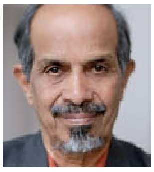
Kosambi and Proper Orthogonal Decomposition

Roddam Narasimha



Roddam Narasimha is at the Jawaharlal Nehru Centre for Advanced Scientific Research at Bangalore. His primary research interests are in fluid dynamics, especially in turbulent and transitional flows, and in applications in aerospace technology and atmospheric science. His most recent work has been concerned with the fluid dynamics of cloud flows, multiple transitions at the leading edge of swept wings on aircraft, and the design of optimal wings for propeller-driven aircraft.

In 1943 Kosambi published a paper titled 'Statistics in function space' in the *Journal of the Indian Mathematical Society*. This paper was the first to propose the technique of statistical analysis often called proper orthogonal decomposition today. This article describes the contents of that paper and Kosambi's approach to the subject. It was only in 1967 that it began to be appreciated that the method that had gained wide currency in several fields under different names was first set out in Kosambi's 1943 paper.

There is a powerful mathematical tool in statistical analysis that is widely used in a variety of disciplines under a variety of different names. It is familiar in fluid dynamics as Proper Orthogonal Decomposition, in meteorology and other related fields as Principal Component Analysis, in more mathematically oriented studies as the Karhunen–Loéve Expansion and in other fields as Empirical Orthogonal Functions. It is not so widely known that the technique was first proposed and described by Damodar Dharmananda Kosambi (1907–1966), in a paper titled 'Statistics in function space' published in 1943 in the *Journal of the Indian Mathematical Society* [1]. (There was an 'exploratory' paper in *Current Science* in 1942 [2].) The reason that it appears under different names – but none acknowledging the original discoverer of the method – is that it seems to have been independently rediscovered by several others during the decade following Kosambi's paper. The tool evidently tackles a problem that has been encountered in numerous disciplines. The present article briefly describes the background to Kosambi's pioneering contribution

Keywords
Proper orthogonal decomposition, Karhunen–Loéve expansion, statistics in function space, characteristic eddies, special calculating machines.

without going into the mathematics of the discovery in any formal detail.

The reason that Kosambi's analysis appears under different names – but none acknowledging the original discoverer of the method – is that it seems to have been independently rediscovered by several others during the decade following Kosambi's paper.

But first let us describe the problem, which typically arises as follows. In a large number of phenomena, occurring in nature or created by technology, there are quantities of interest that vary continuously in time and/or space. For example, if we measure the velocity of water being conveyed in a pipe at some fixed point within it when the flow is turbulent (as it often is), the graph of velocity (say some component of it) versus time is a curve that depicts an apparently random motion. And it appears to demand a statistical description. Similarly if we plot the rainfall or temperature at any meteorological station, or process and control parameters in a chemical plant. All these are examples of stochastic processes in time: what is obtained is a time trace of some apparently random variable.

More generally, we can consider processes where what is observed is best expressed as a function or curve. At the time of Kosambi's paper, the statistics of variables with discrete or continuous values (e.g., heads or tails in tossing a coin, distributions of height among a statistically homogeneous population) had been well studied. What about those processes where the outcome is a curve? Kosambi starts his paper with the sentence: "The main purpose of this note is to develop statistical methods for discrimination between samples consisting of whole curves". Kosambi's question is about how to do statistics if the sample space consists of curves, each of which is a set of infinitely many values constituting a function of some (one or more) independent variables such as time and/or space. G I Taylor at Cambridge had already introduced a statistical theory of turbulence in 1935, and a spectral method of describing stationary stochastic processes had been formulated by Wiener and Khintchine in the early 1930s. Compared to these developments, Kosambi's approach seems more geometrical.

"The main purpose of this note is to develop statistical methods for discrimination between samples consisting of whole curves."
– *Kosambi*

G I Taylor at Cambridge had already introduced a statistical theory of turbulence in 1935, and a spectral method of describing stationary stochastic processes had been formulated by Wiener and Khintchine in the early 1930s. Compared to these developments, Kosambi's approach seems more geometrical.

Its basis is problem-specified.

