ELSEVIER

Journal of Chromatography A, 703 (1995) 3–15

JOURNAL OF CHROMATOGRAPHY A

Sample dimensionality: a predictor of order–disorder in component peak distribution in multidimensional separation

J. Calvin Giddings

Field-Flow Fractionation Research Center, Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA

Abstract

While the use of multiple dimensions in separation systems can create very high peak capacities, the effectiveness of the enhanced peak capacity in resolving large numbers of components depends strongly on whether the distribution of component peaks is ordered or disordered. Peak overlap is common in disordered distributions, even with a very high peak capacity. It is therefore of great importance to understand the origin of peak order/disorder in multidimensional separations and to address the question of whether any control can be exerted over observed levels of order and disorder and thus separation efficacy.

It is postulated here that the underlying difference between ordered and disordered distributions of component peaks in separation systems is related to sample complexity as measured by a newly defined parameter, the *sample dimensionality* s, and by the derivative dimensionality s'. It is concluded that the type and degree of order and disorder is determined by the relationship of s (or s') to the dimensionality n of the separation system employed. Thus for some relatively simple samples (defined as having small s values), increased order and a consequent enhancement of resolution can be realized by increasing n. The resolution enhancement is in addition to the normal gain in resolving power resulting from the increased peak capacity of multidimensional systems. However, for other samples (having even smaller s values), an increase in n provides no additional benefit in enhancing component separability.

1. Introduction

Multidimensional separation techniques constitute a powerful class of methods in which two or more independent separative steps are linked together for separation. For many difficult samples having numerous components, these somewhat complicated techniques provide remarkable gains in resolving power, beyond anything imaginable with one-dimensional (linear) systems such as single chromatographic columns. However, somewhat surprisingly, for other difficult samples whose components are equally numerous and similarly hard to resolve, multidimensional techniques offer no advantages. We conclude that there is some intrinsic property of analytical samples (other than the number m of components) that determines their amenability to multidimensional techniques. It is the thesis of this paper that the key property is a very basic one related to sample variability. We call this property *sample dimensionality*. It is also reasonable to argue that once the sample dimensionality is established, general conclusions can be drawn about important behavioral characteristics of the sample in different multidimensional systems. This relationship suggests that systematic rules can be promulgated in order to better match multidimensional systems to the nature of the sample under consideration.

0021-9673/95/$09.50
SSDI 0021-9673(95)00249-9

The principal advantage of multidimensional separation is that it provides, relative to one-dimensional or linear techniques, a greatly enhanced peak capacity, which represents the maximum number of components theoretically separable in the system. Multidimensional procedures also yield a larger number of parameters for the characterization and identification of components: Whereas for linear systems only one parameter (e.g., the distribution coefficient for chromatographic systems) characterizes separative displacement, in an n-dimensional procedure there are n such parameters associated with the component, which makes its identification much more certain.

While the above advantages are substantial, so are the difficulties of implementing multidimensional methods. Not only are several separation stages necessary, but the effective coupling between stages can be difficult. Consequently, it is important to carefully design the separation system so that it appropriately deals with (i.e., provides an approximate match with) the analytical problem. Both overdesigning and underdesigning can be costly in terms of apparatus complexity, design time, run time, and operator involvement on one hand, and inadequate data and poorly characterized samples on the other. A number of system requirements that relate to the sample are obvious: good selectivity, suitable peak spacing, and adequate peak capacity in every stage along with the substantial independence of successive stages relative to the sample. In this paper we will focus on sample dimensionality as a new sample parameter that strongly influences component resolution in relationship to the dimensionality of the separation system.

Each separation system has its own unique dimensionality n: For linear systems it is one ($n = 1$), for planar beds it is two ($n = 2$), and for coupled column systems it can range from one to three or more. According to the theme of this paper, analytical samples have their own dimensionality, which we label as s. This also can range from one to three or more, but it is commonly higher. Certain generic separation patterns are expected to emerge depending on the relative values of these two dimensionalities. In particular, the relative value of n and s is expected to have a profound effect on the order and disorder of peak spacing, which in turn exerts a strong influence on peak separability. The underlying concepts that give rise to these relationships are explained in the following section.

2. Background concepts

2.1. Order, disorder, and peak capacity

In recent years it has become increasingly clear that the analytical certainty desired from chromatography and related separation methods suffers gravely when the distribution of component peaks is governed by statistical factors [1–6]. For example, a mixture of m components can be completely resolved and characterized in a column of minimum peak capacity, $n_c = m$, providing the component peaks are distributed at uniform intervals (relative to peak width) across the chromatogram. By contrast, if the m components are distributed randomly (as specified generally by a Poisson process), a peak capacity ten times the minimum value m (requiring a hundredfold gain in the number N of theoretical plates) is needed to adequately resolve only 82% of the m components. To resolve 98% of the components, n_c must exceed m by a (usually unattainable) factor of 100, requiring an increase in N of $\sim 100^2$ or 10 000. By way of example, 100 ordered component peaks can be resolved on a column of about 40 000 plates, whereas to resolve only 82 of 100 random component peaks requires $\sim 4 \cdot 10^6$ plates. In general, then, when peak distribution is random rather than ordered, the application of enormously greater effort is inevitably met with substantially inferior results [1].

Despite the importance of the issue of order versus disorder in experimental chromatograms, little work has been done to define the criteria for their development. Davis and Giddings [1] discussed the likely development of disordered chromatograms for *complex samples* (those whose components come from a number of chemical families). The analysis of the ex-

perimental chromatograms of complex samples by these same authors showed the peak spacing to be random, in agreement with theoretical arguments [2]. Guiochon and co-workers have also demonstrated randomness for a variety of systems [3,4], as have Dondi et al. [5]. However, the concept of "complex samples" has not been adequately developed to be quantitatively useful.

