The Science of Fractal Images.
Edited by Heinz-Otto Peitgen & Dietmar Saupe,
New York: Springer, 243-260.

Appendix A

Fractal landscapes without creases and with rivers

Benoit B.Mandelbrot

This text tackles three related but somewhat independent issues. The most basic defect of past fractal forgeries of landscape is that every one of them fails to include river networks. This is one reason why these forgeries look best when viewed from a low angle above the horizon, and are worst when examined from the zenith. This is also why achieving a *fully random* combined model of rivers and of mountains will be a major advance for computer graphics (and also perhaps for the science of geomorphology). The main goal of this text is to sketch two *partly random* solutions that I have developed very recently and Ken Musgrave has rendered.

A second very different irritating defect is specific to the method of landscapes design (called midpoint displacement or recursive subdivision) which is described in [40]. It has gained wide exposure in "Star Trek II," and is fast, compact, and easy to implement. Its main shortcoming is that it generates "features" that catch the eye and, by their length and straightness, make the rendered scene look "unnatural." This defect, which has come to be called "the creasing problem," is illustrated in the body of the present book. It was first pointed out in [69], and its causes are deep, so that it *cannot* be eliminated by cosmetic treatment. This text sketches several methods that avoid creasing from the outset. (We were not aware of the fact that [76] had addressed this problem before us. His method is somewhat similar to ours, but there is no face-to-face overlap.)

A third defect present in most forgeries is that the valleys and the mountains are symmetric because of the symmetry of the Gaussian distribution. Voss's pictures in [69] have shown that this undesirably condition is greatly improved by after-the-fact non linear processing. Furthermore, our "mountains with rivers" involve extreme asymmetry, meaning that this paper tackles this issue head on. But, in addition, this text sketches a different and much simpler path towards the same end. It shows that it suffices to use the familiar midpoint displacement method with displacements whose distribution is itself far from being Gaussian, in fact are extremely asymmetric.

In the search for algorithms for landscape design, hence in this text, a central issue is that of *context dependence.* The familiar midpoint displacements method is context *independent,* which is the source both of its most striking virtue (simplicity) and of its most striking defect (creasing). In real landscapes, however, the very fact that rivers flow down introduces very marked context dependence; it is theoretically infinite, and practically it extends to the overall scale of the rendering.

This introduction has listed our goals in the order of decreasing importance, but this is also the order of increasing complexity, and of increasing divergence from what is already familiar to computer graphics experts. Therefore, the discussion that follows is made easier by moving towards increasing complexity. It is hoped that the new procedures described in this text will inject fresh vitality into the seemingly stabilized topic of fractal landscape design. In any event, landscape design deserves investigation well beyond what is already in the record, and beyond the relatively casual remarks in this paper.

A.1 Non-Gaussian and non-random variants of midpoint displacement

A.1.1 Midpoint displacement constructions for the paraboloids

One of the purposes of this text is to explain midpoint displacement better, by placing it in its historical context. The Archimedes construction for the parabola, as explained in the Foreword, can be expanded to the surface S closest to it, which is the isotropic paraboloid $z = P(x, y) = a - bx^2 - by^2$. Here, the simplest initiator is a triangle instead of the interval, the section of S along each side of the triangle being a parabola. A second stage is to systematically interpolate at the nodes of an increasingly tight lattice made of equilateral triangles. This procedure, not known to Archimedes, is a retrofit to non-random

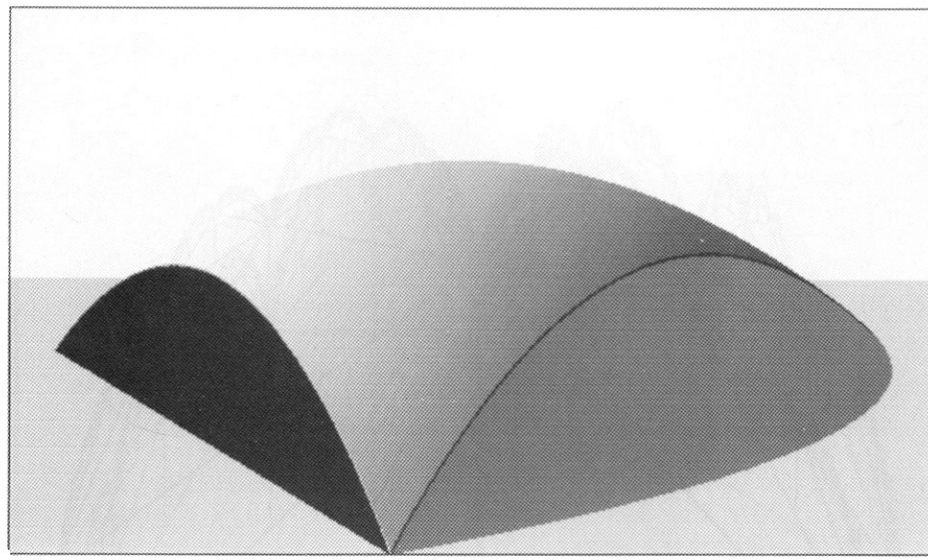

Fig. A.1: Mount Archimedes, constructed by positive midpoint displacements.

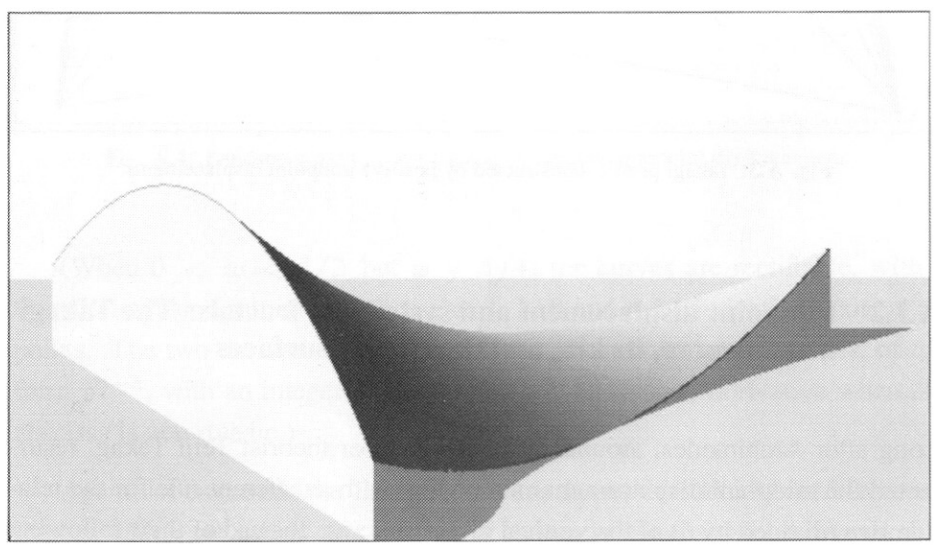

Fig. A.2: Archimedes Saddle, constructed by midpoint displacements.

