Introduction

In recent years, there have been many studies of both exactly self-similar and statistically self-similar fractal curves in space [5, 11]. The property of self-similarity [5] is a central concept in fractal geometry. A curve in 2 space dimensions is exactly self-similar if it can be divided into N smaller copies of itself that are exactly the same as the original except for being scaled down by a factor r in both dimensions. In other words, the curve is invariant under identical changes of scale in two space dimensions. The fractal, or self-similarity, dimension D, which is a measure of the extent to which the curve is space filling [9], is given by the equation $D = (\log N)/\log (1/r)$. A example of an exactly self-similar curve is the Koch curve with D = 1.26 [11]. In contrast, a curve is only statistically self-similar if on division into N smaller copies of itself, each copy is statistically similar to the original object, a coastline being the best-known example.

Fractal curves occur in space. In the case of two dimensional space, a fractal curve will have a fractal dimension less than 2 but greater than 1, for example, the Koch curve with dimension 1.26. A fractal dimension close to 1 means that the curve does not occupy a great deal of the space, that is, it has a low level of wiggliness; a curve with fractal dimension close to 2 occupies most of the space and has a very high level of wiggliness. Thus in 2 dimensions fractal dimension can be usefully interpreted as a wiggliness measure. In the case of a fractal surface in 3 dimensional space, a fractal dimension close to 3 means that the surface is very jagged, whereas a dimension close to 2 means that it is very smooth, so that, in three dimensional space fractal dimension is a measure of the jaggedness of surfaces.

Time functions are quite different from curves in 2 dimensnional space, even though they are frequently represented as space curves on a 2-dimensional medium. A time function may exhibit apparent exact or statistical self-similarity, when so represented. However, this is deceptive, because a time function is always a physical magnitude that is a function of time, for example, the short term interest rate over a period of hours, or the temperature over a period of time, or the price of the S&P 500 index over a period of hours. With such curves, the scales that can be used for the physical magnitude and the time may be varied independently, which is not possible in the case of a geometric curve in 2-dimensional space. For example, in the case of the interest rate time function, when representing the curve on a 2-dimensional medium, in one representation 1 centimeter could be used for 2 basis points of interest on the y-axis, and 10 minutes of time on the x-axis; in another representation, 1 centimeter could represent 5 basis points and 60 minutes. The two curves would nevertheless be the same, although one would appear more wiggly than the other. Thus the concept of fractal dimension cannot be applied to time functions without ambiguity [8, 9, 11].

However, time functions can exhibit self-affinity, as opposed to the self-similarity of geometric curves in two dimensions. An exactly self-similar curve repeats exactly when magnified or scaled by the same factor in both space dimensions, that is, in both the y and x-dimensions. In contrast, a time function has exact self-affinity if it repeats exactly when magnified by one scaling factor r^h in the physical magnitude (y) axis and another scaling factor r in the time (t) axis [9]. Furthermore, a time function exhibits statistical self-affinity [8] if it repeats statistically when magnified by a scaling factor r^h in the physical magnitude axis and a scaling factor r in the time axis. Any time function that follows a random walk or Brownian motion, for example, exhibits statistical self affinity [8].

In the case of the random walk time function, if time t is magnified by a factor r then the amplitude A must be magnified by the factor r^h where h is 0.5. As first pointed out in 1900 by Bachelier [2], in his famous doctoral thesis, with a random walk the standard deviation of amplitude changes in fixed-length time intervals is proportional to the square root of the time interval. It is this statistical property that is preserved under (r, r^h) scaling [8, 11], thus making a random walk statistically self-affine. (Loosely, it takes 4 times as long for the amplitude, on average, to move twice as far, so that if the time axis is scaled by a factor of four the physical amplitude axis must be scaled by a factor of two to have a statistical replica of the original time function.) This fundamental property of a random walk is caused by amplitude changes in fixed time segments being independent, that is, they follow a Gaussian distribution, which leads to zero values for covariances and thus correlation coefficents between sets of amplitude changes in equal time segments. Consider an amplitude change A, made up of a succession of amplitude changes A_1, A_2, ... A_n, in n successive equal time segments. The variance of a large number of such amplitude changes A, denoted by V(A), must obey

$$
\begin{aligned}
V(A) &= V(A_1 + A_2 + \ldots A_n) \\
&= V(A_1) + V(A_2) + \ldots V(A_n) \\
&\qquad + C(A_1 A_2) + C(A_1 A_3) + \ldots C(A_{n-1} A_n)
\end{aligned}
$$

where $C(A_1 A_2)$ is the covariance between sets of A_1, A_2 amplitude changes. This must reduce to

$$V(A) = nV(A_1)$$

since the covariances are zero, because successive changes are independent, and since the variances in equal length time segments are all equal. Thus, if S(A) is the standard deviation:

$$S(A) = (V(A))^{0.5} = Kn^{0.5}$$

so that S(A) is proportional to the square root of the number of time seqments and thus to square root of the time.

