ELSEVIER

Journal of Chromatography A, 734 (1996) 7–47

JOURNAL OF CHROMATOGRAPHY A

Review

Non-linear waves in chromatography II. Wave interference and coherence in multicomponent systems

Friedrich G. Helfferich[a,*], Roger D. Whitley[b]

[a] *Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, USA*

[b] *Adsorption Technology Center, Air Products and Chemicals, Inc., Allentown, PA 18195, USA*

Abstract

This second instalment on non-linear waves examines in detail the effects caused by interactions between different solutes in multicomponent systems. Emphasis is on concepts and qualitative insight, and mathematics is kept at a minimum. To avoid unnecessary complications, the assumption of ideal theory are taken for granted. In the discussion, no restrictions are imposed on the nature of equilibria between the moving and sorbent phases and on the number of components. However, all examples are of two- and three-component systems with competitive sorption equilibria.

Keywords: Reviews; Non-linear chromatography; Wave theory; Multicomponent systems; Coherence theory; Frontal analysis; Elution development; Displacement chromatography; Preparative chromatography; Riemann problems; Shock layer

Contents

*Corresponding author.

0021-9673/96/$32.00
SSDI 0021-9673(95)01067-X

1. Introduction

The most general feature of multicomponent non-linear wave phenomena, in chromatography as elsewhere, is that a single variation of conditions at the entry to the system is propagated in the form of a whole set of waves that travel at different velocities. Elution development and frontal analysis can be viewed as manifestations of this principle. Waves from successive entry variations may interfere with one another. The concept of coherence is introduced for an examination of such phenomena: An arbitrary composition variation, whether imposed at the column entry or arising from wave interference within the column, is in general noncoherent and is resolved into a set of coherent waves that separate from one another as they travel. This insight makes it easier to understand column responses under complex operating conditions and provides a key to efficient mathematics. The use of the tools of coherence theory for the prediction of column responses is discussed and illustrated with examples.

Part I of this series has dealt with fundamental properties of non-linear waves in single-component chromatography [1]. While suited to introduce the concepts of self-sharpening and nonsharpening waves and of shocks and shock layers, this topic is primarily of academic interest. In practice, chromatography as a separation technique is a multicomponent problem almost by definition. The non-linearity in chromatography results from interactions between solute molecules, e.g., from their competition for adsorption sites. It would be naïve to assume that molecules interact only with their own kind, not with those of other solutes. Since a single-component isotherm accounts only for interactions between molecules of the same solute, not with those of others, it does not adequately describe sorption of the respective solute from mixtures in the non-linear range: The non-linear chromatographic response in a multicomponent system cannot be compiled by superposition of non-linear responses calculated for single components.

The present, second part of the series examines specifically the phenomena that arise in chromatography from the presence of several solutes that affect one another's behavior. In terms of wave theory, the key elements are the generation of entire sets of waves by one single composition variation at the column inlet, the interference of waves that originated at different times or locations, and the development of 'coherent' patterns from arbitrary starting variations and from such interferences [2]. As in the preceding part, the discussion will center on concepts, insight, and use of the tools which theory provides. Mathematics will be kept at a minimum. An overview of mathematics, not needed for conceptual understanding, will be given in Appendix A. Examples of practical cases including frontal analysis, elution from a uniformly saturated column, elution of a pulse under overload conditions, displacement development, and elution with a buffer will illustrate applications of concepts and theory. The last of these examples, on elution with a buffer, demonstrates that it is often possible to obtain valuable information from wave theory without recourse to more than a minimum of simple algebra.

So as to steer clear of complications that would distract from the characteristic effects caused by interactions of different solutes, the assumptions of ideal chromatography will be taken for granted. Specifically, these are: (1) local equilibrium between moving and sorbent phases; (2) ideal plug flow; (3) mass transfer in axial direction by convection only; (4) axially uniform volumetric flow-rate of bulk moving phase; (5) isobaric behavior; (6) isothermal behavior; and (7) absence of chemical reactions that transform solutes (except adsorption or chemisorption). The effects of nonidealities are largely the same in multicomponent as in single-component systems and were discussed in Part I (one minor complication arising specifically in multicomponent systems is outlined an appendix, see Appendix B. In essence and with few exceptions, most notably at extremely low mass-transfer rates and in nonisothermal systems, the only effect of nonidealities is to make all waves less sharp than ideal theory predicts. Nonidealities are very important because they affect the sharpness of separation, but they do not alter the general features of the response pattern, as do interactions between different solutes. Thus, nonidealities in non-linear chromatography can usually be accounted for by mere corrections to the results of ideal theory, and fundamental insight can

be gained from the latter alone. Moreover, in many instances, corrections for nonidealities can be applied one wave at a time, and so can be restricted in practice to only the particular wave or waves that affect the result of interest.

Some subtle, complicated points that contribute little or nothing to general qualitative understanding will be stated in square brackets. They may be skipped without loss in continuity. The glossary of symbols and terms gives explanations of symbols and definitions of terms, some of which (such as wave velocities or Riemann problems) might be unfamiliar to the practical chromatographer.

Most of the three-dimensional diagrams showing successive composition profiles were calculated with the VERSE-LC computer program [3,4]. They are for systems with fast but finite mass-transfer rates and accordingly show a slight nonideal effect owing to deviation from local equilibrium. This is because execution times would be prohibitively long for mass-transfer rates high enough to ensure a closer approach to local equilibrium. We have made no attempt to eliminate the slight nonideality, as could easily have been done. The reason is that the concentration profiles are easier to tell apart if the shocks are not ideal discontinuities, and that the slightly nonideal behavior is closer to reality.

The VERSE program is based on integration of the differential mass-conservation equations and does not employ wave theory. However, except for the mentioned slight nonideality, there is good agreement between its results and the composition–route and distance–time diagrams constructed with wave theory. The reader may view this as a demonstration of reliability of VERSE or of soundness of wave theory, depending on which of them he trusts less.

As in Part I, the discussion is in terms of sorption as the underlying equilibrium mechanism. It is equally valid for other equilibrium mechanisms of distribution between the moving and stationary phases, and is mathematically largely analogous for stoichiometric ion exchange [5].

Also as in Part I, the great majority of the findings have been published previously, most of them by other authors. However, no comprehensive review of prior work is presented. Rather, only a small fraction of it is quoted, selected for best illustration or proof of relevant points. We apologize to those who qualify for credit but have not received it.

2. Response signals and variance

To start out from solid ground, let us begin with two phenomena with which the practical chromatographer is thoroughly familiar: standard analytical chromatography (often called elution development) [6–9] and frontal analysis [10–14]. In the former, a very small sample consisting of a mixture of solutes, at concentrations so low that the linear range of the sorption isotherms is not exceeded, is injected ahead of the column into a continuous stream of solvent or carrier gas; the mixture separates into single-component 'peaks' that travel at different velocities, emerging in the sequence of increasing affinity for the sorbent (see Fig. 1). In frontal analysis, a fairly concentrated mixture of solutes is injected continuously and at constant composition into a column initially free of sorbates; what results is successive breakthroughs of the solutes, in the same sequence of increasing affinity for the sorbent (see Fig. 2). That frontal analysis should give rise to such a pattern is easy to see: The sorbent layers near the inlet preferentially take up the solute with highest affinity, largely letting the other solutes pass because they are less capable of competing for sorption sites; these layers thus 'filter out' the solute of highest affinity. Farther downstream, that solute is no longer present, so the process repeats itself with removal of the solute with next-highest affinity in the next layers, and so on – the end result being a pattern with successive 'fronts' of different solutes between 'plateaus' (that is, regions of constant composition or, in the language of mathematics, regions of constant state) that grow in length in proportion to the volume of solution that has entered the column. Given competitive sorption equilibrium and absence of selectivity reversals (e.g., a Langmuir-type multicomponent isotherm), all fronts are shocks.

In practice, such behavior would be observed, for example, if the sorbent in Figs. 1 and 2 is a zeolite, and the solutes 1, 2, and 3 are CO_2, toluene, and hexane, respectively, in air.

