ORDINATION

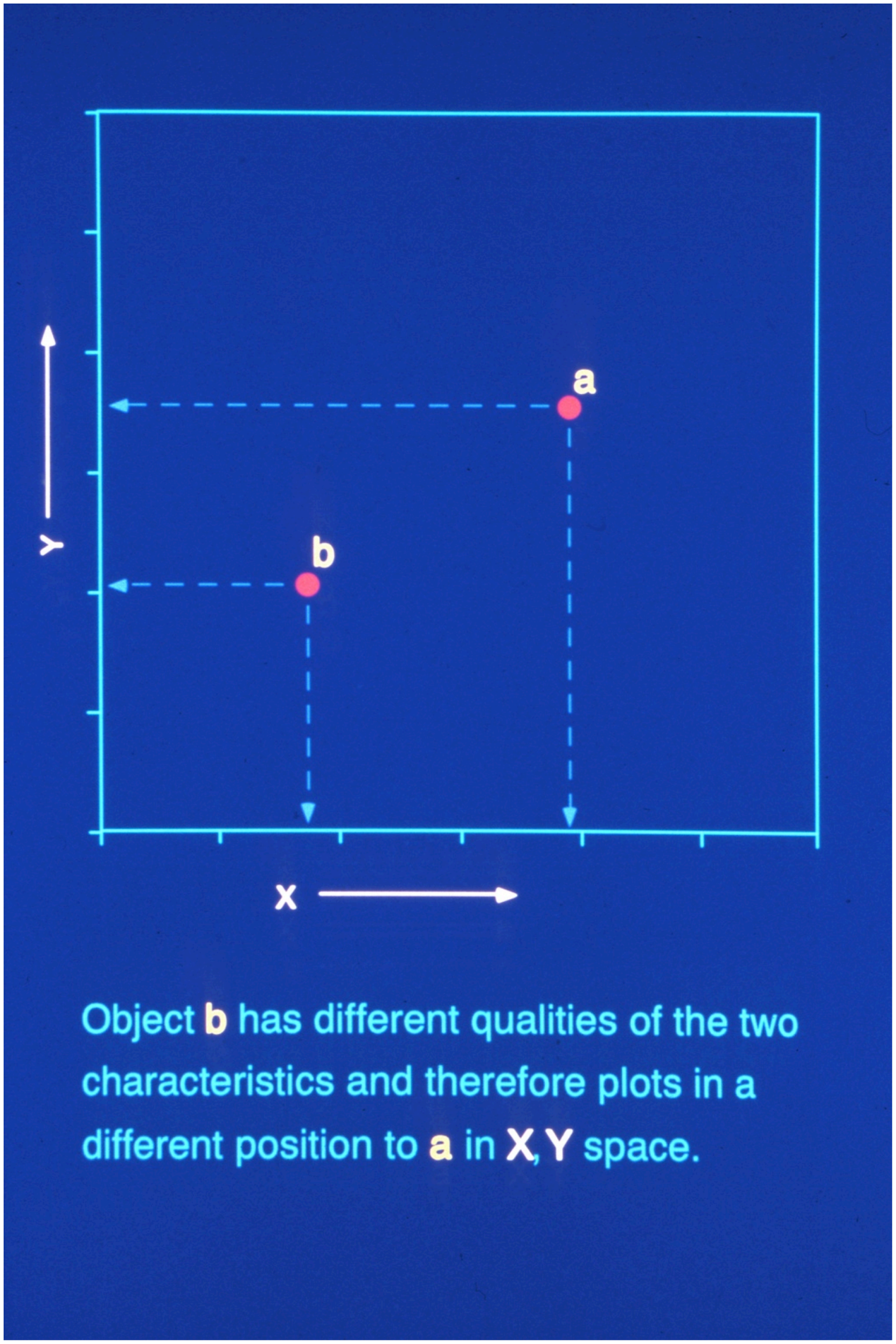
Introduction

Bob Spicer

Centre for Earth, Planetary, Space and Astronomical Research (CEPSAR)
The Open University, UK

r.a.spicer@open.ac.uk

This is a brief non-mathematical introduction to the principles underlying multidimensional ordination techniques of the kind that underpin CLAMP. For a more comprehensive (but again friendly to the non-mathematician) see KENT, M., and COKER,P.,1992, Vegetation and Description: Belhaven Press, London, 363p.



