METHODS FOR ANALYSIS OF DATA REPRESENTING CONCENTRATION PROFILES OF PLATELET ANALOGUES IN BLOOD FLOW

J. Shane Kippenhan, Joachim H. Nagel, Eugene C. Eckstein

Department of Biomedical Engineering
P.O. Box 248294
University of Miami
Coral Gables, Florida 33124

Abstract: Methods for the estimation of particle concentration profiles from numerical data are presented. The estimation techniques described, which involve the use of Fourier transforms, make more efficient use of data than do simple histogram techniques. Additionally, Fourier methods of analysis have been used to test theoretical models of experimental data.

Introduction and Experimental Background

In the flow of blood through small vessels, interactions occur among the red cells and platelets, which are suspended in the blood plasma, that are complex and not completely understood. Recent studies of blood flow have indicated that these types of interactions tend to cause a non-uniform radial distribution of platelets in well-developed flow through capillary-sized tubes [1,2]. Concentrations of simulated platelets have been observed to increase near the wall of the flow vessels. The authors' investigations of this near-wall excess have been conducted with suspensions (of red cells and latex beads) that model the important rheological aspects of human blood. The objectives of the authors' studies have been (1) to obtain concentration profiles for a wide range of blood flow conditions, and (2) to eventually use these results to construct a theoretical model of rheological events in the flow, relating the nature of the near-wall excess to variables such as hematocrit, fluid shear rate, tube diameter, and suspending fluid viscosity.

Briefly, the procedure for experiments was as follows: Fresh human red blood cells, along with platelet-sized fluorescent particles, were suspended in simulated plasma. After preparations to properly set all the controlled variables, steady flow was established in a polyethylene tube for a length of time sufficient for the flow to fully develop. The suspension was then rapidly frozen (by pouring liquid nitrogen over the tube) to preserve the radial positions of particles for later observation. Selected short segments of tube were mounted vertically in a modified microtome assembly which was fixed to the stage of a microscope. Successive cross-sectionally-cut tube surfaces were examined through the microscope and these images were recorded by videocamera. Thus, raw data from these experiments consisted of images of numerous bright, fluorescent particles surrounded by approximately circular tube walls. Analysis of these images was performed in two basic steps. First, distances from individual particles to respective tube walls were calculated. These distances were then compiled and processed to produce representations of the concentration profiles that existed in the tubes at the time of freezing.

Results of these experiments must be carefully evaluated and interpreted on a statistical basis. Individual experiments are time-consuming and difficult to conduct. Since each estimate of the concentration profile for a single experiment is but one sample of the population of concentration profiles for the given experimental conditions, one can appreciate the considerable amount of time required to collect a statistically significant number of samples. An additional statistical consideration is that each of these sample concentration profiles is estimated by considering only a relatively small fraction of the number of particles contained in a particular tube (particle concentrations in these experiments are approximately 200,000/mm^3). For these reasons, a sample size of 1000 observed particle distances (typical for these experiments), though a large number in most statistical settings, is, in the context of these experiments, fairly small. Reasonably accurate and smooth statistical representations of the concentration profiles can be difficult to obtain.

Histogram Techniques

The time-honored histogram approach to the estimation of data behavior has the advantage of being direct and easy to apply, but the process of sorting observations into "bins" often results in unacceptable losses of information. Figure 1, shown below, illustrates some of the weaknesses of the histogram in its application to our data. Each of the three sets of points in the graph represents the same group of raw data, yet obvious differences can be seen in the shape of the three profiles. In each of these three histograms, bin widths have been selected such that each bin represents an equal annular area within the tube lumen, so that the curves represent the actual shapes of concentration profiles.

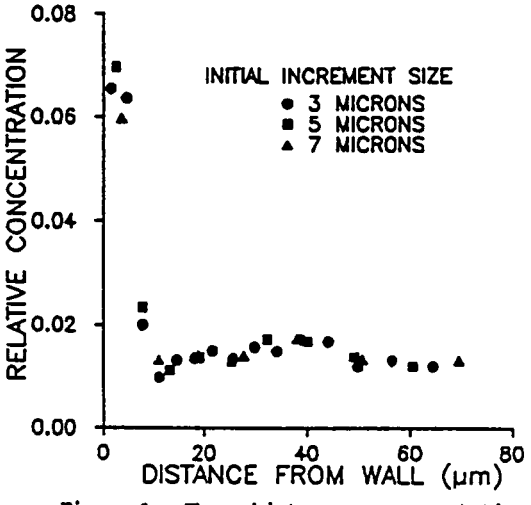

Figure 1: Three histogram representations of concentration profiles of platelet-sized latex particles in a 200-μm-diameter tube. Initial (those closest to the wall) "bin" sizes were chosen as indicated, and subsequent bins represent equal-area annuli in the tube. Abscissa values of the plotted points are at centers of bins.