What is common to both elution development and

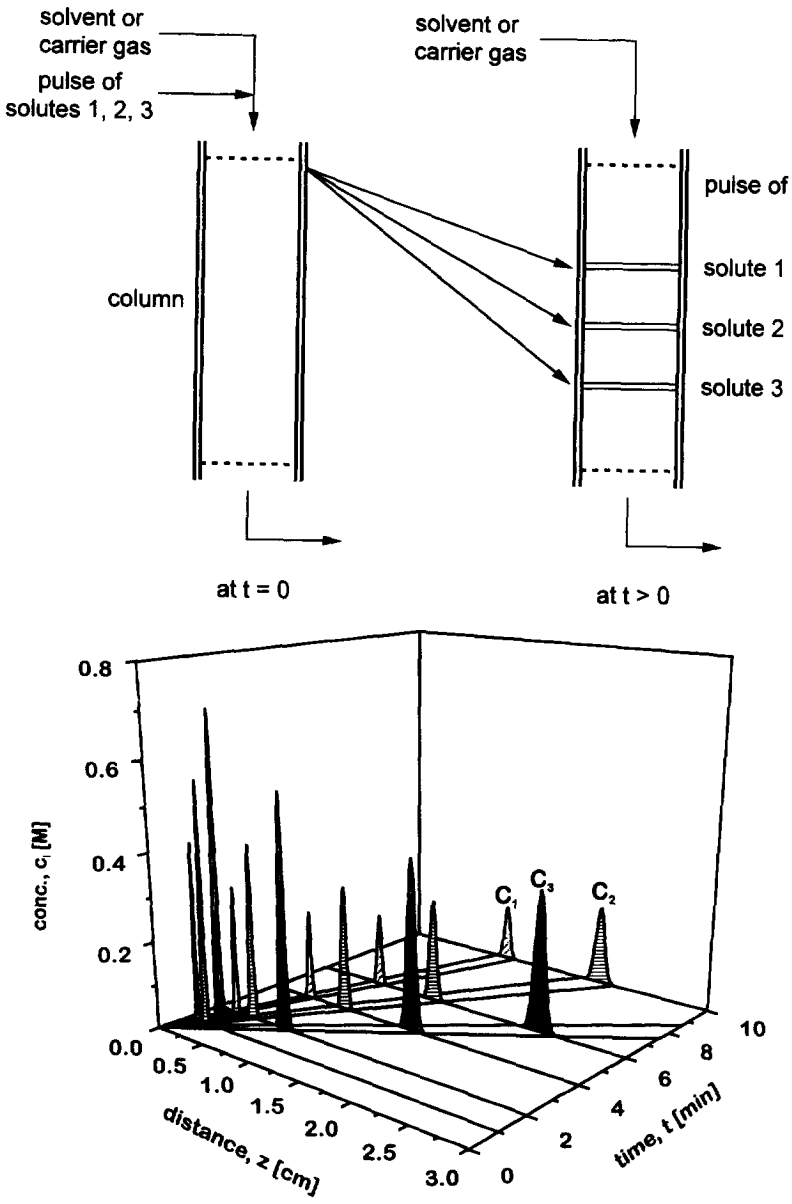

Fig. 1. Elution development: operating procedure (top) and successive composition profiles (bottom). Parameter values $k_1 = 12.0$, $k_2 = 6.0$, $k_3 = 3.0$ (all in cm^3/g) in sorption isotherm equation $q_i = k_i c_i$; $v^0 = 3.1831$ cm/min; $\rho/\epsilon = 2.5$ g/cm^3; injected amounts 1 mmol each.

frontal analysis is that a single, momentary signal (pulse or step) at the column inlet – a composition pulse in elution development, a composition step in frontal analysis – is propagated through the column not as a single signal, but as a set of several that travel at different velocities and become separated by plateaus. This becomes strikingly apparent when one looks at the trajectories which the waves trace in the distance–time field, that is, the bottom planes of the concentration–distance–time diagrams in Figs. 1 and 2. Replotted in Fig. 3, the trajectories – of pulses in elution development, of shock fronts in frontal analysis – are seen to radiate out from a single point, the origin at $z=0$, $t=0$, increasing their distances

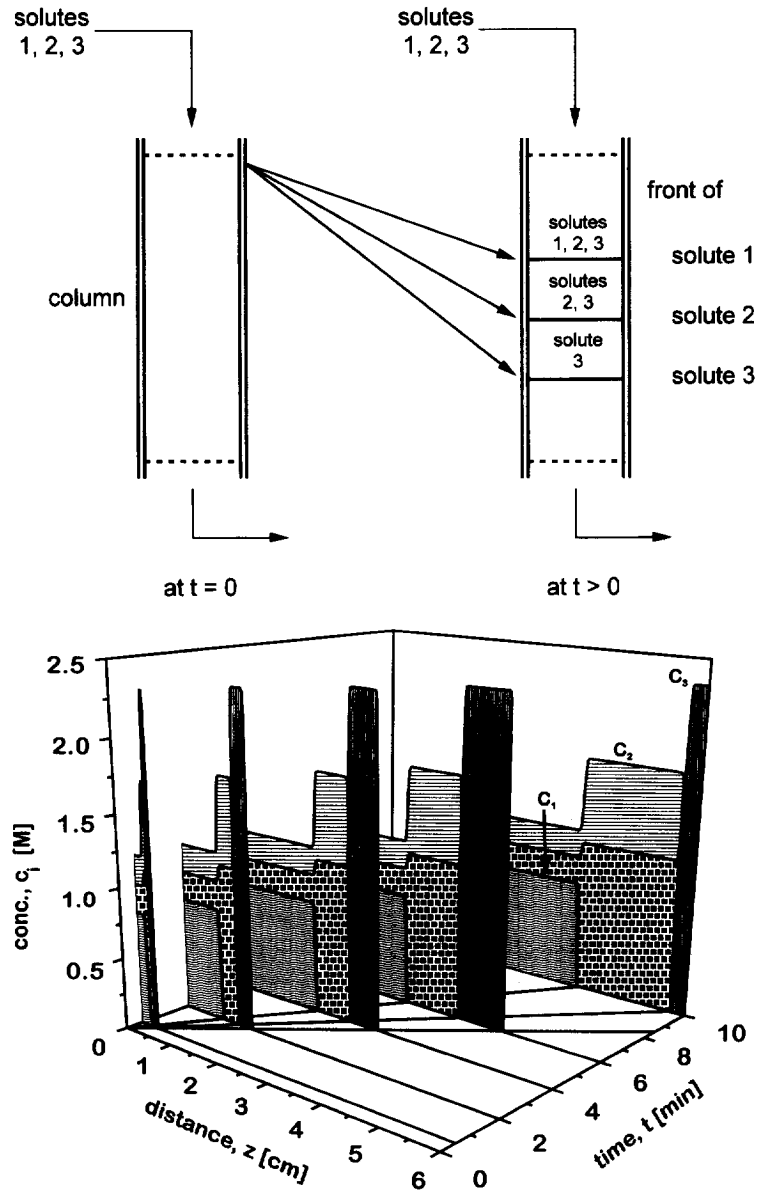

Fig. 2. Frontal analysis: operating procedure (top) and successive composition profiles (bottom). Parameter values $a_1 = 12.0$, $a_2 = 6.0$, $a_3 = 3.0$ (all in cm^3/g), $b_1 = 2.0$, $b_2 = 1.0$, $b_3 = 0.50$ (all in $cm^3/mmol$), in Langmuir sorption isotherm equation $q_j = a_j c_j/(1 + \sum_i b_i c_i)$; $v^0 = 3.1831$ cm/min; $\rho/\epsilon = 2.5$ g/cm^3; entering concentrations $c_1^0 = 0.8$ *M*, $c_2^0 = 1.2$ *M*, $c_3^0 = 1.0$ *M*.

from one another as they advance through the column at different velocities as time progresses. Indeed, that a single variation at the entry of the system is 'resolved' into an entire set of response signals of different velocities is the most fundamental aspect of wave propagation in multivariant systems [15], in chromatography as in many other fields, among them multiphase flow in permeable media [16], sedimentation [17], and convective transport with dissolution and precipitation [18]. Of special interest in the present context is the additional fact that, in non-linear multivariant systems in

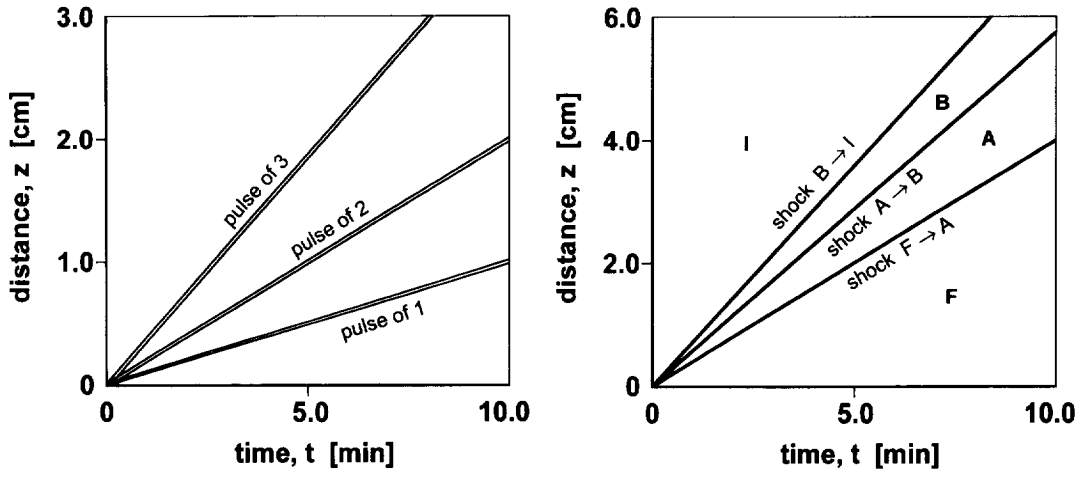

Fig. 3. Distance–time diagrams showing trajectories of pulses and shock fronts of elution development (left) and frontal analysis (right). Conditions as in Figs. 1 and 2, respectively.

general, *each* response signal as a rule involves variations of the values of *all* locally represented dependent variables, even if the entering signal did not. In frontal analysis, for example, every front involves concentration variations of all solutes that are locally present (of solutes 1, 2, and 3 at the front of solute 1; of 2 and 3 at the front of 2; and of 3 alone at its front). That this is so in general will soon become clearer.

