CORRECTING FOR RUNNING MEANS BY SUCCESSIVE SUBSTITUTIONS

By R. N. BRACEWELL*

[*Manuscript received April 22, 1955*]

Summary

The paper discusses a practical procedure which compensates for the effect of taking running means, and a numerical example is worked out. The procedure simply requires the evaluation of further running means and should prove readily applicable in many cases where the need arises. In this procedure, which has already gained importance in more general form in other fields (as the method of successive substitutions), a criterion of convergence has been given ; and it is a major aim of the paper to illustrate, using running means as an instance, that the utility of the method of successive substitutions is wider than is indicated by this criterion. The mathematical theory shows that the proposed procedure leads to a divergent result in the case of running means ; nevertheless the asymptotic nature of the divergence allows results of practical value. An illuminating view of the phenomenon is given from the standpoint of Fourier analysis, which reveals a counterplay of simultaneous deterioration and improvement occurring in different spectral regions.

I. INTRODUCTION

It often happens, when a quantity $f(x)$ has to be observed at $x=x_1$, that the quantity actually measured is a weighted integral of $f(x)$ over an interval surrounding x_1. This interval can sometimes be made so small that further decrease makes no difference, but often this cannot be done, and the problem then arises of allowing for the non-zero width of the interval.

The nature of the weighting function varies with the circumstances. In spectroscopy it is the " apparatus profile ", a more or less smooth humped curve ; in radio astronomy it is the " aerial directional diagram ", a central hump with " side lobes " ; and there are numberless other cases. In many circumstances the weighting function is rectangular—this is the important case where the observation is a simple running mean of the desired quantity. An example of this occurs when a photographic density distribution $f(x)$ is scanned by a slit.

This paper deals with a method of correcting for running means. The method, which in its general form has been presented at length in connexion with the radio-astronomical problem by Bracewell and Roberts (1954), will be referred to as the method of successive substitutions as in the theory of integral equations (e.g. Lovitt 1950). The method seems to have been introduced into

* Division of Radiophysics, C.S.I.R.O., University Grounds, Sydney, and Berkeley Astronomical Department, University of California, Berkeley.

astronomy by van Cittert (1931), but is not well known in some branches of physics where it has immediate application (e.g. Bracewell 1955). It yields a sequence of distributions, each derived from its predecessor by a process involving a further application of the very effect which it is desired to eliminate. In the case of running means the procedure, briefly stated, would be as follows. Take running means of the observations, note the small change produced, and apply this change as a negative correction to the observations. This gives the second member of a sequence of distributions of which the observed distribution itself is the first member. The next distribution is obtained by taking running means of the second, comparing the result so obtained with the observations, and applying the discrepancy as a correction to the second distribution. If the discrepancy tends, in successive stages, towards zero, then the distributions are tending to a form which, when smoothed by running means, agrees more and more closely with the original observations. In other words, the sequence of distributions is converging to a desired solution.

This procedure involves nothing more recondite than repeated running means of the observed data, and, as running means are often easy to get, such a procedure would be very attractive in many applications.

A note of caution now needs to be sounded, for only in favourable cases does the sequence converge. For evidence of this, one need only ponder the case $f(x)=\delta(x)$ and weighting function $\frac{1}{2}\delta(x+1)+\frac{1}{2}\delta(x-1)$, where $\delta(x)$ is the unit impulse function.

Now the criterion of convergence is known and, when applied to running means, reveals that the sequence converges only for initial distributions with a certain peculiar type of spectrum. An understanding of the convergence question is therefore all important in this method of correcting for running means. On the other hand, in radio astronomy, convergence can normally be taken for granted.

The method of successive substitutions is, however, likely to prove of far more general use than its failure to satisfy convergence criteria might lead one to suppose. As explained later in connexion with Figure 1, it may happen that the sequence of distributions, whilst ultimately divergent, nevertheless for a few stages approaches the desired solution asymptotically. As it is probable that, in the bulk of applications, a procedure would not be considered practicable which did not give adequate correction in the first one or two stages, it should be immaterial in practice whether the sequence is, in the end, convergent or divergent. This paper is mainly concerned with showing that excellent results may often be had in spite of divergence, using the case of running means for the purpose.

II. Definition of the Problem

Let $g(x)$ be derived from a function $f(x)$ by averaging over an interval ξ. Then

$$g(x)=\xi^{-1}\int_{x-\frac{1}{2}\xi}^{x+\frac{1}{2}\xi} f(\tau)\mathrm{d}\tau. \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

The problem is, given $g(x)$, to find $f(x)$.

Rewrite equation (**1**) in the form

$$g(x)=\xi^{-1}\int_{-\infty}^{\infty}\Pi\left(\frac{x-\tau}{\xi}\right)f(\tau)\mathrm{d}\tau, \quad \ldots\ldots\ldots\ldots \quad (\mathbf{2})$$

where $\Pi(\tau)$ is the rectangle function of unit height and breadth :

$$\begin{aligned}\Pi(\tau)&=1, \quad -\tfrac{1}{2}<\tau<\tfrac{1}{2},\\ &=0, \quad |\,\tau\,|>\tfrac{1}{2}.\end{aligned}$$

The right-hand side of equation (**2**) may be recognized as a convolution integral, which, for clarity, may be written with the asterisk notation† as

$$g(x)=\Pi_{\xi}*f, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{3})$$

where $\Pi_{\xi}(x)=\xi^{-1}\Pi(\xi^{-1}x)$.

