ON THE LINEAR SEARCH PROBLEM*

BY
ANATOLE BECK

ABSTRACT

A man in an automobile searches for another man who is located at some point of a certain road. He starts at a given point and knows in advance the probability that the second man is at any given point of the road. Since the man being sought might be in either direction from the starting point, the searcher will, in general, have to turn around many times before finding his target. How does he search so as to minimize the expected distance travelled? When can this minimum expectation actually be achieved? This paper answers the second of these questions.

The purpose of this paper is to prove an existence theorem concerning the linear search problem. The problem, which this author has been circulating for some years, is the following. A point t is placed on the real line R according to a known probability distribution F. A search is made for the point by executing a continuous path in R starting at 0. A *search plan* is a program of the the following type: Start in the positive (or negative) direction, and if the point is not found before reaching x_1, turn around and explore the other half of the real line as far as x_2. If this still does not yield the point, turn around again and explore as far as x_3, etc. Thus, we can represent the search plan as a sequence $x = \{x_i\}$ with $\cdots \leqq x_4 \leqq x_2 \leqq 0 \leqq x_1 \leqq x_3 \leqq \cdots$ or $\cdots \leqq x_3 \leqq x_1 \leqq 0 \leqq x_2 \leqq x_4 \leqq \cdots$. In certain cases, as we shall note below, there may be only finitely many entries in a search plan, and one of them (but not more) might be infinite. We use the weak inequalities for technical reasons. If all the inequalities are strong, we call x a *strong search plan*, otherwise it is said to be *weak*. A search plan as so defined represents all the meaningful planning the searcher can do, since no new information is coming in except that t has not yet been found, and this supposition is made in constructing the plan. Thus, assume that F is fixed, and that some plan x has been chosen. For each t, we let $X(x,t)$ be the length of the path from 0 to t according to the search procedure x. For each x, $X(x,t)$ is a random variable, and we define

$$X(x) = E(X(x,t)) = \int_{-\infty}^{\infty} X(x,t)\,dF(t).$$

Received January 29, 1965.

* The research work for this paper was supported by the National Science Foundation (under Grant No. GP 2559 to the University of Wisconsin) and the Wisconsin Alumni Research Foundation.

It has long been known by this author (and probably many others) that $X(x)$ can be made finite for some search plan x exactly if the first moment $M_1(F) = \int_{-\infty}^{\infty} |t| \, dF(t) < \infty$. In that case, we define $x_\delta = \{\delta, -2\delta, 4\delta, \cdots\}$ and then $X(x_\delta) < \infty$ and in fact $\limsup_{\delta \to 0} X(x_\delta) \leqq 9M_1$ (see e.g. [1]). We shall assume throughout that $\int_{-\infty}^{\infty} |t| \, dF(t) < \infty$.

There is an infimum for all the $X(x)$, and we designate it as $m_0 = m_0(F)$ throughout this paper. The main problem is to produce a method for finding a search plan y with $X(y) = m_0$, or at least a y_ε with $X(y_\varepsilon) < m_0 + \varepsilon$ for each $\varepsilon > 0$. A "best answer" would be a formula for the x_i in terms of F, (or at least an algorithm for deciding whether x_1 is positive or negative). In the pursuit of this problem, it is interesting and useful to know whether it is possible to realize the infimum for some search plan.

Under date of October 1963, Wallace Franck, of the University of New Mexico has circulated a preprint (see reference) in which he gives sufficient conditions for the existence of a minimum and also an example in which there is no minimum. The heart of his work is Lemma 1 (pp. 4–5) and an example (pp. 13–14). In this paper, we sharpen these last two results to produce a necessary and sufficient condition for the attainment of the minimum. Aside from the sharpened results (Lemma 3 and Theorem 13 of this paper), the other techniques are reasonably simple and not essentially different from those used by Franck. They are included only for the sake of completeness. Our main theorem is as follows:

1. THEOREM. *Let F be a probability distribution on the real line and let $M_1 = \int_{-\infty}^{\infty} |t| \, dF(t) < \infty$. Define $X(x)$ as in the above discussion, and let $F^+(a) = \limsup_{t \to 0+} \frac{(F(a+t) - F(a^+))}{t}$, $F^-(a) = \limsup_{t \to 0-} \frac{(F(a+t) - F(a^-))}{t}$, Then there is a search plan y with $X(y) \leqq X(x)$ for all search plans x if and only if at least one of $F^+(0)$ and $F^-(0)$ is finite.*