In the 1930s, an obscure accountant, R.N. Elliot, proposed, in a series of vaguely worded articles in Financial World, later published in book form [3], an idealized time function concept as a model for the price-time behaviour of the Dow-Jone-Industrical Average. The concept was a time function that was potentially exactly self-affine since it allowed for an infinite number of replications of an 8-movement or 12345abc sequence, where the successive movements are labelled 1, 2, 3, 4, 5, a, b, and c. [To make the model fit the actual behaviour of the index, Elliot also introduced a large array of, sometimes simple, but more often complex, rules for permissible deviations from the ideal function concept. These complex rules and deviations, together with the issue of the extent to which they model the behaviour of the DJIA are not relevant to this paper.]

This paper is concerned with defining members of a class of precisely-defined exact self-affine time functions that have the common property of being consistent with infinite replication of the basic 8-movement 12345abc sequence. Functions belonging to this class are referred to as Elliot time functions or E(t) functions.

E(t) functions are interesting in their own right because they exibit exact self affinity, and in this paper the main amplitude and time ratios, and scaling factors for what appear to be the more obvious E(t) functions are derived. One of these functions described will be shown to have the unusual property of scaling like a random walk.

Self-affine time function definition

Both fractal curves and self-affine time functions can be defined by means of algorithms. However, it is often more convenient to define them by means of L-systems, also called string-rewriting systems, which were introduced by Lindemayer, and further developed by Prusinkiwicz [12], for computer graphics modelling of the growth of plants. Essentially a first string, or string axiom is given. Then each character from the string axiom is replaced by a further string using a rewriting rule from an allowed set of production rules. In the case of any E(t) function, infinite replication of a 12345abc segment requires the following axiom and production rules:

Ud	axiom
U -> UdUdU	up-impulse decomposition
d -> DuD	down-correction decomposition
D -> DuDuD	down-impulse decomposition
u -> UdU	up-correction decomposition

A sequence of rewritings is:

U	d	1 + 1 = 2
UdUdU	DuD	3 + 5 = 8
UdUdUDuDUdUdUDuDUdUdU	DuDuDUdUDuDuD	13 + 21 = 34
...	...	55 + 89 = 144

The number of characters at each rewriting level follows the Fibonacci sequence. The strings are displayed as a time function in Figure 1. At any level, U corresponds to an upward movement or up-impulse, followed by a correcting downward movement, or down-coorection, d. D corresponds to a downward movement, or down-impulse, followed by correcting upward movement, or upcorrection, u.