“Ordination” is the term given to a family of statistical techniques that order, in this case leaf characteristics or vegetation samples based on attributes that they share.

As with all statistics the aim of the analysis is to simplify the data as an aid to hypothesis generation.

In a simple two variable case objects (e.g. samples) can be positioned in 2D space in terms of the amounts of two attributes (e.g. leaf character states).

In reality natural systems have far more than two attributes.

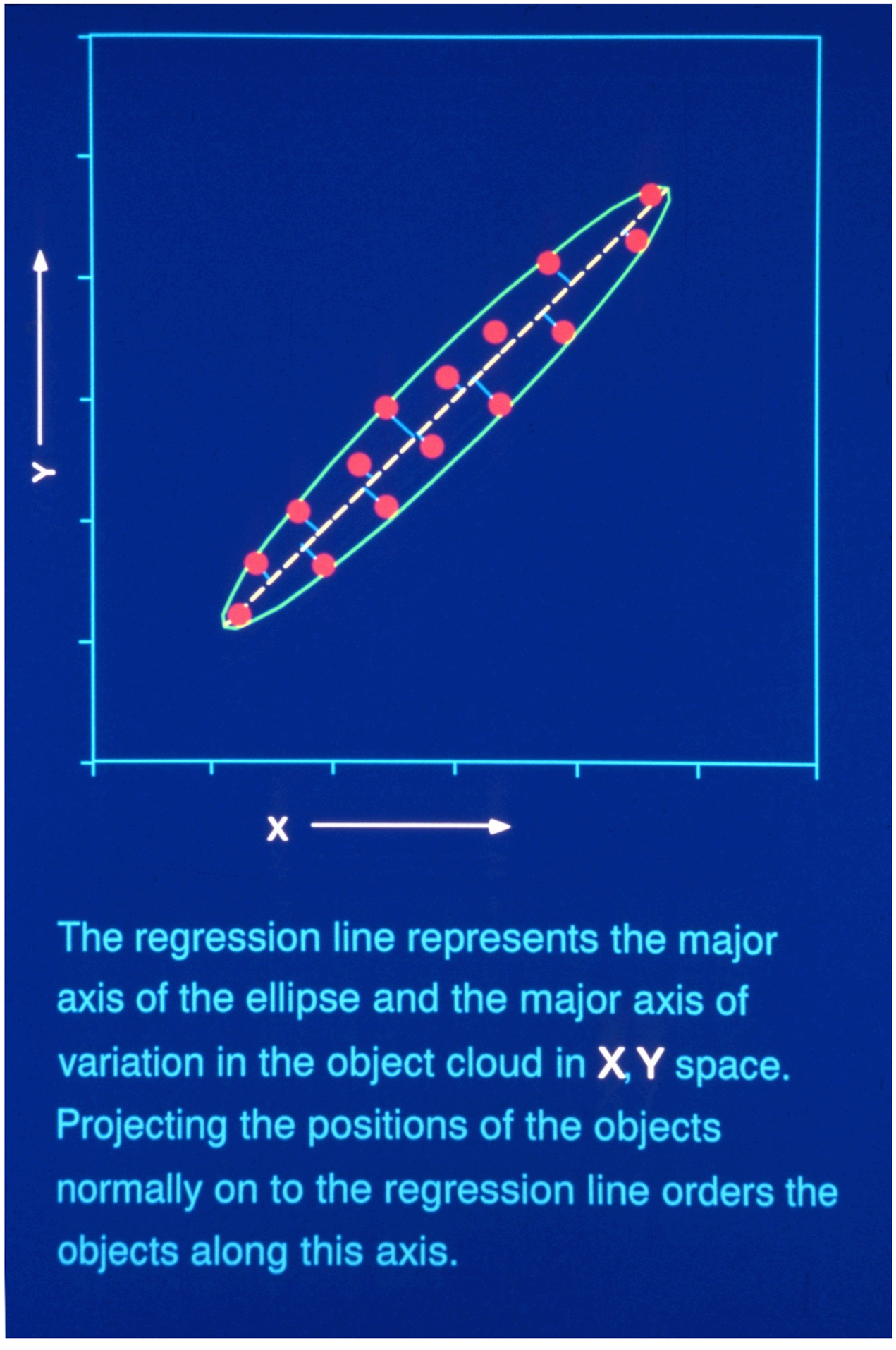
2

“Ordination” is the term given to a family of statistical techniques that order, in this case leaf characteristics or vegetation samples, in terms of