The paper starts by recognizing the importance of defining the 'distance' between any two functions, say $f(x)$ and $g(x)$ in the interval (0,1). Kosambi first proposes a distance $r(f,\, g)$ by the integral

$$\begin{aligned} r(f,g) &= \phi(f-g) \\ &= \int K(s,t)[f(s)-g(s)][f(t)-g(t)]\mathrm{d}s\mathrm{d}t, \end{aligned} \tag{1}$$

where ϕ is a positive definite quadratic form and $K(s,t)$ is a positive definite or semi-definite continuous symmetric kernel. He shows that this definition meets all the criteria for a distance measure. Now a multi-variate normal distribution involving a distance ϕ is the definite integral of

$$(2\pi)^{-k/2}e^{-\phi/2}\mathrm{d}V \tag{2}$$

over a desired region, where k is the number of variables in ϕ and $\mathrm{d}V$ is the associated volume element. A curve is an infinite-dimensional vector, so $k \to \infty$ and (2) is not useful. The key proposal made here by Kosambi is first a choice of independent variables that reduces ϕ to a diagonal form. This is possible for the class of kernels considered, and K can be expanded as a series involving its orthonormal eigenfunctions ϕ_i, with eigenvalues that are positive and so can be denoted by σ_i^2:

$$K(s,t) = \sum \sigma_i^2 \phi_i(s)\phi_i(t)\,. \tag{3}$$

This enables the function $f(t)$ to be expanded in a series of the eigenfunctions ϕ_i with coefficients x_i :

$$f(t) = \sum x_i\phi_i(t), x_i = \int f(t)\phi_i(t)\mathrm{d}t. \tag{4}$$

The x_i here can be thought of as the 'Fourier coefficients' of the function $f(t)$ with respect to the basis ϕ_i.

If the kernel is known the eigenfunctions can be determined, and (2) can be used to define the distribution. If this distribution is normal, so is that of the curves being considered – an assumption necessary for Kosambi's analysis.

The sample of curves can therefore be equivalently described in terms of their uncorrelated Fourier coefficients x_i, with $E(x_i^2) = \sigma_i^2$, and for any finite n, ordinary multivariate statistics can be used.

With kernel K obtained using the canonical expansion (3), Kosambi shows that the covariance between values of the functions at s and t is $K(s,t)$. Thus, if the data are given as a sample of n curves $y_i = f(t_i)$, the matrix of covariances $E(y_i, y_j)$ is the kernel matrix $K(t_i, t_j)$. The best estimate of the population mean and the population kernel are then given respectively by

$$m(t) = \frac{1}{n}\sum f_i(t), \tag{5}$$
$$k(s,t) = \frac{1}{n-1}\sum \left[f_i(t) - m(t)\right]\left[f_j(s) - m(s)\right].$$

Thus if f is known as empirical data, the kernel can be determined as also its eigenfunctions and the associated eigenvalues. The sample of curves can therefore be equivalently described in terms of their uncorrelated Fourier coefficients x_i, with $E(x_i^2) = \sigma_i^2$, and for any finite n, ordinary multivariate statistics can be used. Furthermore, it can be shown by a variational argument that the expansion converges 'as fast as possible' to the total energy (Lumley [3]).

In the above description I have omitted various mathematical conditions that have to be satisfied for formal proofs but I hope that the ideas are clear. Basically what Kosambi succeeded in doing was to show how, from given data on a sample of curves, there is a method by which the statistics can be handled as a multi-variate problem, and the Fourier coefficients of the expansion of the curve in a special basis converge rapidly in a mean square sense.

The importance of this theory in applications arises chiefly from the fact that in many cases observations

The importance of this theory in applications arises chiefly from the fact that in many cases observations can be expressed in terms of a small number of Fourier coefficients.

can be expressed in terms of a small number of Fourier coefficients. Thus the problem is rendered more manageable. In particular the eigenfunction can be ordered in terms of significance, for example by the contribution to the kinetic energy of the fluid motion. The problem therefore leads to a compact description amenable to further analysis.

In the fluid dynamics of turbulent flows these ideas were introduced by John Lumley, then working at the Aerospace Engineering Department, Pennsylvania State University, in a paper read at an international colloquium held at Moscow in 1965 [4]. He listed six conclusions (designated A to F) from the 'proper orthogonal decomposition (POD) theorem' as stated in Loéve's book on *Probability Theory* (1955). All the six conclusions are in Kosambi's 1943 paper [1], the correspondence being as follows: Lumley's A = Kosambi's equation (1.3), B = text following (1.3), C = (1.4), D = text before (1.4), E = (1.3), F = (1.3) and following text. During the discussion following Lumley's paper, the well-known Russian fluid-dynamicist A M Yaglom pointed out that the decomposition attributed to Loéve had earlier been introduced for different purposes by several scientists, in particular by Kosambi (1943), Karhunen (1946), Pougachev (1953) and Oboukhov (1954). He also pointed out that E N Lorenz (well known for his pioneering studies of the chaotic behaviour of the atmosphere) had used POD in studies of atmospheric turbulence (Lorenz incidentally introduced meteorologists to principal components analysis). Lumley has since acknowledged these earlier developments.