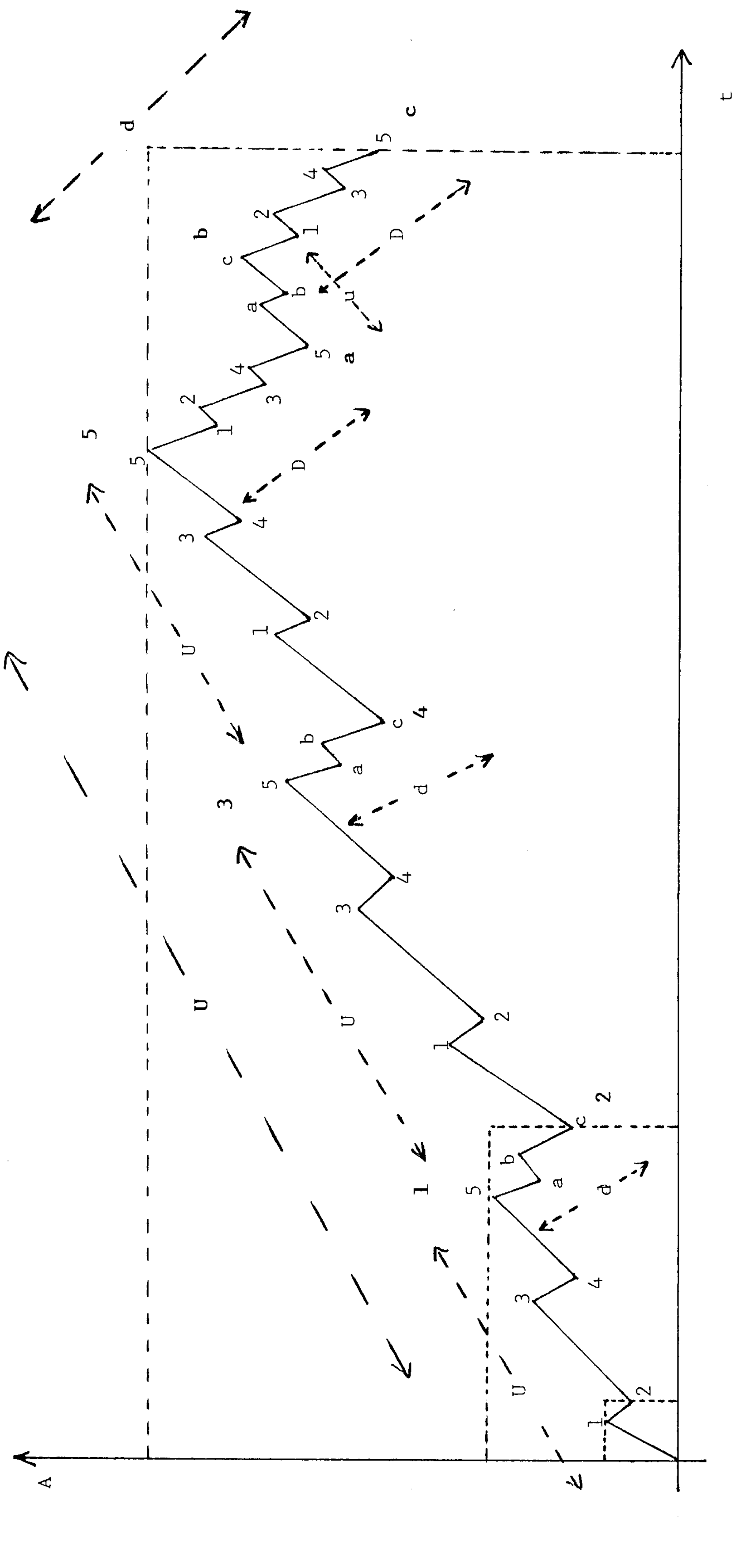

FIGURE 1

An E(t) function is constructed, at all levels except for the axiom level, of sets five movements upwards (UdUdU) or downwards (DuDuD), followed by 3 movements in the reverse direction either downwards (DuD) or upwards (UdU). At any level, it is convenient to label the various movement points on the curve as 0, 1, 2, 3 4, 5, a, b, c. It is evident that an E(t) function can have no derivitive anywhere, and moves in a series of zigzags in an upward trend, in a manner reminiscent of market averages behaviour.

Simple inspection shows that an E(t) function exhibits self-affinity in principle; for example, the 12345abc segments of the function in boxes in Figure 1 could all be exact replicas of each other. Although it is clear from Figure 1 that an E(t) function could replicate itself infinitely many times, in order to precisely define a specific E(t) function, exact information is needed for both amplitude and time scales. For example, with reference to the y-axis, we need to know the precise ratio of the height of any U movement to the subsequent corrective d movement; with reference to the time axis, we need the ratio of the time taken for a U-movement and the time taken for the subsequent d-movement. In this paper we show that for a specific E(t) function, these ratios may be derived from simple assumptions of the extent of amplitude and time movements and from the self-replicating requirement of the function. Since amplitude and time ratios are necessarily independent, they can be taken separately. We begin with possible amplitude ratios.

Amplitude ratio derivations

On examination of any 5-movement sequence of the type found in E(t) functions, it can be seen that there are three major types, depending on the extent of movement 4, as shown in Figure 4. These are:

1. Corrective movement-4 does not correct back as far as the end of movement-1.
2. Corrective movement-4 corrects back exactly to the end of movement-2.
3. Corrective movement-4 correct back beyond the end of movement-2.

These three cases are analysed below.

Case-1: Corrective movement-4 never corrects as far as point 1.

Consider an arbitrary 5-movement segment of an E(t) function, y = E(t), as shown in Figure 2a. It begins at an arbitrary y value, which we set to zero for convenience. Setting it to any other value will not affect the result, but simply make the algebra less readable. The end of movement-1 is set to amplitude level b, of movement-2 to level c, of movement 3 to level a, movement-4 to level x and movement-5 to level m. Thus movement-1 moves a distance b in the y-direction, and movement-5 moves a distance m-x in the y-direction and so on. In this portion of the analysis we are con-

