 1995. Parametric-homogeneous functions, similarity, and fractal function graphs. Technical Report TR/MAT/FMB/95-35, Glasgow Caledonian University, Glasgow, pp. 1–55.

Borodich, F.M., 1997. Parametric-homogeneity and non-classical self-similarity. I. Mathematical background. Acta Mechanica, to be published.

Borodich, F.M., Mosolov, A.B., 1991. Fractal contact of solids. Sov. Phys.-Tech. Phys. 61, 50–54.

Borodich, F.M., Mosolov, A.B., 1992. Fractal roughness in contact problems. J. Appl. Math. and Mech. (PMM) 56, 786–795.

Borodich, F.M., Onishchenko, D.A., 1993. Fractal roughness for problem of contact and friction (the simplest models). Soviet J. of Friction and Wear, No. 3, 452–459.

Borodich, F.M., Onishchenko, D.A., 1997. Multilevel profile with hierarchical structure: self-affinity, fractality and applications to contact problems. Technical Report TR/MAT/FMB-DAO/97-85, Glasgow Caledonian University, Glasgow, pp. 1–43.

Brown, S.R., 1995. Simple mathematical model of rough fracture. J. Geophys. Res. 100 (B4), 5941–5952.

Dubuc, B., Quiniou, J.F., Roques-Carmes, C., Tricot, C., Zucker, S.W., 1989. Evaluation the fractal dimension of profiles. Phys. Rev. A 39, 1500–1512.

Falconer, K.J., 1990. Fractal geometry: mathematical foundations and applications. John Wiley, Chichester.

Greenwood, J.A., 1992. Problems with surface roughness. In: Singer, I.L., Pollock, H.M. (Eds.), Fundamentals of Friction: Macroscopic and Microscopic Processes. Kluwer, Boston, pp. 57–76.

Greenwood, J.A., Williamson, J.B.P., 1966. Contact of nominally flat surfaces. Proc. R. Soc. Lond. A 370, 300–319.

Hansen, A., Plouraboué, F., Roux, S., 1995. Shadows in a self-affine landscape. Fractals 3, 91–98.

Hu, T.Y. Lau, K.S., 1993. Fractal dimensions and singularities of the Weierstrass type functions. Trans. Amer. Math. Soc., 335, 649–665.

Johnson, K.L., 1985. Contact Mechanics. Cambridge University Press, Cambridge.

Kaplan, T., Gray, L.J., Liu, S.H., 1987. Self-affine fractal model for a metal–electrolyte interface. Phys. Rev. B 35, 5379–5381.

Liu, S.H., 1985. Fractal model for the ac response of a rough interface. Phys. Rev. Lett. 55, 529–532.

Longuet-Higgins, M.S., 1957a. The statistical analysis of a random, moving surface. Phil. Trans. R. Soc. A 249, 321–387.

Longuet-Higgins, M.S., 1957b. Statistical properties of an isotropic random surface. Philos. Trans. Roy. Soc. A 250, 157–174.

Lopez, J., Hansali, G., Zahouani, Le Bosse, J.C., Mathia, T., 1995. 3-D fractal based characterization for engineered surface topography. Int. J. Mach. Tools Manufacture 35, 211–217.

Majumdar, A., Bhushan, B., 1990. Role of fractal geometry in roughness characterization and contact mechanics of surfaces. Trans. ASME J. Tribology 112, 205–216.

Majumdar, A., Bhushan, B., 1991. Fractal model of elastic–plastic contact between rough surfaces. Trans. ASME J. Tribology, 113, 1–11.

Majumdar, A., Tien, C.L., 1990. Fractal characterization and simulation of rough surfaces. Wear 136, 313–327.

Mandelbrot, B.B., 1977. Fractals: Form, Chance, and Dimension. Freeman, San Francisco.

Mandelbrot, B.B., 1980. Fractal geometry: what is it, and what does it do? Proc. R. Soc. Lond. A 423, 3–16.

Mandelbrot, B.B., 1983. The Fractal Geometry of Nature. Freeman, New York.

Mandelbrot, B.B., Passoja, D.E., Paullay, A.J., 1984. Fractal character of fracture surfaces of metals. Nature 308, 721–722.

Mauldin, R.D., Williams, S.C., 1986. On the Hausdorff dimension of some graphs. Trans. Amer. Math. Soc. 298, 793–803.

Modenov, P.S., Parkhomenko, A.S., 1965. Geometric Transformations. Vol. 1, Euclidean and Affine Transformations. Academic Press, New York and London.

Moreira, J.G., Kamphorst Leal da Silva, J., Oliffson Kamphorst, S., 1994. On the fractal dimension of self-affine profiles. J. Phys. A: Math. Gen. 27, 8079–8089.

Nayak, P.R., 1971. Random process model of rough surfaces. J. Lubric. Technol. (ASME) 93, 398–407.

Plouraboué, F., Roux, S., Schmittbuhl, J., Vilotte, J.-P., 1995. Geometry of contact between self-affine surfaces. Fractals 3, 113–122.

Roques-Carmes, C., Wehbi, D., Quiniou, J.F., Tricot, C., 1988. Modelling engineering surfaces and evaluating their non-integer dimension for application in material science. Surface Topography 1, 435–443.

Sayles, R.S., Thomas, T.R., 1978. Surface topography as a nonstationary random process. Nature 271, 431–434.

Schmittbuhl, J., Roux, S., Berthaud, Y., 1994. Development of roughness in crack propagation. Europhys. Lett. 28, 585–590.

Schmittbuhl, J., Schmitt, F., Scholz, C., 1995. Scaling invariance of crack surfaces. J. Geophys. Res. 100 (B4) 5953–5973.

Sun, X., Jaggard, D.L., 1990. Wave scattering from non-random fractal surfaces. Optics Communications, 78, 20–24.

Talibuddin, S., Runt, J.P., 1994. Reliability test of popular fractal techniques applied to small two-dimensional self-affine data sets. J. Appl. Phys. 76, 5070–5078.

Thomas, T.R., 1982. Rough Surfaces. Longman, New York.
Tricot, C., 1995. Curves and Fractal Dimension. Springer-Verlag, Berlin.
Tricot, C., Ferland, P., Baran, G., 1994. Fractal analysis of worn surfaces. Wear 172, 127–133.
Vandembroucq, D., Roux, S., 1997. Conformal mapping on rough boundaries. 1. Applications to harmonic problems. Phys. Rev. E 55 (5B), 6171–6185.
Voss, R.F., 1985. Random fractal forgeries. In: Earnshaw, R.A. (Ed.), Fundamental Algorithms in Computer Graphics. Springer-Verlag, Berlin, pp. 805–835.
Wang, S., Komvopoulos, K., 1995. A fractal theory of the temperature distribution at elastic contacts of fast sliding surfaces. Trans. ASME J. Tribology 117, 203–215.
Warren, T.L., Krajcinovic, D., 1995. Fractal models of elastic–perfectly plastic contact of rough surfaces based on the Cantor set. Int. J. Solids Structures 32, 2907–2922.
Warren, T.L., Krajcinovic, D., 1996a. A fractal model for the static coefficient of friction at the fiber–matrix interface. Composites. Part B—Engineering 27, 421–430.
Warren, T.L., Krajcinovic, D., 1996b. Random Cantor set models for the elastic–perfectly plastic contact of rough surfaces. Wear, 196, 1–15.
Warren, T.L., Majumdar, A., Krajcinovic, D., 1996. A fractal model for the rigid-perfectly plastic contact of rough surfaces. Trans. ASME J. Appl. Mech. 63, 47–54.