UNCLASSIFIED

63-2-4

D1-82-0213

BOEING SCIENTIFIC RESEARCH LABORATORIES

Studies in Generalized Random Walks

II. The Expected First Passage Time

A. William Kratzke

Mathematics Research

November 1962

D1-82-0213

STUDIES IN GENERALIZED RANDOM WALKS

II. The Expected First Passage Time

by

A. William Kratzke

Mathematical Note No. 279

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

November 1962

Abstract

A particle moves along a line in steps subject to the following probability law: the particle moves from x to $C_1 + C_2x$ with probability, p or from x to C_3x with probability, $q(= 1 - p)$. Such a movement characterizes a type of random walk and, as such, poses problems associated with random phenomena. One such problem is that of finding the expected number of steps necessary for a particle to cross a pre-assigned level for the first time, having taken its first step from the initial position, x. This expectation is, of course, a function of x and is designated, $D(x)$.

The properties of the function, $D(x)$, are studied and put to use in determining $D(x)$ explicitly. Since the function depends strongly on the choice of the constants, C_2 and C_3, the problem must be treated by cases which are characterized by the relationship between C_2 and C_3. The ease with which solutions are found also depends on the position of the absorbing barrier and several choices are studied.

This paper is a sequel to the paper, Studies in Generalized Random Walks, I. Distribution Functions and Moments, Mathematical Note No. 277. A brief review of the kind of random walk treated in these studies seems in order here.

Consider the motion of a particle in an interval governed by a fixed probability law. The particle position after n steps is given by

$$x_n = \begin{cases} C_1 + C_2 x_{n-1} & \text{with probability } p \\ C_3 x_{n-1} & \text{with probability } q(=1 - p). \end{cases}$$

It is easy to see that the particle is constrained to motion in the interval, $[0, \frac{C_1}{1 - C_2}]$. Having begun its walk in that interval, it can never leave.

An absorbing barrier at some point in the interval, say y, terminates the 'walk.' This means that if x_0, the initial point, is greater than y, the walk never starts and if x_0 is less than y, the motion stops when the particle first passes y. Then, given an initial position, x_0, and an absorbing barrier, y, it is natural to inquire into the problem of finding the number of steps necessary for the particle to pass y. This number is a random variable and finding its expected value is the object of this paper. If each step takes unit time, the problem is a first passage time or absorption time problem. Similarly, if the process is regarded as a game (such as the gambler's analogue discussed in I), the absorption time is just the duration of the game.

II. Expected First Passage Time

1. General expressions and properties.

The most general expression for the expected first passage time may be derived by examining the simple one-step process. If x is the starting point, the next point is either $C_1 + C_2x$ or C_3x with probabilities p and q respectively. Let y be the coordinate at which there is an absorbing barrier and let $D(x)$ be the expected time or number of steps to cross y from below, having started at x. We assume that

$$0 < y < \frac{C_1}{1 - C_2},$$

$$D(x) < \infty,$$

$$D(x) = 0 \quad \text{when} \quad x > y \quad \text{and}$$

$$D(x_1) \geq D(x_2) \quad \text{if} \quad x_1 < x_2.$$

Then, starting at x, if the first step is to $C_1 + C_2x$, the process continues as if $C_1 + C_2x$ were the initial position and if the first step is to C_3x, the process continues as if that were the initial position. Therefore, $D(x)$ may be written in terms of $D(C_1 + C_2x)$ and $D(C_3x)$.

$$\begin{aligned} D(x) &= p(D(C_1 + C_2x) + 1) + q(D(C_3x) + 1) \\ &= pD(C_1 + C_2x) + qD(C_3x) + 1. \end{aligned} \tag{1}$$

This is the basic equation for these absorption time analyses. However, there are several other expressions which prove useful.

If (1) is applied successively to $D(C_3x)$, $D(C_3^2x)$, etc., $D(x)$ may be written as

$$
\begin{aligned}
D(x) &= pD(C_1 + C_2x) + qD(C_3x) + 1 \\
&= pD(C_1 + C_2x) + q[pD(C_1 + C_2C_3x) + qD(C_3^2x) + 1] + 1 \\
&= pD(C_1 + C_2x) + pqD(C_1 + C_2C_3x) + q^2[pD(C_1 + C_2C_3^2x) + qD(C_3^3x) + 1] + 1 + q \\
&= pD(C_1 + C_2x) + pqD(C_1 + C_2C_3x) + pq^2D(C_1 + C_2C_3^2x) + \cdots \\
&\qquad + pq^kD(C_1 + C_2C_3^kx) + \cdots + 1 + q + q^2 + \cdots + q^k + \cdots \\
&= p\sum_{k=0}^{\infty} q^kD(C_1 + C_2C_3^kx) + \sum_{k=0}^{\infty} q^k \\
&= p\sum_{k=0}^{\infty} q^kD(C_1 + C_2C_3^kx) + \frac{1}{p}\,. \qquad (2)
\end{aligned}
$$

Note that $C_1 + C_2x$ may be greater than y or, more generally, it may be that $C_1 + C_2x, C_1 + C_2C_3x, \ldots C_1 + C_2C_3^mx > y$. Then the first $m + 1$ terms of the sum in (2) contribute nothing.