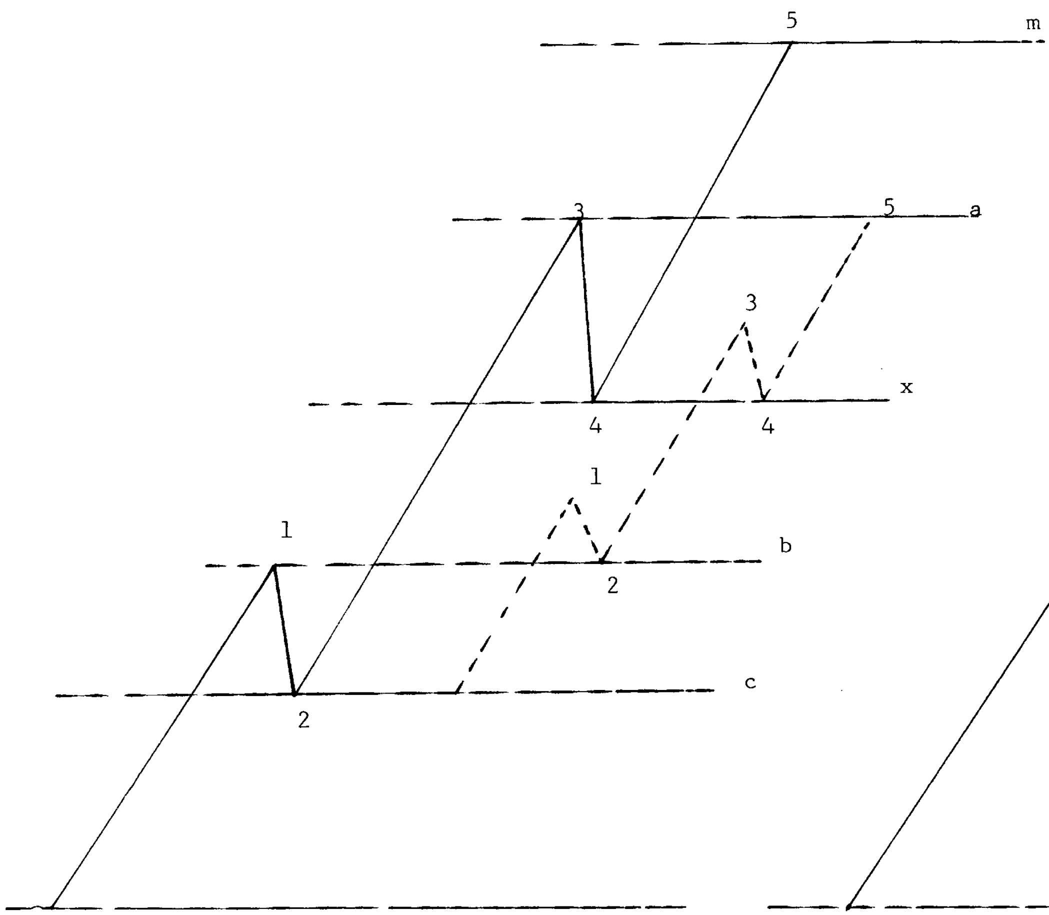

FIGURE 2a Amplitude CASE 1

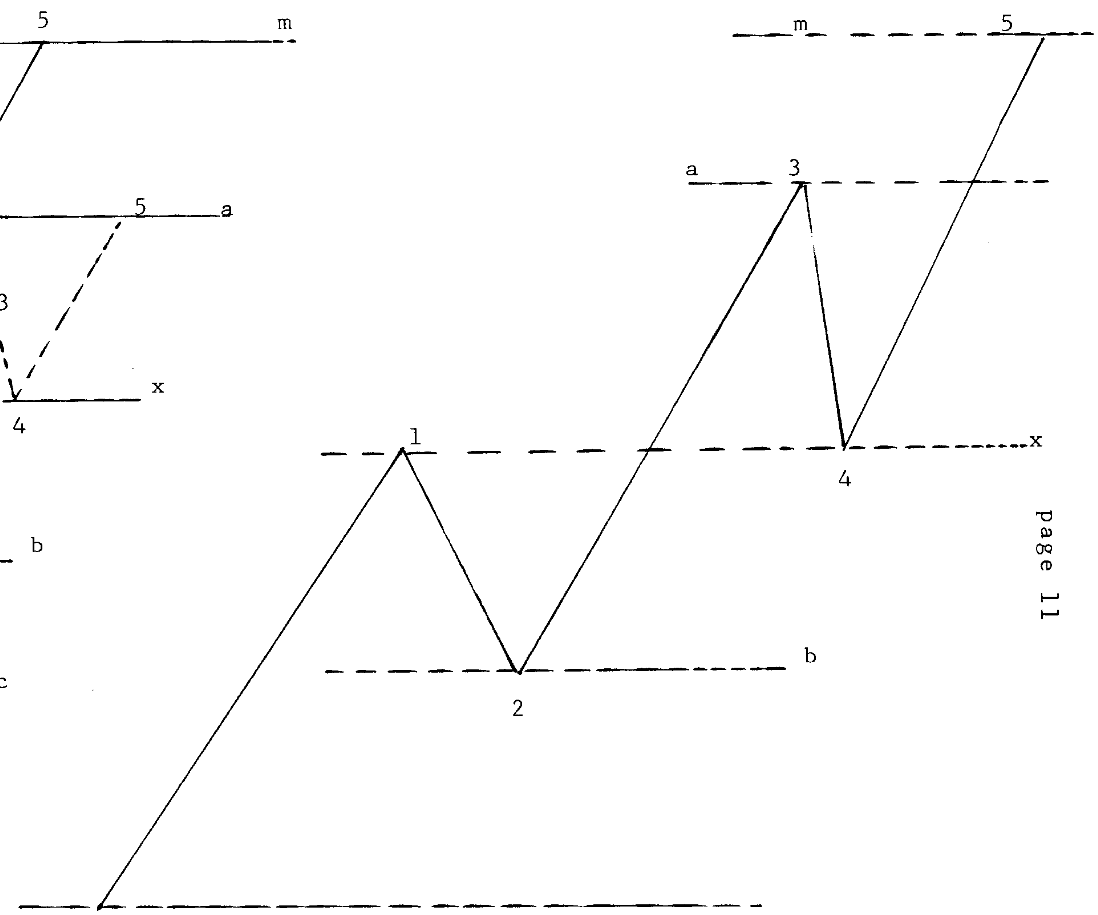

FIGURE 2b Amplitude CASE 2

cerned only with the size and ratios of y-direction movements and not with time.

