

HEAVY-TAILED, NON-GAUSSIAN NATURE OF TERRAIN AND ITS IMPLICATIONS FOR TERRAIN MODELING BY L_1 SPLINES

John E. Lavery
Mathematics Division
Army Research Office, Army Research Laboratory
Research Triangle Park, NC 27709

Shu-Cherng Fang
Department of Industrial Engineering and Operations Research
North Carolina State University
Raleigh, NC 27695

ABSTRACT

This paper presents the first step in establishing a link between the heavy-tailed nature of terrain and a new terrain modeling technique, L_1 splines, that is, splines based on minimizing the L_1 norm rather than the square of the L_2 norm. To establish this link, we focus on the heavy-tailed nature of the second derivatives that occur in the L_1 spline minimization principles. For one urban-terrain data set (Baltimore, Maryland) and two natural-terrain data sets (Killeen, Texas), the second derivatives behave asymptotically like $c|x-\mu|^{-\alpha}$ for $-\alpha$ between -4 and -2 rather than like exponential functions. Similar results for first derivatives minus first differences are presented. The distributions investigated here are not directly related to the spatial frequency spectra that have been the topic of most previous investigations of the heavy-tailed nature of terrain. The heavy-tailed nature of the frequency spectra of terrain has not resulted in any major impact on modeling of large terrain datasets (although it has had significant positive impact on modeling of vehicle-terrain interaction, where the data sets are local and smaller). The investigation of the heavy-tailed nature of the derivatives of terrain will have significant impact by providing the theoretical underpinnings for the current observation that L_1 splines provide better terrain modeling than alternative techniques.

1. INTRODUCTION

The self-similar, fractal nature of many types of terrain has been observed and extensively investigated (Mandelbrot, 1982; Peitgen et al., 2004). Fractally generated terrain can be quite realistic and can be attractively surrealistic, so much so that it is now commonly used in entertainment. However, the widespread use of fractal techniques for generating attractive artificial terrain has not been accompanied by use of fractal techniques for modeling real terrain. Fractal techniques are, in general, not able accurately to

represent huge terrain data sets in a data-compressed, computationally efficient manner and are, therefore, not often used by the terrain modeling community. Currently, the most common technique for producing terrain models from large terrain data sets is piecewise planar surfaces on triangulated irregular networks (TINs) (Mark, 1997). Piecewise planar surfaces on TINs are used by the widely used software packages DYN-TACS, Janus, ModSAF and Bresenham. Piecewise planar surfaces have the advantage of not introducing extraneous oscillation but the disadvantage of needing fine meshes so that the surfaces are accurate. The algorithms for calculating piecewise planar surfaces can be quite efficient but typically do not make direct use of the known characteristics of terrain, including its fractal nature.

Heavy-tailed distributions are distributions with tails that behave asymptotically like $c|x-\mu|^{-\alpha}$ (α positive) rather than like exponential functions $c \exp(-\beta|x-\mu|)$ (exponential distribution, β positive) or $c \exp(-\beta(x-\mu)^2)$ (Gaussian distribution, β positive). A heavy-tailed distribution is the statistical expression of self-similar, fractal phenomena (Dodge, 1987, 1992, 1997, 2002). More precisely stated, the power spectrum of the frequency components of a fractal phenomenon follows a heavy-tailed distribution. In the past, the heavy-tailed nature of the frequency spectrum of terrain has been used in modeling small regions of terrain to determine tire/track and vehicle response to the terrain. However, the heavy-tailed nature of terrain has not been of use in creating methods for modeling large areas of real terrain for visualization.

