MEASURING THE ASSOCIATION OF POINT PROCESSES: A CASE HISTORY

DAVID R. BRILLINGER

1. Introduction. Modern applied statistics typically involves elements of computation, probability theory, statistical theory and collaboration with specialists in the subject matter of some substantive field. In this article I shall describe part of a continuing experience of collaboration with two neurophysiologists from U.C.L.A., H. L. Bryant Jr. and J. P. Segundo. In formal terms, the problem considered is one of measuring the degree of association of points of two different sorts distributed along a straight line in an irregular manner. In real terms, the problem is one of investigating the behavior of a simple nerve cell network in a sea slug (*Aplysia californica*). The paper discusses a summary measure of association that has proved useful in assessing whether two nerve cells are behaving in a related manner or are behaving independently. The experiments by means of which the data were collected are described in Bryant, Ruiz Marcos and Segundo [4], as are the results of preliminary statistical analyses. The paper [4] is representative of the extent to which quantification is now occuring in the life sciences.
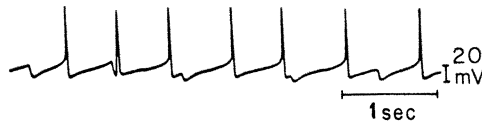

FIG. 1. A typical record of the changing voltage level of a nerve cell.

2. Some neurophysiology. The nerve cell (or neuron) is the basic unit of the animal involved in the transmission of information. Described schematically, it consists of a central cell body (or soma), branches (called dendrites) carrying impulses to the body and a long outgrowth (the axon) conducting impulses from the body. One way information is transmitted through the dendrites and axon is through changes in electrical activity. Figure 1 is an example of the changing voltage recorded when a microelectrode is inserted into a nerve cell. The record is seen to be made up of pulses of large amplitude compared to their duration. Because of its appearance, such a record is often called a spike train.
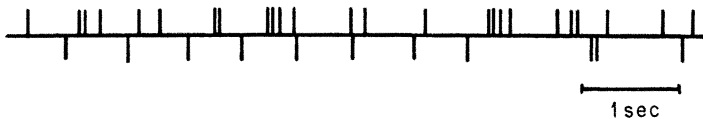

FIG. 2. A record of the times of spikes of two simultaneously firing nerve cells.

The junction whereby one neuron may influence another is called the synapse. When a pulse reaches the terminal point of an axon it provokes the release of a transmitter substance which alters the permeability of the dendrite of the next cell to certain ions. The resulting flow of ions generates a small electric current which moves down the dendrite to the soma. If the junction is excitory, the spike activity of the second cell is increased, if inhibitory it is decreased. Figure 2 is an example of the times of spikes for two nearby cells, the times for one cell corresponding to spikes above the line and for the other corresponding to spikes below. In practise, given two neurons, it may not be known whether either is influencing the other and it may be of interest to determine if there is some influence or association. Doing this by eye from records such as those of Figure 2 can be very difficult. Researchers have therefore been led to compute summary values from the records (see Griffith and Horn [9] for example) and this is the concern of the present paper.

The data discussed is recorded simultaneously on cells L3 and L10 of the sea hare. This

16

particular animal and these particular cells were used because the cells may be identified in different specimens and consequently experiments may be repeated. The experimental methods are described in detail in [4]. Further information concerning neurons and synapses may be found in Eccles [7].

3. Some probability theory. Commonly in his work on applied problems, a statistician brings the apparatus of probability theory into use. This involves his asking the experimentalists and himself whether or not it is reasonable to talk about random outcomes and probabilities of events connected with outcomes. The statistician seeks to bring probability theory into a problem because it provides a precise means of defining parameters and models and it allows him to interpret and assess various manipulations of experimental data. Not all problems of data analysis require the introduction of probability theory, but many seem to benefit from its appearance — among the latter are problems concerning nerve cell spike trains.

The branch of probability theory concerned with entities like irregular spike trains is that of stochastic point processes. A stochastic point process is a random, non-negative, integer-valued measure. If I is an interval of the real line and ω is a random element, then the values of this measure may be denoted by $N(I, \omega)$, with $N(I, \omega)$ denoting the number of points in the interval I for the realization corresponding to ω. Here the atoms of the measure $N(I, \omega)$ correspond to the times of spikes of a particular spike train. Repeating the experiment would most likely yield a different set of spike times and consequently a different measure $N(I, \omega')$. In this sense N is a random measure. (We remark that in many problems one can suppress the dependence of N on ω, however, it is an essential element of the approach.) Point processes were considered recently in the MONTHLY by Chung [5] and are discussed in Cox and Lewis [6] and in a volume [11] edited by Lewis, for example.