Because the concepts underlying the development of order and disorder in component peak distribution have been neglected, it follows that no means are known to gain control over disorder. While it is unlikely that significant gains in resolving power can be realized by controlling disorder for *highly complex samples*, advances in this area for *less complex samples* are imaginable and would be quite useful. The gain or loss of such control is found here to depend on the dimensionality n of the separation system.

Our approach to this problem starts with the concept of *sample dimensionality*, s, and the derivative dimensionalities, s' and s''. Parameter s is proposed as a measure of sample complexity. Most importantly in the present context, arguments are developed to suggest that s (or s') is a predictor of component peak disorder in chromatograms. However, disorder is not expected to depend solely on the sample parameter s, but rather on the relationship of s to the system parameter n, the system dimensionality. Values of n are subject to some control with the flexible use of multidimensional separation methods. This advances the possibility that variations in n can be used in certain cases to limit disorder in the distribution of component peaks and thus to greatly improve the quality of analytical results in these cases.

2.2. System dimensionality

Multidimensional separation techniques are those in which different separation steps or stages, based on different mechanisms, are linked according to certain criteria (see antecedent article [7]; also Ref. [8]). The number of different stages can be defined as the dimensionality, n. For planar beds, $n = 2$ (i.e., the system is two dimensional or 2D) providing different separation mechanisms are employed along the two axes [9]. (If the planar separation follows solvent extraction, another stage and another dimension is added to the system as a whole, making $n = 3$.) For coupled column systems, the dimensionality can be two, three, or higher depending on the number of successive stages linked together [7]. (The essential equivalence of a 2D separation on a plane and a 2D separation in a coupled column system is outlined in the antecedent article [7].)

Multidimensional techniques have been recognized as providing a greatly enhanced peak capacity (n_c) relative to linear systems, thus substantially reducing *peak saturation* (m/n_c) and the resulting statistical overlap of component peaks from complex samples [7–10]. In effect, multidimensional systems provide more "space" than 1D systems, allowing component peaks to spread out across additional coordinates, thus reducing peak overlap. Fig. 1 shows component peaks distributed (rather randomly) over a 2D plane, which has sufficient space to minimize the overlap of individual peaks (represented by spots in the figure). When all these components are compressed onto a single axis (as would be the

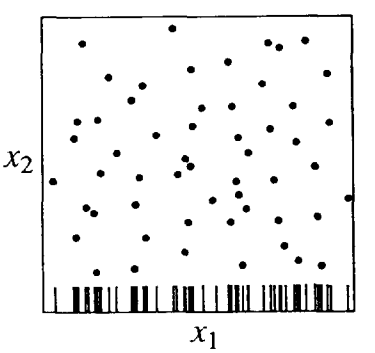

Fig. 1. Component peaks or spots distributed randomly over a two-dimensional plane. The 2D plane has more "space" and thus more peak capacity than either of its two axes. Thus components are able to spread out and peak overlap is reduced in the 2D space compared to that in a 1D separation. The crowding together of peaks in a single dimension (representing a 1D separation) is illustrated in the figure by projecting the peak positions in the plane onto the single axis x_1. In this projection onto x_1, each peak is represented by a line; the crowding of lines along axis x_1 leads to excessive peak overlap.

case for a 1D separation), the peaks (whose centers are now represented by lines on the x_1 axis) become very crowded and will seriously overlap one another due to finite band broadening. In general, a reduction in peak crowding (saturation) and thus peak overlap will be realized whenever system dimensionality is increased, independently of whether the gain in n is achieved by using a planar system or a coupled column arrangement.

We develop evidence here to suggest that component separation for some samples may be greatly improved, *independently of reduced saturation*, with increasing system dimensionality. The improvement is expected solely as a consequence of increasing order in the peak distribution. Conversely, for less complex samples, no gains in resolving power for either of the two reasons are expected with increasing system dimensionality.

2.3. Sample dimensionality

The critical sample parameter that characterizes a sample's varied response to the system dimensionality n is the sample dimensionality s. The parameter s is defined here as the number of independent variables that must be specified to identify the components of the sample. It is assumed (as part of this definition) that the properties of the components, including chromatographic retention parameters, vary in some systematic way with the s variables. (A systematic dependence implies that a definitive trend exists, whether or not the theoretical basis of the trend is understood.) The variables under consideration can generally be taken as structural factors that collectively yield molecular identity.

A few simple examples will illustrate the nature of s. If we know that our sample consists entirely of saturated straight-chain fatty acids, then the components of the mixture can be fully specified in terms of one variable, which can be chosen as either carbon number or molecular mass. (This variable is referred to later as a sample parameter.) This one-variable specification serves to define a one-dimensional (1D) sample. If now we choose a fatty acid sample of variable carbon number in which the straight chain has one substituent group or double bond, the sample acquires a second dimension: the position of the group or double bond.

Going a step further, a sample in which the fatty acid molecules may have 0, 1, or 2 substituted groups of a given kind in different places is a four-dimensional sample. One coordinate is carbon number, another specifies the number of groups (0, 1, or 2), a third may be chosen to give the position of the group closest to the carboxyl end, and a fourth to locate the most distant group. (The dimensionality is reduced to three if the variables describing group position are assigned the value "zero" for the absence of a group. In this case specification of the number of groups is redundant. However, this lower dimensionality is not consistent with the requirement for systematic behavior as imposed above.)