It is the purpose of this paper to present the first step in establishing a link between the heavy-tailed nature of terrain and a new terrain modeling technique, L_1 splines. To establish this link, we will focus on the heavy-tailed nature not of the frequency spectrum of terrain but rather of the derivatives of terrain.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 00 DEC 2004	2. REPORT TYPE N/A	3. DATES COVERED -	
4. TITLE AND SUBTITLE Heavy-Tailed, Non-Gaussian Nature Of Terrain And Its Implications For Terrain Modeling By L1 Splines		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mathematics Division Army Research Office, Army Research Laboratory Research Triangle Park, NC 27709; Department of Industrial Engineering and Operations Research North Carolina State University Raleigh, NC 27695		8. PERFORMING ORGANIZATION REPORT NUMBER	
		10. SPONSOR/MONITOR'S ACRONYM(S)	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
		12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited	
13. SUPPLEMENTARY NOTES See also ADM001736, Proceedings for the Army Science Conference (24th) Held on 29 November - 2 December 2005 in Orlando, Florida.			
14. ABSTRACT			
15. SUBJECT TERMS			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	
19a. NAME OF RESPONSIBLE PERSON			

2. HEAVY-TAILED NATURE OF SECOND DERIVATIVES USED IN L_1 SPLINES

Conventional splines are generated by minimizing the L_2 -norm-based functional

$$\iint_D \left[\left(\frac{\partial^2 z}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 z}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 z}{\partial y^2} \right)^2 \right] dx dy \quad (1)$$

Conventional splines are not used for terrain modeling because they have extraneous, nonphysical oscillation (Lavery, 2001). L_1 splines are surfaces $z = z(x,y)$ that are generated by minimizing the L_1 -norm-based functional

$$\iint_D \left[\left| \frac{\partial^2 z}{\partial x^2} \right| + 2 \left| \frac{\partial^2 z}{\partial x \partial y} \right| + \left| \frac{\partial^2 z}{\partial y^2} \right| \right] dx dy \quad (2)$$

(or related functionals, see (Lavery, 2001)) over C^1 -smooth piecewise cubic functions on rectangles or triangles that interpolate or approximate the given data. L_1 splines provide excellent terrain modeling free of the extraneous oscillation of conventional splines (Gilsinn and Lavery, 2002; Lavery, 2001, Lavery and Gilsinn, 2000) and free of the discontinuous derivatives of piecewise planar surfaces on TINs. At the 22nd and 23rd Army Science Conferences, L_1 splines were shown to be superior to conventional splines for urban terrain (Lavery and Gilsinn, 2000) and to be competitive with conventional splines for line of sight in natural terrain (Champion and Lavery, 2002). The rationale for the ability of L_1 splines to provide enhanced terrain representation has been outlined in (Lavery, 2001, Sec. 6). The ‘‘shape preservation’’ capabilities of univariate L_1 splines have been investigated on local scales (Cheng et al., 2004). However, previous results have not established the link between L_1 splines and the heavy-tailed nature of terrain.

For heavy-tailed distributions, minimization of L_1 norms (sums of absolute values) has been shown to be more accurate and robust for estimating the center of the distribution than minimization of (squares of) L_2 norms (sums of squares) (Dodge, 1987, 1992, 1997, 2002). However, this fact alone is not sufficient to establish the link between L_1 splines and the heavy-tailed nature of terrain, because the heavy-tailed nature of terrain, as the term has been used in the past, refers to the heavy-tailed nature of the power spectrum of the frequency components of terrain, not to the second derivatives that appear in functional (2). To establish the link that we desire, we must present evidence that the second derivatives that appear in functional (2) follow heavy-tailed distributions. To establish this conclusively for large classes of terrain is beyond the scope of this paper. However, we will establish this for one urban-terrain

data set and two natural-terrain data sets. The urban-terrain data set is a set of DTED5 (1m spacing) data with 1000^2 points for Baltimore, Maryland provided by the Joint Precision Strike Demonstration Project Office (JPSD PO) Rapid Terrain Visualization (RTV) ACTD. The natural-terrain data sets are a set of DTED2 (30m spacing) data with 901^2 points and a set of DTED1 (100m spacing) data with 1201^2 points, both for Killeen, Texas. These two data sets were provided by the National Geospatial-Intelligence Agency. The DTED5 data for Baltimore and the DTED1 data for Killeen were previously used in (Lavery and Gilsinn, 2000) and (Lavery, 2001), respectively.