For nerve cell trains, it is appropriate to assume that the point process is without multiple points; that is, the spike times are isolated, separated by positive distances. Because the spikes proceed from no inherent origin, it also seems appropriate to assume that the point process is stationary in time in the sense that the probability distribution of the random vector

$$\{N(I_1, \omega), \cdots, N(I_k, \omega)\}$$

is the same as that of the shifted vector

$$\{N(I_1 + t, \omega), \cdots, N(I_k + t, \omega)\}$$

for all t and $k = 1, 2, \cdots$, where $I + t$ denotes the interval $(a + t, b + t)$ if $I = (a, b)$.

Important parameters of a stationary point process N include the **mean intensity,** p_N, and the **second-order product density,** $p_{NN}(u)$, given by

(1) $$p_N = \lim_{h \downarrow 0} \text{Prob}\{\text{point in the interval } (t, t + h)\}/h$$

and

(2) $$p_{NN}(u) = \lim_{h, h' \downarrow 0} \text{Prob}\{\text{point in } (t + u - h, t + u + h) \text{ and point in } (t - h', t + h')\}/(4hh')$$

$-\infty < u < \infty$, respectively when these limits exist.

In fact we shall be concerned with two different types of points, say M points and N points, with $M(I, \omega)$ referring to the number of M points in the interval I and $N(I, \omega)$ the number of N points in the interval I. We denote the mean intensity of M points by p_M and the second-order product density of M points by $p_{MM}(u)$. We also define a **cross-product density,** $p_{MN}(u)$, by

(3) $$p_{MN}(u) = \lim_{h, h' \downarrow 0} \text{Prob}\{M \text{ point in } (t + u - h, t + u + h) \text{ and } N \text{ point in } (t - h', t + h')\}/(4hh').$$

The parameters in (1), (2), (3) do not depend on t because the process is stationary.

The first thing that one tends to notice when examining a spike train is whether there are a lot of spikes or only a few. The mean intensity of the process gives information in this connection. Expression (1) implies that the probability of there being an N point in a small interval of length h is approximately $p_N h$. The next thing that one tends to notice is the relative positioning of pairs of spikes of a single train or from one train to another. Expressions (2) and (3) give information in this connection. From expression (3), for example, we have

(4) $\text{Prob}\{M \text{ point in } (t + u - h, t + u + h) \text{ and } N \text{ point in } (t - h', t + h')\}$

$$\sim p_{MN}(u)4hh'$$

for h, h' non-negative and small. Using the definition of conditional probability and (1), this implies that

(5) $\text{Prob}\{M \text{ point in } (t + u - h, t + u + h) \text{ given an } N \text{ point at } t\}$

$$\sim 2h\, p_{MN}(u)/p_N.$$

In the case that the M points are distributed independently of the N points, the probability referred to in expression (5) is just $\text{Prob}\{M \text{ point in } (t + u - h,\ t + u + h)\}$ and so

(6) $$p_{MN}(u)/p_N = p_M \quad \text{or} \quad p_{MN}(u) = p_M p_N$$

for all u. This last suggests that the function $p_{MN}(u)$, and related functions such as

(7) $$\frac{p_{MN}(u)}{p_M p_N} \quad \text{or} \quad \sqrt{\frac{p_{MN}(u)}{p_M p_N}}$$

might prove useful measures of the degree of association of points of the M process with points of the N process. They are identically 1.00 in the case of independence.

We remark that, since we have assumed the points of the processes to be isolated, we can replace the probabilities of expressions (1)–(3) by expected values, for example we could write for (3)

$$p_{MN}(u) = \lim_{h,h' \downarrow 0} E\{N(t + u - h, t + u + h)\, N(t - h', t + h')\}/(4hh').$$

4. Some statistical theory. The preceding section described a mathematical idealization that could be of use in examining the degree of relationship of two given spike trains. The idealization suggested the definition of parameters p_M, p_N, $p_{MN}(u)$ based on the probabilities of certain events. In order to make concrete use of these parameters we need to have some idea of their values for the spike trains at hand.