attributes that they share. This ordering may then be used to speculate on any causal mechanisms such as environmental trends that may explain any emerging pattern.

As with all statistics the aim of the analysis is to simplify the data as an aid to hypothesis generation.

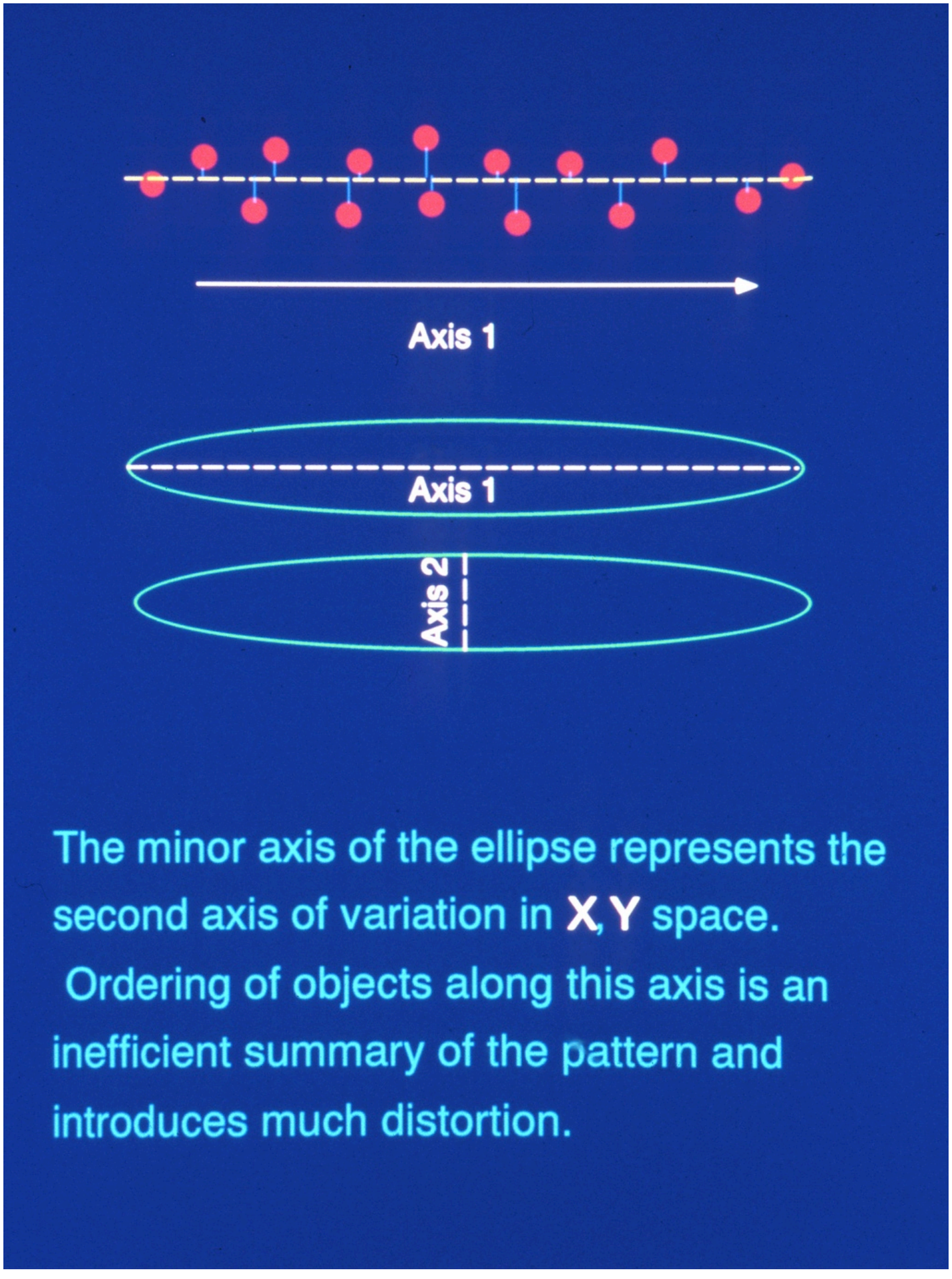
To start with we ca take a simple two variable case where objects (e.g. samples) can be positioned in 2D space in terms of the amounts of two attributes (e.g. leaf character states). The position of sample a is determined by the amounts of the attributes X and Y that it has. Similarly is sample B has different amounts of attributes X and Y it will plot is a different position.