classical geometry of the method Fournier et al. use to generate fractal surfaces. When all the interpolations along the cross lines go up, because $b > 0$ (resp., down, because $b < 0$), we converge to *Mount Archimedes* (Figure A.1) (resp., to the *Archimedes Cup*, which is the *Mount* turned upside down.) Suitable slight changes in this construction yield the other paraboloids. In reduced coordinates, a paraboloid's equation is $z = P(x, y) = a - bx^2 - cy^2$, with $b > 0$ in all cases, and with $c > 0$ if the paraboloid is elliptic (a cap), but $c < 0$ if it is hyperbolic (a saddle). By suitable "tuning" of the up and down midpoint displacements, one can achieve either result. The *Archimedes saddle* is shown on Figure A.2.

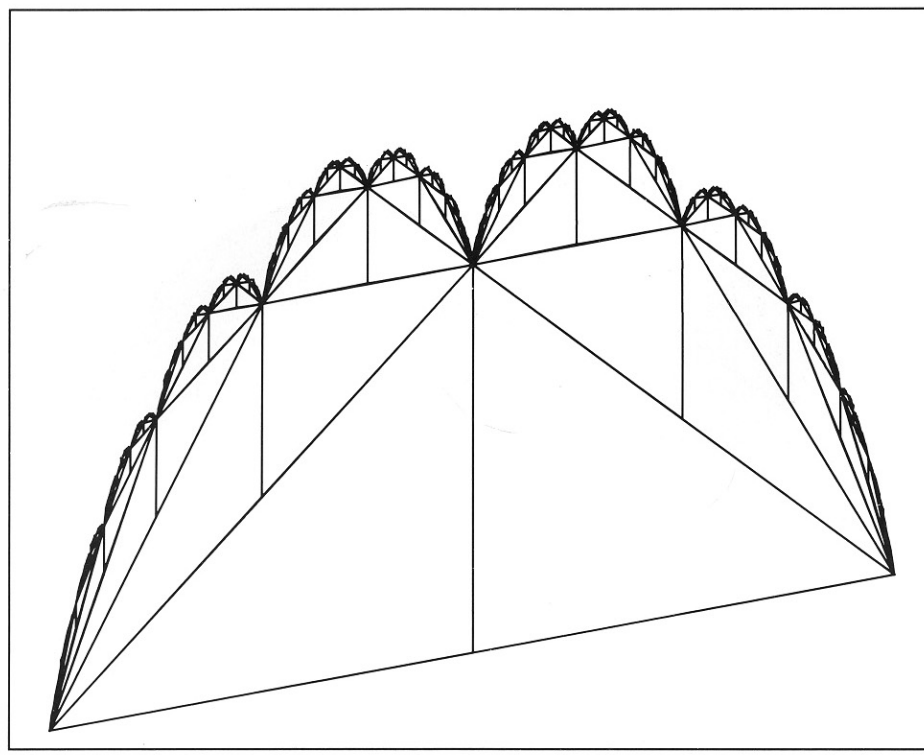

Fig. A.3: Takagi profile, constructed by positive midpoint displacements.

A.1.2 Midpoint displacement and systematic fractals: The Takagi fractal curve, its kin, and the related surfaces

Long after Archimedes, around 1900, the number theorist Teiji Takagi resurrected the midpoint displacement method, but with a different rule for the relative size (divided by δ) of the vertical displacements: Instead of their following the sequence 1 (once), 4^{-1} (twice), 4^{-2} (2^2 times) ... 4^{-k} (2^k times), etc., he made the relative sizes be 1 (once), 2^{-1} (twice), 2^{-2} (2^2 times) ... 2^{-k} (2^k times), etc.. The effect is drastic, since the smooth parabola is then replaced by the very unsmooth Figure A.3, which is an early example of a function that is continuous but non differentiable anywhere. Then a mathematician named Landsberg picked w arbitrarily in the interval $1/2 < w < 1$ and performed midpoint displacements of 1 (once), w (twice), ... w^k (2^k times), etc. ... Figure A.4 corresponds to $w = .6$.

These curves have the fractal dimension $2 - |\log_2 w|$. As w ranges from $1/2$ to 1, this dimension ranges from 1 (a borderline curve of dimension 1, but of logarithmically infinite length) to 2.