An obvious trend, well illustrated by this comparison, is the general trade-off that occurs between smoothness and resolution of the peak location. The larger the bin width, the greater will be the number of observations per bin. These larger samples will be less subject to statistical scatter, resulting in smoother curves. The use of large bins, however, will result in significant losses of sample information. Figure 1 illustrates the fact that histograms with smaller bins can more accurately estimate the location

TH0206-3/87/0000 - 167 $1.00

of the peak concentration, but there is always a practical bin size limit below which the amount of bounce in the histogram becomes unacceptable.

Density Estimation by the Kernel Method

Several non-parametric methods for probability density estimation have been presented in recent statistical literature [3]. One of the most powerful and best understood techniques is that known as the kernel method. In this approach, each observation within a sample group, rather than being sorted into bins of a histogram, is represented by a function which is a miniature form of a symmetric probability density function. The center of symmetry is located at the abscissa value corresponding to the value of the observation, and the value of the density estimate at any abscissa value is equal to the sum of the contributions of the miniature function values for all the observations in a given sample.

Standard Kernel Estimation

The above description is represented by the following equation:

$$f(x)=(1/nh) \cdot \sum_{i=1}^{n} K((x-X_i)/h) \qquad (1)$$

where $f(x)$ = the density estimate as a function of the abscissa,
X_i = the ith sample observation,
n = the number of observations in the sample,
K = the kernel function, and
h = the "window width" or "smoothing parameter".

Graphically, the kernel estimate can be thought of as the sum of a number of "bumps" placed at the abscissa values of the observations.

The kernel estimate can be made either by performing the above summation in the spatial domain, or, more efficiently, by using Fourier transform techniques. Once it is realized that the kernel estimate is actually the convolution of the data with a selected probability density function, this convolution can be performed by multiplying, in the spatial frequency domain, the transform of the data by the transform of the appropriate density function. The fast Fourier transform (FFT) provides a very efficient method for this process.

The estimation process begins with the construction, from a list of observed distances, of a discrete sequence of numbers by a process in which each element of the sequence is assigned a value representing the relative frequency of observations at that element location. The number of elements in this sequence is chosen to be a power of two (usually 128 in our case) to enable the subsequent use of the FFT. This stage is somewhat similar to the formation of a histogram with very small bins, but compared to the grouping of data for a histogram, more sample information is retained at this point, because weights of observations that fall between any two of the element locations are proportionally allocated to these adjacent elements. The sequence of weights is transformed, using the FFT, and the transform of a selected kernel function is applied to the transformed sequence. A plot of the inverse Fourier transform of this conditioned sequence provides a smoother representation of the distribution of the data than would be provided by a simple histogram. This representation can more accurately estimate features of the distribution, such as the mode, than would a histogram subjected to simple smoothing with non-statistical techniques.

The choice of a window width h is analogous to the choice of a histogram bin width. Larger window widths allow smoother representations of the density function, but significant bias is often introduced in this way. The use of more narrow windows, while reducing the bias, will cause the curves to be less smooth. Choice of the window width used for estimating a group of data can be made to depend upon the number and spread of observations, or it can be set equal to a constant for the purpose of reliably comparing multiple data sets.

Adaptive Kernel Estimation

A modification of the standard kernel method is a technique known as the adaptive kernel method, in which the window width is made to automatically depend inversely upon the local density of data. The local density must be approximated by an initial estimate, which can be a standard kernel estimate. Higher density regions are estimated with more narrow windows, while lower density regions are estimated with wider windows. It is not practical to apply Fourier transform methods in this technique, so this method is not as computationally efficient as the standard kernel method, but it does result in smoother estimates of regions in which observations are relatively scarce. Silverman [3] gives an excellent detailed description of this technique.