In reality natural systems have far more than two attributes.



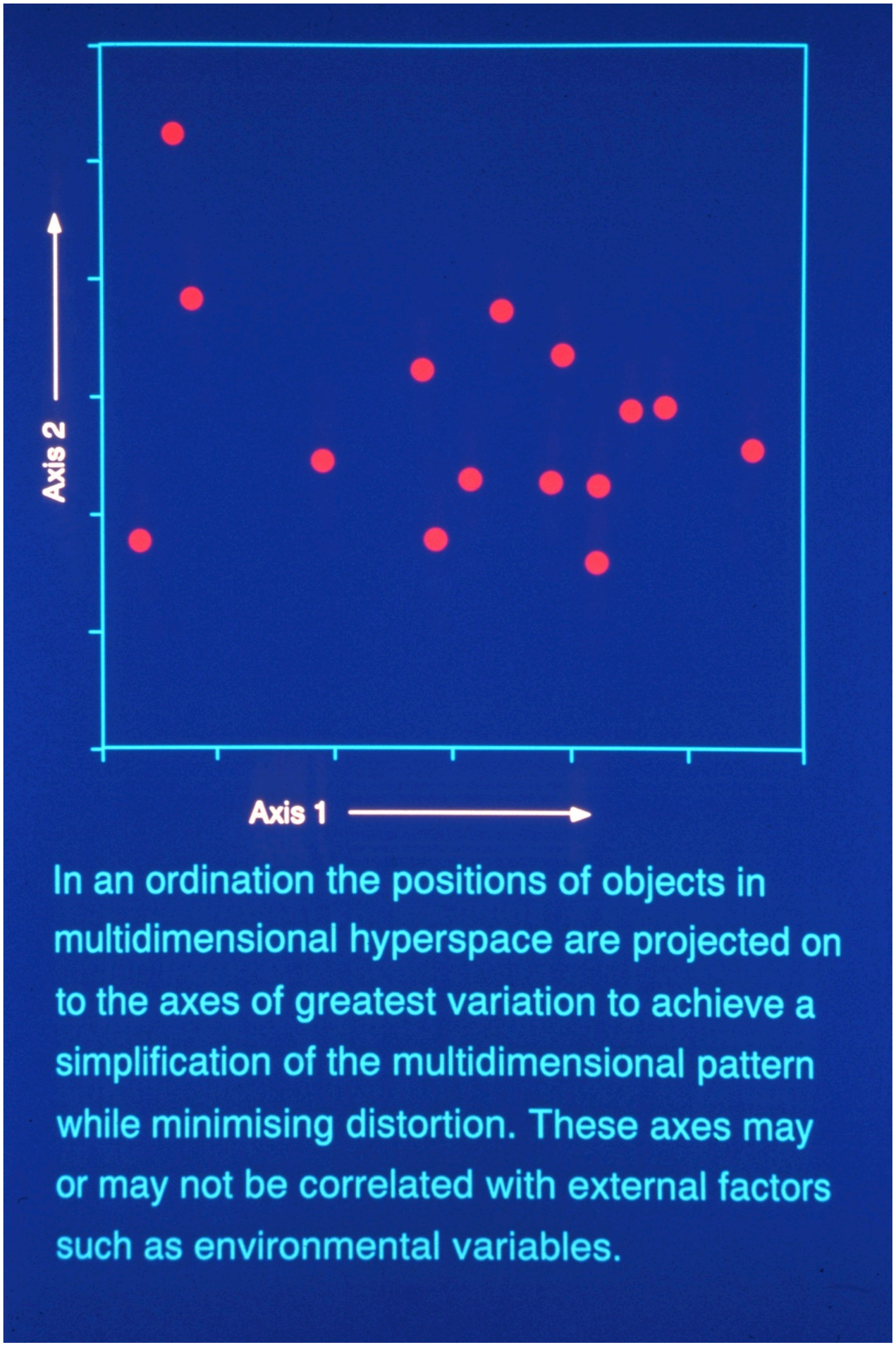
As more samples are added they may well form a cloud that has a major axis and a minor axis.