Statistical theory has long been concerned with the problem of estimating the probability of an event given experimental results. In elementary situations one estimates the probability of an event A by n_A/n, where n_A denotes the number of times the event A occurred out of n times when it might have occurred. Let us use this approach to construct estimates of $p_M, p_N, p_{MN}(u)$.

Suppose that spike trains M and N are observed throughout the time interval $(0, T)$. Let $s_1 < s_2 < \cdots < s_{M(T)}$ be the observed times of M spikes and $t_1 < t_2 < \cdots < t_{N(T)}$ be the observed times of N spikes where we have observed $M(T)$ M spikes and $N(T)$ N spikes in all. Let h be small and imagine the interval $(0, T)$ divided into T/h intervals of length h. The number of times the event "M spike in small interval of length h" occurred is $M(T)$. It might have occurred T/h times. This suggests estimating $p_M h$ by $M(T)/(T/h)$ and so estimating p_M by

(8) $$\hat{p}_M = M(T)/T.$$

Likewise we could estimate p_N by $\hat{p}_N = N(T)/T$.

Next we consider the estimation of $p_{MN}(u)$. For small h, let

(9) $n(u, h) =$ the number of s_i such that $t_j + u - h < s_i < t_j + u + h$ for some j,

then the probability of expression (5) may be estimated by $n(u, h)/N(T)$. This suggests the estimation of $p_{MN}(u)$ by

$$(10) \qquad \hat{p}_{MN}(u) = \frac{n(u, h)\hat{p}_N}{N(T)2h} = \frac{n(u, h)}{2hT} .$$

This estimate may be used in turn to construct estimates of the functions of expression (7) if so desired. We must not forget, however, that we have not measured $p_{MN}(u)$ exactly. Rather we have constructed an expression that should be near it, especially in the case that T is large. We must also remember that were we to repeat the experiment, almost certainly, we would obtain a different value for $\hat{p}_{MN}(u)$. This last variation is called sampling fluctuation. Figure 3 is a graph of $\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)$ for the nerve cell data described in Section 2. Here $N(T) = 1232$, $M(T) = 816$, $\hat{p}_N = 1.77$ events/sec. and $\hat{p}_M = 1.17$ events/sec.. The Figure suggests that the probability of occurrence of an M event is depressed for about 5 seconds after the occurrence of an N event. However, before we can come to a belief that the outputs of the two nerve cells are in fact related, we must first have some confidence that the deviation of the function from the value 1.00 is not due simply to sampling fluctuations.
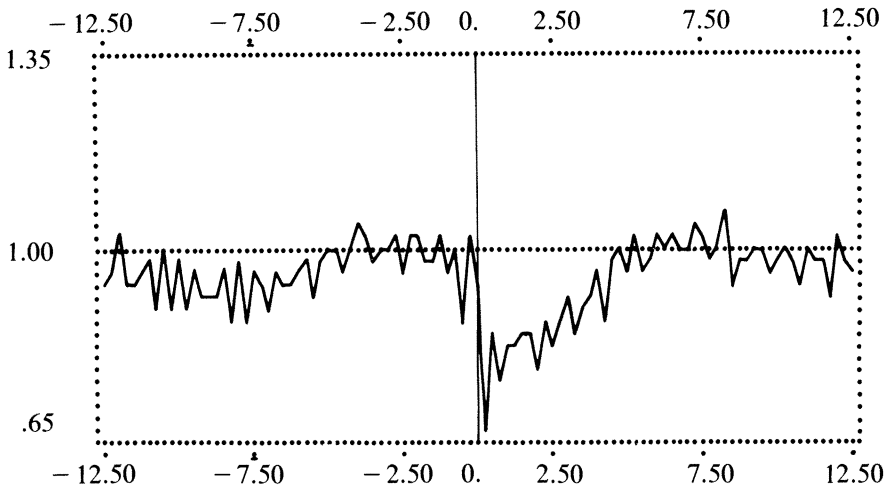

FIG. 3. An example of $\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)$ for cells L3 and L10 of *Aplysia californica*. The horizontal axis gives u in seconds.