Lumley listed six conclusions from the 'proper orthogonal decomposition (POD) theorem' as stated in Loéve's book on *Probability Theory* (1955). All the six conclusions are in Kosambi's 1943 paper.

The power spectrum of turbulent kinetic energy is usually expressed in terms of its density as a function of wave number or frequency (the integral giving the total energy). Using POD, it can also be expressed in terms of the energy in the modes, also called 'characteristic eddies'; this description often accounts for most of the

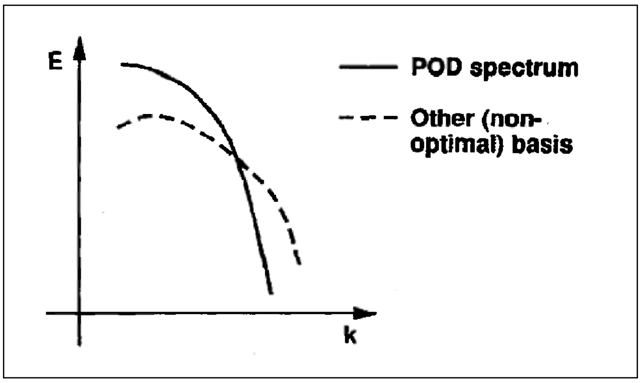


***Figure 1. Schematic of energy spectra in POD and non-POD bases (Berkooz et al 1993* [6]).**

energy in relatively few modes, and is more compact (*Figure* 1). In the turbulent flow in a channel, for example, it has been found that the dominant eddy (i.e., the first term in the expansion) accounts for 76% of the total energy [5] . POD provides us therefore with an effective expansion of the velocity field in organized modes (coherent structures) with random coefficients.

Another major application has been to provide low-dimensional dynamical-system models for understanding turbulent flows.

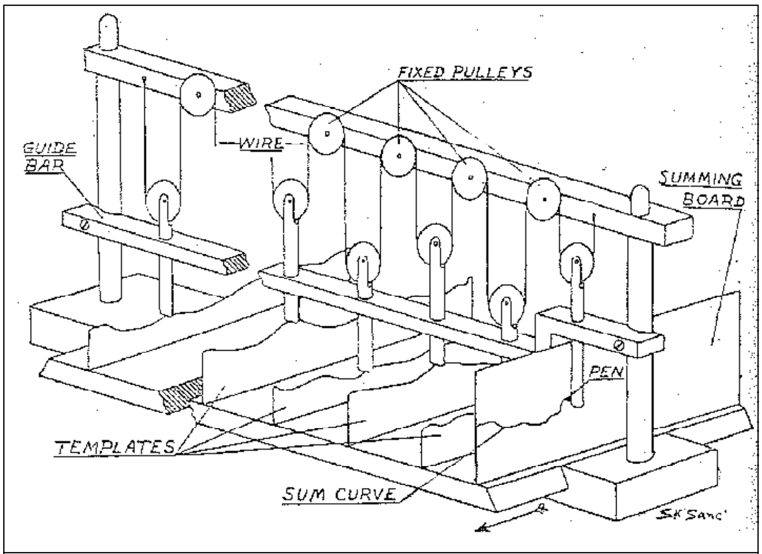
In spite of all this, the problem of turbulence remains unsolved, because POD, while an effective tool for analysis, demands prior data on covariances. To make it a powerful tool for dyanmical prediction has not been easy.