Another expression for $D(x)$, similar to (2), may be found by successive applications of (1) to $D(C_1 + C_2x)$, $D(C_1 + C_1C_2 + C_2^2x)$, etc.

$$
\begin{aligned}
D(x) &= pD(C_1 + C_2x) + qD(C_3x) + 1 \\
&= p[pD(C_1 + C_1C_2 + C_2^2x) + qD(C_3C_1 + C_3C_2x) + 1] + qD(C_3x) + 1 \\
&= p^2[pD(C_1 + C_1C_2 + C_1C_2^2 + C_2^3x) + qD(C_3C_1 + C_3C_1C_2 + C_3C_2^2) + 1] \\
&\qquad + qpD(C_3C_1 + C_3C_2x) + qD(C_3x) + 1 + p \\
&= qD(C_3x) + qpD(C_3(C_1 + C_2x)) + qp^2D(C_3(C_1 + C_1C_2 + C_2^2x)) + \cdots \\
&\qquad + qp^kD\left(C_3\left(C_1\frac{1 - C_2^k}{1 - C_2} + C_2^kx\right)\right) + 1 + p + p^2 + \cdots + p^k. \qquad (3)
\end{aligned}
$$

Since $y < \frac{C_1}{1 - C_2}$,

$$C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^k x > y$$

for all k greater than some n. Then, because $D(u) = 0$ for all $u > y$, (3) has a finite number of terms. If by $[a]$ is meant the greatest integer in a, the highest power of p appearing in (3) is given by

$$N = \left[\frac{\ln(\frac{C_1}{1 - C_2} - y) - \ln(\frac{C_1}{1 - C_2} - x)}{\ln C_2} \right],$$

and it is obvious that $N = m - 1$. Equation (3) can now be written in a more concise form.

$$D(x) = q \sum_{k=0}^{N} p^k D\left[C_3(C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^k x)\right] + \sum_{k=0}^{N} p^k$$

$$= q \sum_{k=0}^{N} p^k D\left[C_3(C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^k x)\right] + \frac{1 - p^{N+1}}{q}. \qquad (4)$$

Having thus established the functional relations, (1), (2), and (4), some general properties of the expected duration function can now be derived.

It is expedient to introduce here the operator notation used in I. The operators, T_1, T_2, and T are defined by

$T_1 x = C_1 + C_2 x$

$T_2 x = C_3 x$

Tx = either $T_1 x$ or $T_2 x$.

One obvious property of these operators is nonlinearity. To be sure, T_2 is a linear operator but any combination of operations that includes at least one T_1 operation is nonlinear.

$$T^m(x_1 + x_2) < T^m x_1 + T^m x_2, \qquad (5)$$

where at least one of the operators is a T_1. Now suppose $T^m x = y$. Then $T^m(x + \Delta x) > y$ and one is led to suspect that the number of steps necessary for a particle to go beyond y, having started at x, is surely at least one greater than the number necessary if the particle had started at $x + \Delta x$. In particular, if only forward steps are considered (assume that the probability of a T_1 operation is one), then (5) implies that the expected duration would have discontinuities at $\frac{y - C_1}{C_2}$, $\frac{y - C_1(1 + C_2)}{C_2^2}$, etc. and the jumps would be of unit amplitude. That there are, in general, discontinuities in $D(x)$ is demonstrated in the following theorem.

Theorem 1.

$C_1 > 0$, $0 < C_2, C_3 < 1$.

There exists discontinuities in the expected duration function.

Proof: First note that the point, y (the absorbing barrier point) is a point of discontinuity.

$$\begin{aligned} &D(y) = 1 + qD(C_3 y) \\ &D(y+) = 0 \qquad (6) \\ &D(y) - D(y(+)) = 1 + qD(C_3 y) > 0, \end{aligned}$$

since $D(x) \geq 0$ for all $x \epsilon [0, \frac{C_1}{1 - C_2}]$. Similarly,

$$D(\frac{y - C_1}{C_2}) = pD(y) + qD(C_3 \frac{y - C_1}{C_2}) + 1$$

$$D(\frac{y - C_1}{C_2} +) = 1 + qD(C_3 \frac{y - C_1}{C_2} +) \qquad (7)$$

$$D(\frac{y - C_1}{C_2}) - D(\frac{y - C_1}{C_2} +) = q[D(C_3 \frac{y - C_1}{C_2}) - D(C_3 \frac{y - C_1}{C_2} +)] + pD(y).$$

Continuing in this way, it is easy to show that many more such discontinuities exist. In fact, there is such a discontinuity at each of the points,

$\frac{1}{C_2^k}(y - C_1\frac{(1 - C_2^k)}{(1 - C_2)})$, $k = 0, 1, 2, \ldots n$, where $n = \left[\frac{\ln(1 - \frac{1 - C_2}{C_1}y)}{\ln C_2}\right]$

(read greatest integer in $\frac{\ln(1 - \frac{1 - C_2}{C_1}y)}{\ln C_2}$).

That there may be other points of discontinuity is also easy to see. For example if $C_3 y > \frac{y - C_1}{C_2}$, then surely there is a discontinuity at $\frac{y - C_1}{C_2 C_3}$ generated by the one at $\frac{y - C_1}{C_2}$, which in turn is generated by the discontinuity at y. This completes the proof of the Theorem.

Extending the results of Theorem 1 provides some coarse bounds on the functions at the points of discontinuity. We will cite only two examples. Consider equation (5). Because of the monotonicity property,

$$D(y) = 1 + qD(C_3 y) \geq 1 + qD(y)$$

$$D(y) \geq \frac{1}{p} \,. \qquad (8)$$

Again using the monotonicity of $D(x)$ and (6),

$$D\left(\frac{y - C_1}{C_2}\right) = pD(y) + qD\left(C_3 \frac{y - C_1}{C_2}\right) + 1$$

$$\geq pD(y) + qD\left(\frac{y - C_1}{C_2}\right) + 1 \qquad (9)$$

$$D\left(\frac{y - C_1}{C_2}\right) \geq D(y) + \frac{1}{p} \geq \frac{2}{p} .$$

Furthermore, it can be shown that

$$D\left(\frac{y - C_1 \frac{(1 - C_2^k)}{(1 - C_2)}}{C_2^k}\right) \geq \frac{k + 1}{p} . \qquad (10)$$

The inequality (10) gives bounds for the values of the function at discontinuity points. In addition, something of the nature of the jump magnitudes can be determined using the same ideas. Consider the saltus at $\frac{y - C_1}{C_2}$.

$$D\left(\frac{y - C_1}{C_2}\right) - D\left(\frac{y - C_1}{C_2} +\right) = p(D(y) - D(y+)) + q\left(D\left(\frac{C_3(y - C_1)}{C_2}\right) - D\left(\frac{C_3(y - C_1)}{C_2} +\right)\right). \qquad (11)$$