As more samples are added they may well form a cloud that has a major axis and a minor axis.



By ordering the samples along the longest axis of variation the data can be simplified while simultaneously retaining most of the important information.

The samples can be ordered along the longest axis of the cloud. This summarizes the main characteristic of the cloud (i.e. its length) and the positions of the samples relative to that characteristic.

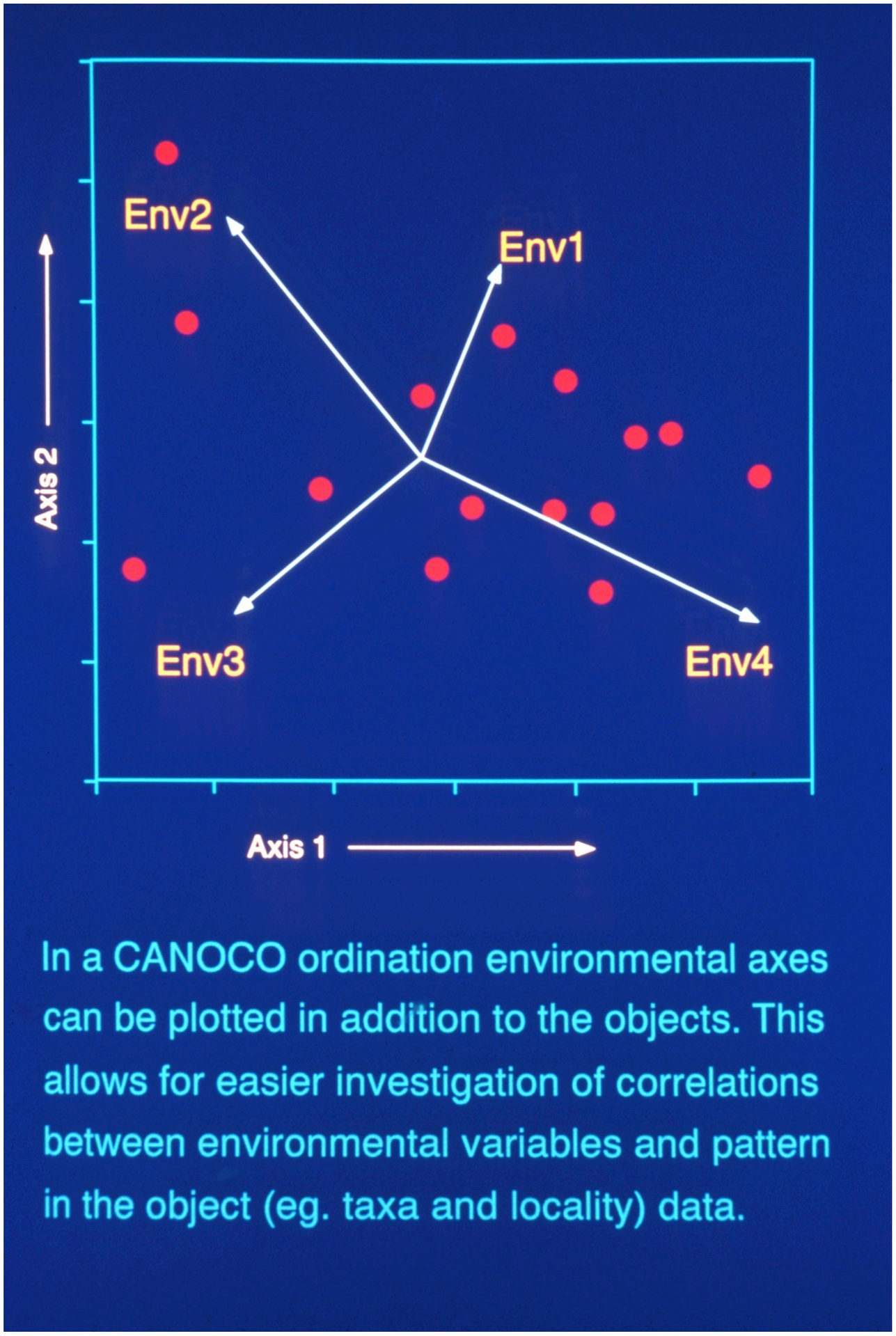


Ordination plot of Axis 1 (dimension of greatest variation) against Axis 2 (dimension of next greatest variation).

Several statistical techniques use such multidimensional ordering. Two commonly used methods are Principal Components Analysis (PCA) and Correspondence Analysis (CA) (sometimes called Reciprocal Averaging).