Four basic equations about the size of the movements can be deduced. These are as follows:

$$c/b = x/m \qquad (1)$$

$$(x-c)/(a-c) = x/m \qquad (2)$$

$$b = m-x \qquad (3)$$

$$(a-c)/m = (x-b)/(x-c) \qquad (4)$$

The first two equations follow from the self-affinity property of the function. Each impulse movement and subsequent correction must replicate or be congruent to every other impulse movement and subsequent correction. Thus the movement from level zero to level m corrects back to level x and this must be congruent to the movement from level zero to level b correcting back to level c; thus the ratio of x to m must equal the ratio of c to b, giving equation (1). Similarly movement-3 followed by correcting movement-4 gives equation (2). Equation (3) follows from an assumed symmetry property, which dictates that the size of movement-5 equal that of movement-1.

Finally, equation (4) comes from the decomposition of movement-3 into a 5-movement replica of the movement from level zero to level m. Thus, appealing to the congruence in the situation, equation (4) is one of several ratios that must hold.

The equations are solved as follows:

Combining (4) and (2):

$$(x-b)/(x-c) = (x-c)/x \qquad (5)$$

Combining (5) with (1) and simplifying:

$$x(1-b/m)^2 = x - b \qquad (6)$$

Combining (6) with (3) gives:

$$(x/m^2)(m - m + x)^2 = x - m + x$$

or: $$(x/m)^3 = 2(x/m) - 1$$

The ratio x/m is the key amplitude ratio in the function. If we let this be k, then the amplitude ratio in the E(t) function obeys the cubic equation:

$$k^3 - 2k + 1 = 0$$

This can be factored as:

$$(k - 1)(k^2 + k - 1) = 0$$

It has one non trivial positive solution:

$$k = (-1 + 5^{0.5})/2 = 0.618$$

This result enables the relative magnitudes in decomposition movements to be computed, and thus fully characterizes the essentials of the upward impulse movements of the function. Referring back to Figure 1, for this case the (vertical) scaling ratio in each replication of the basic 12345abc segment is m/b, which computes to 2.618. To see this recall that x = 0.618m, from the definition of k and the solution for k, and that b = m - x from equation 3, so that b = m(1 - 0.618). Hence m/b is 1/(1 - 0.618) or 2.618.

Although the above equations require that the vertical scaling factor operating in Figure 2a always be 2.618, the equations tells us nothing about the actual scale itself used along the y-axis. It could be a linear scale, or a logarithmic scale, or something else, depending on the application. For example, if the function were to be used as a basis for a model of the price-time behaviour of the Dow-Jones Industrial Index, or any other index, it should clearly be logarithmic, since it is percentage changes that matter over a long period [1, 6, 10]. In that case, the basic scaling ratio would be:

$$\text{Log } (P_5) - \text{Log } (P_0)/(\text{Log } (P_1) - \text{Log } (P_0)) = 2.618$$

where P_0 is the index level at the beginning of movement 1, P_1 is the level at the end of movement 1, and P_5 is the level at the end of movement 5.

Case-2: Corrective movement-4 corrects exactly to point 1.

Consider an arbitrary 5-movement segment of an E(t) function, y = E(t), as shown in Figure 2b. It begins at an arbitrary y value, which we set to zero for convenience, as before. The end of movement-1 is set to level x, of movement-2 to level b, of movement-3 to level a, movement-4 to level x and movement-5 to level m. Thus movement-1 moves a distance x in the y-direction, and movement 5 moves a distance m-x in the y-direction, and so on. We are again concerned only with the size and ratios of y-direction movements and not with time.

The following equations hold:

$$b/x = x/m \qquad \ldots\ldots (7)$$

$$(x - b)/(a - b) = x/m \qquad \ldots(8)$$

$$x = m - x \qquad \ldots(9)$$

Equations (7) and (8) follow from corrective movements always being the same fraction of impulse movements. Equation (9) follows from movement-5 being equal to movement-1. The case is relatively trivial. From (8) it follows that x is m/2, so that, from (7) b/x is 0.5, so that each corrective movement corrects half of each impulse movement.

Referring back to Figure 1, for this case, the (vertical) scaling ratio in each replication of the basic 12345abc segment is clearly 2.0.

Case-3: Corrective movement-4 corrects past point 1.

This case was analysed in the same manner as the previous two. The major assumption was that all corrective movements must be the same fraction of impulse movements. However, this time, no solution for the key ratio was possible. The reader is left to attempt it as an exercise.