Few practical chromatographers habitually think in terms of distance–time diagrams. However, it is well worthwhile to train oneself to do so because the diagrams are invaluable for seeing what may and may not happen in complex phenomena of wave interference. The simple translation of the trajectories in the bottom planes of Fig. 1 and Fig. 2 into the distance–time diagrams in Fig. 3 may help the unaccustomed reader to get used to this tool.

In the two examples above, the number of response signals generated by a single variation at the column inlet equals the number of sorbable solutes. This is in general true in standard chromatography based on isothermal adsorption or partitioning between a moving phase and a stationary sorbent and will be taken for granted for the time being. More accurately, the number of such response signals is given by a quantity that has been called variance [19] and can be viewed as the number of degrees of freedom of the system.[1] The variance, in turn, equals the number of dependent variables (solute concentrations in standard chromatography) minus the number of mathematical constraints. For instance, the variance in standard n-component isothermal adsorption chromatography is n; that in n-component non-isothermal adsorption chromatography is $n + 1$ because temperature is an additional dependent variable; and that in stoichiometric n-component isothermal ion exchange with dilute solutions is $n - 1$ because the requirement of stoichiometric exchange (a consequence of the requirement of conservation of electroneutrality in both the moving phase and the ion exchanger) imposes a constraint. Of course, it is always possible that one or more of the response signals of a set remain vanishingly small, so that fewer than indicated by the variance are actually observed. [In exceptional cases, more response signals may be generated than the variance indicates. This is possible, for example, if the entering composition variation involves a selectivity reversal [20,21].]

[1] This quantity is not related in any way to the variance of statistics, often used in linear non-ideal chromatography to characterize the width of a Gaussian peak.

3. Waves and wave velocities

For the benefit of the reader who has not studied Part I and does not intend to do so, a brief summary of the definition and properties of waves is given here.

A wave is defined as a monotonic variation of the dependent variables (concentrations in standard isothermal chromatography). For instance, each of the advancing fronts in frontal analysis is a wave; each peak in elution development consists of two waves: its front and its rear. Thus, in elution development the number of response waves is twice the variance of the system, but they are generated by two entering waves: the front and rear of the injected pulse containing the sample.

In non-linear chromatography a wave may be self-sharpening (compression wave in the language of physics of fluids) or nonsharpening (dispersive wave or rarefaction wave in that language). If initially diffuse, a self-sharpening wave sharpens into a shock and then continues its travel without further change in its profile. A nonsharpening wave, whether sharp or diffuse initially, spreads as it travels. In ideal chromatography, a self-sharpening wave sharpens into, or remains, an ideal shock, that is, a composition discontinuity; a nonsharpening wave spreads linearly with traveled distance.

In ideal chromatography, a given concentration c_i of a solute i within a diffuse wave advances at its 'natural velocity', given by[2]

$$v_{c_i} = \frac{v^0}{1 + (\rho/\epsilon)\mathrm{d}q_i/\mathrm{d}c_i} \tag{I.4}$$

and a shock advances at the velocity

$$v_{\Delta c_i} = \frac{v^0}{1 + (\rho/\epsilon)\Delta q_i/\Delta c_i} \tag{I.6}$$

(For definitions, see Glossary of Symbols). These 'wave equations' are derived without further assumptions from the differential equation for conservation of mass for the solute in question

$$\frac{\rho}{\epsilon}\left(\frac{\partial q_i}{\partial t}\right)_z + \left(\frac{\partial c_i}{\partial t}\right)_z + v^0\left(\frac{\partial c_i}{\partial z}\right)_t = 0 \tag{I.2}$$

and so remain valid even if other solutes are also present.

4. 'Coherence'

In Part I of this series, the wave Eqs. I.4 and I.6 had proved highly useful for deriving properties of waves and gaining insight into cause and effect in single-component non-linear chromatography. In multicomponent systems the equations remain valid but, unfortunately, do not by themselves provide a sufficient description. The trouble is that the sorbent-phase concentration q_i of the solute in question now depends not only on the concentration c_i of that solute in the moving phase, but also on those of all other solutes:

$$q_i = q_i(c_1, \ldots, c_n) \tag{II.1}$$

Even if the concentrations of all solutes are given, the derivative $\mathrm{d}q_i/\mathrm{d}c_i$ in Eq. I.4 is ill-defined and can in principle assume any positive or negative value. This is because the isotherm of solute i, described by Eq. II.1, now is no longer a curve in two dimensions (coordinates q_i and c_i), but an n-dimensional hypersurface in the $(n + 1)$-dimensional space with coordinates $q_i, c_1, \ldots, c_n$. The slope of a tangent to that hypersurface at a given point depends on the direction in which it is measured, that is, on the variations of *all* moving-phase concentrations, not just on that of solute i alone. Thus, the wave equation I.4 by itself does not contain enough information. Even if all concentrations $c_1, \ldots, c_n$ are given, the wave velocity v_{c_i} of a solute i can have any value: it is not a 'point quantity.' Similarly, for shocks, even if the moving-phase concentrations c_i of the respective solute are known on both sides of the shock, the shock velocity $v_{\Delta c_i}$ cannot be calculated from Eq. I.6 alone because the values of q_i on both sides depend on the concentrations of the other solutes also.

Of course, it is always possible to solve any given mathematical problem of chromatography without invoking wave theory, that is, by numerical integration of the differential mass-conservation equations

[2] Equations, tables and figures from Part I are refered to by 'I' followed by the number of the respective items.

I.2 of all solutes over distance and time (analytical solutions exist only for very few, exceptionally simple cases). However, in complex situations a very large amount of computation is likely to be required, especially if the differential equations turn out to be stiff. Moreover, the results of any such brute-force integration are valid only for the conditions that were specified. Unless we are willing to invest in the calculation and evaluation of an enormously large number of cases, this approach – with the column a 'black box' – teaches us little about cause and effect and does not lend itself to gaining much insight that would allow us to state general rules and make predictions without detailed calculations. For that, we need a better *conceptual* understanding. A key concept here is that of 'coherence' [2,22–24].

If we look again at the frontal analysis experiment in Fig. 2, we can say that it really involves four waves rather than only three: The composition variation between the pure solvent or carrier gas initially in the column and the entering mixture containing all three solutes, although existing for only one fleeting instant at start, also qualifies as a wave by our definition because it is a variation of the dependent variables, the concentrations. It could be called 'unstable' because it cannot persist as a single wave, but instability in phenomena involving fluid flow has all kinds of connotations not intended here. Therefore, a different terminology has been created: In the language of coherence theory, the initial wave is *noncoherent* and is resolved into a set of *coherent* waves. The latter, by definition, travel without further break-up; they may sharpen or spread, but retain their integrity in all other respects – except that interference with other waves or a change in conditions may again produce a local and temporary state of noncoherence, as will be seen later (see Section 9).[3]

For frontal analysis and any other problems with a column of uniform initial composition and constant composition of the entering fluid (so-called Riemann problems) the resolution into coherent waves may appear trivial. Only a perfectionist might demand proof that a particular wave or moving-phase composition $\{c_1, \ldots, c_n\}$ has traveled to its current position from the column inlet rather than having arisen at the present moment from a shift of solute concentration profiles relative to one another. Indeed, the early theoreticians of non-linear chromatography (except Baylé and Klinkenberg [26]), preoccupied with Riemann problems, took such eminently plausible behavior as self-evident, to the point that they failed to realize that noncoherence is possible at all. The coherence concept comes into its own when conditions are more complex, in particular, if noncoherence arises or persist over finite time and space intervals. Examples are interferences occurring when faster coherent waves generated later catch up with slower ones generated earlier (e.g., in displacement development), and gradual approach to coherence from a gradual composition variation at the column inlet (e.g., in gradient elution) or from a gradual initial composition variation in a column (e.g., at start of regeneration of an incompletely exhausted adsorption column). Such matters will be taken up later (see Sections 8 and 9). What can be said at the present stage is:

- A noncoherence does not know or care how, when, and where it came into being – at start and column inlet, or later and farther downstream – and whether it is sharp or diffuse; it is always resolved into a set of coherent waves by the same rules.

This rule is a key to understanding the response behavior of even quite complex systems.