Equation (**3**) is strictly analogous to the aerial smoothing equation discussed by Bracewell and Roberts, and it will be seen that Π_{ξ} is already correctly normalized in the sense that

$$\int_{-\infty}^{\infty}\Pi_{\xi}(x)\mathrm{d}x=1.$$

The sequence of distributions yielded by the method of successive substitutions is therefore as follows :

$$\left.\begin{aligned} &g,\\ f_1&=2g-\Pi_{\xi}*g,\\ f_2&=3g-3\Pi_{\xi}*g+\Pi_{\xi}*\Pi_{\xi}*g,\\ &\ldots\ldots\\ &\ldots\ldots\end{aligned}\right\} \quad \ldots\ldots\ldots\ldots \quad (\mathbf{3a})$$

It has now to be considered whether the members of this sequence approach the desired distribution f, or, if not, whether any of them is an improvement on g.

III. Condition for Convergence

According to Bracewell and Roberts, the sequence converges if, and only if,

$$|\,1-\overline{\Pi}_{\xi}(s)\,|<1$$

for all s for which $\bar{g}(s)\neq 0$, where $\overline{\Pi}_{\xi}(s)$ and $\bar{g}(s)$ are respectively the Fourier transforms of Π_{ξ} and g. In general the sequence converges not to f, but to the function whose transform is

$$\frac{\bar{f}(s)}{1+\delta(\overline{\Pi}_{\xi})}.$$

Now

$$\overline{\Pi}_{\xi}(s)=\frac{\sin \pi\xi s}{\pi\xi s}. \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\mathbf{4})$$

† In this notation $f*g\equiv\int_{-\infty}^{\infty}f(x-u)g(u)\mathrm{d}u$.

Hence $|1-\bar{\Pi}_\xi(s)|\geqslant 1$ for all s such that

$$n\xi^{-1}\leqslant s\leqslant(n+1)\xi^{-1},$$

where n is any odd integer.
Consequently the sequence converges only in the cases where $g(x)$ has no spectral components in the shaded zones of Figure 1.

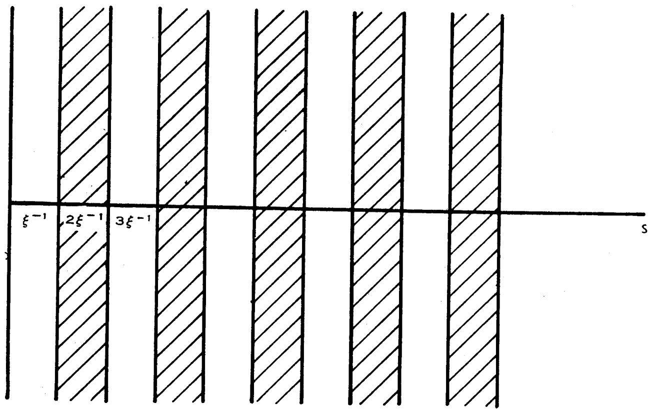

Fig. 1

Few distribution functions in practice could be expected to comply with such a severe condition. Therefore, in general, the sequence could not be relied on to converge. However, there remains the possibility of asymptotic representation and this may be illustrated with a numerical example.

We adopt the symmetrical distribution $g(x)$, as given in Table 1 and graphed in Figure 2, constructed from the assumed distribution $f(x)$ by taking running

Table 1

AN EXAMPLE OF ONE STAGE OF CORRECTION

x	$f(x)$	$g(x)$	$\Pi\xi * g$	f_1	Discrepancy
0	8800	8400	8032	8768	—32 (½%)
	8600	8200	7880	8520	—80 (1%)
2	8000	7680	7440	7920	—80 (1%)
	7000	6920	6760	7080	80 (1%)
4	6000	6000	5920	6080	80 (1%)
	5000	5000	4992	5008	8 (2%)
6	4000	4000	4040	3960	—40 (1%)
	3000	3040	3120	2960	—40 (1%)
8	2000	2160	2280	2040	40 (2%)
	1200	1400	1560	1240	40 (3%)
10	600	800	984	616	16 (3%)
	200	400	560	240	40 (20%)
12	0	160	280	40	40
	0	40	120	—40	—40
14	0	0	40	—40	—40
	0	0	8	— 8	— 8

means over five values. The column $\Pi_\xi * g$ is derived from $g(x)$ in the same way that $g(x)$ is derived from $f(x)$. The column $f_1 = 2g - \Pi_\xi * g$ gives the result of one stage of correction, and the last column gives the discrepancy between the corrected distribution $f_1(x)$ and the known original distribution $f(x)$.

Figure 2 shows the agreement between the corrected distribution (-----) and the known original function $f(x)$, an agreement which for many practical purposes would clearly be very acceptable.