For linear polymers (e.g., polystyrene) where components are specified by one variable, namely chain length, $s = 1$. For a copolymer of A and B groups arranged in the form A_{n_1}–B_{n_2} the dimensionality is two, corresponding to the two variables n_1 and n_2.

Returning to the one-dimensional sample of saturated unsubstituted fatty acids, as long as the separative displacement (measured by retention time, electrophoretic mobility, and so on) varies systematically with carbon number, giving an ordered distribution of component peaks, the sample can be fully separated in a 1D system of adequate peak capacity. Importantly, there is little to be gained in employing a two-dimensional (2D) or higher dimensional system to deal with this 1D sample, despite their larger peak capacities. If we choose a 2D approach in which each of the two separative displacements of the 1D sample depends upon carbon number in a systematic way, then the two displacements will be strongly correlated and there will be little if any increase in separation power [9]. When strong correlation exists between displacements, component peaks tend to fall in a narrow band in the 2D space (see Fig. 2). Correlation therefore prevents the utilization of the expanded space available in the 2D system [9]. In effect, there is no significant gain in peak capacity over that available in a 1D system and the separation

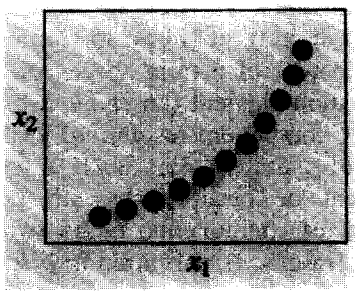

Fig. 2. Component displacements along the two axes of a 2D separation system are strongly correlated for a 1D sample. Consequently, components fail to expand into the full 2D space and as a result the large peak capacity available in 2D systems is not utilized by 1D samples.

shows little (if any) improvement over that achievable along a single separation axis (equivalent to a 1D system).

Once we choose a 1D separation system, it is important, of course, to match the resolving power insofar as possible to the analytical demands. For the fatty acid one will normally wish to separate every homolog (or its derivative); if the number m of homologs is not large, separation can usually be achieved with relatively few theoretical plates. In such a case the peak capacity n_c need not be much larger than m. However, for 1D samples with large m, such as linear polymer samples, all the homologs *cannot* (and normally need not) be resolved in a 1D chromatographic system. Again, a 2D system does not significantly improve the situation.

For the monosubstituted fatty acids, $s = 2$, any given separative displacement will normally exhibit a systematic (although not equal) dependence on both variables: chain length and group position. In this case two dimensions (two separation mechanisms) are justified for the separation system provided that one can achieve independent migration (combined with reasonably high selectivity) along the two separation coordinates. This simply means that the two different sample variables influence the two separative displacements dissimilarly or at least that the two sample factors are weighted differently in their effect upon the two displacements. Ideally but infrequently, one factor (carbon number) controls one displacement and the other factor (group position) controls the second displacement.

The same considerations apply to polymer analysis. To characterize linear polystyrene or polyethylene we need only determine the chain length or molecular mass distribution. This is best approached using a one-dimensional separation system such as size exclusion chromatography or thermal field-flow fractionation. While we must be careful to generate enough resolving power to satisfy our analytical needs, nothing is gained by utilizing a second dimension.

If we must fully characterize the copolymer sample A_{n_1}–B_{n_2} (generally necessary if the relevant properties depend simultaneously on chain length and on the relative amount of the two constituent polymers), a 1D system, no matter what its resolving power, will no longer suffice. Such a problem requires two dimensions. Random copolymers, by contrast, are more complex and have higher dimensionality; they cannot be fully characterized in a 2D system. Nonetheless, for reasons to be discussed later, useful information is still forthcoming on such polymers from a 2D system under the right circumstances.

If we attempt to separate the components of a 2D material in a linear (1D) system, the resulting band pattern will generally be disordered; by exhibiting a mixed dependency on both sample properties it fails to display a systematic distribution based on either of the sample properties. Likewise, a sample material requiring three or more dimensions for its characterization will not ordinarily develop a systematic pattern in a 2D system. Generally, if the dimensionality of the sample exceeds that of the system, components of the sample will not be systematically resolved in the system. The resulting retention (separation) pattern is disordered and may be termed "chaotic."

3. Theory

We consider an n-dimensional separation system in which the final position of a component zone (or its center of gravity) is specified by the individual distances along the coordinate axes:

$x_1, x_2 \ldots x_n$. These distances may be expressed in terms of length of migration path, retention time, or a column index number in coupled column system [7]. The sample is fully characterized by the s sample parameters: $p_1, p_2, \ldots p_s$. This means that the molecular identity of a component is fully established by specifying the s values of the p's. Generally the p space can be occupied only at discrete intervals because of the discreteness of molecular structure.

Since the set of s values of p uniquely establishes molecular identity, it also specifies the position X_i along each of the separative axes x_i and thus the final position of the component in the system. In mathematical form

$$X_i \equiv X_i(p_1, p_2, \ldots p_s) \quad (1)$$

The relationship expressed by Eq. (1) will normally be a complicated one, particularly in chromatography where the intermolecular interactions controlling the displacement distances $X_1, X_2, \ldots$ are not rigorously calculable. Many efforts have been made to establish the nature of these relationships and various approximate equations have been developed (e.g., Ref. [11]). However, our arguments do not depend on the specific nature of the relationships, but only on the assumption that such relationships exist and that they are systematic.

We observe that there are n equations (one for each system dimension) of the type shown in Eq. (1). Given a specific component and its associated p values, all n of the coordinate positions are fixed by the n equations. However, given an observed set of coordinate values, $X_1, X_2, \ldots X_n$, the p values (and thus the component identity) may or may not be established by the n expressions of Eq. (1). If $n \geq s$, the s values of p will be fixed. If $n < s$, then the component properties cannot be obtained because there are fewer equations (n) than unknown parameters (the s values of p) that must be obtained from the equations.