In general terms, the coherence concept resembles that of equilibrium and steady state. Each is a state a system strives to attain, and will approach arbitrarily closely if not disturbed again. That state is equilibrium if the system is closed; it is a steady state if we allow the system to be open, but impose fixed boundaries and constant boundary values; it is a state of coherence if we relax these latter two restrictions also. Thus, coherence can be viewed as a general concept of 'stability' that includes equilibrium and steady state as special cases. In the real world, coherence is only approached asymptotically, but so are equilibrium and steady state. This does not

[3] In the language of mathematics of non-linear waves, the coherence principle can be formulated as follows: An arbitrary starting variation, if embedded between sufficiently large regions of constant state, is resolved into simple waves or shocks, between which new regions of constant state arise [25].

impair the utility of the concepts. In fact, they are most useful in situations far from the final, 'stable' state since they tell us in which direction the system will move under complex, transient conditions. This is true for coherence as much as for equilibrium and steady state.

The notion of 'stable' modes other than equilibrium and steady state is, of course, an old one in physics and mechanics. Examples that come to mind are the harmonic vibrations of an oscillator or a musical string and the flutter of an aircraft wing or a flag in high wind. A perhaps closer analogy to coherence is a behavior observed by the famous Dutch physicist Christian Huygens (1629–1695). He had mounted his collection of pendulum clocks on a wall board and noticed a few days later that all their pendulums were swinging in perfect synchronization (see Fig. 4), and this although the clocks separately had slightly different frequencies. They had adapted to one another through vibrations transmitted by the board to which they were fastened. Typically, the 'stable' states in all such mechanical systems are characterized by eigenvalue solutions, as is coherence. The world we live in is full of eigenvalues. It is only in physical chemistry and chemical engineering that our preoccupation with thermodynamics – which really should be called thermostatics – has let us neglect dynamics and be slow to recognize and appreciate the fact that 'stable' states exist beyond equilibrium and steady state.

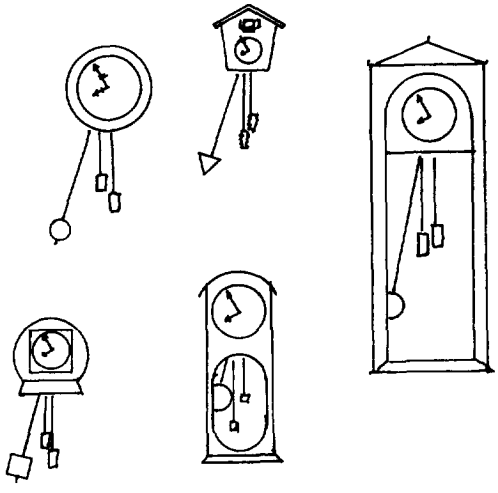

Fig. 4. Huygens' clocks with pendulums swinging in synchronization.

A general proof for development toward coherence, independent of the exact form of the differential equations of motion, is so far lacking – as, for that matter, is a proof for development toward equilibrium (the free-energy argument of thermodynamics only replaces one unproved axiom by another). Proofs for development toward coherence in systems with hyperbolic differential equations have been given [25,27], but the concept has shown itself to be valid for some other types of systems as well. What can be shown without mathematics is that the principle is plausible: As mentioned above, coherent behavior appears trivial if there is only a single, instantaneous variation of the composition of the injected fluid, so that all waves must have originated from a single distance–time point (i.e., under Riemann-type conditions). It also appears plausible that the same wave pattern should result if, instead, the injection variation were extended over a finite, but exceedingly short time interval (in the language of mathematics, this is so if the problem is 'well-posed'). But the time over which the injection is varied is a relative quantity, and would constitute a large fraction of the total time in a very much shorter experiment with a very much shorter column. That latter experiment can, in turn, be viewed as a scale-down of one with a long column and an injection variation that extends over a significant length of time [28]. This is illustrated in Fig. 5, which shows trajectories of compositions and shocks in the distance–time plane. Thus, unless the minute change in conditions from discontinuous to very fast but continuous injection variation is to produce significant and long-lasting effects, development toward coherence must be expected. This is a plausible argument rather than a proof, for nature is full of systems that are not well-posed (see the 'butterfly effect' of meteorology, in which a butterfly's bat of a wing could conceivably set a chain of events in motion that much later culminates in a typhoon thousands of miles away). However, there is absolutely no reason to suspect such behavior in the types of systems of interest here.

Coherence is a very simple concept, as simple as

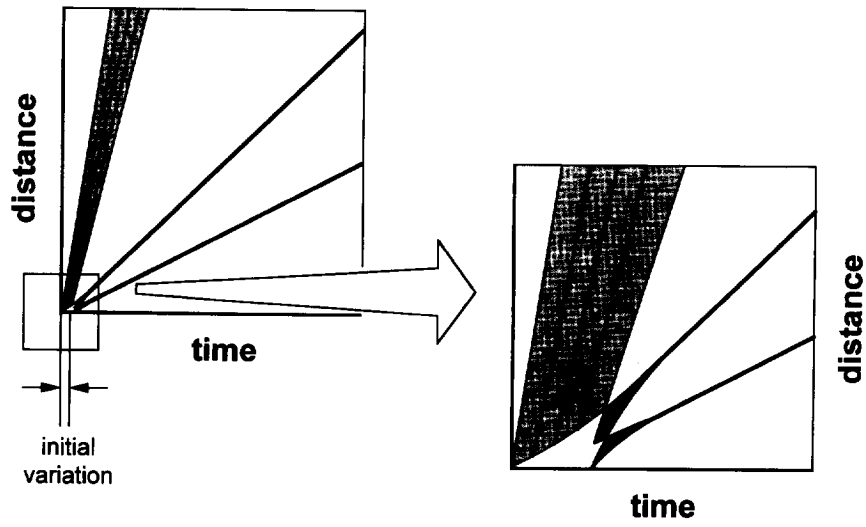

Fig. 5. Distance–time diagrams illustrating attainment of coherence under non-Riemann conditions, granted well-posed behavior. Trivariant noncoherence producing one nonsharpening wave (bundle of trajectories of given compositions $\{c_1, \ldots, c_n\}$, shown as thin lines within shaded wave) and two shocks (trajectories shown as heavy lines). Left: coherent pattern from very short continuous composition variation; right: enlargement of small area near origin in left diagram, with coherent pattern arising from composition variation that extends over significant fraction of total time. (From Helfferich [28].)

equilibrium or steady state. We all have long become used to applying the latter two without asking for proof that our systems will settle down to these states if not further disturbed. If we accept coherence on these same terms, as a state a dynamic system will strive to attain, all we have to realize is that in wave propagation an arbitrary composition variation (wave) usually is not in such a state, and therefore breaks up into a set of several waves that are and that separate from one another as they travel. Frontal analysis is the simplest manifestation of such behavior, but the principle is more general in that any noncoherent wave, whether sharp or diffuse and regardless of its origin and position, breaks up in this same manner. The greatest hurdle to an understanding of coherence is to accept the simplicity of the idea, to overcome the notion that great complexities must lurk behind. The greatest difficulty in explaining coherence is, the more said in explanation, the more that notion is reinforced. As Maurice Dirac once said, "you don't *learn* a theory, you get used to it" – as we all got used to the ideas of equilibrium and steady state, by applying them with success.

It is true, however, that application of the coherence concept requires thinking along lines the average chromatographer is not used to. Conventional thinking has mostly been in terms of mathematically deriving and understanding effluent composition histories, given the initial and entering compositions. In contrast, coherence theory starts with examining what will develop from local and momentary concentrations and their gradients, regardless of where and when they exist and how they arose. Coherence, or the lack of it, is a property of a wave, not of an entire column (although the latter can be called coherent if all the waves it contains are coherent). The conventional approach is eminently satisfactory as long as the operating conditions are simple, as in traditional elution development and frontal analysis. It does run into difficulties, however, when the column initially contains composition gradients or the composition of the entering fluid is varied gradually or repeatedly, as is often the case in modern preparative chromatography. Imagine, for example, a column as in Fig. 2 run to breakthrough of butane (solute 3) and now to be stripped while still containing the fronts of toluene and CO_2 (solutes 2 and 1). Or consider the problem of predicting the required column length for resolution of a multicomponent mixture of solutes by displacement de-

velopment, where the wave sets produced by introduction of the mixture and, later, of the development agent interfere with one another in the column (transient profiles in displacement development will be examined in more detail in Section 9). Here, a knowledge of what develops locally makes it possible to unravel the resulting behavior with relative ease. For any given type of sorption isotherm, sets of rules can be deduced for what may and may not happen in the column, as will be shown in Parts III and IV of this series. Thus, if the initial and entry conditions are simple, the concept has little to offer, but it comes into its own in more complex situations. Coherence theory and its tools will never replace trial or numerical calculation, but where the variety of potential operating schemes and conditions is overwhelming, it can help to decide *what* to try or to calculate, and can provide insight from which innovative ideas might spring.

To adjust to the different way of thinking, although not at all difficult, requires an effort and is harder for the expert than for the novice. No argument can persuade the practical chromatographer that it will be worth his time. Only his successes with the approach can show him that it was.