Now, if $C_3\left(\frac{y - C_1}{C_2}\right)$ is not a point of discontinuity, the saltus at $\frac{y - C_1}{C_2}$ is just p times the saltus at y. Whether or not $C_3\left(\frac{y - C_1}{C_2}\right)$ is a discontinuity point is determined by the constraints involving C_2 and C_3. The cases $C_2 + C_3 \leq 1$ and $C_2 + C_3 > 1$ must, in general, be handled separately. In an earlier paper (Studies in Generalized Random Walks, I. Distribution Functions and Moments) it was shown that if $C_2 + C_3 \leq 1$, the path that carries a particle from x to y is unique. Therefore $C_3\left(\frac{y - C_1}{C_2}\right)$ is not a point of discontinuity (by point of discontinuity is meant one of the points reached by starting at y and using inverse transformations, $\frac{y - C_1}{C_2}$, $\frac{y - C_1}{C_3C_2}$, etc.) and the second term on the right of (11) vanishes. Similarly,

$$D(\frac{y - C_1}{C_3C_2}) - D(\frac{y - C_1}{C_3C_2}+) = qp(D(y) - D(y+)) = qpD(y). \qquad (12)$$

Having established the fact that $D(x)$ has discontinuities, it remains to determine the nature of the solutions between discontinuities. One is strongly led to suspect that the function remains constant there. Heuristically, the argument for this is that any path that carries x beyond y but does not carry a point in a neighborhood of x beyond y must necessarily have an infinite number of steps or one of the two points would be a point of discontinuity. The paths with an infinite number of steps contribute nothing to the expected value since they have zero probability. Theorem 2, below, demonstrates this more rigorously.

Theorem 2. If the expectation function has a derivative at x, the derivative vanishes there.

Proof: If $D'(x)$ exists, (1) may be differentiated and $D'(x)$ must then satisfy

$$D'(x) = pC_2D'(C_1 + C_2x) + qC_3D'(C_3x). \qquad (13)$$

The derivates of $D(x)$ are

$$\overline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x}, \qquad \overline{\lim_{x_1 \uparrow x}} \frac{D(x_1) - D(x)}{x_1 - x},$$

$$\underline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x}, \text{ and } \underline{\lim_{x_1 \uparrow x}} \frac{D(x_1) - D(x)}{x_1 - x}.$$

Note that each of the derivates satisfies (13), even at the discontinuities, but at those points they are not necessarily equal so the derivatives do not exist there. Let $D'(x)$ be one of the derivates and define M_m and δ as

$$M_m = \{x | D'(x) > -m\} \text{ and } \quad \delta = \inf_M D'(x). \tag{15}$$

Choose x_ϵ such that

$$D'(x_\epsilon) < \delta + \epsilon.$$

Then

$$pC_2 D'(C_1 + C_2 x_\epsilon) + qC_3 D'(C_3 x_\epsilon) = D'(x_\epsilon) < \delta + \epsilon < D'(C_1 + C_2 x_\epsilon) + \epsilon,$$

$$qC_3 D'(C_3 x_\epsilon) \leq (1 - pC_2) D'(C_1 + C_2 x_\epsilon) + \epsilon \tag{16}$$

$$\frac{qC_3}{1 - pC_2} D'(C_3 x_\epsilon) < D'(C_1 + C_2 x_\epsilon) + \frac{\epsilon}{1 - pC_2} .$$

$$pC_2 D'(C_1 + C_2 x_\epsilon) + qC_3 D'(C_3 x_\epsilon) = D'(x_\epsilon) < \delta + \epsilon < D'(C_3 x_\epsilon) + \epsilon$$

$$pC_2 D'(C_1 + C_2 x_\epsilon) < (1 - qC_3) D'(C_3 x_\epsilon) + \epsilon \tag{17}$$

$$\frac{pC_2}{1 - qC_3} D'(C_1 + C_2 x_\epsilon) < D'(C_3 x_\epsilon) + \frac{\epsilon}{1 - qC_3} .$$

Multiply both sides of (16) by $\frac{pC_2}{1 - qC_3}$ and then apply (17).

$$\frac{pC_2 qC_3}{(1 - pC_2)(1 - qC_3)} D'(C_3 x_\epsilon) < \frac{pC_2}{(1 - qC_3)} D'(C_1 + C_2 x_\epsilon) + \frac{pC_2 \epsilon}{(1 - qC_3)(1 - pC_2)}$$

$$< D'(C_3 x_\epsilon) + \frac{\epsilon}{(1 - qC_3)(1 - pC_2)} , \tag{18}$$

$$(pC_2 + qC_3 - 1) D'(C_3 x_\epsilon) < \epsilon .$$

The same procedure using $\frac{qC_3}{1 - pC_2}$ as a multiplier on (17) yields

$$\frac{pC_2qC_3}{(1 - pC_2)(1 - qC_3)} D'(C_1 + C_2x_\epsilon) < \frac{qC_3}{(1 - pC_2)} D'(C_3x_\epsilon) + \frac{qC_3\epsilon}{(1 - qC_3)(1 - pC_2)}$$

$$< D'(C_1 + C_2x_\epsilon) + \frac{\epsilon}{(1 - qC_3)(1 - pC_2)} \qquad (19)$$

$(pC_2 + qC_3 - 1)D'(C_1 + C_2x_\epsilon) < \epsilon$.

Since $p + q = 1$ and $0 < C_2, C_3 < 1$, $pC_2 + qC_3 - 1 < 0$. The monotonicity of $D(x)$ insures that $D'(x) \leq 0$. Therefore

$$(pC_2 + qC_3 - 1)D'(C_3x_\epsilon) \geq 0$$
$$(pC_2 + qC_3 - 1)D'(C_1 + C_2x_\epsilon) \geq 0. \qquad (20)$$

Since ϵ was chosen arbitrarily and $D'(x)$ may be any of the derivates, the conclusion is that

$$\delta = 0. \qquad (21)$$

Finally, the set of x for which (21) is true is simply

$$\{x | D'(x) = 0\} = \bigcup_m M_m \qquad (22)$$

The usefulness of Theorem 2 depends, of course, on the existence of derivatives. The next theorems demonstrate the existence of right-hand and left-hand derivatives.