Let us dispose immediately of the trivial cases. If the probability that t is negative (resp. positive) is 0, then the only reasonable way to search is to proceed to the right (resp. left) until t is found. Because of cases like this, we require only that t be found almost surely (i.e. with probability 1), and we allow the possibility of a search plan with only finitely many entries, one of which might be $\pm\infty$. Otherwise, if we were to require that every point be searched, even simple cases like this would have no minima, thus ruining an otherwise interesting problem.

Let x^- and x^+ be so designated that $F(t) = 0, \forall t < x^-$, $F(t) = 1, \forall t > x^+$, and $0 < F(t) < 1, \forall x^- < t < x^+$. In any case where $x^- > -\infty$ or $x^+ < +\infty$, we allow the possibility that search procedures may be finite and may have $\pm\infty$ for the last entry.

Whenever we reach a point, we shall assume it has been searched, and thus we normalize F by assuming that F is continuous from the right in the positive half of the real axis and from the left in the negative half. The jump at 0, if there is one, is unimportant, as 0 is searched immediately at the outset. Thus, our answer will be the same whether we consider the given probability distribution or the conditional distribution on the hypothesis $t \neq 0$. The distribution function for the conditional probability is continuous at 0, and thus it will not affect the generality of our result if we assume that F is continuous at 0.

Suppose we choose a sequence $x^{(n)} = \{x_i^{(n)}\}$ with $X(x^{(n)}) \to m_0$. A reasonable procedure would be to define $y_i = \lim_{n\to\infty} x_i^{(n)}$ and prove that $X(y) = m_0$, where $y = y_i$. In fact, where possible, that is exactly our procedure. We must be wary, however of the possibility that $x_i^{(n)} \to 0$, $\forall i$, and this unhappy circumstance can occur regardless of the nature of F, since for any search plan x, arbitrarily many points can be added at the beginning arbitrarily close to 0, such that the change in $X(x)$ is as small as we like. Furthermore, $x_i^{(n)}$ can diverge to $\pm\infty$ but not, as we shall see, if $x^- = -\infty$ and $x^+ = +\infty$. So let us take this case first.

2. Lemma. *If $x^- = -\infty$, $x^+ = +\infty$, then we can find a sequence $\{b_i\}$ such that $|x_i| < b_i < \infty$, $\forall i = 1, 2, \cdots$, holds for every search plan x with $X(x) < 2m_0$.*

Proof. Let $P_1 = \min(Pr(t<0), Pr(t>0))$. Assume that $x_1 > 0$; the other case is dual. In this case,

$$2|x_1| \cdot P_1 \leqq \int_{-\infty}^{0} 2|x_1|\,dF(t) \leqq \int_{-\infty}^{0} X(x,t)\,dF(t) < \int_{-\infty}^{\infty} X(x,t)\,dF(t) < 2m_0.$$

Thus $|x_1| < (m_0/P_1) = b_1$. In the same way, let $P_2 = \min(Pr(t < -b_1), Pr(t > b_1))$. Then

$$2|x_2| \cdot P_2 \leqq \int_{b_1}^{\infty} 2|x_2|\,dF(t) \leqq \int_{b_1}^{\infty} X(x,t)\,dF(t) < 2m_0,$$

so that $|x_2| < (m_0/P_2) = b_2$. In general, then, we choose $P_n = \min(Pr(t < -b_{n-1}), Pr(t > b_{n-1}))$ and $b_n = (m_0/P_n)$, and this sequence gives us the desired bounds. Q.E.D.