5. Coherence conditions

A sufficient and necessary condition for a wave to be coherent is readily deduced. A multicomponent wave can be viewed as a composite of single-component waves, one for each solute that is present. For example, the front of solute 1 in the frontal analysis experiment of Fig. 2 involves concentration variations of all three solutes 1, 2, and 3 and so is a '1 wave' as well as a '2 wave' and a '3 wave'. The wave equation I.4 for a concentration within a diffuse wave, or Eq. I.6 for the concentration step of a shock, is valid for each solute that is present. If the solutes had different wave or shock velocities, their concentration variations would separate from one another, that is, the wave would break up into several waves and thus be noncoherent by definition. Accordingly, for a wave to be coherent, the wave velocities must be the same for all solutes at a given point in space and time:

· Coherence requires all components that are locally present to have the same wave or shock velocity at any point in space and time within the wave.

[A more general definition requires all represented dependent variables, including e.g. temperature and pressure, to have the same wave velocity.] Reference to the wave equations I.4 and I.6 shows the corresponding conditions for coherence to be

$$\text{for diffuse waves:} \quad dq_i/dc_i = \lambda \quad \text{for all } i \tag{II.2}$$

$$\text{for shocks:} \quad \Delta q_i/\Delta c_i = \Lambda \quad \text{for all } i \tag{II.3}$$

where λ or Λ has the same value for all solutes i. The conditions are known as the differential coherence condition and integral coherence condition, respectively. (In mathematics of waves, the conditions are called the fundamental differential equation of Riemann's problem and the compatibility condition, respectively.) [Regarding the nature of the total differential in condition II.2, see Appendix A]

[An interesting corollary of the differential coherence condition II.2 is that concentrations $c_1, \ldots, c_n$ that coexist at the same point in space and time within a coherent wave will necessarily remain in each other's company because they travel at the same velocity. Accordingly, coherence can also be defined as the conservation of complete sets of values of dependent variables (here, the concentrations of the solutes). This alternative definition makes it most clearly apparent that coherence includes equilibrium and steady state as special cases: In the latter two, the values of the dependent variables do not change with time, so that the condition of conservation of sets is automatically obeyed. In comparison, coherence is seen to be the more general concept because it allows the sets of variables to move in space as time progresses.]

In a case as simple as frontal analysis or, for that matter, in any Riemann problem, the coherence conditions may appear trivial. Yet, they are the key to efficient mathematics. We had encountered the fundamental difficulty that, in principle, the wave equation I.4 allows a given concentration c_i at given composition $\{c_1, \ldots, c_n\}$ to have any arbitrary positive or negative velocity. Now, the differential

coherence condition II.2 helps by imposing itself as a constraint. The result is that the wave velocity, which in a coherent wave is common to all solutes according to condition II.2 or II.3, can only assume a few distinct values. Mathematically, the velocity now is given by eigenvalues. These are the natural velocities at which, at given moving-phase flow-rate v^0, the composition can travel if within a coherent wave:

$$v_c = \frac{v^0}{1 + (\rho/\epsilon)\lambda} \qquad \text{(II.4)}$$

They will here be called eigenvelocities. In a system with variance n, any composition $\{c_1, \ldots, c_n\}$ has n such eigenvelocities v_c.

Perhaps more importantly, to each of these eigenvalues belongs an eigenvector $\{dc_1, \ldots, dc_n\}$. The eigenvectors indicate how the concentrations $c_1, \ldots, c_n$ may vary relative to one another across a coherent wave. An arbitrary composition variation is unlikely to be in the direction of an eigenvector, and cannot constitute a coherent wave. Only certain distinct types of composition variations can occur across coherent waves. Since all composition variations in the column tend to sort themselves out into coherent waves, a knowledge of the types of variations compatible with coherence is an invaluable aid in predicting column behavior even under complex operating conditions. How this can be done will be discussed in the next section.

The mathematically disinclined reader should not be discouraged by the references to eigenvalues and eigenvectors and the use of the term eigenvelocities. The essential concepts can be grasped without background in higher mathematics. At this point it will suffice to accept three facts. The first is that coherent behavior is a state the system is comfortable with and strives to attain or at least approach from any arbitrary initial state; this is true for coherent waves in chromatography as much as for Huygens' pendulums, the harmonic vibration of a musical string, or the flutter of an aircraft wing. The second fact is that, in chromatography, coherence implies equal local wave or shock velocities of all represented components. The third fact is that coherence is compatible with only certain distinct composition variations (namely, those characterized by the eigenvectors). For the interested reader, Appendix A provides an outline of the general mathematical formulation of the eigenvalue problem.

6. Composition paths and path grids

The obvious objective of multicomponent theory is to predict the behavior of chromatographic columns under specified initial and entry conditions. In the context of coherence theory, the first key tool for this purpose is the so-called composition path grid. The present section explains the nature of such grids. The application to the prediction of response behavior will be shown in the next three sections. The presentation here is based on the methods set forth in the book Multicomponent Chromatography [2], to which the reader is referred for details (see also a more readable and practice-oriented survey [29]). However, the book mainly addresses systems with stoichiometric ion exchange rather than non-stoichiometric adsorption. While the two phenomena are mathematically equivalent and the methods employed are substantially the same, there are differences in physical interpretation and graphical representation (for adsorption systems, see also work by Glückauf [30,31], Rhee [32,33], and Guiochon [14,34] and their co-workers).

Composition variations that are compatible with the differential coherence condition II.2 can be mapped as curves in the composition space, that is, a space with the concentrations c_i as coordinates. [Stated in more general mathematical terms, variations compatible with coherence can be mapped in the hodograph space, that is, the space with the dependent variables as coordinates.] The differential coherence condition is obeyed along these so-called composition paths (characteristics in the hodograph space), and only along them. Any composition variation that cuts across the grain of these paths is noncoherent and will be resolved into coherent variations that follow the paths. One might say, the paths are the 'grooves' into which the system wants to settle.

It is important to note:

- The composition path grid is uniquely given by the equilibrium equations and the values of their

coefficients, and is independent of the initial and entry conditions.

This makes it possible to construct the path grid for a system with given sorption equilibrium properties once and for all and then use it to predict response behavior under any initial and entry conditions. Establishing response behavior with use of the path grid is much like plotting the best route for a car trip on a road map.

The path grid of a two-component Langmuir system is shown in Fig. 6. The coordinates are the moving-phase concentrations c_1 and c_2 of the two solutes. (This is a practical choice; stationary-phase concentrations q_1 and q_2 could be taken as well.) In accordance with the variance 2 of the system there are two families of paths, shown as solid and dashed lines. The entire composition plane is covered by an infinite number of paths of both families, of which only some at regular intervals are shown. Each composition point is at an intersection of two paths, one from each family. At each point, the solid line corresponds to the path with the lower eigenvelocity; the dashed line, to that with the higher eigenvelocity. For short, the paths are called 'slow' and 'fast.'

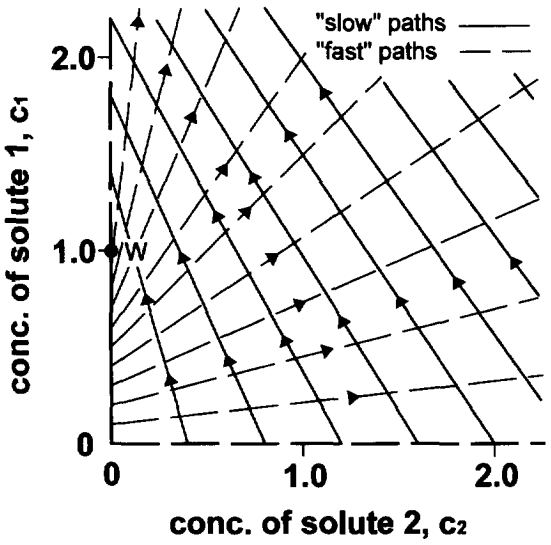

Fig. 6. Composition path grid of typical two-component sorption system, calculated with $a_1 = 6.0$, $a_2 = 3.0$ (both in cm^3/g), $b_1 = 1.0$, $b_2 = 0.5$ (both in $cm^3/mmol$) in Langmuir equation $q_j = a_j c_j/(1 + \sum_i b_i c_i)$. Arrowheads indicate direction in which eigenvelocity increases along respective path. W=watershed point.

Along both the 'slow' and 'fast' paths the eigenvelocity increases in the direction in which the concentration c_1 of the more strongly adsorbed solute increases; in Fig. 6 this direction is indicated by arrowheads along the paths. (Note that solutes are numbered 1, 2,... in the sequence of decreasing affinity for the sorbent.) The axes themselves are composition paths: The c_2 axis is a 'fast' path, and the c_1 axis is divided by the 'watershed point' W into a 'fast' and a 'slow' path; paths into the interior are 'slow' if starting from an axis that is a 'fast' path, and are 'fast' if starting from one that is a 'slow' path. This makes any point on an axis also an intersection of a 'fast' and a 'slow' path.

In a system with Langmuir adsorption isotherms (or constant separation factors in ion exchange) the paths are straight lines. This is not in general true for other types of equilibrium isotherms.