If the properties represented by the p's were continuous variables (not discretized by molecularity), then for $n < s$ a given coordinate position would correspond to any of a number of different combinations of p values. Because p values are in fact discrete, it is best to consider a small volume element of system space corresponding to a resolution element (an element barely large enough for component resolution) located in the intervals $\delta x_1, \delta x_2, \ldots \delta x_n$. Such a volume element might be occupied by any one of an assortment of components, or by two or more unrelated (insofar as permitted in p space) components (an example of statistical peak overlap), depending on the phase of different arrays of discrete points in the vicinity. The possibility of finding unrelated components in the same resolution element would make components identification based on coordinate position untenable and would be conducive to the random distribution of peaks in the system.

Some of the implications of the n relationships expressed by Eq. (1) are best understood by examining a single system coordinate, x_i. Given a systematic dependence of the displacement along this coordinate on the p_j's, then a one-dimensional (1D) sample characterized by only one variable p_1 will have component peaks distributed along x_i according to an ordered pattern. This does not mean that the spacing between successive peaks is equal, but only that the changes are systematic.

Normally, the incremental changes in X_i due to incremental changes in $p_1, p_2, \ldots p_s$ will bear no rational relationship to one another; that is, $\Delta X_i/\Delta p_j$ will be unrelated to $\Delta X_i/\Delta p_k$, and so on. For example, for monounsaturated fatty acids the separation caused by a unit increase in carbon number will be unrelated to that caused by a single shift in the position of the double bond. Accordingly, a sample subject to both variations will no longer exhibit any consistent order in the spacing, ΔX_i, between successive peaks [1,12]. Thus the peak distribution along a single dimension x will often appear to be disordered (see Fig. 3). Any residue of order will be further degraded by variations in additional sample dimensions, that is, by incremental changes along another p-axis. Thus the distribution of components along a given separation coordinate will be most highly ordered for $s = 1$ and increasingly disordered as s increases.

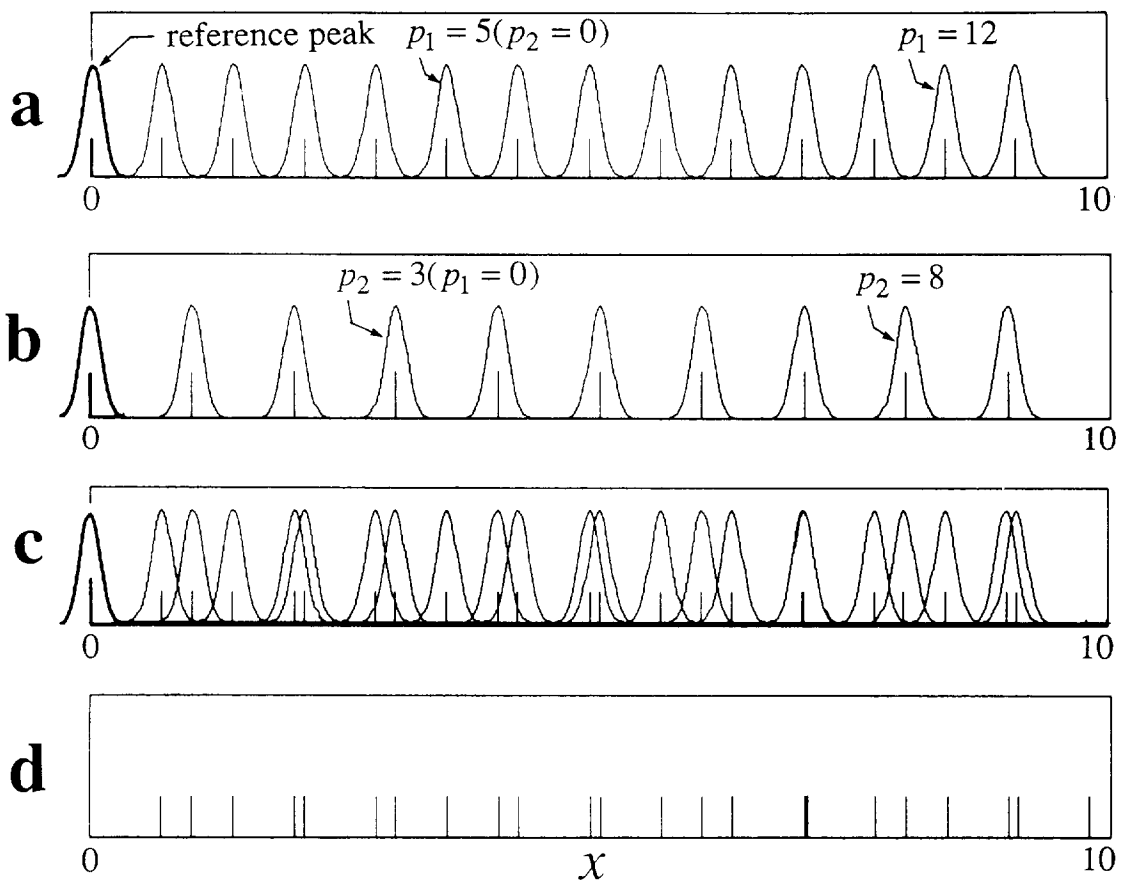

Fig. 3. Hypothetical peak sequences resulting from incremental changes in two different sample parameters (such as carbon number and substituent position) of sets of related molecules. (a) Systematic (and, in this hypothetical case, uniform) sequence of peaks resulting from unit variations in a single sample parameter (structural factor) p_1. (b) Sequence resulting from unit incremental increases in a second sample parameter p_2. (Note that the spacing between peaks is uncorrelated between a and b.) (c) Superposition of two uncorrelated peak sequences shown in a and b illustrates how easily order is destroyed when more than one sample parameter is subject to variation. (More complexity and disorder would be introduced if p_1 and p_2 varied simultaneously.) (d) Sequence of lines ("line chromatogram") positioned at the center of gravity of the various peaks in c clearly illustrates the considerable disorder in spacing between peaks.