Time ratio derivations

At any decomposition level, for any 5-movement upward impulse movement and subsequent 3-movement downward corrective movement of any E(t) function, as far as the time taken for each movement is concerned, there are two major possibilities. At any level of decomposition the major possibilities are:

Case-1: The time taken for each impulse movement is the same, as is the ratio of the time for an impulse movement and the ensuing corrective movement.

Case-2: The time taken for each impulse movement may differ, although the ratio of the time for an impulse movement and the ensuing correction is always the same.

The first case is the simplest.

Case-1 All impulse movements take the same time, as do all corrective movements

Figure 3a shows shows an impulse movement of time length I that decomposes into 5 movements, followed by a corrective movement of time length C that decomposes into 3 movements, without regard to the amplitudes of the movements. All impulse movements in the decomposition are assumed to take the same time i, and all corrective movements in the decomposition are assumed to take the same time c. The following must hold:

$3i + 2c = I$ ....(10)

$2i + c = C$ ....(11)

$c = ki$ ....(12)

$C = kI$ ....(13)

Equations 10 and 11 follow from the decomposition. Equations 12 and 13 follow from the ratio of the time for a corrective movement being a fixed proportion k of an impulse movement. Solving for k:

$$k(3i + 2ki) = 2i + ki$$

giving $k^2 + k - 1 = 0$

which has the positive solution

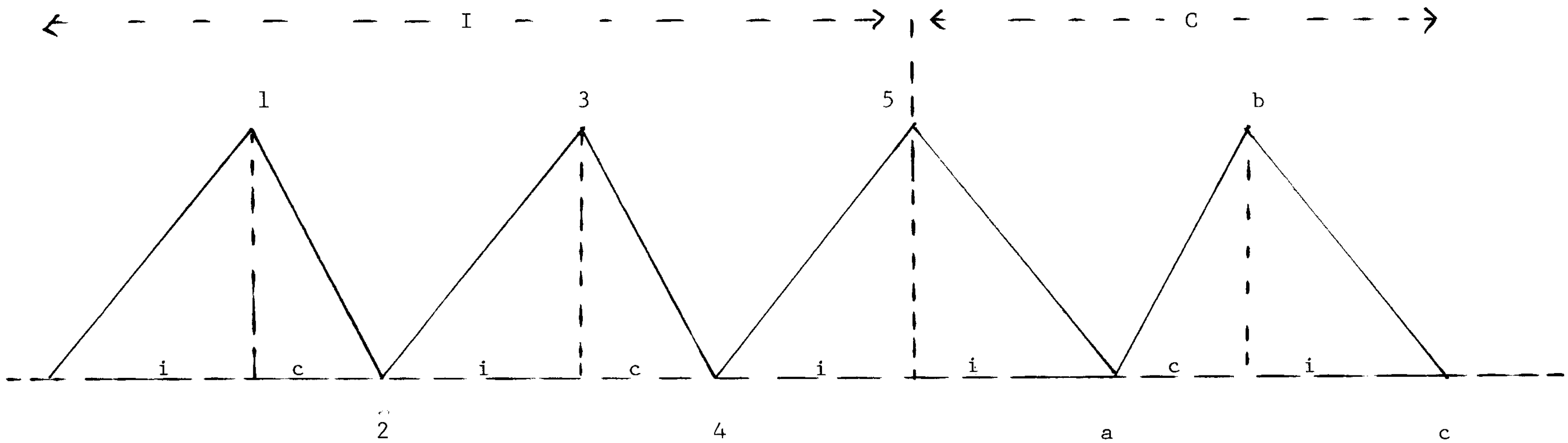

FIGURE 3a Time CASE 1

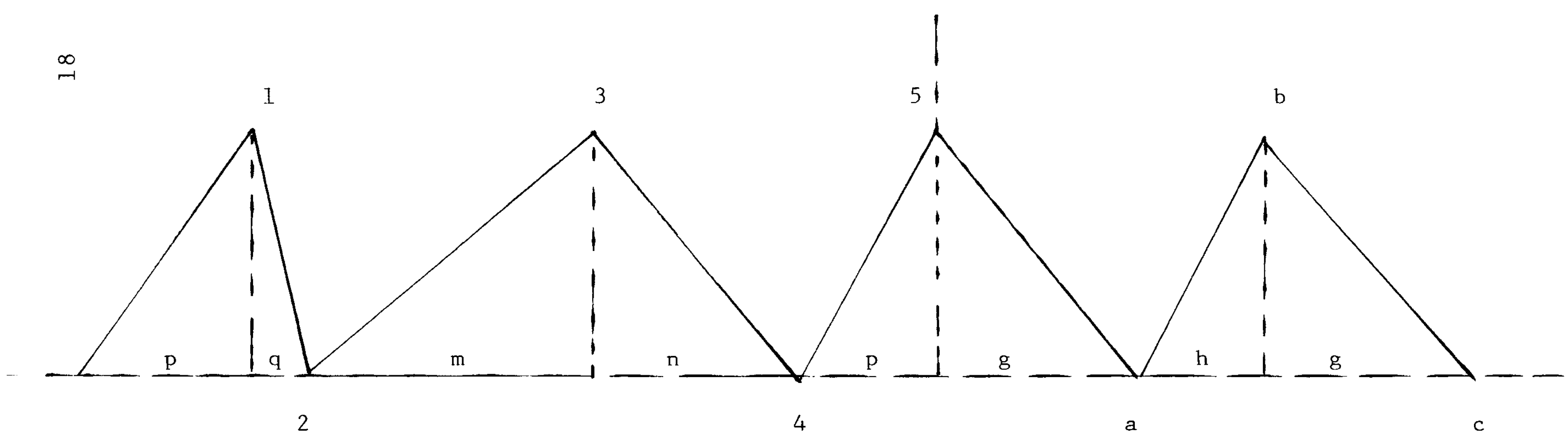

FIGURE 3b Time CASE 2

$$k = (-1 + 5^{0.5})/2 = 0.618$$

Referring back to Figure 1, for this case the (horizontal) scaling ratio in each replication of the basic 12345abc segment is clearly

$$(I + C)/(i + c)$$
or $(3i + 2c + 2i + c)/(i + c)$

or $(5i + 3ki)/(i + ki)$

or $(5 + 3k)/(1 + k)$

or 6.854/1.818, or 4.236

which is different from the ratios derived for Cases 1 and 2 with amplitude scaling ratios.

Case-2 All impulse movements do not take the same time, nor do all corrective movements

There are clearly many possibilities here. We analyse only the obvious one, shown in Figure 3b. Following the relative amplitudes of impulse movements the time p for impulse movement 1 is assumed to be the same as that for impulse movement 5, but differs from the time m for impulse movement 3. The time g for downward impulse movement a is also assumed to be the same as for downward impulse movement c. The three corrective movements are all assumed to have different times q, n, and h. However, everywhere in the function, the time for a corrective movement is assumed to be k times the time for an impulse movement. The following must hold:

$q = kp$ ....(14)

$n = km$ ....(15)

$h = kg$ ....(16)

$2g + h = k(2p + q + n + m)$ .....(17)

$p = kg$ ....(18)

$p = kn$ ....(19)