Even if the hypotheses of Lemma 2 are not satisfied, something can yet be salvaged, as we shall see later. What of the possibility that $x_i^{(n)} \to 0$ as $n \to \infty$? We shall modify the proof of Franck's Lemma 1 to show

3. Lemma. *If $F^-(0) < D < \infty$, then we can find a $K > 0$ such that for all sequences x with $x_2 < 0 < x_1$ and $x_3 - x_4 < K$, we can form the sequence y by removing x_1 and x_2 from x ($y_i = x_{i+2}$, $\forall i$), in which case $X(y) \leqq X(x)$.*

Proof. It is easy to see (or cf. [1] p. 3 formula (2)) that

$$\begin{aligned} X(x) - X(y) = {} & 2[\,|x_1|(1 - |F(x_1) - F(0)|) + |x_2|(1 - |F(x_2) - F(x_1)|) \\ & + |x_3|(1 - |F(x_3) - F(x_2)|) - |x_3|(1 - |F(x_3) - F(0)|) \\ = {} & 2[(x_1 - x_2)(1 - (F(x_1) - F(0))) - (x_3 - x_2)(F(0) - F(x_2))]. \end{aligned}$$

If K is small enough, this difference is positive. Indeed, let $K > 0$ be chosen so that

1° $F(K) - F(0) < \frac{1}{2}Pr(t > 0)$,

2° $(F(t) - \frac{F(0)}{t} < D, \ \forall -K < t < 0$,

3° $K < \frac{1}{2D}$.

Note that K is dependent only on F, and not on x. Assume $x_3 - x_4 < K$. In this case, we have. $x_1 \leqq x_1 - x_2 \leqq x_3 - x_4 < K$, so that $1 - (F(x_1) - F(0)) \geqq 1 - (F(K) - F(0)) > \frac{1}{2}$. Also, $|x_2| \leqq x_1 - x_2 \leqq x_3 - x_4 < K$, so that $F(0) - F(x_2) < D|x_2|$. Finally, $x_3 - x_2 \leqq x_3 - x_4 < K$, so that

$$X(x) - X(y) > 2[|x_2|\tfrac{1}{2} - K \cdot D|x_2|] > 2[\tfrac{1}{2}|x_2| - \tfrac{1}{2}|x_2|] = 0.$$

Q. E. D.

4. Lemma. *If* $F^{-}(0) < \infty$, $\varepsilon > 0$, *and* x *is any strong search plan, then we can find a search plan* y *such that* $X(y) < X(x) + \varepsilon$, $y_1 > 0$, *and* $y_3 - y_4 \geqq K$, *where* $K = K(F)$ *is defined in the proof of the previous lemma.*

Proof. Perhaps $x_1 \geqq 0$. If not, let $z = \{z_i\}$ be chosen with $0 < z_1 < x_2$ and $z_i = x_{i-1}$, $\forall i \geqq 2$. If z_1 is small enough, then $X(z) < X(x) + \varepsilon$. If $x_1 > 0$, let $z = x$. If $z_3 - z_4 \geqq K$, we are through. If not, then $z^{(1)} = \{z_i^{(1)}\}$ defined by $z_i^{(1)} = z_{i+2}$ has $X(z^{(1)}) \leqq X(z) < X(x) + \varepsilon$. Defining $z^{(n)}$ by $z_i^{(n)} = z_{i+2}^{(n-1)}$, we have $X(z^{(n)}) \leqq X(z^{(n-1)})$ whenever $z_3^{(n-1)} - z_4^{(n-1)} < K$. Since $z_3^{(n)} = z_{3+2n}$ must eventually exceed K (Recall that $F(K) - F(0) < \frac{1}{2}Pr(t > 0)$.), we must come to a first n for which $z_3^{(n)} - z_4^{(n)} \geqq K$. Then if we set $y = z^{(n)}$ we have $X(y) < X(x) + \varepsilon$.

Q. E. D.

5. Lemma. *If* $X(x) < 2m_0$, $x_1 > 0$, $x_3 - x_4 \geqq K$, and $x^{-} < a < 0 < b < x^{+}$, *then* $x_i \in [a, b]$ *for only* n_0 *values of* i *at most, where* n_0 *depends only on* F, a, b, *and* K.