Theorem 3a. The right-hand derivative of $D(x)$ exists for all x.

Proof: Consider the first and third derivates of (14). Each of these satisfy (13) and is 0 or $-\infty$. If Δ is the difference quotient, it is only necessary to show that $\overline{\lim}\Delta \leq \underline{\lim}\Delta$, since the other inequality is obvious. Take a point, x, in the interval and choose a sequence, ξ, which goes to zero from above $(\xi_n \downarrow 0)$, such that

$$\lim_{n\to\infty} \frac{D(x+\xi_n) - D(x)}{\xi_n} = \overline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x} = 0. \tag{23}$$

Choose a second sequence, $\eta_n \downarrow 0$, such that

$$\lim_{n\to\infty} \frac{D(x+\eta_n) - D(x)}{\eta_n} = \underline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x} = 0. \tag{24}$$

Choose a subsequence ξ_{n_k} such that $\xi_{n_k} > \eta_k$.

Then,

$$0 \geq \frac{D(x+\eta_k) - D(x)}{\eta_k} \geq \frac{D(x+\xi_{n_k}) - D(x)}{\eta_k}$$

$$= \frac{\xi_{n_k}}{\eta_k} \frac{D(x+\xi_{n_k}) - D(x)}{\xi_{n_k}} \geq \frac{D(x+\xi_{n_k}) - D(x)}{\xi_{n_k}}. \tag{25}$$

Combining (23), (24), and (25)

$$\overline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x} \leq \underline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x}, \tag{26}$$

which, since $\overline{\lim}\Delta \geq \underline{\lim}\Delta$, demonstrates that

$$\overline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x} = \underline{\lim_{x_1 \downarrow x}} \frac{D(x_1) - D(x)}{x_1 - x} \tag{27}$$

and the theorem is proved.

Theorem 3b. The left-hand derivative of $D(x)$ exists for all x.

Proof: The proof follows exactly the proof for Theorem 3a.

The last general property considered is that of the uniqueness of the solution. Assuming the existence of functions which satisfy (1), (2), and (4), the following theorem demonstrates that there is only one such solution.

Theorem 4. Let $D_1(x)$ and $D_2(x)$ be two solutions of (1). If $\delta(x) = D_1(x) - D_2(x)$ and $\delta(x) = 0$ for $x > y$, then $\delta(x) = 0$ for all $x \in [0, \frac{C_1}{1 - C_2}]$.

Proof: Let x_1 be a point at which $\delta(x)$ assumes a maximum value. Then

$$\begin{aligned} &p\delta(C_1 + C_2x_1) + q\delta(C_3x_1) = \delta(x_1) \geq \delta(C_1 + C_2x_1) \\ &q\delta(C_3x_1) \geq (1 - p)\delta(C_1 + C_2x_1) \\ &\delta(C_3x_1) \geq \delta(C_1 + C_2x_1). \end{aligned} \tag{28}$$

Similarly,

$$\begin{aligned} &p\delta(C_1 + C_2x_1) + q\delta(C_3x_1) = \delta(x_1) \geq \delta(C_3x_1) \\ &\delta(C_1 + C_2x_1) \geq \delta(C_3x_1) \end{aligned} \tag{29}$$

and, therefore

$$\delta(C_1 + C_2x_1) = \delta(C_3x_1) = \delta(x_1). \tag{30}$$

Then, if x_1 is a maximum point, $C_1 + C_2x_1$ and C_3x_1 are maximum points. Furthermore $T_1^2x_1$, T_1^3, etc. are also maximum points, and for all k greater than

$$\left[\frac{\ln(\frac{C_1}{1 - C_2} - y) - \ln(\frac{C_1}{1 - C_2} - x_1)}{\ln C_2}\right],$$

$$T_1^kx_1 > y.$$

By hypothesis,

$$\delta(T_1^kx_1) = 0,$$

so the conclusion is that

$$\max_{0 \leq x \leq \frac{C_1}{1 - C_2}} \delta(x) = 0. \tag{31}$$

A similar argument may be used to show that

$$\min_{0 \leq x \leq \frac{C_1}{1 - C_2}} \delta(x) = 0, \tag{32}$$

so that, finally

$$\delta(x) = 0 \quad \text{for all} \quad x\varepsilon[0, \frac{C_1}{1 - C_2}]. \tag{33}$$

Although the properties of $D(x)$ set forth in the preceding theorems describe the general behavior of the function, there is no method of solution yielding a function that satisfies (1), (2), and (4) for arbitrary values of the constants, C_1, C_2, and C_3. Therefore, it is necessary to treat cases separately, each case being characterized by a different constraint on the constants.

The first two cases perhaps should not be part of this discussion, since the constant, C_3, is taken to be zero and one for the two cases, and this violates the general constraint, $0 < C_1, C_2, C_3 < 1$. However, the solutions are limiting cases and, as such, give some insight into the kind of solutions to expect.

When $C_3 = 0$, equation (1) becomes

$$D(x) = pD(C_1 + C_2x) + qD(0) + 1. \tag{34}$$

let

$$k = \left[\frac{\ln(\frac{C_1}{1 - C_2} - y) - \ln(\frac{C_1}{1 - C_2} - x)}{\ln C_2}\right] + 1.$$

Applying (34) to itself k times yields, for $D(x)$

$$D(x) = p^kD(C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^kx) + (1 + qD(0))(1 + p + p^2 + \ldots + p^{k-1}). \tag{35}$$

Since $D(x) = 0$ for $x > y$, the first term on the right vanishes.

$$D(x) = \frac{1 - p^k}{q}(qD(0) + 1). \tag{36}$$

Let r be the least number of forward steps $(x \to C_1 + C_2x)$ necessary to carry a particle from zero to a point beyond y.

$$r = \left\lceil \frac{\ln(\frac{C_1}{1 - C_2} - y) - \ln(\frac{C_1}{1 - C_2})}{\ln C_2} \right\rceil + 1.$$

Then $D(0)$ is given by

$$D(0) = \frac{1 - p^r}{qp^r} \tag{37}$$

and finally, $D(x)$ becomes

$$\begin{aligned} D(x) &= \frac{1 - p^k}{q}(qD(0) + 1) \\ &= \frac{1 - p^k}{q}\left(\frac{1 - p^r}{p^r} + 1\right) \\ &= \frac{1 - p^k}{qp^r} . \end{aligned} \tag{38}$$

The other limiting case of interest is that one for which $C_3 = 1$. Here a particle in a one-dimensional random walk never moves backward, but may, with probability q, remain at its present position. The problem of finding the expected number of steps until first passage may be solved directly because the first passage distribution is derivable and the sum defining the first moment is expressible in closed form. Later on, the same result will be obtained using the functional equation (1).