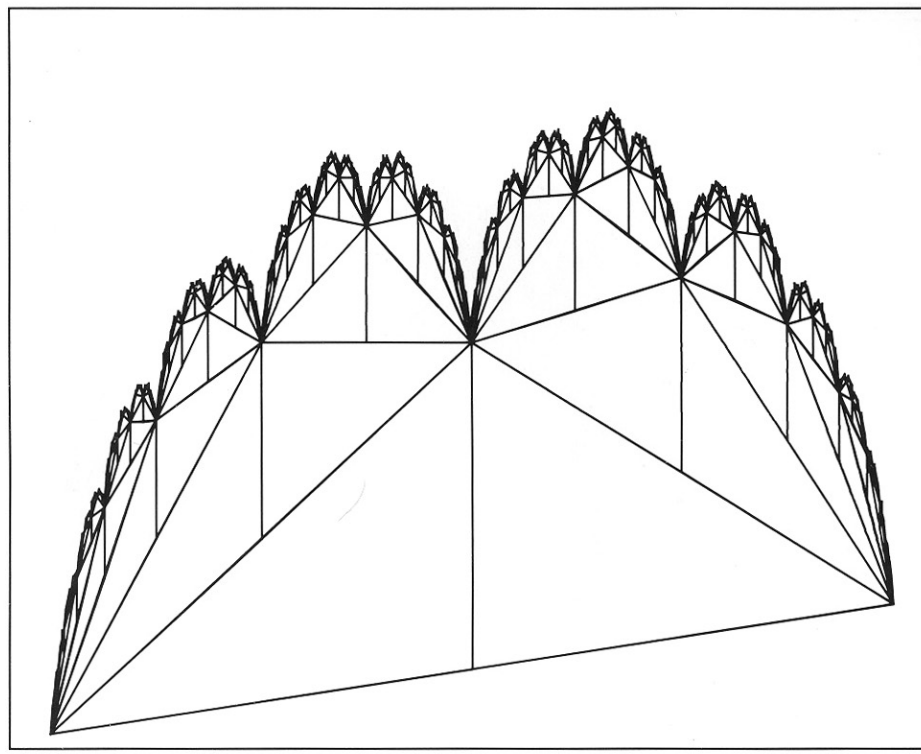

Fig. A.4: Landsberg profile, constructed by positive midpoint displacements.

(When $0 < w < 1/2$ but $w \neq 1/4$, the curves are rectifiable, with a fractal dimension saturated at the value 1. Left and right derivatives exist at all points. The two derivatives differ when the abscissa is dyadic, that is, of the form $p2^{-k}$, with an integer p, but there is a well defined derivative when the abscissa is non-dyadic.)

Again, let us retrofit the Takagi curve à la Fournier et al.. Its surface counterpart, which we call *Mount Takagi,* is shown as Figure A.5. A remarkable fact is revealed when we examine this surface from a low enough angle, relative to the vertical scale in the picture, and from a horizontal angle that avoids the lattice directions. At ridiculously low intellectual cost, we have delivered a realistic first-order illusion of a mountain terrain (other viewers may be more strongly taken by botanic illusions). However, seen from a high vertical angle, Mount Takagi reveals a pattern of valleys and of "cups" without outlets, which is totally unrealistic. From a horizontal angle along the lattice directions, the illusion also collapses altogether: the original triangular grid used for simulation overwhelms our perception. Inspecting Mount Takagi from a randomly chosen low angle yields a surprising variety of terrains that do not immediately seem to be all the same.

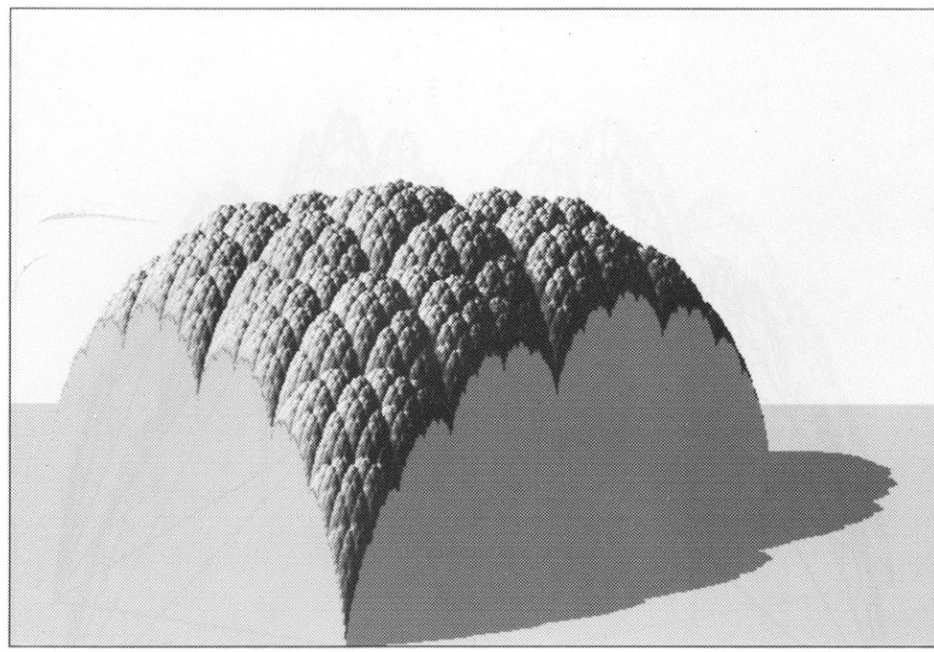

Fig. A.5: Mount Takagi, constructed by positive midpoint displacements.

The lesson we just learned happens to be of much wider consequence. Every one of the fractal landscapes explored thus far is best when viewed at low angle.

(The question of what is a high or a low angle, compared to the vertical, is related to the counterpart of the quantity δ in the definition of Mount Archimedes, hence to the fact that our various reliefs are self-affine, rather than self-similar.)

A.1.3 Random midpoint displacements with a sharply non-Gaussian displacements' distribution

In random midpoint displacement (Fournier et al.), the structure of valleys and ridges is less accentuated than it is with Mount Takagi, yet it remains apparent and bothersome. It can be attenuated by modifying the probability distribution of the random midpoint displacements. The Gaussian distribution is symmetric, hence a fractional Brownian surface can be flipped upside down without changing statistically. Actual valley bottoms and mountain crests *are not* symmetric. This difficulty had led Voss to introduce non linear after the fact transformations. The resulting non Gaussian landscapes (Plates 8 and 10) are stunning, but contrived. If one abandons the straight and true fractional Brown ideal, one may as well replace the Gaussian by something else.