The distribution of interest is the joint distribution of the second derivatives that occur inside the absolute value signs in expression (2), that is, the joint distribution of $\partial^2 z / \partial x^2$, $\partial^2 z / \partial x y$, $\partial^2 z / \partial y^2$ (repeated, because the absolute value of this term is multiplied by 2 in expression (2)) and $\partial^2 z / \partial y^2$. We estimated the second derivatives in the x and y directions by 3-point central differences. We estimated the cross derivative $\partial^2 z / \partial x y$ by 4-point differencing using the four corners of each grid cell. We divided the range from the smallest to the largest of the second derivatives into 100 ‘‘bins’’ and calculated the number of second derivatives in each bin. Finally, we plotted on a log-log scale the density (number of second derivatives in each bin) vs. the absolute value of the midpoint of the bin. We did not plot points for bins further out in each tails than the first bin containing no second derivatives because the information for those bins has reduced reliability.

For heavy-tailed distributions with tails that behave asymptotically like $d = c|x-\mu|^{-\alpha}$ (d = density, x = position, α positive, $\mu = 0$ for our DTED data sets), the points on the log-log plots are expected to lie approximately on a line with slope $-\alpha$. In contrast, points on log-log plots for exponential and Gaussian distributions should be on exponential curves that, as x increases, decrease faster than any linear function (or any polynomial function).

The log-log plots for the second derivatives of the DTED5, DTED2 and DTED 1 data are given in Figs. 1, 2 and 3.

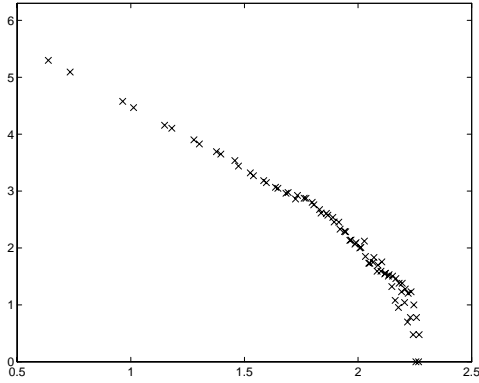


Fig. 1. Log-log plot of density of second derivatives for DTED5 Baltimore data set