Let k be the least number of forward steps necessary to carry a particle from the point, x, to a point greater than y. This is the same k as the one used in the previous problem. If the number of steps necessary to cross y for the first time having started at x, is N, and $P(i) = \Pr\{N = i\}$, then

$$P(k) = p^k. \tag{39}$$

For $N = k + 1$, consider the paths that have $k + 1$ steps, start at x, and terminate beyond y. Initially, and after each forward step, one move that goes from x to $x(x \to C_3x = 1)$ may take place. There are just k ways for this to happen so

$$P(k + 1) = kqp^k. \tag{40}$$

Recall that the number of ways one can put C indistinguishable objects into M numbered boxes with no restrictions on the number of objects in any one box is

$$\frac{(M + C - 1)!}{(M - 1)!C!}.$$

If the objects correspond to the steps, $x \to C_3x = x$ and the boxes correspond to forward step positions given by

$$C_1 \frac{1 - C_2^i}{1 - C_2} + C_2^i x, \; i = 1, 2, \ldots k,$$

then the number of ways that the particle can experience n 'standing still' steps at k positions, with no restriction on the number at each position is

$$\frac{(k + n - 1)!}{(k - 1)!n!}.$$

Therefore,

$$P(n + k) = \frac{(k + n - 1)!}{(k - 1)!n!} q^n p^k. \tag{41}$$

Then, it remains but to multiply (41) by $(n + k)$ and sum over all n to find $D(x)$. If $\bar{N}$ is the first moment of N,

$$D(x) = \bar{N} = p^k \sum_{n=0}^{\infty} \frac{(n + k - 1)!}{(k - 1)!n!}(n + k)q^n$$

$$= p^k k \sum_{n=0}^{\infty} \frac{(n + k)!}{k!n!}q^n \tag{42}$$

$$= \frac{kp^k}{(1 - q)^{k+1}} = \frac{k}{p}.$$

The other method for obtaining the solution (42) involves use of the basic equation (1). When $C_3 = 1$, that equation becomes

$$D(x) = pD(C_1 + C_2x) + qD(x) + 1$$
$$D(x) = D(C_1 + C_2x) + \frac{1}{p}. \tag{43}$$

Again, one applies (43) to itself k times, i.e., until the argument of the first term on the right exceeds y.

$$D(x) = D(C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^kx) + \frac{k}{p}$$
$$= \frac{k}{p}. \tag{44}$$

As expected, the results of (42) and (44) agree.

These limiting cases are useful when only bounds are needed for the function. Note that for any combination of constants subject to the condition, $0 < C_1, C_2, C_3 < 1$, $D(x)$ is bounded above and below by the solutions for $C_3 = 0$ and $C_3 = 1$.

$$\frac{k}{p} < D(x) < \frac{1 - p^k}{qp^r}. \tag{45}$$

Considered next are the solutions of (1) when $C_2 + C_3 \leq 1$. In the interests of simplicity, let the absorbing barrier first be taken as one of the points reached by forward steps only, having started at the origin, i.e.

$$y = C_1 \frac{(1 - C_2^k)}{1 - C_2} \quad \text{for some} \quad k.$$

First, let $y = C_1$. Then for all x in the interval $(0, C_1]$

$$\begin{aligned} D(x) &= 1 + qD(C_3 x) \\ &= 1 + q(1 + qD(C_3^2 x)) \\ &= 1 + q + q^2 + \ldots + q^n D(C_3^n x). \end{aligned}$$

Since, by hypothesis, $D(0) < \infty$, the last term vanishes, and

$$D(x) = \sum_{i=0}^{\infty} q^i = \frac{1}{p} . \tag{46}$$

From (1),

$$D(0) = D(C_1) + \frac{1}{p}, \tag{47}$$

and (1) is then solved for all $x \varepsilon [0, C_1]$.

The solution (46) is available immediately, if it is recognized that

$$D(x_1) = D(x_2), \quad 0 < x_1, x_2 \leq C_1 . \tag{48}$$

To see this, note that every path that carries x_1 beyond y also carries x_2 there. Note, also, that, in this one instance, it matters not what the constraint relating C_2 to C_3 is, just so long as $0 < C_3 < 1$. If (48) obtains,

$$\begin{aligned} D(x) &= 1 + qD(C_3 x) = 1 + qD(x) \\ D(x) &= \frac{1}{p}, \end{aligned} \tag{49}$$

and then (46) and (48) agree.