The first suitable non symmetric distribution we have played with is the asymmetric binomial of parameter p: $X = 1$ with the probability p and

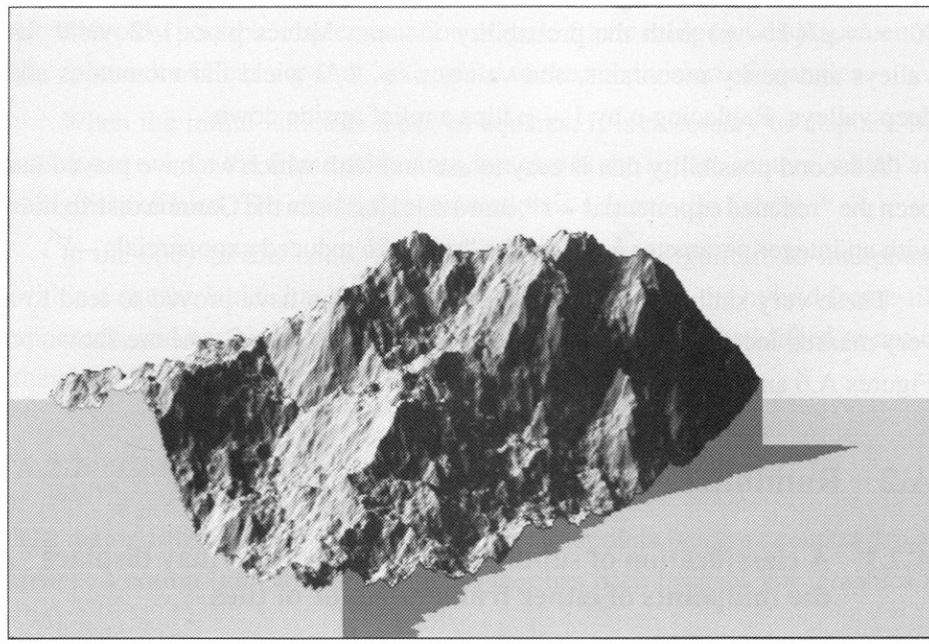

Fig. A.6: First fractal landscapes constructed by random midpoint displacements with a very skew distribution.

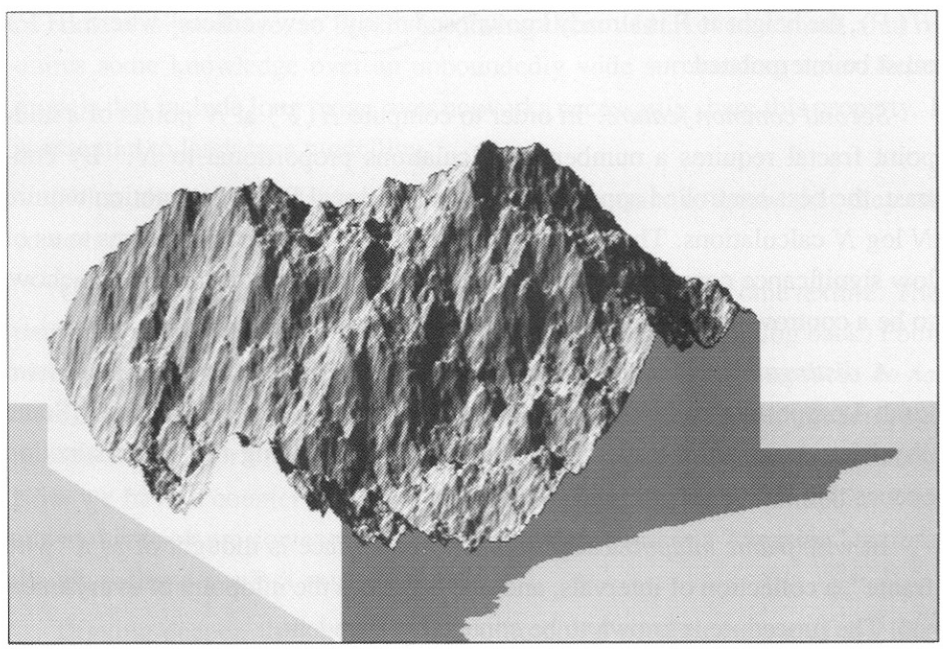

Fig. A.7: Second fractal landscapes constructed by random midpoint displacements with a very skew distribution.

$X = -p/(1-p)$ with the probability $1-p$. Values $p < 1/2$ yield flat valleys and peaky mountains, and values $p > 1/2$ yield flat mountains and deep valleys. Replacing p by $1-p$ flips a relief upside down.

A second possibility that is easy to use and with which we have played has been the "reduced exponential -1", and a third has been the Gamma distribution with an integer parameter b, that is the "sum of b reduced exponentials $-b$".

These very simple changes beyond the Gaussian have proved to lead to a very marked increase in realism. Examples due to Réjean Gagné are shown on Figures A.6 and A.7.

A.2 Random landscapes without creases

A.2.1 A classification of subdivision schemes: One may displace the midpoints of either frame wires or of tiles

Irrespective of the issue of distribution, a second development begins by carefully examining the different published implementations of random midpoint displacement, to see what they have in common and how they differ.

First common feature. Starting with an initial lattice of nodes or vertices, one "nests" it within a finer lattice, which contains both the "old vertices" where $H(P)$, the height at P is already known, and many "new vertices" where $H(P)$ must be interpolated.

Second common feature. In order to compute $H(P)$ at N points of a midpoint fractal requires a number of calculations proportional to N. By contrast, the best-controlled approximations to fractional Brownian motion require $N \log N$ calculations. The factor $\log N$ in the cost of computing seems to us of low significance compared to the cost of rendering. But this seems, somehow, to be a controversial opinion.

A distinguishing feature. Various methods of interpolation fall into two kinds I propose to call *wire frame* displacement and *tile* displacement. (Some physicists among the readers may be helped by knowing that the distinction echoes that between *bond* and *site* percolation).

In *wire frame midpoint displacement,* the surface is thought of as a "wire frame", a collection of intervals, and one displaces the midpoint of every interval. The procedure is known to be context independent.

In *tile midpoint displacement,* the surface is thought of as a collection of tiles, and one displaces points in the "middle" of every tile. The procedure will be seen to context dependent.

The only possible implementation of pure frame midpoint displacement is the approach of Fournier et al., as applied to the subdivision of triangles.

When the initial lattice is made of squares, it is necessary to displace the relief, not only at the frame midpoints, but also at the square centers, which are midpoints of tiles. The resulting construction is a frame-tile hybrid.

In the body of this book, the desire for a cleaner algorithm to interpolate within a square lattice has pushed the authors away from the hybrid frame-tile approach, and has led them to a tile midpoint procedure. (I had failed to notice they had done so, until told by Voss.)

A.2.2 Context independence and the "creased" texture

The major difference between frame and tile midpoints involves *context dependence*, a notion from formal linguistics that has been brought into graphics in [96].