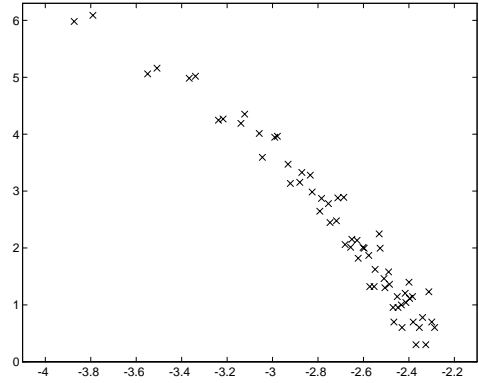


Fig. 3. Log-log plot of density of second derivatives for DTED1 Killeen data set

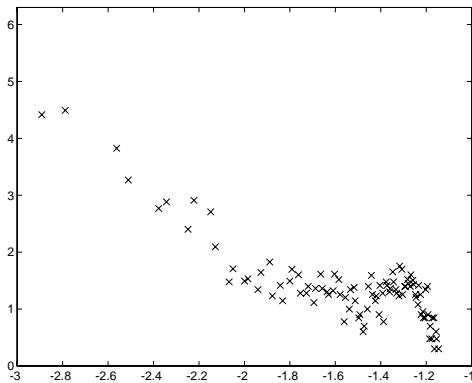


Fig. 2. Log-log plot of density of second derivatives for DTED2 Killeen data set

The points in Fig. 1 and Fig. 3 are well approximated by straight lines with slopes -2 and -4 , respectively. The fact that the urban (Baltimore) data set of Fig. 1 has larger tails than the natural-terrain (Killeen) data set of Fig. 3 is consistent with expectations, since urban data sets typically have more discontinuities (large changes over small distances) than do natural-terrain data sets. The points in the left half of Fig. 2 are well approximated by a straight line with slope -3.5 while the points in the right half of this figure are approximated by straight lines with slope 0 (for all points in the right half except those in the rightmost portion of this half of the plot) and -15 (rightmost portion). If the data were Gaussian rather than heavy-tailed, the values plotted in the figures would be close to a concave curve that decreases exponentially as one proceeds from left to right. The more rapid local decreases at the right ends of the data in all of Figs. 1, 2 and 3 may be due to the limited nature of the data sets and to the numerical procedures used to calculate the second derivatives. See Sec. 4 for a discussion of these issues. Overall, the second derivatives are clearly from heavy-tailed distributions, not from distributions with small, exponentially decreasing tails.

3. HEAVY-TAILED NATURE OF FIRST DERIVATIVES MINUS FIRST DIFFERENCES

Splines based on first derivatives have not yet been developed. As a basis for potential development of such splines, we investigated the structure of the tails of the distributions of first derivatives minus first differences. Specifically, we investigated the structure of the tails of the joint distribution of $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$, where $\Delta_x z$ and $\Delta_y z$ are, respectively, the x - and y -slopes of the piecewise planar interpolant of the data. We estimated the first derivatives $\partial z/\partial x$ and $\partial z/\partial y$ at the midpoints of the line segments connecting the data points by 4-point central differencing. We divided the range from the smallest to the largest of the quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ into 100 bins and calculated the number of these quantities in each bin. Finally, we plotted on a log-log scale the density (number of quantities in each bin) vs. the absolute value of the midpoint of the bin. We did not plot points for bins further out in each tails than the first bin containing no quantities.

The log-log plots for the quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ of the DTED5, DTED2 and DTED 1 data are given in Figs. 4, 5 and 6.

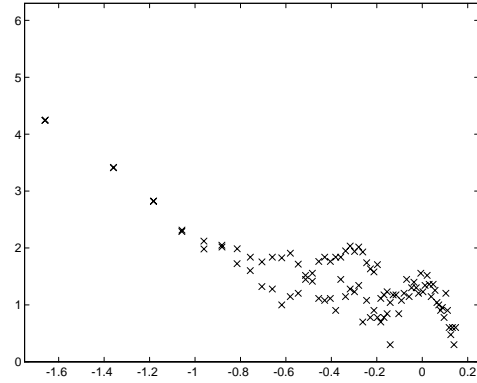


Fig. 5. Log-log plot of density of quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ for DTED2 Killeen data set

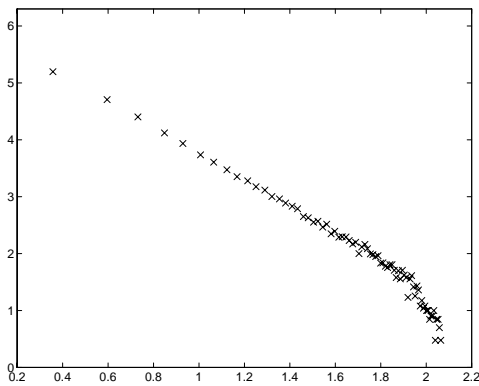


Fig. 4. Log-log plot of density of quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ for DTED5 Baltimore data set

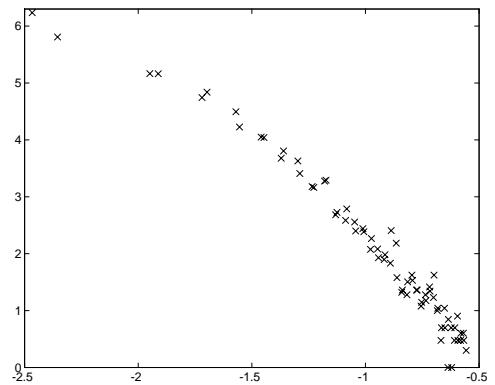


Fig. 6. Log-log plot of density of quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ for DTED1 Killeen data set

The structure of the points in Figs. 4–6 is nearly the same as that in Figs. 1–3. The quantities $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ are, just like the second derivatives, heavy-tailed.