Regardless of the type of isotherm, each additional component adds a dimension to the composition space and a family of paths. For instance, the composition space of a four-component, isothermal adsorption system is four-dimensional, with coordinates c_1, c_2, c_3, and c_4, and is filled with four families of paths; each point in the composition space is at an intersection of four paths, one from each family.

[In Langmuir systems with three or more solutes, paths of any two families lie on common (planar) surfaces in the composition space [35]. This makes it possible to travel on a circular route along paths of only two families: first on a path of family (j), then on one of family (k), then on another one of family (j), and back to the starting point on another one of family (k) (see Fig. 7, left). The common planes of (j) and (k) paths are intersected by paths of all other families. Grids with these properties are 'orthogonalizable' [36]. Non-Langmuir grids are not necessarily orthogonalizable. Instead, a route alternating between paths of families (j) and (k) might lead the traveler to a point from which he would have to take paths of other families to get back to his starting point, as though he had descended or ascended to a different floor on a spiral staircase (see Fig. 7, right). The problem of a finding a sufficient or necessary condition a multicomponent isotherm must obey to form an orthogonalizable grid has not yet been solved. However, numerical examination of a num-

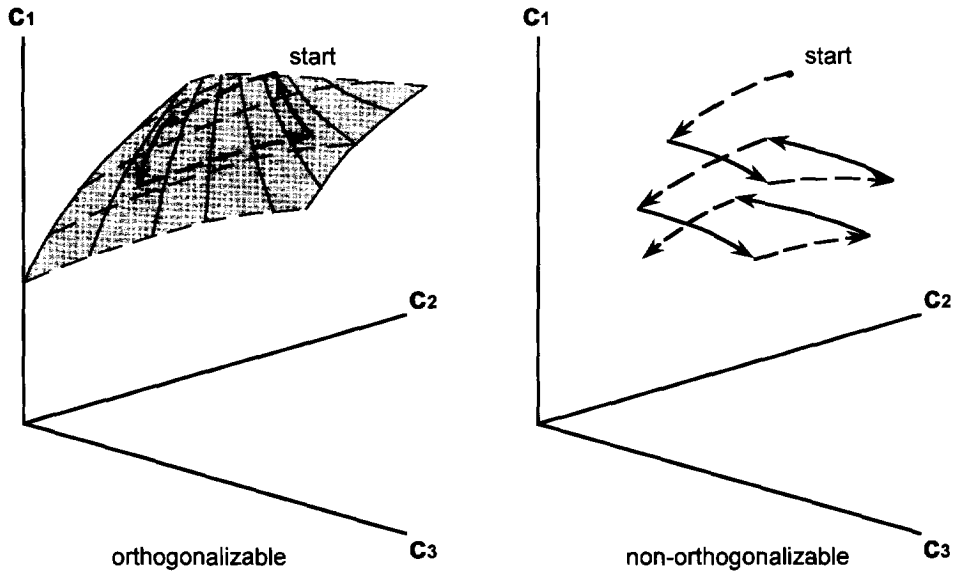

Fig. 7. Circular route along composition paths in trivariant non-Riemann systems. Left: orthogonalizable grid, route remains on common surface (shaded) of intersecting 'slow' and 'fast' paths; right: non-orthogonalizable grid, no common surface exists, paths of all three families needed for return to starting point.

ber of non-Langmuir grids has shown their deviations from orthogonalizability to be quite small, if at all real [37].

In principle, the path grid of any system can be constructed step by step as follows [16]. First, the eigenvalue problem is solved at a selected starting point. Then a small distance in the direction of one eigenvector is marked off and the eigenvalue problem is solved again at the new composition point. This procedure is repeated over and over again, always with the eigenvector of the same family, to trace a curve which at each of its points is in the direction of that eigenvector. After construction of the first path is complete, the entire procedure is repeated for other paths of the same family (same eigenvector) and for paths of all other families (other eigenvectors). A computer program for tracing paths in this fashion has recently been described [38].

Obviously, such step-by-step procedure requires a lot of calculation. However, the topology and general structure of the path grid depend only on the form of the equilibrium equations, not on the values of the coefficients, and with their knowledge an approximate grid can usually be sketched in after calculation of only a few characteristic points or curves (singular points, loci of selectivity reversal, etc.). For example, the grid of a two-component Langmuir system as in Fig. 6 can be constructed from the solutions of only two simple algebraic equations, as shown in an appendix (see Appendix C). Other examples will be seen in Parts III and IV of this series. An even greater simplification that makes the construction of the path grid superfluous can be achieved in adsorption systems with Langmuir isotherms and ion-exchange systems with constant separation factors by use of a mathematical transformation, as will be shown in Part III.

Many multicomponent adsorption systems that do not obey the Langmuir equation nevertheless are 'Langmuir-like' in that their equilibria are competitive and without selectivity reversals. In entirely competitive equilibria, an increase in the concentration c_j of any solute j decreases the distribution coefficients q_i/c_i of *all* solutes. The respective geometrical properties of the path grid allow a large number of rules for coherent waves in such systems to be derived, as will be shown in Part III. In that context it will also become possible to attribute some typical deviations from these rules to specific non-Langmuir equilibrium properties.

Of course, multidimensional grids can no longer be shown in simple diagrams. However, this does not impair their abstract mathematical use. Moreover, it is not all that difficult to train one's imagination to handle more than three dimensions. For an excellent introduction to such thinking, the interested reader is referred to a nineteenth-century classic, Edwin Abbott's "Flatland" [39], a delightful book that should be required reading for any student of multivariant problems.

7. Composition routes

A key step in the prediction of column responses is the construction of composition routes. Composition paths, discussed in the previous section, are curves in composition space along which the differential coherence condition is obeyed; that is, they map composition variations that may occur across coherent waves. In contrast, a composition route maps the actual composition variations encountered in the column or effluent under specified conditions. A route may or may not follow composition paths. Where it does, and only where it does, the corresponding wave is coherent. By convention, a route is usually shown as a sequence of arrows, one for each wave. A profile route indicates the sequence of compositions found along the column from inlet to outlet at fixed time, with the arrows pointing in the direction of flow. A history route indicates the sequence of compositions encountered at a fixed position, usually the column outlet, as a function of time; if arrows are used, they should logically point in the direction of increasing time. In work to date, profile routes are so much more common that a 'route' is taken to be a profile route unless specifically declared a history route.

As a simple example, Fig. 8, left, shows the profile route for a two-component, Riemann-type case (uniform initial composition of the column and constant composition of the entering fluid). The path grid is that of Fig. 6, and both solutes 1 and 2 are present in the entering fluid as well as initially in the column. The route follows the 'slow' path from point F (entering fluid) until, at point A, it switches to the 'fast' path that leads to point I (initial state of column). The arrow F→A corresponds to the slower of the two coherent response waves; the arrow A→I, to the faster one; and point A, to the plateau between the two waves. Along the slow path through F and A the eigenvelocity decreases from F to A (i.e., in the direction of flow), so the wave F→A is a shock; the distance–time diagram shows its trajectory as a single, heavy line. In contrast, on the fast path through A and I the eigenvelocity increases from A to I, so that wave is nonsharpening and diffuse; in the distance–time diagram it appears as a bundle of divergent trajectories of compositions within the wave (thin lines).

The route in Fig. 8 runs entirely along paths. A more general rule can be stated:

- The composition route of a Riemann problem runs entirely along composition paths.

This is so because the waves of a Riemann problem all originate from one single distance–time point ($z = 0$, $t = 0$), namely, column inlet and start of operation. All concentration and shock trajectories radiate out from that point, and none intersect (see Fig. 8, right). There is no wave interference, and behavior is entirely coherent from the start (granted the conditions of ideal chromatography). The behavior of non-Riemann systems is more complex, as will be seen in Section 9.