Next, let $y = C_1(1 + C_2)$. When x is in the interval $(0,C_1]$

$$D(x) = pD(C_1 + C_2x) + qD(C_3x) + 1$$

$$= p^2D(C_1(1 + C_2) + C_2^2x) + pqD(C_3(C_1 + C_2x) + qD(C_3x) + 1 + p.$$

$$= 0 + pqD(C_3(C_1 + C_2x)) + qD(C_3x) + 1 + p. \qquad (50)$$

Using the same method as that used in the previous problem, i.e. iterating (50)

$$D(x) = pqD(C_3(C_1 + C_2x)) + qD(C_3x) + 1 + p$$

$$= p^2q^2D(C_3C_1(1 + C_2C_3) + C_2^2D_3^2x) + pq^2D(C_3^2(C_1 + C_2x))$$

$$+ pq^2D(C_3(C_1 + C_2C_3x)) + q^2D(C_3^2x) + (1 + p)(1 + q + pq)$$

$$= (1 + p)(1 + (1 - p^2) + (1 - p^2)^2 + \ldots) \qquad (51)$$

$$= (1 + p) \sum_{i=0}^{\infty} (1 - p^2)^i$$

$$= (1 + p) \frac{1}{1 - (1 - p^2)}$$

$$= \frac{1 + p}{p^2}$$

$$= \frac{1}{p^2} + \frac{1}{p}.$$

Again, the finiteness of $D(x)$ assures the vanishing of all but the last term in iterative scheme. If x is in the other interval, $(C_1,C_1(1 + C_2)]$,

$$D(x) = 1 + qD(C_3x).$$

But $C_3x < C_1$. So

$$D(x) = 1 + q(\frac{1}{p^2} + \frac{1}{p})$$

$$= 1 + \frac{q}{p^2}(1 + p) \qquad (52)$$

$$= 1 + \frac{1 - p^2}{p^2} = \frac{1}{p^2}.$$

Note that the argument following (48) applies here, too. However, the constraint $C_2 + C_3 \leq 1$ is necessary so that $C_3x < C_1$ for all x in $[0,y]$. Following the method for obtaining (49),

$$D(x) = pqD(C_3(C_1 + C_2x)) + qD(C_3x) + 1 + p \qquad 0 < x \leq C_1$$

$$D(x) = \frac{1 + p}{1 - pq - q} = \frac{1 + p}{p^2} = \frac{1}{p^2} + \frac{1}{p}. \qquad (53)$$

Of course, the same thing applies for $x \varepsilon (C_1, C_1(1 + C_2)]$. Now, one can generalize for

$$y = C_1 \frac{1 - C_2^k}{1 - C_2}.$$

First let $x \varepsilon (0, C_1]$. The method is as follows: iterate the first term of (1) until the argument of the first term is greater than y.

$$D(x) = pD(C_1 + C_2x) + qD(C_3x) + 1$$

$$= p^2D(C_1(1 + C_2) + C_2^2x) + pqD(C_3(C_1 + C_2x)) + qD(C_3x) + 1 + p$$

$$= p^kD(C_1 \frac{1 - C_2^k}{1 - C_2} + C_2^kx) + q\left\{p^{k-1}D(x_{k-1}) + p^{k-2}D(x_{k-2}) + \ldots + D(C_3x)\right\}$$

$$+ 1 + p + p^2 + \ldots + p^{k-1}, \qquad (54)$$

where $0 < x_i \leq C_1$ for $i = 1,2,\ldots k - 1$. If we designate C_3x by x_0, a more compact form for (53) is available. In fact, a special case of equation (4) obtains. The first term on the right side of (53) vanishes and

$$D(x) = q\sum_{i=0}^{k-1} p^i D(x_i) + \sum_{i=0}^{k-1} p^i$$
$$= q\sum_{i=0}^{k-1} p^i D(x_i) + \frac{1-p^k}{q} . \tag{55}$$

Again, remembering that $D(x) < \infty$, iterating (54) yields

$$D(x) = \frac{1-p^k}{q}\left(1 + (1-p^k) + (1-p^k)^2 + \ldots\right)$$
$$= \frac{1-p^k}{(1-p)p^k}$$
$$= \frac{1 + p + p^2 + \ldots + p^{k-1}}{p^k} \tag{56}$$
$$= \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \ldots + \frac{1}{p^k} .$$

A simpler, more direct solution is again available if the same heuristic argument is used as that used for finding (49) and (52). All the arguments in the sum in (54) are in the interval $(0,C_1]$ so we may equate all of the functions $D(x_i)$ to $D(x)$.

$$D(x) = D(x)q\sum_{i=0}^{k-1} p^i + \frac{1-p^k}{q}$$
$$D(x) = \frac{1-p^k}{q} \frac{1}{1 - q\left(\frac{1-p^k}{q}\right)} \tag{57}$$
$$= \frac{1-p^k}{qp^k} ,$$

which agrees with (55). For $x \varepsilon \left(C_1 \frac{1-C_2^{n-1}}{1-C_2}, C_1 \frac{1-C_2^n}{1-C_2}\right]$, $n = 2,3, \ldots k$, the solution follows the same methods as for the derivation of (52). Then,

$$D(x) \begin{cases} = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots + \frac{1}{p^k} & 0 < x \leq C_1 \\ = \frac{1}{p^2} + \frac{1}{p^3} + \cdots \frac{1}{p^k} & C_1 < x \leq C_1(1 + C_2) \\ = \frac{1}{p^3} + \frac{1}{p^4} + \cdots \frac{1}{p^k} & C_1(1 + C_2) < x \leq C_1 \frac{(1 - C_2^3)}{(1 - C_2)} \\ \vdots & \\ = \frac{1}{p^k} & C_1 \frac{(1 - C_2^{k-1})}{1 - C_2} < x \leq C_1 \frac{1 - C_2^k}{1 - C_2}. \end{cases} \tag{58}$$

Next, we shall derive $D(x)$ for the case $C_2 + C_3 \leq 1$ but

$$y \neq C_1 \frac{1 - C_2^k}{1 - C_2} .$$

Again, the simplest case will be solved first. If $0 < y \leq C_1$, the earlier results apply and

$$D(x) = \frac{1}{p} \qquad 0 \leq x \leq y. \tag{59}$$

When y is not one of the points, $C_1 \frac{1 - C_2^k}{1 - C_2}$, the solutions become much more involved. The derivations follow the same pattern as before but the intervals which divide the domain of the function, $D(x)$, become more complicated. Only one case is presented here but all other possible choices for y lead to the same type of derivation. This case is, then, more of an example than a useful result since the choice of position for the absorbing barrier is not very far out on the line of admissible values. In fact, y is chosen to lie between the first and second forward step, starting at the origin, i.e., $C_1 < y \leq C_1(1 + C_2)$. It will become evident later on that a further subdivision on the values y may assume is necessary and to that end, the following inequality must be proved.