The estimate (10) was proposed in Griffith and Horn [9]. Its direct computation involves the comparison of $M(T) \cdot N(T)$ values. In many of the experiments referred to $M(T)$ and $N(T)$ are both about 1000, so fairly clearly a high speed computer must be used in its computation. Hugh Bryant, Jr. has noted that if the spike times are recorded by increasing time, with a simple indicator to say whether a spike was an M or an N, then there exists a direct algorithm for computing (9) with one pass through the data. Suppose the data is denoted (u_j, α_j), $j = 1, 2, \cdots$ where $u_1 < u_2 < u_3 \cdots$ and u_j is an s_k if $\alpha_j = 0$, u_j is a t_k if $\alpha_j = 1$. The algorithm is the following: (1) initialize $n(lh, h)$ to 0 for $l = 0, \pm 1, \pm 3, \cdots$, (2) for $j = 1, 2, \cdots$ and $k = j + 1, j + 2, \cdots$ if $\alpha_j = 1$, $\alpha_k = 0$ compute $l = [(u_k - u_j)/(2h)]$ and set $n(lh, h) = n(lh, h) + 1$ or if $\alpha_j = 0$, $\alpha_k = 1$ compute $l = -[(u_k - u_j)/(2h) + 1/2]$ and set $n(lh, h) = n(lh, h) + 1$. (Here $[x]$ means the integral part of the number x.)

5. Some more probability theory. Figure 3 is a graph of an estimate of $p_{MN}(u)/(p_M p_N)$ rather than the function itself. Part of the irregular nature of the figure is undoubtedly due to sampling fluctuations. Before we can come to a reasonable decision that the two spike trains are related, with

a spike of the N train associated with an apparent depression in the rate of M spikes for example, we must assess the magnitude of sampling fluctuations. The key random variate appearing in $\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)$ is $n(u, h)$ given by expression (9). Let us attempt to approximate the distribution of this variate for large T.

$n(u, h)$ is a counting variate. As h is small it is counting rare events. Now in many situations, counts of rare events are approximately Poisson (see for example Feller [8] p. 282). Volkonski and Rozanov [12] demonstrate the related result that if $N^T(I, \omega)$, $T = 1, 2, \cdots$ is a sequence of point processes with mean intensities $p_N^T \to 0$ as $T \to \infty$, then under a further regularity condition, the sequence of processes with rescaled time, $N^T(I/p_N^T, \omega)$, $T = 1, 2, \cdots$ tends to a Poisson process. Perhaps $n(u, h)$ here is approximately Poisson with mean $2hTp_{MN}(u)$.

Supposing $h = L/T$, with L constant, we define a sequence of processes $N^T(I, \omega)$ by saying that $N^T(\cdot, \omega)$ has a spike at t, if $N(\cdot, \omega)$ has a spike at t and if $M(\cdot, \omega)$ has a spike in the interval $(t + u - L/T, t + u + L/T)$. The mean intensity of the process $N^T(\cdot, \omega)$ is $\sim p_{MN}(u) \, 2L/T \to 0$. In Brillinger [2] it is shown that the result of Volkonski and Rozanov [12] may be applied to conclude that for large T the process $N^T(IT, \omega)$ is approximately Poisson and in particular $n(u, T) \sim N^T((0, T), \omega)$ is approximately Poisson with mean $2hTp_{MN}(u)$.

6. Some statistical inference. A common procedure that a statistician employs to communicate an interval of plausible values for an unknown parameter, in the light of data collected, is a confidence interval. For example, a 95 per cent confidence interval has the formal interpretation that 95 is the long run percentage of such intervals that actually contained the true parameter value. At this point we could use Table 40 of Biometrika Tables, based on the Poisson distribution, to construct a confidence interval for $p_{MN}(u)$ and by division through by $\hat{p}_M\hat{p}_N$ a confidence interval for $p_{MN}(u)/(p_Mp_N)$.