Figure 2 shows a comparison of a variable kernel estimate and a histogram estimate for the same set of data. The curves in this figure do not represent the kind of actual concentration profile shapes shown in Figure 1. The linear decrease of annular area (for fixed-width annuli) with distance from the tube wall has not been compensated for here, with the result that the essentially uniform concentration in the central tube region is represented here by a roughly triangular shape that decreases to zero. This "waveform" shape is necessary to satisfy the assumption of periodicity inherent in the FFT methods described here. This comparison clearly illustrates the superiority of the kernel estimate's definition of the concentration profile's near-wall peak. While the histogram vaguely shows a peak somewhere within the first 6-7 microns away from the wall, the kernel estimate clearly indicates that a concentration maximum occurs between 3 and 4 microns from the wall.

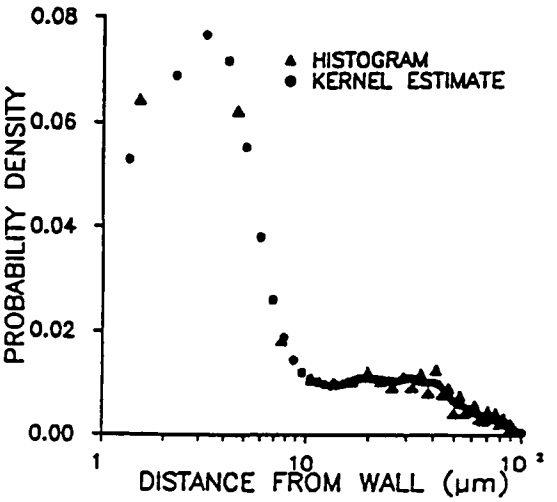

Figure 2: A comparison of histogram and adaptive kernel estimation techniques. Data is plotted on a log scale to facilitate comparison in the near-wall region. The kernel estimate is smoother than the histogram, yet gives better resolution of the near-wall peak location.

Most of the methods described here could easily be used in any application that requires estimation of probability density from data. Many medical imaging aplications, for example, could benefit from the increased spatial resolution and signal-to-noise ratio available with the processing techniques shown here.

Parameter Estimation Methods

The procedures described thus far have been methods to estimate distribution shapes non-parametrically, i.e., with no previous assumptions about data behavior. One of the goals of the authors' research, however, is an understanding of the underlying principles governing the shape of the concentration profiles. An important part of this goal is to search for an analytical description of the authors' concentration curves that is based on a few parameters. An example of this type of description is a distribution known as the Inverse Gaussian, which has a shape similar to that of the authors' data distributions, and is completely defined by two parameters. Extensive comparisons of data with this distribution, however, have revealed repeatable inconsistencies in its fit.

The technique of applying Fourier transforms to the data, as introduced earlier, has led to interesting developments in the direction of parameter estimation. Methods to describe the data in terms of their Fourier transforms presently have more promise than past attempts to describe the curves in the spatial domain. These past attempts have included attempts to fit data to analytical functions, as well as calculations of the first few moments of the data (mean, variance, skewness, and kurtosis) for the purpose of comparisons between data sets.

The Fourier transform of a probability density estimate completely describes all existing moments of the data. An analytical function that accurately represented this Fourier transform would be the "characteristic function" of probability theory, from which all existing moments can be calculated. This function, defined as the Fourier transform of the probability density function, can describe the data as completely as can any function in the spatial domain. The inverse Fourier transform of this characteristic function, performed analytically, would yield a closed-form analytical expression, completely describing the radial particle distribution.

The most promising results of this approach to date are illustrated in Figures 3 and 4. Figures 3a and 3b show the real and imaginary parts, respectively, of the complex sequence that is the Fourier transform of the density estimate shown in Figure 2.

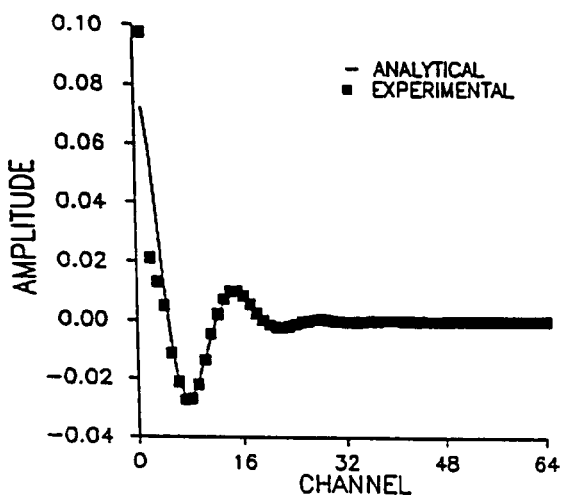

Figure 3a: The real part of the complex sequence that is the Fourier transform of the kernel estimate shown in Figure 2. An exponentially damped cosine function has been fit to the transformed sequence.