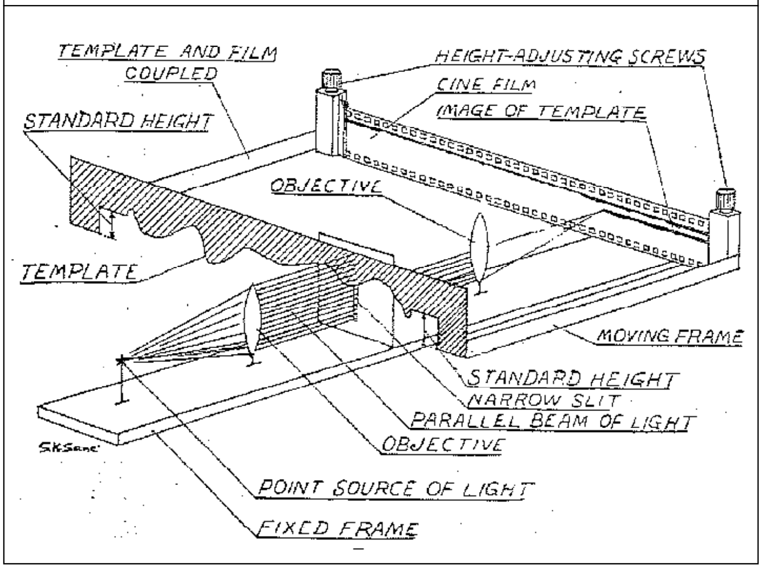
In meteorology, analysis of the principal components of rainfall time-series at different stations enables one to define geographical clusters where the proximity of the principal components helps identify regions of coherent rainfall.

In the turbulent flow in a channel, for example, it has been found that the dominant eddy (i.e., the first term in the expansion) accounts for 76% of the total energy.

There are two interesting features of Kosambi's paper that reflect his views of mathematics. First, applications are never far from his mind: the areas he mentions – again characteristically – are meteorology and anthropology (both important in India). (Perhaps his interest in the latter subject accounts for the geometrical flavour

***Figure 2* (*top*). *Design for a mechanical calculating machine (Kosambi,* 1943 [1]).**
***Figure 3* (*bottom*). *Design for an optical calculating machine* (*Kosambi,* 1943 [1]).**





Applications are never far from Kosmabi's mind: the areas he mentions – again characteristically – are meteorology and anthropology.

of his work; he talks about measurements from the ear orifice to the profile at fixed angles, for example.) Secondly, he proceeds to describe how the calculations can actually be carried out using special calculating machines, for which he offers two designs of his own, one mechanical and the other optical (*Figures* 2,3). The design of the mechanical calculator seems to have been inspired by Kelvin's tidal machine, with templates and pulleys. Kosambi shows how it is possible to perform additions, subtractions, divisions and summation of squares in his machine with the assistance of a pantagraph and measurement of torque and moments. He remarks

Calculating machines, under the circumstances that now limit my activity, cannot go beyond the stage of design. The fundamental ideas will be made clearly by the two schematic figures appended here in the hope of doing service to some more fortunately situated experimenter.

It will be seen that in those pre-digital days what Kosambi designed were analogue computers. Prof Dani mentions in his article in this issue of *Resonance* a letter by Homi Bhabha that refers to electro-mechanical calculating machines made by Kosambi with the help of an RAF engineer; it is not clear whether they were related to the designs shown in *Figures* 2 and 3.

Proper orthogonal decomposition, by whatever name it may be called in each discipline, is now a standard mathematical tool in the analysis of a variety of stochastic processes. Is it not interesting that Indian scientists learnt how to use it from Lumley, Lorenz *et al.*, rather than directly from Kosambi?

Is it not interesting that Indian scientists learnt how to use POD from Lumley, Lorenz *et al*, rather than directly from Kosambi?

Suggested Reading

[1] D D Kosambi, Statistics in function space, *J. Ind. Math. Soc.*, Vol.7, pp.76–88, 1943.

[2] D D Kosambi, A test of significance for multiple observations, *Curr. Sci.*, Vol.11, pp.271–274, 1942.

[3] J L Lumley, *Stochastic Tools in Turbulence*, Academic Press, New York, 1970.

[4] J L Lumley, The structure of inhomogeneous turbulence, In *Atmospheric Turbulence and Wave Propagation*, edited by A M Yaglom and V l Tatarstki, Nauka, Moscow, pp.166–78, 1967.

[5] P Moin and R D Moser, Characteristic eddy decomposition of turbulence in a channel, *J. Fluid Mech.*, Vol.200, pp.471–509, 1989.

[6] G P Berkooz, P Holmes and J L Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, *Annual Review of Fluid Mechanics*, Vol.25, pp.539–75, 1993.

Address for Correspondence
Roddam Narasimha
Jawaharlal Nehru Centre for Advanced Scientific Research
Jakkur
Bangalore 560 064, India.
Email:
roddam@caos.iisc.ernet.in