In a similar vein, the systematic relationships expressed by Eq. (1) assure us that a two-dimensional sample will exhibit an ordered distribution in a 2D separation system. A three-dimensional sample, however, will generally exhibit disorder in 2D system space; an ordered distribution in most cases requires three dimensions of displacement.

The above principles can be further illustrated by computer calculations of the distribution of points (corresponding to imaginary components) in a plane and their subsequent projection onto one of the edges (axes) of the plane to represent a single dimension of displacement (see Fig. 4). The sample is assumed to have dimensionality $s = 2$. To imitate the complexities of real systems, the governing expressions are arbitrarily composed and the coefficients (or their multiples) unrelated. (Specifically, one coefficient utilizes an irrational number, in the present case π, to avoid the exact coincidence of points along a given axis for different sets of p values.) The

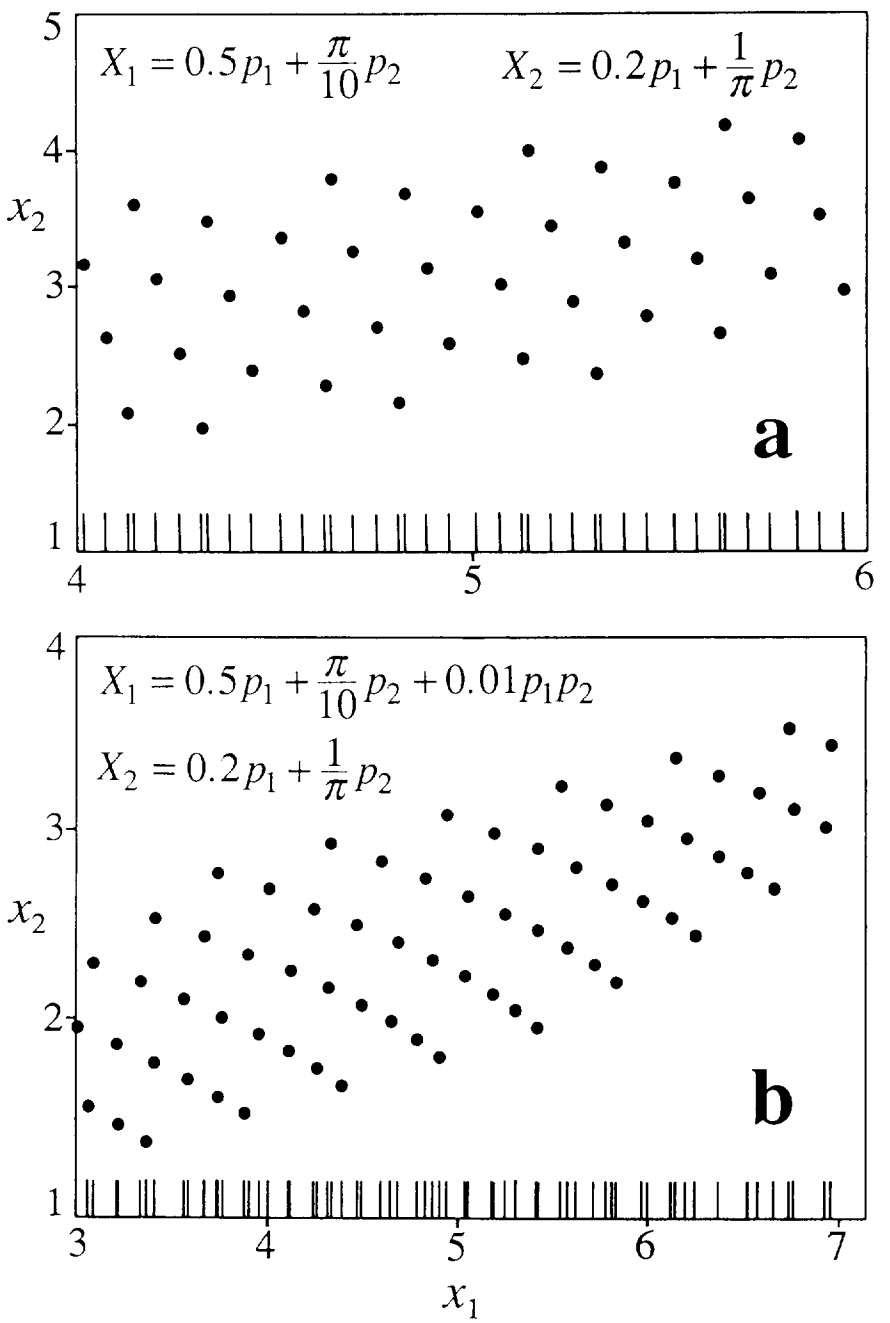

Fig. 4. Systematic distribution of points (corresponding to peak centers) on a plane and of same points projected onto the x_1 axis of the plane. The points occur at the positions $x_1 = X_1$ and $x_2 = X_2$ found when index numbers (sample parameters) p_1 and p_2 (e.g., carbon number) are integers. (a) Distribution of points given by Eqs. (2) and (3), shown at the top of the figure. (b) Distribution given by Eqs. (3) and (4), also shown at the top of the figure.

assumed expressions for component displacements X_1 and X_2 along axes x_1 and x_2 in Fig. 4a are of a simple additive form

$$X_1 = 0.5p_1 + (\pi/10)p_2 \quad (2)$$

$$X_2 = 0.2p_1 + (1/\pi)p_2 \quad (3)$$

where the sample parameters p_1 and p_2 assume integer values only. We observe that the plotted points in Fig. 4a are distributed as an ordered periodic array in 2D space. However, when these points are projected along the x_1 axis (corresponding to Eq. 2) much of the order breaks down.