A less troublesome way in which to proceed is to take advantage of the fact that if P is a Poisson variate with mean μ, then $\sqrt{P}$ is approximately a normal variate with mean $\sqrt{\mu}$ and standard deviation $1/2$ (see pp. 88–96 in Kendall and Stuart [10]). The application of this square root transformation has two advantages; tables of the normal distribution are widely available for constructing confidence intervals, and the approximate standard deviation does not depend on the unknown parameter. All of this suggests a consideration of the estimate $\sqrt{\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)}$ and of the approximation of its distribution by a normal with mean $\sqrt{p_{MN}(u)/(p_Mp_N)}$ and standard deviation $1/(2\sqrt{2hTp_Mp_N})$. Approximate 95 per cent confidence intervals are constructed by adding and subtracting $1.96/(2\sqrt{2h\hat{p}_M\hat{p}_N})$ along the curve.
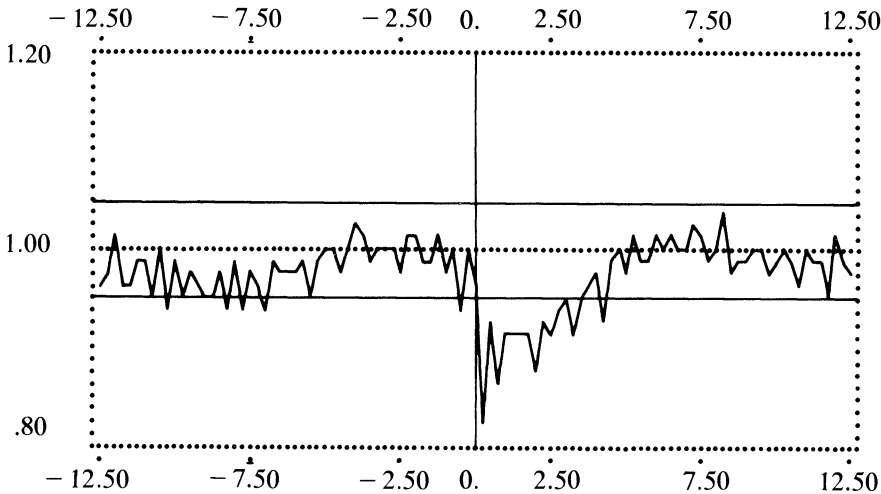

FIG. 4. $\sqrt{\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)}$ and 95 per cent confidence limits about the level 1.00.

In the case of independence of the two processes $p_{MN}(u)/(p_M p_N) = 1$ and this hypothesis may be easily checked into by plotting horizontal lines at the levels $1.00 \pm 1.96/(2\sqrt{2hT\hat{p}_M\hat{p}_N})$. This has been done in Figure 4, for the data of Figure 3. This new Figure is strongly suggestive of the association of a reduced rate of M spikes for a period after the occurrence of an N spike.
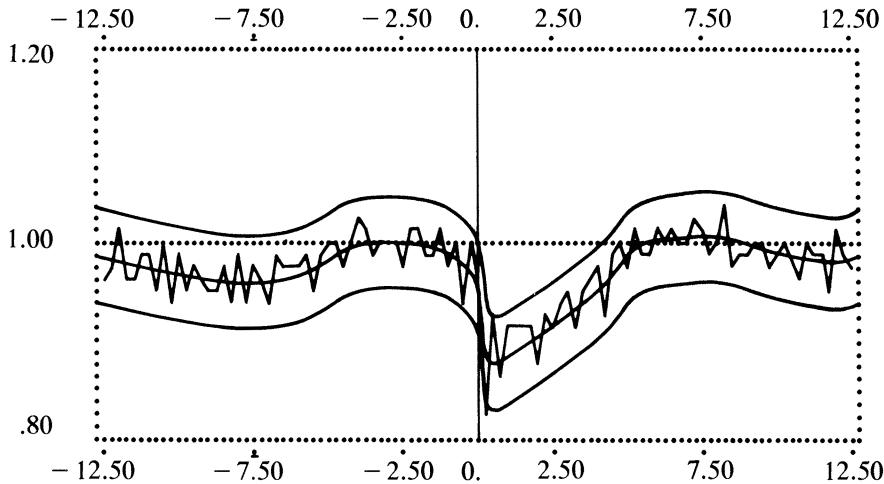

FIG. 5. $\sqrt{\hat{p}_{MN}(u)/(\hat{p}_M\hat{p}_N)}$ and 95 per cent confidence limits about a smoothed version of these values.

We might have chosen to indicate sampling fluctuations in the manner of Figure 5 where we have plotted $\pm 1.96/(2\sqrt{2hT\hat{p}_M\hat{p}_N})$ limits around a heavily smoothed version of the estimate. The points where the upper line is below 1.00 might be assessed significant. Figure 5 is also suggestive of a direction of causation for the two cells, namely the N spikes seem to be associated only with later M spikes (i.e., at positive u). This last is consistent with the neurophysiologists' understanding of the relation of the two particular cells for which this data was collected.