$$C_1 < \frac{C_1}{1 - C_2C_3} \leq C_1(1 + C_2). \tag{60}$$

Since both C_2 and C_3 are less than one, the left side of the inequality is obvious. We need only the constraint, $C_2 + C_3 \leq 1$ to show the other.

$$\begin{aligned}
\frac{C_1}{1 - C_2C_3} &\leq \frac{C_1}{1 - C_2(1 - C_2)} \\
&= C_1(1 + C_2(1 - C_2) + C_2^2(1 - C_2)^2 + C_2^3(1 - C_2)^3 + \ldots) \\
&= C_1(1 + C_2) + C_1(- C_2^2 + \sum_{i=2}^{\infty} C_2^i(1 - C_2)^i) \\
&= C_1(1 + C_2) + C_1(- C_2^2 + \frac{C_2^2(1 - C_2)^2}{1 - C_2(1 - C_2)}) \\
&= C_1(1 + C_2) - \frac{C_2^3}{1 - C_2(1 + C_2)} \\
&\leq C_1(1 + C_2).
\end{aligned}$$

First, take the case,

$$\frac{C_1}{1 - C_2C_3} < y \leq C_1(1 + C_2). \tag{61}$$

From (61), we note that

$$\begin{aligned}
\frac{y - C_1}{C_2} &= \frac{y - C_1}{C_2} + C_3y - C_3y \\
&= \frac{y - C_1 - C_2C_3y}{C_2} + C_3y \\
&= \frac{y(1 - C_2C_3) - C_1}{C_2} + C_3y \\
&\geq \frac{C_1 - C_1}{C_2} + C_3y = C_3y.
\end{aligned} \tag{62}$$

The following sketch is an aid to remembering the ordering of pertinent points:

$$0 \qquad C_3y \qquad \frac{y - C_1}{C_2} \qquad C_1 \qquad \frac{C_1}{1 - C_2C_3} \qquad y \qquad C_1(1 + C_2) .$$

Now, to return to the problem. When $0 < x \leq \frac{y - C_1}{C_2}$,

$$\begin{aligned} D(x) &= pD(C_1 + C_2x) + qD(C_3x) + 1 \\ &= pqD(C_3(C_1 + C_2x)) + qD(C_3x) + 1 + p. \end{aligned} \qquad (63)$$

From (62) and the condition on x

$$C_3(C_1 + C_2x) \leq C_3y \leq \frac{y - C_1}{C_2},$$

so the arguments of both of the $D(\cdot)$ functions on the right side of (63) are to the left of x. The solution of (63) then reduces to that of (53) and

$$D(x) = \frac{1}{p} + \frac{1}{p^2}. \qquad (64)$$

When $\frac{y - C_1}{C_2} < x \leq y$,

$$D(x) = 1 + qD(C_3x).$$

But

$$C_3x \leq C_3y \leq \frac{y - C_1}{C_2},$$

so

$$D(x) = 1 + q\left(\frac{1}{p} + \frac{1}{p^2}\right) = \frac{1}{p^2}. \qquad (65)$$

The remaining part of the interval, $(C_1, C_1(1 + C_2)]$, cannot be taken as the set from which possible values for y may be drawn and get a single result for $D(\cdot)$. To see this note that different points of discontinuity (relative to y) arise when different values of y are taken from $(C_1, \frac{C_1}{1 - C_2C_3}]$. (The latter set is simply

$$(C_1, C_1(1 + C_2)] - (\frac{C_1}{1 - C_2C_3}, C_1(1 + C_2)].)$$

The method for finding the values of y which admit of a single solution (in terms of y) is to determine the number of points of discontinuity to allow and then find the range of y which yields that number of such points. Then by appealing to Theorems 2 and 3, we may take $D(\cdot)$ to be constant in each interval between adjacent points of discontinuity. For example, if only three regions are wanted, proceed as follows: Take y as the first point of discontinuity, and since $y > C_1$, surely $\frac{y - C_1}{C_2}$ is also such a point (Theorem 1). The point,

$$T_2^{-1}(\frac{y - C_1}{C_2}) = \frac{y - C_1}{C_2C_3}$$

may be the last point of discontinuity or it may generate more points. To insure that $\frac{y - C_1}{C_2C_3}$ is the last such point, neither

$$T_1^{-1}(\frac{y - C_1}{C_2C_3}) \text{ nor } T_2^{-1}(\frac{y - C_1}{C_2C_3})$$

may exist. The range of y must be such that

$$T_1^{-1}(\frac{y - C_1}{C_2C_3}) \leq 0 \text{ and}$$

$$T_2^{-1}(\frac{y - C_1}{C_2C_3}) > y,$$

or

$$C_3 y < \frac{y - C_1}{C_2 C_3} \le C_1 ,$$

and the range of y is, then,

$$\frac{C_1}{1 - C_2 C_3^2} < y \le C_1 (1 + C_2 C_3), \qquad (66)$$

and these points are in $\left(C_1, \frac{C_1}{1 - C_2 C_3}\right]$.

$$\left(\frac{C_1}{1 - C_2 C_3^2}, C_1(1 + C_2 C_3)\right] \subset \left(C_1, \frac{C_1}{1 - C_2 C_3}\right].$$

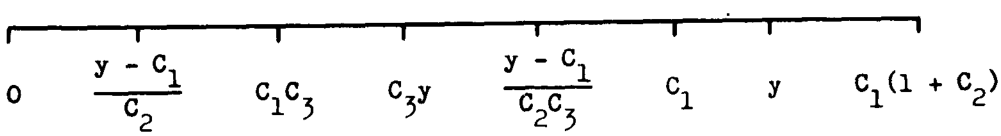
Now, one can proceed to iterate the functional equation. When $0 < x \le \frac{y - C_1}{C_2}$,

$$\begin{aligned} D(x) &= pD(C_1 + C_2 x) + qD(C_3 x) + 1 \\ &= pqD(C_3(C_1 + C_2 x)) + qD(C_3 x) + 1 + p. \end{aligned} \qquad (67)$$

But

$$C_3(C_1 + C_2 x) > C_1 C_3 > \frac{y - C_1}{C_2}.$$

Here, it is convenient to sketch the ordering of points and label the regions between discontinuities.