Frame midpoint fractals are context independent, in the sense that the only external inputs that affect the shape of a surface within a triangle are the altitudes at the vertices. Everyone in computer graphics knows this is an extraordinary convenience.

On the other hand, the fractional Brown surfaces incorporate an infinite span of context dependence, so that interpolating them within even a small area requires some knowledge over an unboundedly wide surrounding area. Relief models that include long range river networks necessarily share this property. It is essential to landscape modelling.

What about tile midpoint fractals? It is important to know that they are context dependent.

Unfortunately, frame midpoint fractals have a most unwelcome texture. They reminded me of folded and crumpled paper, and (after first arguing back) Fournier et al. have acknowledged my criticism. Fractional Brown surfaces *do not* have this undesirable "creasing" texture. This discrepancy has led me long ago to wonder whether or not context dependence is the sole reason for creasing. Now we have a counter example: the implementations of tile midpoint fractals in this book are context independent, yet they possess a "creasing" texture, though perhaps to a less extreme degree.

Digging deeper, the undesirable texture could perhaps be traced to either of the following causes. A) "The Course of the Squares": The programmers' and the physicists' "folklore" suggests that spurious effects due to the lattice structure tend to be strongest in the case of square lattices in which privileged

directions number two. This suggests trying triangular or hexagonal lattices, in which privileged directions number three. B) "The Nesting of the Frames". The strongest texture problems always appear along the frames that bound the largest tiles. Let us dwell on this point. As has been previously observed, the *tile vertices* at the k-th stage of construction are always nested among the tile vertices at all later stages. But, in addition, triangular tiles have the far stronger property that the network of *tile sides* at the k-th stage of construction is nested in (i.e., part of) the network of tile sides at the $(k+2)$-th stage, the $(k+4)$-th, etc.. I believe this feature is responsible for creasing and propose to demonstrate how it can be avoided.

My own recent renewal of active interest in fractal reliefs happens to provide tests of both of the above explanations of creasing. A first approach avoids squares and uses the classical triangular tiles, whose frames *are* nested. A second approach uses tiles that are based on the hexagon, but are actually bounded by fractal curves; their frames *are not* nested.

A.2.3 A new algorithm using triangular tile midpoint displacement

With triangles, the aim is to construct a fractal surface whose heights at the vertices of an initial triangular lattice are independent identically distributed random variables, while its values within each tile are interpolated at random. The result is said by physicists to have a "cross-over scale" that is precisely equal to the initial triangles' side: above this scale, the surface is statistically stationary, and below this scale it is more like fractional Brownian motion.

In the first stage, Figure A.8a, the height at the center of the triangle ABC in the first stage lattice is given the value

$$\frac{1}{3}[H(A)+H(B)+H(C)] + \text{a random midpoint displacement.}$$

This determines $H(P)$ at the vertices of a second order triangular lattice, which is turned by $30°$ and whose sides are smaller in the ratio $1/\sqrt{3}$. The second construction stage, Figure A.8b, takes averages in every tile and adds a suitable midpoint displacement. Note that two values will have been interpolated along each side of ABC, and that they will depend on initial values beyond the triangle ABC. The third stage lattice, Figure A.8c is again triangular and rotated by $30°$, hence it nests the first stage lattice. The third construction stage within ABC is also affected by initial values beyond this triangle.

Conceptually, one must also combine interpolation with extrapolation. The latter amounts to changing the character of the values of $H(P)$ on the nodes