7. Final remarks. In this article I have sought to describe some of the stages involved in a modern applied statistics problem. These include: (i) the experimenter collects interesting data, (ii) the experimenter recognizes relevent scientific parameters to estimate, (iii) the experimenter consults a statistician as to whether or not his estimates are significant, (iv) the statistician suggests means of assessing sampling fluctuations and possibly suggests transformations in order that the data be more simply described and (v) the experimenter and statistician collaborate to determine and fit a statistical model and to design future experiments to confirm that model. The distinction between these stages is not always apparent nor is it clear whose ideas are whose. The experimenter must learn a fair amount of statistical methodology and the statistician must learn a fair amount of the experimenter's subject matter before real progress can be made.

Previous to my involvement in this work and the carrying out of the research leading to [3], another applied statistician, Peter Lewis, was involved. He suggested, [4], some clever and widely applicable alternate procedures for assessing the significance of the estimate (10). These were: (i) look at the variations of $n(u, h)$ for $|u|$ large enough that the interactions of the two cells are independent, (ii) substitute for the N train, a spike train of $N(T)$ independent spikes and examine the variation of $n(u, h)$ here, (iii) split the observed trains into J (say 20) pairs of shorter records and examine the variation of $n(u, h)$ when computed for the shorter records. The easiest of these procedures to carry out is (i). When done, it does lead to an estimate of the variation of the same order of magnitude as that of this paper.

Reference [3] is a joint   paper carrying the analysis of the data described in this paper considerably further forward. I should like to thank Hugh Bryant, Jr. and Jose Segundo for the great pleasure I have derived from working with them on this problem.

References

1. Biometrika Tables for Statisticians, Vol. I, Third Edition (eds. E. S. Pearson and H. O. Hartley) Cambridge Univ. Press, Cambridge, 1966.

2. D. R. Brillinger, Estimation of the second-order intensities of a bivariate stationary point process, J. Royal Statistical Soc. B December (1975).

3. D. R. Brillinger, H. L. Bryant, Jr. and J. P. Segundo, The identification of neuronal interactions, To appear in J. Neurophysiol. (1975).

4. H. L. Bryant, Jr., A. Ruiz Marcos and J. P. Segundo, Correlations of neuronal spike discharges produced by monosynaptic connections and by common inputs, J. Neurophysiol., 36 (1973) 205–225.

5. K. L. Chung, Crudely stationary counting processes, this MONTHLY, 79 (1972) 867–877.

6. D. R. Cox and P. A. W. Lewis, The Statistical Analysis of Series of Events, Methuen, London, 1966.

7. J. Eccles, The synapse, Scientific American, 212 (1965) 56–66.

8. W. Feller, An Introduction to Probability Theory and its Applications, Vol. I, 3rd ed., Wiley, New York, 1968.

9. J. S. Griffith and G. Horn, Functional coupling between cells in the visual cortex of the unrestrained cat. Nature, 199 (1963) 876, 893–895.

10. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol. 3, Griffin, London, 1966.

11. Stochastic Point Processes (ed. P. A. W. Lewis), Wiley, New York, 1972.

12. V. A. Volkonskii and Yu. A. Rozanov, Some limit theorems for random functions, I. Theory Probability and its Applications, 4 (1959) 178–197.

DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720.

ADDENDUM TO "STRONG DERIVATIVES AND INVERSE MAPPINGS"

ALBERT NIJENHUIS

8. Further comments. An excellent discussion of strong derivatives and other properties of functions appears in Gleason's lectures [8], which also make reference to [7].

A treatment of the subject of the first four sections of the present paper, ostensibly finite dimensional but general in spirit, appears in [9]. Their use of the "homeomorphism theorems" 5.1.6–7, is very clarifying.

A correction: formula (4.7) should be

$$(4.7) \qquad\qquad \varphi(y, z) = x_0 + \varphi^*(\lambda^{-1}(y - y_0), z).$$

Additional References

7. E. Asplund and L. Bungart, A first course in integration, Holt, Rinehart and Winston, New York, 1966.

8. L. K. Durst (ed.), Proceedings of the CUPM Geometry Conference, Part II: Geometry in other subjects, Math. Assoc. of Amer., 1967.

9. J. M. Ortega and W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PA 19174.