Because of Theorems 2 and 3,

$$D(x) = D(x_i), \; x \in X_i .$$

Then the iterated equation becomes

$$\begin{aligned} D(x_1) &= pqD(x_2) + qD(x_1) + 1 + p \\ &= pq^2 D(x_1) + qD(x_1) + 1 + p + pq \end{aligned} \qquad (68)$$

$$D(x_1) = \frac{1 + p + pq}{1 - q - pq^2} = \frac{1}{p^2(1 + q)} + \frac{1}{p} .$$

Similarly, for $\frac{y - C_1}{C_2} < x \le \frac{y - C_1}{C_2 C_3}$,

$$D(x_2) = 1 + qD(x_1) \tag{69}$$
$$= 1 + q(\frac{1}{p^2(1 + q)} + \frac{1}{p})$$
$$= \frac{1}{(1 + q)}(\frac{1}{p^2} + \frac{1}{p} - 1),$$

and for $\frac{x - C_1}{C_2 C_3} < x \le y$,

$$D(x_3) = 1 + qD(x_2) \tag{70}$$
$$= 1 + q + q^2(\frac{1}{p^2(1 + q)} + \frac{1}{p}$$
$$= \frac{1}{p^2(1 + q)} .$$

The two problems solved above, are, of course, special cases. Some generalization is possible so that the arduous task of seeking points of discontinuity is not always necessary. The first of the two problems solved above is especially easy to generalize. Consider those values of y which lie in the intervals given by

$$C_1 \frac{(1 - C_2^{k-1})}{(1 - C_2)(1 - C_2^{k-1}C_3)} < y \le C_1 \frac{1 - C_2^k}{1 - C_2}, \quad k = 1,2,3,\ldots .$$

The functional equation (4) may be used now to determine the function. When

$$0 < x \le \frac{1}{C_2^{k-1}}(y - C_1 \frac{(1 - C_2^{k-1})}{1 - C_2}),$$

$$D(x) = q \sum_{i=0}^{k-1} p^i D(C_3(C_1 \frac{1 - C_2^i}{1 - C_2} + C_2^i x)) + \frac{1 - p^k}{q} \tag{71}$$

$$= \sum_{i=0}^{k-1} p^i (1 + qD(C_3(C_1 \frac{1 - C_2^i}{1 - C_2} + C_2^i x)).$$

Now each of the arguments of the $D(\cdot)$ function is less than

$$\frac{1}{C_2^{k-1}}(y - C_1 \frac{(1 - C_2^{k-1})}{1 - C_2}),$$

so, again appealing to Theorems 2 and 3, (71) may be written

$$D(x) = \frac{1 - p^k}{q}(1 + qD(x)) \qquad (72)$$

and, finally,

$$D(x) = \frac{1 - p^k}{qp^k} . \qquad (73)$$

The same method applies when x is in the next interval, viz.,

$$\frac{1}{C_2^{k-1}}(y - C_1 \frac{(1 - C_2^{k-1})}{1 - C_2} < x \leq \frac{1}{C_2^{k-2}}(y - C_1 \frac{(1 - C_2^{k-2})}{1 - C_2}).$$

Let $x = x_1$ when x is in

$$(0, \frac{1}{C_2^{k-1}}(y - C_1 \frac{(1 - C_2^{k-1})}{(1 - C_2)})]$$

and $x = x_2$ when x is in

$$(\frac{1}{C_2^{k-1}}(y - C_1 \frac{(1 - C_2^{k-1})}{(1 - C_2)}), \frac{1}{C_2^{k-2}}(y - C_1 \frac{(1 - C_2^{k-2})}{(1 - C_2^{k-2})})].$$

Then the functional equation becomes

$$\begin{aligned} D(x_2) &= \frac{1 - p^{k-1}}{q}(1 + qD(x_1)), \qquad (74) \\ &= \frac{1 - p^{k-1}}{q}(1 + \frac{1 - p^k}{p^k}) \\ &= \frac{1 - p^{k-1}}{qp^k} . \end{aligned}$$

The previous examples suffice to demonstrate the method for finding $D(x)$ when $C_2 + C_3 \leq 1$. There are always a finite number of points of discontinuity between zero and y. For some n,

$$C_3 y < T^{-n} y \leq C_1 \tag{75}$$

and then $T^{-n}y$ is the last point of discontinuity found by applying inverse operations, T_1^{-1} and T_2^{-1}. Note that $C_3 u < C_1$ for all

$$u \varepsilon [0, \frac{C_1}{1 - C_2}] \text{ and } C_2 + C_3 \leq 1, \text{ so } C_3 u \leq \frac{C_3 C_1}{1 - C_2} \leq C_1.$$

The reader may verify the fact that no more than a finite number of inverse operations are necessary to assure that $T^{-n} y \varepsilon (C_3 y, C_1]$, using the arguments of Section 2 of the paper, "Studies in Generalized Random Walks, I. Distribution Functions and Moments." If there are m points of discontinuity, exclusive of zero, and they are denoted by ξ_i, let $D(x) = D(x_i)$ for $\xi_{i-1} < x \leq \xi_i$. Equations (2) or (4) may then be used to set up a system of equations,

$$D(x_i) = f(x_1, x_2, \ldots x_m, b_1, b_2, \ldots) \tag{76}$$

where $f(\ldots)$ is a function of the x_i's and constants, $b_i, b_2, \ldots$, which involve powers of p and q. The solution of (76) is then the complete solution for $D(\cdot)$.

No simple method has been devised for finding $D(x)$ when $C_2 + C_3 > 1$. There are, however, several approximation techniques. These approximations will be discussed in another paper in connection with the applications. The conditions that are imposed by some of the applied problems make the solutions quite special and are, therefore, more logically discussed with those problems.