Proof. Let $P = \min\{Pr(x^{-} < t < a),\ Pr(b < t < x^{+})\}$. Then if, for example, $a \leqq x_{2n} \leqq 0$, we see immediately that

$$2(n-1)KP \leqq 2\int_{-\infty}^{x_{2n}} (n-1)K\,dF(t) \leqq \int_{-\infty}^{x_{2n}} [2|x_3| + 2|x_4| + \cdots + 2|x_{2n-1}| + 2|x_{2n}|]\,dF(t)$$

$$< \int_{-\infty}^{x_{2n}} X(x,t)\,dF(t) < \int_{-\infty}^{\infty} X(x,t)\,dF(t) < 2m_0,$$

so that $n < (m_0/KP) + 1$. Similarly if $0 \leqq x_{2n+1} \leqq b$. Q. E.D.

6. Theorem. *Let* F *be our given distribution. If* $x^{-} = -\infty$, $x^{+} = +\infty$, and $F^{-}(0) < \infty$, *then there exists a search plan* y *with* $X(y) = m_0$.

Proof. The proof given is very similar to that of Franck. First note that for any weak search plan x, we can find a strong search plan z with $X(z) \leqq X(x)$. We do this by first finding the least value j of i with $x_j \neq 0$. If we define w by $w_i = x_{i-j+1}$, we see that $X(w) = X(x)$. Now let k be the greatest value of i with $w_k = 0$. Define $v_i = w_{i-k}$. Then $X(v) \leqq X(w)$. Finally, whenever $v_i = v_{i+2}$, eliminate v_i and v_{i+1} from v, thus giving us a new search plan z. It is clear that z is a strong search plan and $X(z) \leqq X(v) \leqq X(x)$.

Let $x^{(n)} = \{x_i^{(n)}\}$ be chosen for each n in such a way that $X(x^{(n)}) \to m_0$. Let K be chosen as before, and let a sequence $\{\varepsilon_n\}$ be chosen with $\varepsilon_n > 0$, $\varepsilon_n \to 0$ as $n \to \infty$. By Lemma 4 and the comments above, we can choose a strong search plan $z^{(n)}$ based on $x^{(n)}$ with $z_1^{(n)} > 0$, $\forall n$, $z_3^{(n)} - z_4^{(n)} \geqq K$, and $X(z^{(n)}) < X(x^{(n)}) + \varepsilon_n$. Then $X(z^{(n)}) \to m_0$, and for each i, $\{z_i^{(n)}\}$ is a bounded sequence. Using the diagonal method, we can extract a subsequence $\{z^{(n_j)}\}$of $\{z^{(n)}\}$ with $\{z_i^{(n_j)}\}$ convergent for each i as $j \to \infty$. For convenience, we assume $\{z^{(n)}\} = \{z^{(n_j)}\}$. Let $y_i = \lim_n z_i^{(n)}$, $\forall i$, and $y = \{y_i\}$. Then y is a search plan (possibly weak). We wish to show that $X(y) = m_0$. Let $w_i^{(n)}$be chosen so that it has the same sign as $x_i^{( )}$ and so that $|w_i^{(n)}| = \max\{|x_i^{(n)}|, |y_i|\}$. Then $w_i^{(n)} \to y_i$ as $n \to \infty$, $\forall_i$ and $|w_i^{(n)}| \geqq |y_i|$.

Choose any $k > 0$, $\delta > 0$, and let n_0 be chosen so that $|x_i^{(n)} - y_i| < \delta$, $\forall n > n_0$, $\forall i = 1, \cdots, 2k$. Then for every $y_{2k} < t < y_{2k+1}$, we have

$$X(w^{(n)}, t) \leqq X(x^{(n)}, t) + 2k \cdot 2\delta.$$

Thus

$$\int_{y_{2k}}^{y_{2k-1}} X(w^{(n)}, t)dF(t) \leqq \int_{y_{2k}}^{y_{2k-1}} X(x^{(n)}, t)dF(t) + (y_{2k-1} - y_{2k})2k \cdot 2\delta.$$

On the other hand, as $n \to \infty$, we have

$$X(w^{(n)}, t) \to X(y, t) \text{ for every } t \in R.$$

Thus, for n large enough, say $n > n_1$, we have

$$\left| \int_{y_{2k}}^{y_{2k-1}} X(y, t)dF(t) - \int_{y_{2k}}^{y_{2k-1}} X(w^{(n)}, t)dF(t) \right| < \delta$$