For real systems governed by more complex equations and assorted fluctuations, very little residue of order is expected in a 1D (one axis) representation of a 2D ($s = 2$) sample. Thus if we add to Eq. (2) a small correlation term such that

$$X_1 = 0.5p_1 + (\pi/10)p_2 + 0.01p_1p_2 \quad (4)$$

the results shown in Fig. 4b emerge: The 2D distribution is still fully ordered whereas the 1D projection has become largely disordered.

The linear expressions of Eqs. (2) and (3) are obviously oversimplified. These expressions have been replaced by exponential equations more representative of the dependence of chromatographic migration on $\exp(-\Delta\mu^0/RT)$, where the chemical potential change $\Delta\mu^0$ is a consequence of partitioning. Plots based on this exponential form (Fig. 5) also produce an organized pattern of points in 2D system space. Again, the projection of such points on one axis (not shown) has few vestiges of order left.

While ideal separative displacements will be subject to the systematic relationships of Eq. (1), many real displacements will be subject to small secondary variations superimposed on the background order. For example, the distribution coefficient underlying the chromatographic migration of components having hydrocarbon chains of various lengths can be approximated by using the concept of additive free energies [13]

$$\Delta\mu^0 = C + p_c\left(\Delta\mu^0_{CH_2}\right) \quad (5)$$

where $\Delta\mu^0$ is the standard chemical potential associated with phase transfer, C is a constant

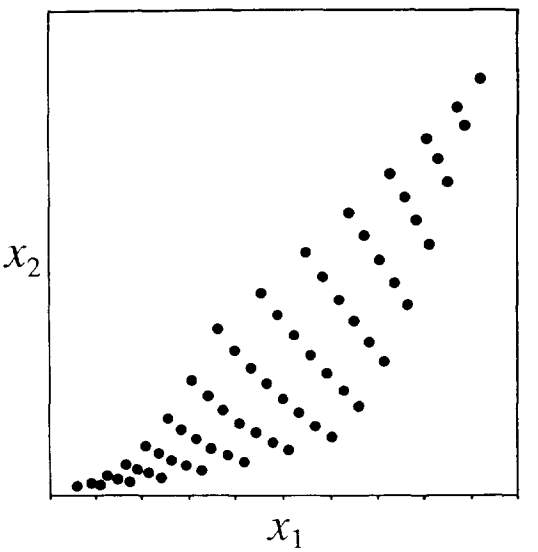

Fig. 5. Organized pattern of points (peak locations) in 2D space resulting from unit variations in sample parameters p_1 and p_2. The displacement positions X_1 and X_2 along axes x_1 and x_2 are governed by exponential relationships much like those controlling retention ratios or R_F values in chromatography. The assumed relationships are $X_1 = 1/[1 + 0.1\ \exp(\theta_1)]$ and $X_2 = 1/[1 + 0.1\ \exp(\theta_2)]$ where $\theta_1 = (-\Delta\mu_1^0/RT) = 0.5p_1 + (\pi/10)p_2$ and $\theta_2 = (-\Delta\mu_2^0/RT) = 0.2p_1 + (1/\pi)p_2$.

characteristic of the chemical family, and $\Delta\mu^0_{CH_2}$ is the $\Delta\mu^0$ increment for a single CH_2 unit in the hydrocarbon chain. While Eq. (5) leads to a systematic relationship between chromatographic retention and chain length parameter p_c, actual measurements will reveal small second order departures, both systematic and erratic, relative to this equation. Second order systematic perturbations can presumably be described by more sophisticated equations; erratic disturbances, however, lead to small uncertainties in separative displacements. This may constitute a fertile area for the application of fuzzy theory [14,15].

4. Sample dimensionality: further considerations

The sample dimensionality s, as noted, can be considered as the number of independent variables requiring specification in order to identify the components of a sample. It is not clear that a

unique value of s exists for truly complex samples (or for what one perceives to be the full range of variability of the sample), but logical designations emerge (based largely on the requirement that displacements be systematically dependent on the p's as expressed by Eq. 1) providing the sample is relatively simple. We have already outlined a means for obtaining the dimensionality of simple samples like the substituted fatty acids. However, some ambiguity may exist as illustrated by the fact that the finite number of discrete points (corresponding to discrete molecules) in the s-dimensional sample space of a complex sample can be numbered according to some arbitrary sequence and individual components can then be specified in terms of one variable, namely by their position within this sequence. This ambiguity is largely resolved, as noted above, by the condition that observable sample properties (and thus displacements) must be systematically dependent on the chosen variable set.

A somewhat different definition for sample dimensionality that incorporates the foregoing condition equates s to the number of distinguishable sample variables or parameters systematically influencing sample properties, such as distribution coefficients and electrophoretic mobilities, that in turn control separative displacements. Several difficulties can be expected to arise in applying either definition, particularly to high dimensional (large s) samples. For one thing, the full variability of the sample must be known, which is presently unlikely for complex samples (e.g., plant extracts). On the other hand, the dimensionality of such samples is probably in the hundreds or thousands (perhaps much higher), a value so high that its precise determination has no practical implications because there is no hope of achieving separative order.