Equations 14-17 follow from the time for a corrective movement always being k times the time for an impulse movement, no matter what the impulse movement. Equations 18 and 19 follow from symmetry considerations. These are solved for k as folows:

From equations 16 and 18 it follows that $h = p$. From equations 18 and 19, it follows that $g = n$. From equation 17 we therefore have:

$$2n + p = k(2p + kp + n + n/k)$$

Applying equation 19 to this, we get:

$$2n + kn = k(2kn + k^2n + n + n/k),$$

which reduces to $2 + k = 2k^2 + k^3 + k + 1,$

giving $k^3 + 2k^2 - 1 = 0,$

which, taking fractions, gives:

$$(k + 1)(k^2 + k - 1) = 0$$

This cubic has one real positive solution:

$$k = (-1 + 5^{0.5})/2, \text{ or } 0.618.$$

From this the relative times in Figure 3b can be deduced. To simplify ratio calculations, it is convenient to (arbitrarily) set p to unity, so that we get:

$$q = 0.618,\ m = 2.618,\ n = 1.618,\ g = 1.618,\ h = 1$$

Referring back to Figure 1, for this case of unequal impulse move times, the (horizontal) scaling ratio in each replication of the basic 12345abc segment is clearly

$$(p + q + m + n + p + g + h + g)/(p + q)$$

which is 11.09/1.618 or 6.854, or 2.618^2. This ratio is again different from either of the ratios for amplitude (vertical) replication obtained earlier.

Time function constructions

The combination of the amplitude Case-1 construction with time Case-1 construction gives a definition of a time function with specific amplitude and time ratios. Similarly we can combine amplitude Case-1 with time Case-2 to define another function. Combining amplitude Case-2 with time Case-1 gives another time function, and amplitude Case-2 with time Case-2 gives yet another, for a total of four distinct E(t) time functions for all the combinations of the cases considered in this paper. Probably many other distinct E(t) functions could be constructed from development of new amplitude and time ratio cases. In each of the four functions defined here the amplitude scaling or replication ratio is different from the time scaling or replication ratio. The important point is that the existence of such differing ratios means that each of the four functions is an exact self-affine time function.

Of the four functions that can be defined from the cases considered here, one of them, the one constructed by combining amplitude Case-1 with time Case-2, has an unexpected property. The amplitude scaling or replication ratio is 2.618. The time scaling or replication ratio is 2.618^2. Remembering that we are dealing with an exact self-affine function that has no derivitive anywhere, these amplitude and time scaling ratios mean that the structure (and the wiggliness) of the curve is preserved under scaling the time by an arbitrary factor r, provided the amplitude is also scaled by a factor r^h where h is 0.5. But recall that in the case of the random walk time function, if time t is magnified by a factor r then the amplitude must be magnified by the factor r^h where h is 0.5.