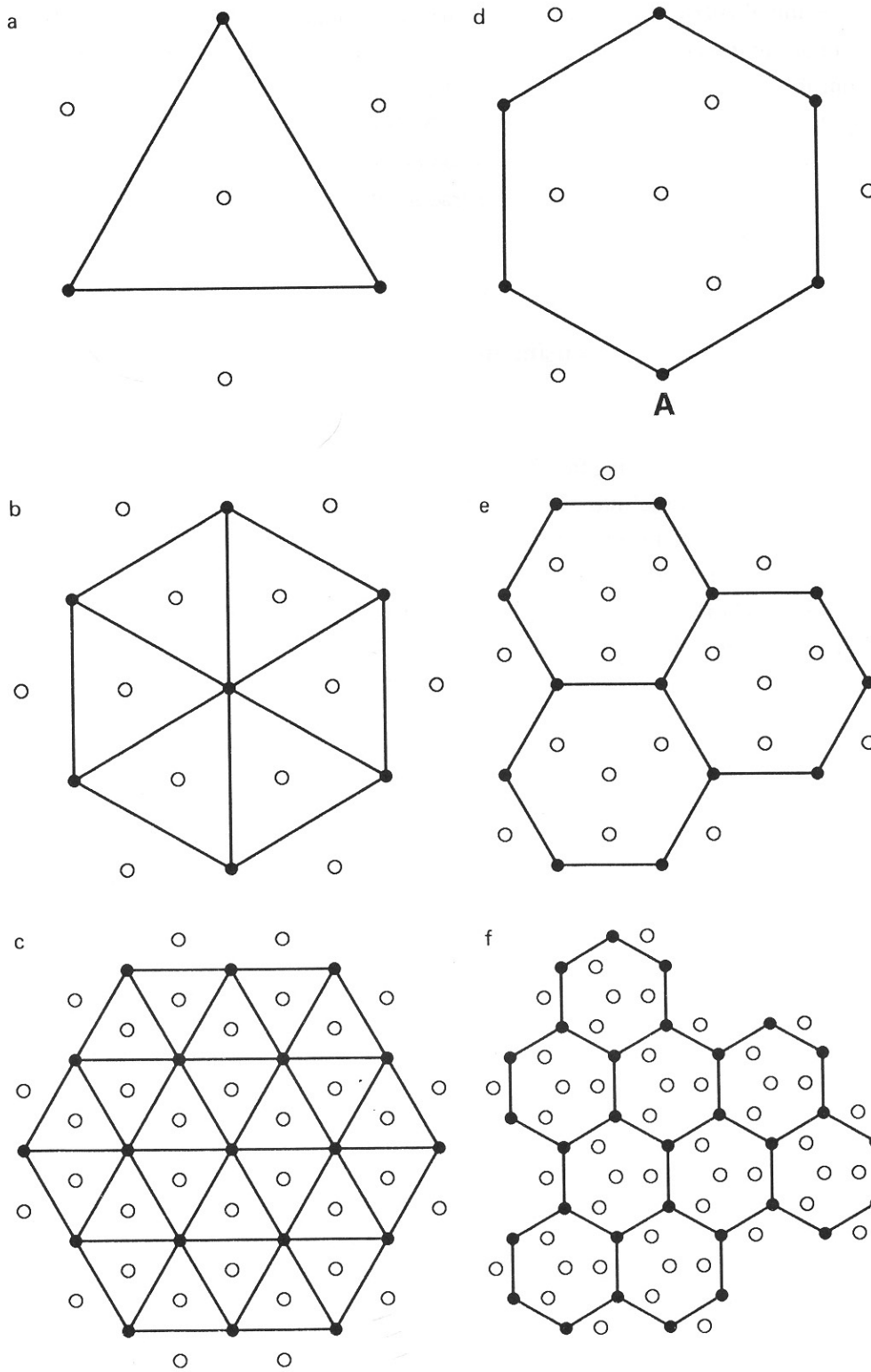

Fig. A.8: Triangular (left) versus hexagonal (right) tiles and midpoint displacements.

of the initial lattice. Instead of viewing them as stationary, extrapolation views them as having been interpolated, during a previous negative-numbered stage, from the values on the vertices of a looser lattice. If extrapolation continues indefinitely, the result is a surface in which $H(P')$ and $H(P'')$ are slightly but not negligibly correlated even when the points P' and P'' are very far from each other. This is precisely what holds for fractional Brownian motion.

A.2.4 A new algorithm using hexagonal tile midpoint displacement

The main novelty is that tiles are chosen so that their frames do *not* nest. The reason lies in the following familiar fact. While a square is subdivided into four squares and a triangle into nine triangles after two stages, a hexagon *cannot* be subdivided exactly into smaller lattice/hexagonal hexagons. However, an approximate subdivision can be done by using tiles that are near-hexagonal, but whose boundaries are fractal curves. The simplest example is illustrated by Plate 73 of [68] where it is shown being drained by an especially simple river tree that we shall encounter again in Section A.3. This tile divides into 3 subtiles of smaller size but of identical shape. (Plate 71 of [68] shows a different such tile, divisible into sevens).

Starting with a regular hexagon, the first stage of construction of this tile is shown by comparing the bold lines on Figures A.8d and A.8e: here the side that one meets first (when moving from A clockwise) is rotated by 30°; it may have been rotated by either +30° or −30° and the remainder of this construction stage is fully determined by this single choice. At the same time, Figure A.8d, $H(P)$ is interpolated at the center of the initial hexagon by the equal weights average of six neighbors displacement. But $H(P)$ is also interpolated by suitable averages at three other points off-center, where three sub-hexagons meet (hence "*midpoint*" is not synonymous with "*center*".) The averaging weights can be chosen in many different ways, with considerable effect on the texture. Displacements are added to all these weighted averages. The next interpolation stage by comparing the bold lines, shown on Figures A.8e and A.8f, again involves an arbitrary decision between +30° and −30°, thus the rotation is −30°. At each stage, the boundary of the modified hexagon becomes increasingly crumpled, and it never nests in the combination of higher stage boundaries. While an undesirable texture may well be present along these crumpled curves, the fact will matter little, because these curves fail to catch the eye, contrary to straight lines that stand out.

A.3 Random landscape built on prescribed river networks

The idea behind the two landscape algorithms to be described now is to first construct a map, with its river and watershed trees, and then to "fill in" a random relief that fits the map. The two implementations available today have the weakness that the maps themselves fail to be random.

A major ingredient of this approach (beyond the current implementation) amplifies on the already discussed asymmetry between valleys and mountains, by drawing very asymmetric rivers and watersheds. The relief along a watershed can and should go up and go down, with local minima at the sources, and this requirement can be (and will be in this text) implemented by a randomized procedure of positive midpoint displacements. On the other hand, the relief along a river necessarily goes downstream, hence must follow a fractal curve that differs profoundly in its structure from anything that has been used so far in this context.

A.3.1 Building on a non-random map made of straight rivers and watersheds, with square drainage basins

Our first example of prescribed river and watershed trees is very crude, but worth describing, because it is easy to follow. It is suggested by Figure A.9, which reproduces a quarter of Plate 65 of [68], rotated by 45° clockwise. One sees thin black triangles: one is of length 1 and runs along the vertical diagonal, two are of length 2^{-1}, eight of length 2^{-2}, etc. ... Each triangle is meant to represent a straight river flowing towards its very short side. Similarly, one sees thin white triangles, meant to represent a straight watershed. One builds up a non-random highly regular map by making all those triangles become infinitely thin.

Relief along the watershed. This is the most conspicuous feature in the rendering of the previous relief models. This is also where the midpoint displacement curves are least controversial, and we may as well use them to model relief here. The local minima along this relief are sharp cusps, they will be the rivers' sources, and the map prescribes their positions. The simplest implementation uses positive displacements, and consists in modeling the relief along the watersheds by the Landsberg functions defined in Section I. One may also multiply the positive midpoint displacement by a random factor that is positive, hence does not displace the sources.