One complication in unequivocally predicting order based on the relationship of the dimensionalities s and n is that the major properties of a sample may depend only very weakly on a given sample variable (e.g., p_j). In such a case, $\Delta X_i/\Delta p_j$ will be relatively small for such a weakly expressed sample variable. The pattern resulting from a sufficiently weak dependence will then be washed out by resolution limits or by erratic second order departures (as described above) from the assumed systematic dependence of displacement on the major variables. For example, for mono-unsaturated fatty acids, chromatographic retention will normally be strongly dependent on carbon number and only weakly dependent on the double bond position. In this case a 1D chromatogram will display apparent order despite $s > n$. When the sample is subjected to separation in two dimensions, it will not spread out in the 2D space available and thus will not enjoy the advantageous peak capacity afforded by 2D systems. The sample will, instead, tend to accumulate in a narrow band as shown in Fig. 6. Thus the behavior of a sample with a weakly expressed variable is much like that of a sample of lower dimensionality. The lower dimensionality may be thought of as an *apparent dimensionality* s', equal to the number of variables expressed strongly enough to produce suitable resolution. Thus a sample may assume an ordered (or pseudo-ordered) arrangement in a system for which $s' = n$ despite the fact that $s > n$. The problem is that an observed "point" (within resolution limits) in such an ordered array may correspond to a number of components differing in the value of the weakly expressed variable(s). Consequently a point

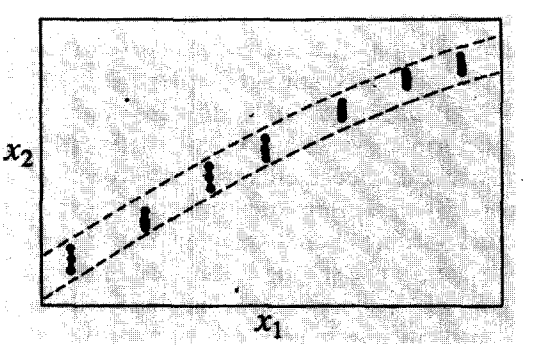

Fig. 6. The components of a two-dimensional sample tend to fall in a narrow band (shown as the region between the broken lines) in a 2D system when one of the sample variables or parameters is weakly expressed. The weak expression of the one sample variable will greatly diminish the effective peak capacity (see discussion in text).

would not specify a single component, but only a family of components in which the members differ in the weakly expressed variable or parameter. Thus a peak observed in a normal 1D chromatogram of monounsaturated fatty acids may represent a given carbon number but could include components with different double bond positions. The resolution of the latter components would require using a separative mechanism sensitive to double bond position and, in most cases (see exception in next paragraph), a separation system of dimensionality $n = 2$ in place of $n = 1$.

A different pattern may emerge for sample variables of intermediate expression, with $\Delta X_i / \Delta p_j$ large enough to produce resolution but small enough to localize components separated according to property p_j. Such components will often appear as satellite peaks next to primary peaks separated from one another by a stronger mechanism. The satellite peak structure can be repeated for successive primary peaks. Thus there is a secondary order superimposed on a primary pattern of repeating peaks. In such circumstances order can emerge for samples with dimensionality greater (by one unit) than n. In theory a tertiary structure could arise for samples with $s = n + 2$, but extraordinary resolution would be needed to make it observable. Secondary peak structures, however, are often produced by high-efficiency GC where the high resolution accompanying a change in carbon number leaves gaps to accommodate a secondary pattern caused by changes in group position, isomer expression, etc.

Still another facet of sample dimensionality is that an analyst may not require the resolution of components along all s sample coordinates. In the foregoing example it may be found, depending on the goals of analysis, that only differences in carbon number, not in group or double bond position, need be resolved. Thus a *required dimensionality* s'' can be defined as the number of sample variables that must be determined for purposes of analysis. In an ideal system the unneeded sample variables will also be weakly expressed variables, giving $s'' = s'$. In this case, and generally only in this case, the sample can be systematically characterized in a system of dimensionality $n = s''$.

An example is found in copolymer analysis. The general structure A_{n_1}–B_{n_2} represents a two-dimensional sample, as noted earlier, and can be systematically distributed in a two-dimensional system. However, the random copolymer A-B-B-B-A-A-B-A-A . . . has a dimensionality of N, the maximum number of chain elements. In many cases one only needs N and the A/B ratio for the purposes of analysis; the required dimensionality, s'', is therefore two. Size exclusion chromatography, for all essential purposes, provides N. Adsorption or precipitation chromatography gives the A/B ratio providing Eq. (5) is applicable in the expanded form

$$\Delta\mu^{0} = C + p_1 \Delta\mu_A^0 + p_2 \Delta\mu_\beta^0 \qquad (6)$$

This equation implies that the sequence of A and B groups in the copolymer is unimportant. If true, sample dimensions expressing group sequence will be weakly (in this case negligibly) expressed. If this is the case, the desired 2D pseudo-ordered distribution ($s'' = 2$) of random copolymers can be obtained by combining the above two chromatographic mechanisms. A number of authors have pursued copolymer characterization using such a combination of mechanisms [16–20].

4. Conclusions

Sample dimensionality s (or apparent dimensionality s'), in conjunction with system dimensionality n, appears to provide some predictive capabilities with regard to ordered versus disordered component distribution following multidimensional separation. It also indicates whether the full peak capacity of multidimensional systems can be substantially exploited. In short, when $s > n$ (or more generally when $s' > n$) the component peak distribution is predicted to be largely disordered, thus hindering separation. When $s < n$, the component distribution is ordered but the greatly enhanced peak capacity of the multidimensional system is not utilized.