4. DISCUSSION

One of the challenges in estimating the structure of the tails is accuracy of the data and of the numerical procedures. Raw terrain elevation data is not generally available at the nodes of a uniformly spaced tensor-product grid such as that used for DTED data. Data on the tensor-product grid are produced typically by a local averaging process, which introduces error. The 3- and 4-point differentiation formulas, which are based on local approximation by polynomials, introduce additional error, especially for derivatives with large absolute value, that is, the derivatives at the ends of the tails (the rightmost points in the figures). Both local averaging and numerical differentiation introduce extraneous smoothing and lead to results that are less heavy-tailed than reality. The behavior at the rightmost portions of the point sets in Figs. 1–6, which are steeper than the other portions of these point sets, may be due to the smoothing effects of local averaging and numerical differentiation.

Data sets with on the order of 1000^2 points, such as the three data sets used in this paper, are large for geometric modeling purposes. However, they are not large for purposes of characterizing the structure of the tails of the distributions. Much further work in this area needs to be done. Many large terrain databases have to be analyzed in the process of confirming the heavy-tailed nature of the second derivatives of terrain. Different classes of natural and urban terrain and different spacings of the data may yield different exponents $-\alpha$ describing the tails.

When designing L_1 splines or any geometric modeling procedure, the question of whether the data are anisotropic (different behavior in different directions) can be of considerable importance. Because of the presence of large structures, for example, stadiums in urban terrain and drainage basins in natural terrain, anisotropies can occur. However, we did not note any significant differences among the distributions of $\partial^2 z/\partial x^2$, $\partial^2 z/\partial xy$ and $\partial^2 z/\partial y^2$ or among the distributions of $\partial z/\partial x - \Delta_x z$ and $\partial z/\partial y - \Delta_y z$ for the three data sets used in this paper. For reference, we present in Figs. 7–9 the log-log plots for $\partial^2 z/\partial x^2$, $\partial^2 z/\partial xy$ and $\partial^2 z/\partial y^2$ for the DTED5 Baltimore data set. The slopes of the lines fitting the points in Figs. 7–9 are all approximately -2 , the same as in Fig. 1.

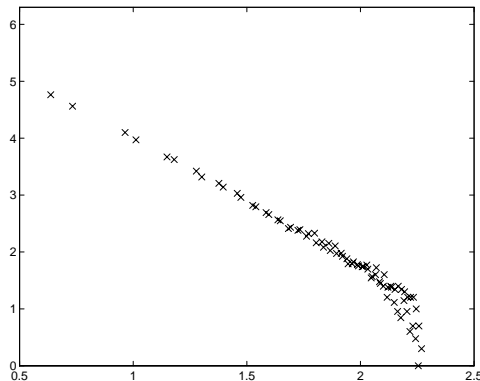


Fig. 7. Log-log plot of density of $\partial^2 z/\partial x^2$ for DTED5 Baltimore data set

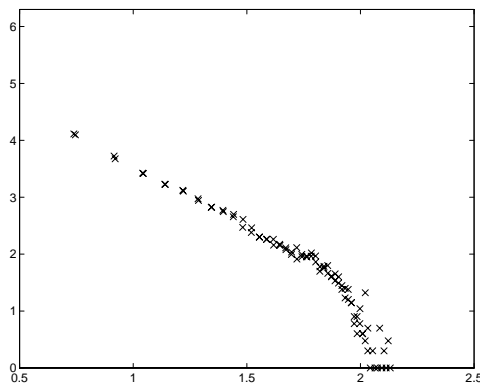


Fig. 8. Log-log plot of density of $\partial^2 z/\partial xy$ for DTED5 Baltimore data set