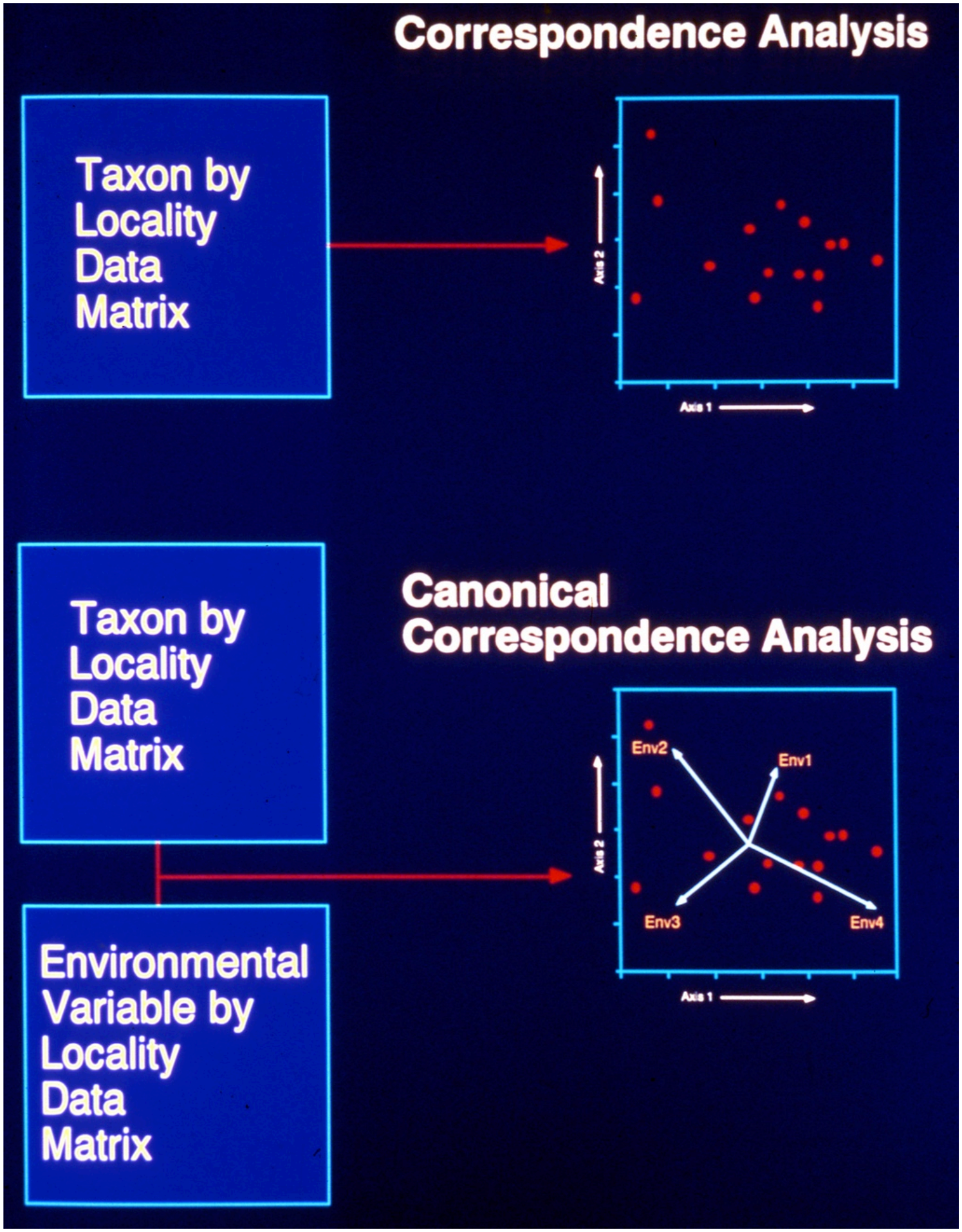
Because environmental trends within the multidimensional cloud of samples are determined subjectively both PCA and CA methods are termed *indirect ordination* methods

Instead of plotting the samples in terms of the attributes X and Y we can plot them relative to the axes of variation. This is the basis of an ordination. Axis 1 represents the axis of greatest variation, axis 2 the next and so on. This simplifies the multidimensional nature of the structure in the most efficient manner. Trends in the pattern along axis 1 (or 2 etc) may be explained by environmental trends. If this is done subjectively the methods are termed indirect.



Canonical Correspondence Analysis (CANOCO) is a *direct ordination* method that positions environmental trends (vectors) on the ordered samples using environmental data. In the case of CLAMP it is climate data.

Canonical Correspondence Analysis (CANOCO) is a direct ordination method that positions environmental trends (vectors) on the ordered samples mathematically using environmental data. In the case of CLAMP it is climate data.



CANOCO is the most popular and powerful method in plant ecology for analyzing complex patterns of plant distributions in terms of environmental influences. This is because it has few constraints in terms of the distribution of variables (it tolerates non-normally distributed data), missing data, and correlations between variables.

CANOCO (ter Braak, 1986) is the most popular and powerful method in plant ecology for analyzing complex patterns of plant distributions in terms of environmental influences. This is because it has few constraints in terms of the distribution of variables (it tolerates non-normally distributed data), missing data, and correlations between variables. CANOCO underpins CLAMP.