When the dimensionalities are equal, $s = n$, the best possibility exists to fully exploit the power of multidimensional separation without the severe disadvantages of disordered peak distribution. These characteristics are summarized in Table 1.

It is probable that some of these concepts extend to other analytical methods, such as NMR and optical spectroscopy. It is unlikely, for instance, that a 2D system will provide any significant advantages over a 1D system with a 1D sample, no matter what combination of analytical techniques is employed.

Since system dimensionality n is subject to experimental control, it would appear that one could always design an analytical system properly balanced with respect to the sample. Unfortunately, multidimensional systems tend to be complex in design and operation; this complexity increases rapidly with dimensionality n. Huber [21] has fully utilized a coupled column system with $n = 3$ with great success (over 6000 peaks separated), but the difficulty of the project suggests that higher n values would provide extraordinary challenges except in probing local regions of system space for specific components (see Ref. [7]). Unfortunately, most s values will be larger; for complex samples they are expected to be in the hundreds or beyond. In these cases there is no reasonable hope for designing a system with $n = s$. Thus our realization of the advantages of the $n = s$ condition appears to be limited to samples with components simply related to one another.

Table 1
Order and disorder in separation patterns emerging from different relationships between sample dimensionality s, apparent sample dimensionality s', and system dimensionality n

s, s', n relationship	Separation pattern
$s > s' > n$	Disordered[a]
$s = s' > n$	Disordered[a]
$s > s' = n$	Pseudo-ordered
$s = s' = n$	Ordered
$s > s' < n$	Pseudo-ordered
$s = s' < n$	Ordered

[a] May be pseudo-ordered or ordered with secondary pattern when $s' = n + 1$.

While the ideal conditions in which the two dimensionalities balance (specified by $n = s$) may not be practically achievable for most complex samples, the present analysis offers additional insight into the origins and conditions underlying disordered peak distributions. Since chaotic distributions place extraordinary demands on separation systems, it is important to have criteria (namely $s > n$ or $s' > n$) for predicting their existence.

It must be emphasized that the above analysis is preliminary in nature. Many additional issues must be addressed to evaluate the scope, significance, and limitations of this approach.

List of symbols

C	Constant in Eq. (5) or (6)
m	Number of sample components
m/n_c	Peak saturation
n	System dimensionality
n_c	Peak capacity
N	Number of theoretical plates
p	Sample parameter
p_c	Chain length parameter
R	Gas constant
s	Sample dimensionality
s'	Apparent sample dimensionality
s''	Required sample dimensionality
T	Temperature
x_i	Displacement axis
X	Component displacement position along axis x_i
$\Delta\mu^0$	Chemical potential change of phase transfer

Acknowledgements

The author acknowledges Michael C. Giddings for generating Fig. 3 and the 2D plots of Figs. 4 and 5. This work was supported by Public Health Service Grant GM10851-37 from the National Institutes of Health. The basic concepts developed here were first presented at the 26th

Eastern Analytical Symposium in New York, September 13–18, 1987 (see report by V. Berry, *LC · GC*, Vol. 6, No. 1, p. 65).

References

[1] J.M. Davis and J.C. Giddings, *Anal. Chem.*, 55 (1983) 418.

[2] J.M. Davis and J.C. Giddings, *Anal. Chem.*, 57 (1985) 2178.

[3] D.P. Herman, M.-F. Gonnord and G. Guiochon, *Anal. Chem.*, 56 (1984) 995.

[4] M. Martin, D.P. Herman and G. Guiochon, *Anal. Chem.*, 58 (1986) 2200.

[5] F. Dondi, Y.D. Kahle, G. Lodi, M. Rernelli, P. Reschiglian and C. Bighi, *Anal. Chim. Acta*, 191 (1986) 261.

[6] J.M. Davis, in P.R. Brown and E. Grushka (Editors), *Advances in Chromatography*, Vol. 34, Marcel Dekker, New York, 1994, pp. 109–176.

[7] J.C. Giddings, in H.J. Cortes (Editor), *Multidimensional Chromatography: Techniques and Applications*, Marcel Dekker, New York, 1990, Ch. 1.

[8] J.C. Giddings, *J. High Resolut. Chromatogr. Commun.*, 10 (1987) 319.

[9] J.C. Giddings, *Anal. Chem.*, 56 (1984) 1258A.

[10] G. Guiochon, M.-F. Gonnord and M. Zakaria, *Chromatographia*, 17 (1983) 121.

[11] R. Kaliszan, *Anal. Chem.*, 64 (1992) 619A.

[12] A. Felinger, L. Pasti and F. Dondi, *Anal. Chem.*, 62 (1990) 1846.

[13] A.J.P. Martin, *Ann. Rev. Biochem.*, 19 (1950) 517.

[14] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Application*, Academic Press, New York, 1980.

[15] M. Otto and H. Bandemer, *Anal. Chim. Acta*, 191 (1986) 193.

[16] B.G. Belenkii and E.S. Gankina, *J. Chromatogr.*, 141 (1977) 13.

[17] H. Inagaki and Y. Tanaka, *Pure Appl. Chem.*, 54 (1982) 309.

[18] G. Glöckner, J.H.M. van den Berg, N.L.J. Meijerink, T.G. Scholte and R. Koningsveld, *Macromolecules*, 17 (1984) 962.

[19] G. Glöckner, J.H.M. van den Berg, N.L.J. Meijerink, T.G. Scholte and R. Koningsveld, *J. Chromatogr.*, 317 (1984) 615.

[20] S. Mori, *Anal. Chem.*, 60 (1988) 1125.

[21] J.F.K. Huber, private communication, 1987.