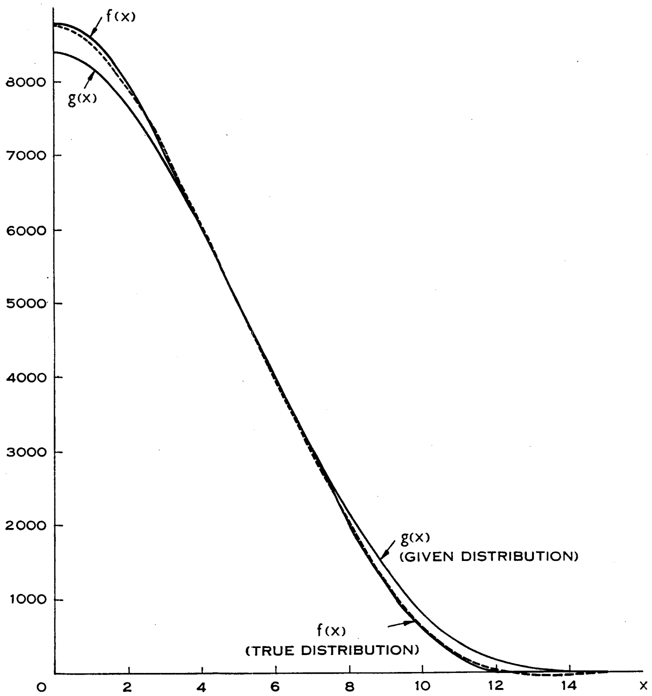

Fig. 2

IV. Spectral Aspect of Asymptotic Phenomena

Applying the convolution theorem to (**3**) we have

$$\bar{g} = \bar{\Pi}_\xi \bar{f}.$$

Hence, from (**4**),

$$\bar{g} = \frac{\sin \pi\xi s}{\pi\xi s} \bar{f}.$$

This relationship between the spectra $\bar{f}$ and $\bar{g}$ is illustrated qualitatively in Figure 3. Now consideration of equation (**3a**) will show that the transforms of the successive distributions fall as in the broken line of Figure 3. In the non-shaded strips, the broken line is closer to $\bar{f}$ than $\bar{g}$ is, but in the shaded strips it is not, and tends to depart further with succeeding stages.

If then a stage of restoration is to improve the agreement with f, the improvement occurring in the unshaded strips must outweigh the deterioration setting in in the shaded strips. This will occur when the spectrum of f is weak in the shaded strips, as may be shown to be the case in the example of Figure 2.

V. Criteria for Practical Use

We have now seen that satisfactory results can follow from asymptotic sequences, and an illuminating explanation of what happens has been possible. In practice, however, the results cannot be verified in the way used here, nor can Fourier transforms be considered in detail. A criterion is therefore required by which it can be decided whether the corrected distribution $f_1(x)$ represents an improvement over $g(x)$ as an approximation to the solution $f(x)$.

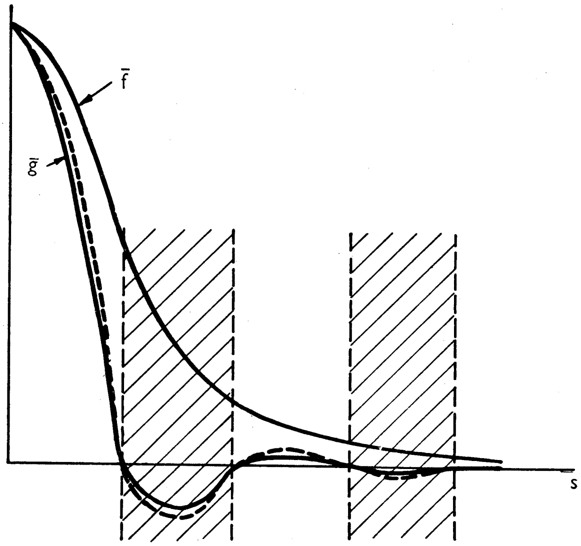

Fig. 3

Fortunately the method itself gives an indication, for, in cases of convergence, successive corrections become smaller and smaller. Now it is not reasonable to attempt to assess the quality of a distribution numerically. For example, a corrected distribution might show improvement in some respects (say general shape) and deterioration in others (presence of ripple) ; and in some circumstances this might be on the whole " better " and in other circumstances " worse ". The required criterion must therefore be a qualitative one involving the user's needs ; thus as a practical test, it would seem to be sufficient that $\Pi\xi{*}f_1$ should agree with g much " better " (for the existing purpose) than g does with f_1.

VI. References

Bracewell, R. N. (1955).—A method of correcting the broadening of X-ray line profiles. *Aust. J. Phys.* **8** : 61.

Bracewell, R. N., and Roberts, J. A. (1954).—Aerial smoothing in radio astronomy. *Aust. J. Phys.* **7** : 615.

van Cittert, P. H. (1931).—Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. *Z. Phys.* **69** : 298.

Lovitt, W. V. (1950).—" Linear Integral Equations." (Dover : New York.)