Thus a fundamental property of this particular exact self-affine time function is that it scales like a random walk. However it is clear that it is not a random walk. The scaling is a result of a specific well-ordered structure (a 12345abc segment) that is continually replicated using an amplitude scaling ratio of 2.618 and a time scaling ratio of 2.618^2. Preservation of this property under (r, r^h) scaling makes the E(t) function exactly self affine. In contrast with a random walk the (r, r^h) scaling, with h = 0.5, is due to the standard deviation of amplitude changes in fixed-length time intervals being proportional to the square root of the time interval. It is this statistical property that is preserved under r, r^h scaling, thus making a random walk statistically self-affine.

The existence of an E(t) function with random walk scaling leads to speculation about the possible existence of an E(t) function that is a random walk forgery [13]. In principle, such a function might be constructed as follows. For each impulse movement, instead of the single possibility allowed by an exact self affine E(t) function, allow one of a large number to be selected at random, perhaps as a result of a set of coin tosses, so that the impulse movements selected would distribute according to the binomial distribution, or if the pool were large enough, according to the Gaussian distribution. Corrective movements would be selected similarly. Things could thus conceivably be arranged so that on average the ratio between 12345abc amplitude replications is 2.618, and on average the ratio between 12345abc time replications is 2.618^2. With such a function the 12345abc structures would replicate with random walk scaling only **on average**, however. But, in addition, because of the Gaussian distribution of impulse and corrective movements, the standard deviation of movements in fixed time periods should be proportional to the square root of the time period, as in a random walk. Such a function would have to be classified as a statistically self affine E(t) function. The casual observer of such a function, particularly with segments that are many levels down, where the replicating structures are not apparent, could easily, as a result of measurement, conclude that the function was a random walk, when in reality it was an E(t) function that was a random walk forgery. The proof that such a function could exist, although currently lacking, would be an algorithm for its generation [4]. This speculative line of reasoning leads to a final speculation, namely that in spite of the construction of very large financial data bases and careful statistical research in the past forty years, showing that the major market averages essentially follow a random walk (or at least a stochastic) [1, 2, 6, 7, 9], the existence of an exact self-affine E(t) function with random walk scaling also raises the possibility that these averages have more inherent order than can be detected by conventional statistical tests and are in reality random walk forgeries.

Conclusions

Self-affine time functions have much in common with fractals in two-dimensional space but are different from fractals in that they scale differently in the amplitude and time axis. An exact self-affine time function replicates exactly when scaled by appropriate ratios in the amplitude and time axes. A statistical self affine time function replicates only statistically when scaled appropriately in the amplitude and time axis, the best known example being a random walk, where the time scaling factor is the square of the amplitude scaling factor. The existence of at least four exact self affine time functions, called Elliot or E(t) functions, that allow for infinite number of exact replications of 12345abc structures, has been demonstrated. These E(t) functions are defined algorithmically and have no derivitive anywhere. One of these E(t) functions has the unexpected property of scaling like a random walk. This leads to speculation that it might be possible to construct a statistically self-affine E(t) function that would be a random walk forgery.

REFERENCES

1. Alexander, S. S., 1961. Price movements in speculative markets: Trends or random walks, Ind. Management Review, 2 (2), pp. 7-26.

2. Bachelier, L., 1900. Theorie de la speculation, Ann. Sci. Ecole Norm, Sup. (3), No. 1018, Paris, Gauthier-Villars.

3. Elliot, R.N., 1938. The Wave Principle, available in "The Major Works of R.N. Elliot", R. Prechter (ed), New Classics Library, Gainesville, Georgia.

4. Fournier, A. et al, 1982. Computer rendering of stochastic models, Comm. of ACM, 25, pp 53-60.

5. Hutchinson, J., 1981. Fractals and self-similarity, Indiana University Journal of Mathematics, 30, pp 713-747.

6. Kendal, M.G., 1953. Analysis of economic time series, Part 1: Prices, J. Royal Statistical Society, 96 (1), pp 11-25.

7 Mandelbrot, B., 1963. The variation of certain speculative prices, J. of Business, 36 (4), pp. 394-419.

8. Mandelbrot, B. B., Van Ness, V. W., 1968. Fractional Brownian motion, fractional noises and applications, SIAM Review, 10 (4), pp 422-437.

9. Mandelbrot, B. B., 1985. Self-affine fractals and the fractal dimension, Physica Scripta, 32, pp. 257-260.

10. Osbourne, M. F. M., 1959. Brownian motion in the stock market, Operations Research, 7 (2), pp 145-173

11. Peitgen, H., Saupe, D., editors, 1988. The Science of Fractal Images, Springer-Verlag, Berlin.

12. Prusinkiwicz, P., 1986. Graphical applications of L-systems, Proc. Graphics Interface, 1986, pp 247-253.

13. Voss, R.F., 1985. Random fractal forgeries, in Fundamental Algorithms for Computer Graphics, R. A. Earnshaw (ed), Springer-Verlag, Berlin, pp. 805-835.