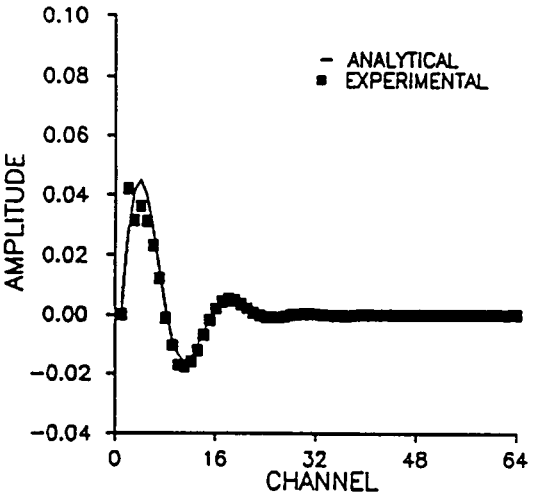

Figure 3b: The imaginary part of the complex sequence that is the Fourier transform of the kernel estimate shown in Figure 2. An exponentially damped sine function has been fit to the transformed sequence.

Also shown in these figures are curves which have been fit to the real and imaginary parts of the transformed data. The equation which was fit to the real sequence was

$$f(s) = A \cdot \exp(-Bs) \cdot \cos(Cs). \qquad (2)$$

The equation fit to the imaginary sequence was

$$g(s) = A \cdot \exp(-Bs) \cdot \sin(Cs). \qquad (3)$$

Figure 4 shows the inverse transform of the complex sequence whose real and imaginary parts are defined by equations (2) and (3), respectively. Though some information pertaining to the shape of the density estimate has obviously been lost, the shape of the peak has been fairly well reproduced.

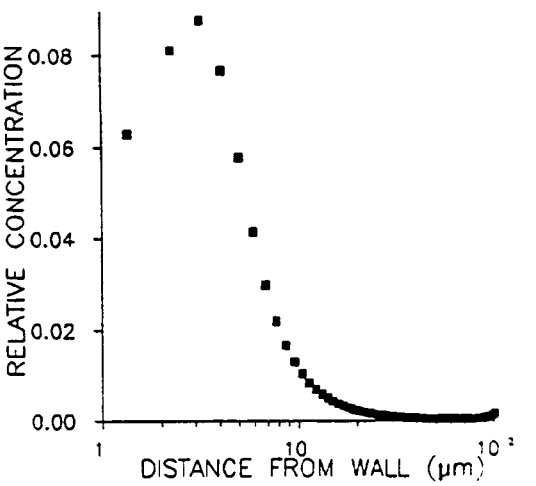

Figure 4: The inverse Fourier transform of the complex sequence whose real and imaginary parts are defined by the analytical functions plotted in Figures 3a and 3b, respectively. The curve is plotted on a log scale to facilitate examination of the near-wall peak.

According to Euler's identity, equations (2) and (3) can be represented by the following complex exponential equation:

$$h(s) = A \cdot \exp((-B+iC)s). \qquad (4)$$

The real importance of the type of approach illustrated here in terms of an overall theoretical model has yet to be evaluated, but it would appear to have some usefulness. Equation (4) seems to predict a significant part of the data's behavior, and it does so on the basis of three parameters. Further work with transformed data, in combination with presently ongoing efforts to formulate governing equations based on appropriate physical concepts, may eventually enable comprehensive theoretical modeling of the function and transport of platelets in blood flow.

Acknowledgements

This research was supported by Grant HL33100 from the National Heart, Lung and Blood Institute. The authors would like to express their appreciation to Mr. Christopher Waters and Mr. David Bilsker, whose work produced the data analyzed in this study.

References

[1] G. J. Tangelder and H. C. Teirlinck, "Distribution of blood platelets flowing in arterioles," in Amer. J. Physiology, vol. 248, H318-H323, 1985.

[2] A. W. Tilles and E. C. Eckstein, "The near-wall excess of platelet-sized particles in blood flow: its dependence on hematocrit and wall shear rate," in Microvascular Research, vol. 33, 211-223, 1987.

[3] B. W. Silverman. Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall, 1986.