Also, in a Riemann problem, at any moment a wave that is faster must be downstream of one that is slower; and it must have emerged from the column, or reached a given location in the column, sooner than one that is slower. This is because the faster wave has traveled farther in the elapsed time, and has reached the column exit or a given location earlier. This allows the following general rules for the sequence of path segments along the route to be formulated:

- The profile route of a Riemann problem runs from the composition point of the entering fluid to the composition point of the initial state of the column. On its way it takes the path of lowest eigenvelocity from the starting point, switches successively each time to a path of next higher eigenvelocity, and reaches the end point on a path of highest eigenvelocity.
- The history route of a Riemann problem

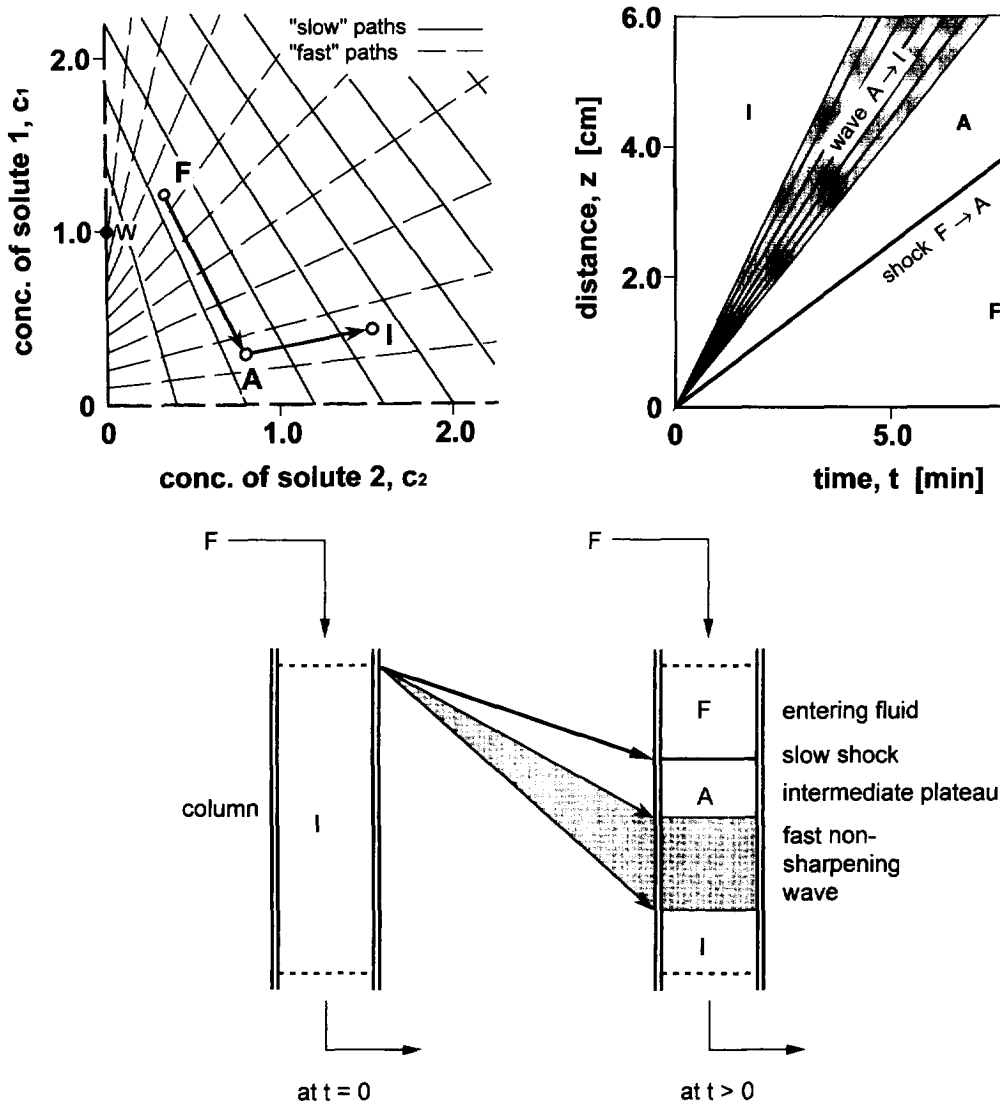

Fig. 8. Development of a response pattern in a two-component system with uniform initial composition of column and constant composition of entering fluid. Top left: composition route; top right: distance–time diagram; bottom: resulting column behavior; grid as in Fig. 6. I=initial composition in column; F=composition of entering fluid.

coincides with the profile route, but runs from the point of initial state to the point of entering fluid, with path segments in opposite sequence.

To illustrate the application of the sequence rule, Fig. 9 shows the profile route of a three-component frontal analysis case (for distance–time diagram see Fig. 3). The route runs along the 'very slow' path from point F and eventually reaches point I on a 'fast' path. In between, it switches at point A from the 'very slow' to a 'slow' path, and at point B from the latter to the 'fast' path that leads to point I. The arrows F→A, A→B, and B→I correspond to the three coherent response waves; the points A and B, to the plateaus between them. Note that all waves are shocks, that each involves concentration variations of all solutes locally present, and that the concentration of each solute increases across each wave in the

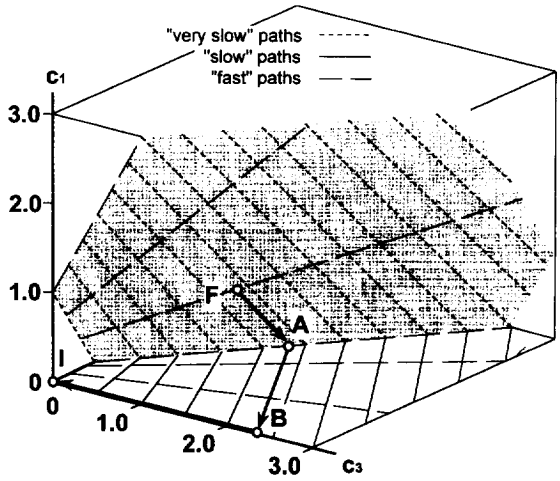

Fig. 9. Composition route of three-component frontal analysis system of Fig. 2, shown on grid of common plane of 'very slow' and 'fast' paths through F (shaded) and of 'slow' and 'fast' paths in c_2, c_3 plane. I=initial composition in column, F =composition of entering fluid.

direction of flow until it drops to zero at the respective front. This is true in general for frontal analysis with any number of solutes, granted sorption equilibria are competitive and without selectivity reversals. For comparison, Fig. 10 shows an observed effluent concentration history of frontal analysis [40] (note that the sequence of waves with progressing time in a history is the opposite of that with increasing distance in the corresponding profile).

The 'bunching up' of molecules of a solute to concentrations higher than in the entering fluid is characteristic of frontal analysis with competitive sorption equilibria. Wave theory makes it easy to understand the physical cause of this effect. For simplicity, consider the case of only two solutes. The leading band contains only the solute of lower affinity for the sorbent. As molecules of that solute outrun those of the solute of higher affinity to form the leading band, they cease to have competition by

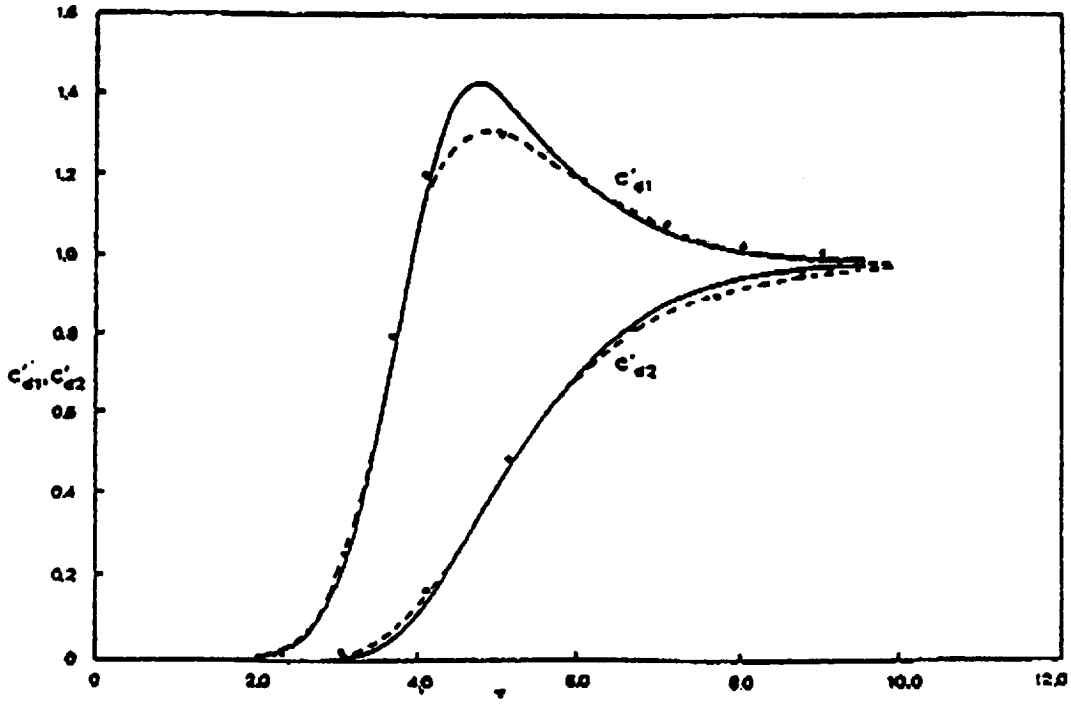

Fig. 10. Observed effluent composition histories of frontal analysis of 2-butanol (c_1) and *tert.*-amyl alcohol (c_2) on carbon. Plotted are normalized effluent concentrations vs. dimensionless time; ●=experimental points; solid and dashed curves are for different values of mass-transfer coefficient in liquid phase. (From A.I. Liapis and D.W.T. Rippin [40], reproduced with permission of Chem. Eng. Sci.).

them for sorption sites. As a result, they are more strongly sorbed in the leading band than behind it, and therefore are more strongly retarded in the leading band. (More precisely put, the lack of competition in the leading band increases their distribution coefficient q_i/c_i, and this reduces their particle velocity; see Eq. I.5 of Part I.) This is what causes the concentration to increase, just as traffic will bunch up when cars ahead are forced to slow down.