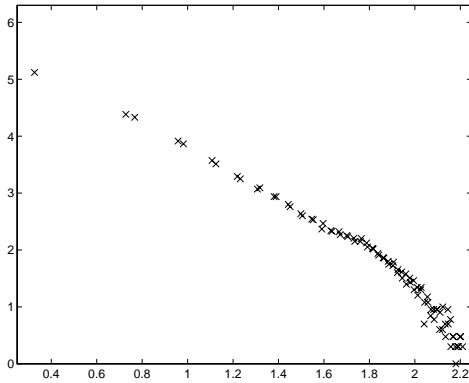


Fig. 9. Log-log plot of density of $\partial^2 z / \partial y^2$ for DTED5 Baltimore data set

5. CONCLUSION

The distributions of the second derivatives and of expressions involving the first derivatives that have been investigated in this paper are not directly related to the spatial frequency spectra that have been the topic of most previous investigations of the heavy-tailed nature of terrain. In terrain just like in imaging, frequency-based methods (such as JPEG) are recognized to be limited in accuracy and in compression capability. The heavy-tailed nature of the frequency spectra of terrain has not resulted in any major impact on modeling of large terrain datasets (although it has had significant positive impact on modeling of vehicle-terrain interaction, where the data sets are local and smaller). The investigation of the heavy-tailed nature of the derivatives of terrain will have significant impact by providing the theoretical underpinnings for the current observation that L_1 splines provide better terrain modeling than alternative techniques.

REFERENCES

Champion, D.C. and Lavery, J.E., 2002: Line of sight in natural terrain determined by L_1 -spline and conventional methods, *Proc. 23rd Army Science*

- Conf.*, Department of the Army, Washington, DC, OP-09.
- Cheng, H., Fang, S.-C. and Lavery, J.E., 2004: Shape-preserving properties of univariate cubic L_1 splines, *J. Comput. Appl. Math.*, to appear.
- Dodge, Y. (ed.), 1987: *Statistical Data Analysis Based on the L_1 -Norm and Related Methods*, North-Holland, New York.
- Dodge, Y. (ed.), 1992: *L_1 -Statistical Analysis and Related Methods*, North-Holland, New York.
- Dodge, Y. (ed.), 1997: *L_1 -Statistical Procedures and Related Topics, IMS Lecture Notes and Monographs Series, 31*, Berkley, California.
- Dodge, Y. (ed.), 2002: *Statistical Data Analysis Based on the L_1 Norm and Related Methods*, Birkhäuser, Basel.
- Gilsinn, D.E. and Lavery, J.E., 2002: Shape-preserving, Multiscale fitting of Bivariate Data by Cubic L_1 Smoothing Splines, in Chiu, C.K., Schumaker, L.L. and Stoeckler, J. (eds.): *Approximation Theory X: Wavelets, Splines, and Applications*, Vanderbilt University Press, Nashville, Tennessee, 283–293.
- Lavery, J.E., 2001: Shape-preserving, multiscale interpolation by bi- and multivariate cubic L_1 splines, *Comput. Aided Geom. Design*, **18**, 321–343.
- Lavery, J.E. and D.E. Gilsinn, 2000: Multiresolution representation of urban terrain by L_1 splines, L_2 splines and piecewise planar surfaces, *Proc. 22nd Army Science Conf.*, Department of the Army, Washington, DC, 767–773.
- Mandelbrot, B., 1982: *The Fractal Geometry of Nature*, W.H. Freeman, New York.
- Mark, D.M., 1997: The History of Geographic Information Systems: Invention and Re-Invention of Triangulated Irregular Networks (TINS), *Proceedings, GIS/LIS'97*, Bethesda, Maryland.
- Peitgen, H.-O., Jurgens, H. and Saupe, D., 2004: *Chaos and Fractals, New Frontiers of Science*, 2nd Ed., Springer-Verlag, New York.