Thus, for $n > \max\{n_0, n_1\}$, we have

$$\begin{aligned}\int_{y_{2k}}^{y_{2k-1}} X(y, t)dF(t) &< \int_{y_{2k}}^{y_{2k-1}} X(w^{(n)}, t)dF(t) + \delta \\ &\leqq \int_{y_{2k}}^{y_{2k-1}} X(x^{(n)}, t)dF(t) + (y_{2k-1} - y_{2k}) \cdot 2k \cdot 2\delta + \delta \\ &\leqq \int_{-\infty}^{+\infty} X(x^{(n)}, t)dF(t) + (y_{2k-1} - y_{2k})2k \cdot 2\delta + \delta\end{aligned}$$

Since $X(x^{(n)}) \to m_0$, we have for each $\delta > 0$

$$\int_{y_{2k}}^{y_{2k-1}} X(y,t)dF(t) \leqq m_0 + (y_{2k-1} - y_{2k})2k \cdot 2\delta + \delta,$$

so that

$$\int_{y_{2k}}^{y_{2k-1}} X(y,t)dF(t) \leqq m_0, \quad \forall k > 0,$$

and thus

$$X(y) = \int_{-\infty}^{+\infty} X(y,t)dF(t) \leqq m_0.$$

Since y is a search plan (weak or strong), $X(y) \geqq m_0$, so that $X(y) = m_0$. Furthermore, we could make y a strong search plan (by removing no more than three zeroes at the beginning) and it would then still have this property.

Q.E.D.

7. Corollary. *If, in the above theorem, the hypothesis $F^-(0) < \infty$ is replaced by $F^+(0) < \infty$, then the same conclusion follows.*

Proof. Clear by symmetry.

We deal now with the thornier case in which $x^- > -\infty$ or $x^+ < +\infty$. Let us consider first the case $x^- > -\infty$, $x^+ = +\infty$. As we noted before, the case $x^- \geqq 0$ is trivial.

8. Theorem. *Let F be our given distribution. If $-\infty < x^- < 0$, $x^+ = +\infty$, and $F^-(0) < \infty$, then there is a y with $X(y) = m_0$.*

Proof. Let $x^{(n)}$ be a sequence of search plans with $X(x^{(n)}) \to m_0$ as $n \to \infty$. As before, choose $z^{(n)}$ so that each $z^{(n)}$ is a strong search plan with $z_1^{(n)} > 0$, $z_3^{(n)} - z_4^{(n)} > K$, and so that $X(z^{(n)}) \to m_0$ as $n \to \infty$. If $\{z_i^{(n)}\}$ is bounded for each i as $n \to \infty$, then the proof given for Theorem 6 will apply here. In the other case, let k be the least value of i for which $\{z_i^{(n)}\}$ is unbounded as $n \to \infty$. Choose a subsequence $\{z^{(n_j)}\}$ of $\{z^{(n)}\}$ such that $\{z_i^{(n_j)}\}$ converges for each $1 \leqq i < k$, while $\{z_k^{(n_j)}\} \to +\infty$, and such that $X(z^{(n_j)}) < 2m_0$, $\forall j$. We shall save notational difficulties by assuming that $\{z^{(n)}\}$ itself is this subsequence. We first show that $z_{k-1}^{)} \to x^-$. Let $u_n = z_k^{(n)}$ and $v_n = z_{k-1}^{(n)}$. Then we have

$$2|u_n|(Pr(t < v_n)) = \int_{x^-}^{v_n} 2|u_n|\,dF(t) \leqq \int_{x^-}^{v_n} X(z^{(n)},t)dF(t)$$

$$\leqq \int_{-\infty}^{\infty} X(z^{(n)},t)dF(t) < 2m_0,$$

so that

$$Pr(t < v_n) < \frac{m_0}{u_n} \to 0 \text{ as } n \to \infty, \text{ since } u_n = z_k^{(n)} \to \infty.$$

Thus $z_{k-1}^{(n)} = v_n \to x^-$, as asserted.

Letting $y_i = \lim_{n\to\infty} z_i^{(n)}$, $\forall i = 1, \cdots, k-1$, we have as before

$$\int_{x^-}^{u_n} X(y,t)dF(t) \leqq m_0, \ \forall n \text{ and thus}$$

$y = (y_1, \cdots, y_{k-1}, +\infty)$ is a search procedure with $X(y) = m_0$. Q.E.D.