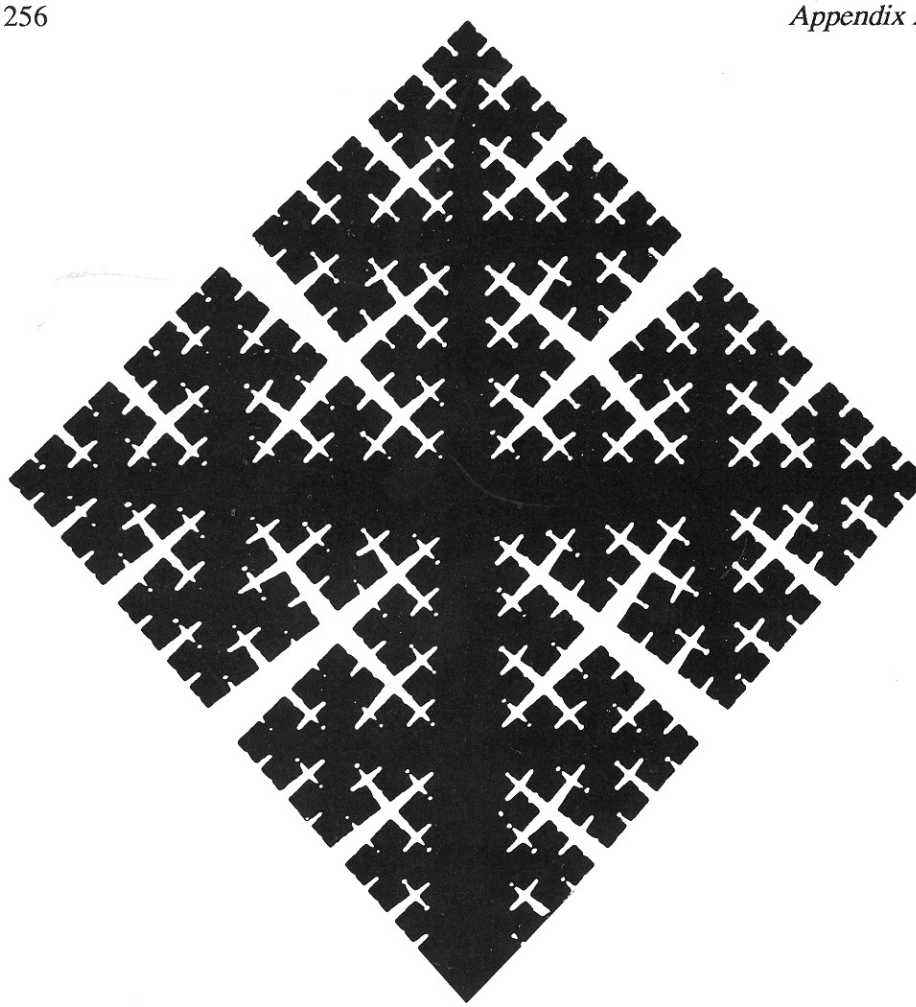

Fig. A.9: Straight rivers and watersheds and square drainage basins.

Relief along rivers. This is where we must bring in something new, namely fractal curves that never go up. The simplest implementation is called binomial non-random of proportion $p < 1/2$. Here, Figure A.10a a river's altitude loss between its source and its mouth is apportioned in the ratio $1-p$ in the upstream half, and $p < 1-p$ in the downstream half. Similar apportioning, Figure A.10b on, is carried out along every segment of a river between points whose distances from the source are of the form $q2^{-k}$ and $(q+1)2^{-k}$. This yields for each river profile a scalloped line that is illustrated on the top line of Figure A.10. Each river is flat near its mouth and vertical near its spring, a feature that exaggerates on everyone's perception of reality, but does not do it massive injustice.

Needless to say, the fixed p may be replaced by a variable randomly selected multiplier.

Fractal structure of the saddle points. These points are exemplified by the point $(1/8, 1/8)$ on Fig. A.9, after the figure has been mentally transformed

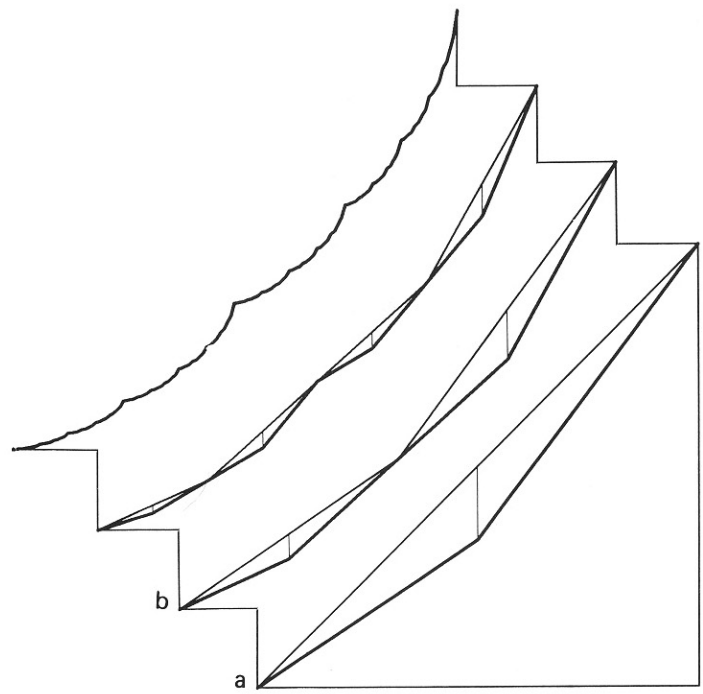

Fig. A.10: River profile constructed by negative proportional midpoint displacements.

so all the triangles are infinitely thin. Along each of the two watersheds (which become lines of slopes 45° or −45° on the map), there is a local minimum at this saddle point. The four rivers are orthogonal to the watersheds (have slopes 0 or 90° on this map), and on them the saddle point is a source and an absolute maximum.

Fractal structure of the points where rivers meet. Three rivers that meet flatten just upstream of their point of confluence and the river that emerges has a steep "waterfall" just downstream of this point. Each watershed reaches a minimum.

The initiator of the construction. The initial altitudes at the four vertices of the square must be selected to satisfy the following constraints: the mouth has the lowest altitude (think of it as sea level), the diagonally opposed point has some altitude $h > 0$, and the two remaining points have altitudes that exceed ph, but are otherwise arbitrary. Join each vertex to its neighbors, not only on the map but in relief, and join the mouth to the opposite vertex, again in relief. The resulting initial prefractal relief is a single drainage basin which projects upon two straight isosceles triangles joined by their hypotenuses.

The first stage of construction. The first step of this stage is to break each of the triangles of the initiator into two, by moving the altitude above the midpoint

of the square down, from the initiator's altitude $h/2$ to the lower value ph. The effect of this step is to draw two rivers that are orthogonal to the original one, and slope towards the center. The next step is to move the midpoints of the sides of the initial squares up by a quantity w. The effect is to split each straight watershed interval into two intervals, and to add four additional watershed intervals, each going from one of the four new sources to the midpoint of the initial river. As a result, each triangle in the initiator will have split into four. The outcome of this first stage of construction consists in four squares on the map, and on each square a drainage basin that has the properties we have required of the initiator relief.

The second stage of construction. It can proceed exactly as the first. The final relief is defined by the two parameters p and w. Needless to say, again, everything can be randomized.

A.3.2 Building on the non-random map shown on the top of Plate 73 of "The Fractal Geometry of Nature"

This map in [68] was already referenced in this Appendix, when describing the second construction of a creaseless fractal forgery without rivers. Now we construct rivers and water sheds simultaneously, as on Figure A.11. The map of the initiator, Figure A.11a, is a hexagonal drainage basin, crossed by a river that joins the mouth S_0 to the source S_2. One supposes that the altitudes $H(S_0)$, $H(S_2) > H(S_0)$ and $H(S_4) > H(S_2)$ are known.

The first stage of construction. First, the altitudes at the midpoints S_1, S_3 and S_5 along the watershed are interpolated by positive midpoint displacement (either the Landsberg procedure, or a randomized version of it). This yields $H(S_1)$, $H(S_3)$ and $H(S_5)$. Next, the originally straight river is replaced by a Y-shaped river, as marked. Figure A.11b, the altitude at the midpoint of the hexagon is determined in the same spirit as in the case of square drainage basins: by dividing the difference or altitude between the original endpoints S_0 and S_2 into the following ratios: $1 - p$ in the upstream half, and $p < 1 - p$ in the downstream half.

If the construction were stopped at this stage, the last step of the first stage would be to represent the resulting relief by six triangles. But if the construction it to continue, the last step is to draw the three smaller drainage basins shown on Fig.A.11c.

The second stage of construction. Each or the basins at the end of stage one has one "mouth" vertex and five "source" vertices, and otherwise each satisfies

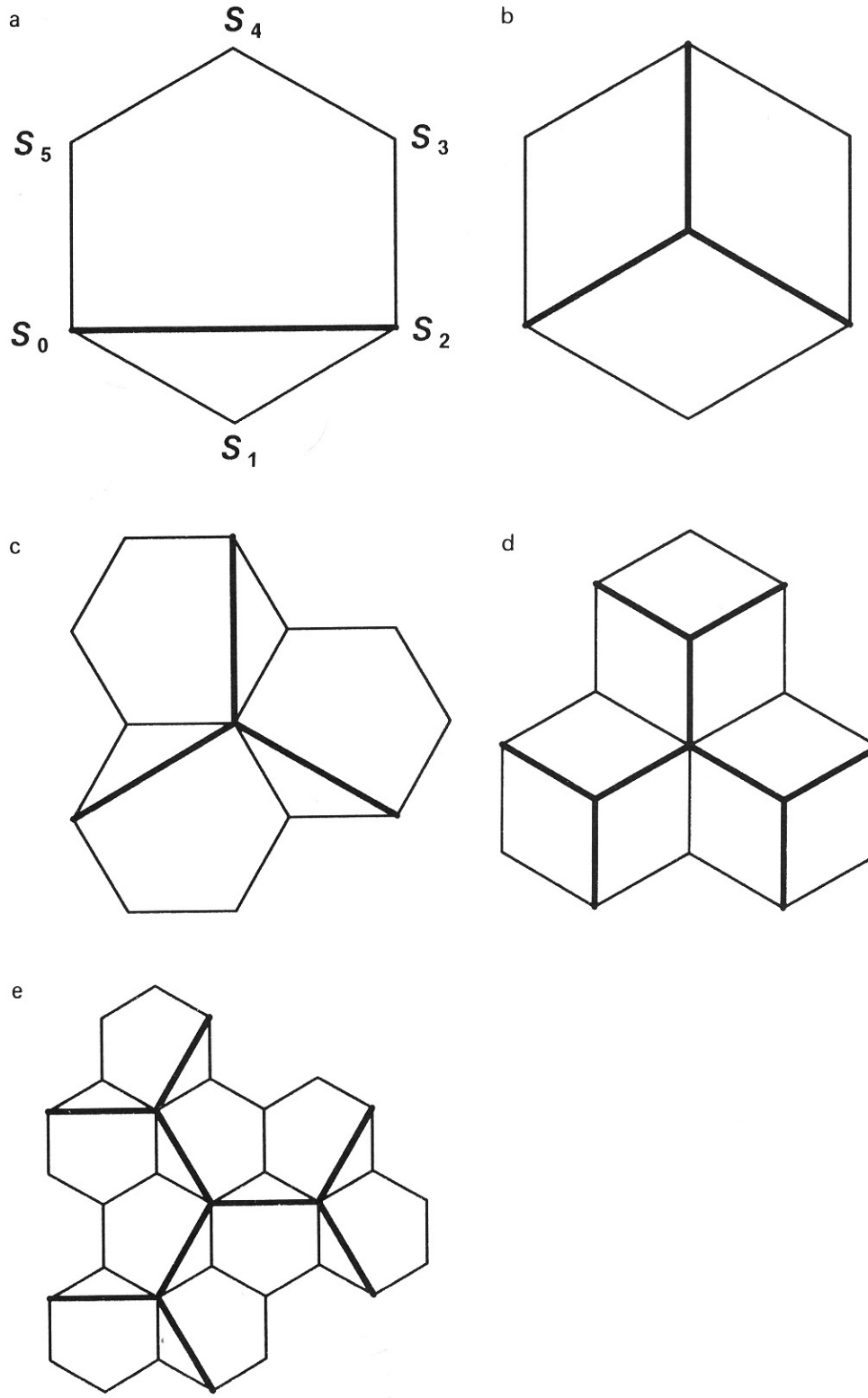

Fig. A.11: Hexagonal tiles with rivers.

the conditions demanded of the initiator. Therefore, one might think that the construction can proceed as in the first stage. However, a major complication arises: some of the new hexagon's source vertices are shared between two neighboring hexagons. This injects context dependence, and demands a new rule to break the deadlock. Between several possibilities, the rule we have adopted is this: when two interpolators of altitude are in conflict, choose the higher.

Exploratory implementations of the above novel algorithms by F.Kenneth Musgrave and Réjean Gagné are stunning, and suggest that the procedures advanced in this text may revive the topic of fractal landscapes.