9. Corollary. *If, in the above theorem, the hypothesis $\bar{F}^-(0) < \infty$ is replaced by $\bar{F}^+(0) < \infty$, then the same conclusion follows.*

Proof. The proof is identical except that $z_1^{(n)} < 0$ in each $z^{(n)}$, which is an inessential difference.

10. Corollary. *If in Theorem 8 or Corollary 9, the hypothesis $-\infty < x^- < 0$, $x^+ = +\infty$ is replaced by $x^- = -\infty, 0 < x^+ < +\infty$, then the conclusions still hold.*

Proof. Clear by symmetry.

11. Theorem. *Let F be our given distribution. If $-\infty < x^- < 0 < x^+ < +\infty$, and $\bar{F}^-(0) < \infty$, then we can find a search plan y with $X(y) = m_0$.*

Proof. Define $x^{(n)}$ and $z^{(n)}$ as before. If the sequence $z_i^{(n)}$ is bounded away from x^- and x^+ for every i as $n \to \infty$, then the proof is the same as in Theorem 6. Otherwise, let k be the smallest value of i for which $\{z_i^{(n)}\}$ violates this condition. Assume $\{z_k^{(n)}\}$ has x^+ for a limit point; the other case is dual. As previously, we can assume each $\{z_i^{(n)}\}$ converges, $\forall 1 \leqq i < k$, to a point of (x^-, x^+), and $z_k^{(n)} \to x^+$. Then let $y_i = \lim_{n\to\infty} z_i^{(n)}$, $\forall 1 \leqq i < k$, and set $y = (y_1, \cdots, y_{k-1}, x^+, x^-)$. Then as before, $X(y) = m_0$. Q.E.D.

12. Corollary. *If, in Theorem 11, the assumption $\bar{F}^-(0) < \infty$ is replaced by $\bar{F}^+(0) < \infty$, then the same conclusion follows.*

Proof. Clear by symmetry.

13. Theorem. *Let F be a probability distribution with $\int_{-\infty}^{\infty} |t| dF(t) < \infty$. Suppose that $\bar{F}^-(0) = \bar{F}^+(0) = \infty$. Let x be any search procedure. Then there is a search procedure y such that $X(y) < X(x)$.*

Proof. Since $\bar{F}^-(0) = \bar{F}^+(0) = \infty \neq 0$, $x^- < 0 < x^+$. Thus, any search procedure has at least two entries. Assume $x_1 > 0$; the other case is dual. Choose any y_1 with $x_2 < y_1 < 0$ and $y_1^{-1}(F(y_1) - F(0)) > (1/x_1)$, so that $F(y_1) - F(0) < (y_1/x_1)$, and define $y_i = x_{i-1}$, $\forall i \geqq 2$. Then

$$\begin{aligned} X(y) - X(x) &= 2[y_1(F(0) - F(y_1) - 1) + x_1(F(y_1) - F(0)] \\ &< 2\left[y_1(F(0) - F(y_1) - 1) + x_1 \cdot \frac{y_1}{x_1} \right] \\ &= 2y_1(F(0) - F(y_1)) < 0, \end{aligned}$$

so that $X(y) < X(x)$. Q.E.D.

Proof of Theorem. 1. Direct consequence of Theorems 6, 8, 11, and 13 and Corollaries 7, 9, 10 and 12.

Q.E.D.

In fact, Theorem 12 shows a little more than promised. It shows that if $F^-(0) = \infty$ (resp. $F^+(0) = \infty$), then there is no minimal search procedure with $x_1 > 0$ (resp. $x_1 < 0$). Thus, if $F^-(0) = \infty$, $F^+(0) < \infty$ (resp. the reverse), then we see that there is a minimal search procedure, x, and $x_1 < 0$ (resp. $x_1 > 0$). Thus, in this limited case, we have an indication of the direction of the first entry.

Reference

1. Franck Wallace, *On the optimal search problem*, Technical Report No. 44, University of New Mexico, October, 1963.

The Hebrew University of Jerusalem,
and
The University of Wisconsin