As a second example, Fig. 11 shows the behavior upon elution of a column that had been uniformly

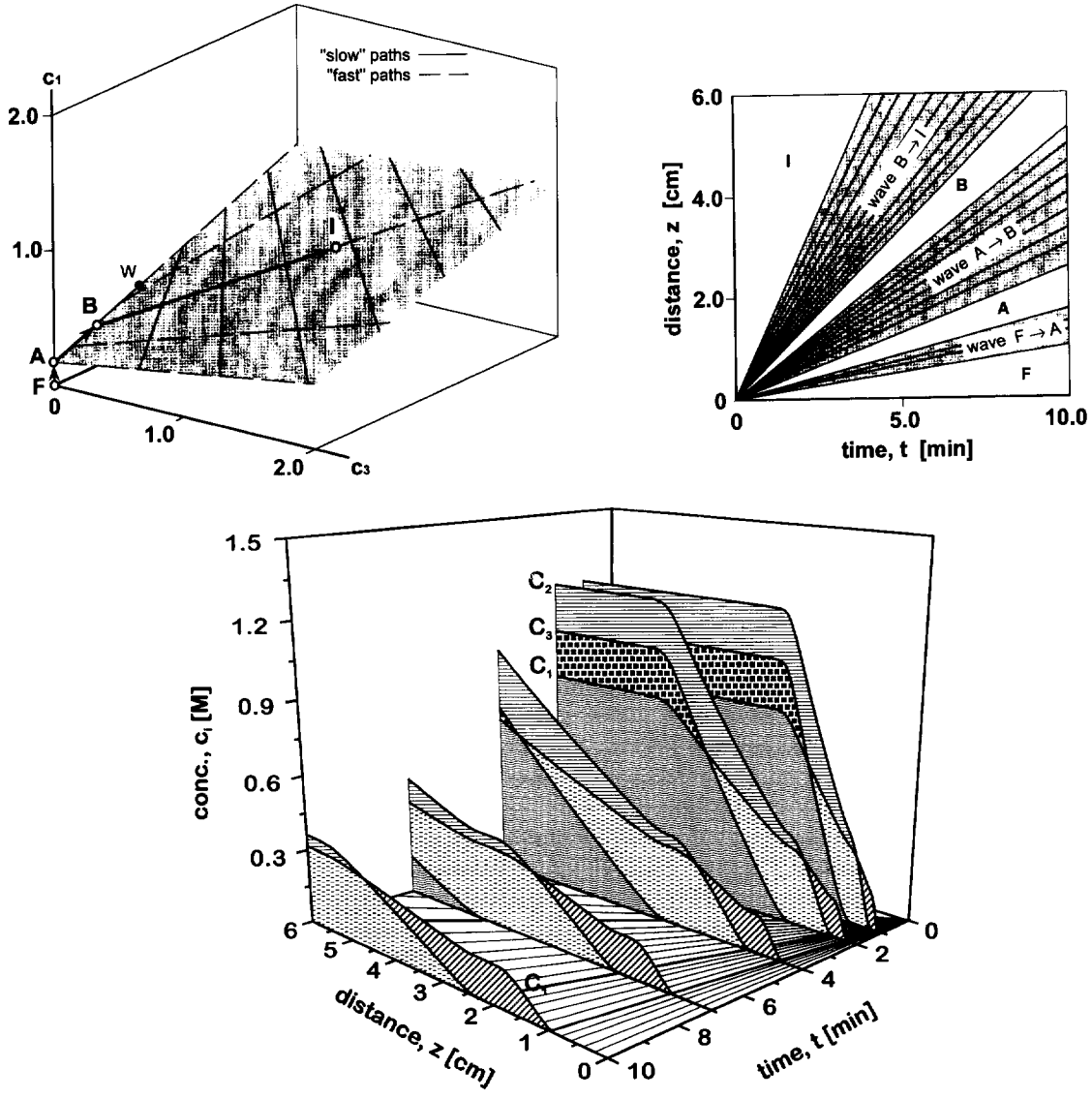

Fig. 11. Elution from a uniformly loaded column: composition route (top left), distance–time diagram (top right), and successive composition profiles (bottom). Common plane of 'slow' and 'fast' paths in route diagram is shaded; W=watershed point on that plane. Conditions as in Figs. 2 and 9, except initial and entering compositions are interchanged. (Note inverted time and distance scales of composition profile diagram.)

saturated with the three-component mixture used in the frontal analysis experiment just described: The conditions are the same as stated for Fig. 2, except the compositions of the initial and entering fluids are interchanged. As in the frontal analysis example, three waves are generated but, given competitive sorption equilibrium, all now are nonsharpening. As always, the route takes paths in the sequence of increasing velocity, but since its end points are interchanged it runs a different course. Another difference between the two cases is that the solute concentrations are lower in the plateaus than initially: no 'bunching up' as in frontal analysis! Here, as molecules of the high-affinity solute are outrun by the others and left behind in a band of their own, they no longer face competition by the other, and so are more strongly sorbed and retarded. Because of this slowing-down the stragglers become spread out more thinly.

For systems with variances greater than 3, we run out of dimensions to plot path grids and routes on a sheet of paper, except if the absence of some solutes from portions of the pattern allows us to work with projections onto sub-spaces with fewer dimensions (e.g., see the constructions for displacement development in Section 9 and other examples in Ref. [2]). This largely deprives us of the convenience of graphical constructions, but does not impair the abstract mathematical use of path grids.

In all Riemann systems with two components, and in some with n components of which several are absent initially or from the entering fluid, the route can be traced immediately. This includes frontal analysis. In more complex Riemann systems the paths between the 'slowest' and 'fastest' may have to be found by trial and error. Route construction for non-Riemann systems requires an understanding of wave interference and will be taken up in the next two sections.

[The discussion above will suffice for a qualitative understanding of the significance of composition routes. For quantitative constructions, one fundamental difficulty remains to be addressed in the general case with curved composition paths, even under Riemann-type conditions. Fig. 12 shows schematically a curved composition path – by definition a curve along which the differential coherence condition II.2 is met – and the locus of points such as B_1, B_2, and

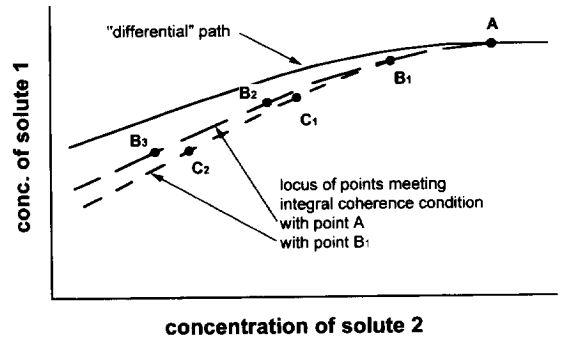

Fig. 12. Non-obeyance of integral coherence condition by pairs of points on curved differential composition path. Locus of points meeting integral coherence condition with A is curve through B_1, B_2, B_3 and deviates from differential path through A; locus of points meeting that condition with B_1 is curve through C_1, C_2 and deviates farther (schematic).

B_3 which meet the integral coherence condition II.3 with a point A on the path. These points are not on the path through A! [41–43]. At first glance one might think that an 'integral' composition path grid with curves meeting condition II.3 could be constructed, but this is not true either: While that condition is met for pairs such as AB_1, AB_2, and AB_3, it is not met for pairs B_1B_2, B_1B_3, B_2B_3, etc. To illustrate this, Fig. 12 also shows the locus of points such as C_1 and C_2 that meet the condition with point B_1. Accordingly, any 'integral' path constructed as locus of points obeying condition II.3 with a starting composition A is valid only for shocks with that particular composition A on their upstream or downstream side. A procedure commonly employed is (1) to construct an approximate route along (differential) composition paths, (2) check which waves are shocks, and (3) correct the route by use of the algebraic condition II.3 for such shocks [44]. A simpler procedure can be used if all waves are shocks, as in frontal analysis, for example. Each wave then obeys condition II.3, and the compositions of the plateaus in the response pattern can be obtained by the solution of the n respective simultaneous algebraic equations and the constraint of having the shocks in the sequence of increasing velocity. To avoid this distracting difficulty, systems with straight-line paths were chosen here for the examples: Any pair of points on a straight-line path

also automatically obeys the integral coherence condition, so that the entire route of a Riemann problem runs along (differential) paths even if some or all of the waves are shocks.]

8. Wave interference

Without asking as yet when such a phenomenon will occur, let us examine now what will happen if a faster wave finds itself upstream of a slower one. Fig. 13 shows the profile route and distance–time diagram of such a situation, with both waves being shocks. As time passes, the 'fast' shock A→B catches up with the 'slow' shock B→C. At the moment this happens, the two shocks merge into a single concentration discontinuity A→C. The composition variation from A to C across this discontinuity is not along a path, and thus is noncoherent. It is immediately resolved into two new coherent waves A→D and D→C, in exactly the same way as it would be if it had been generated by introduction

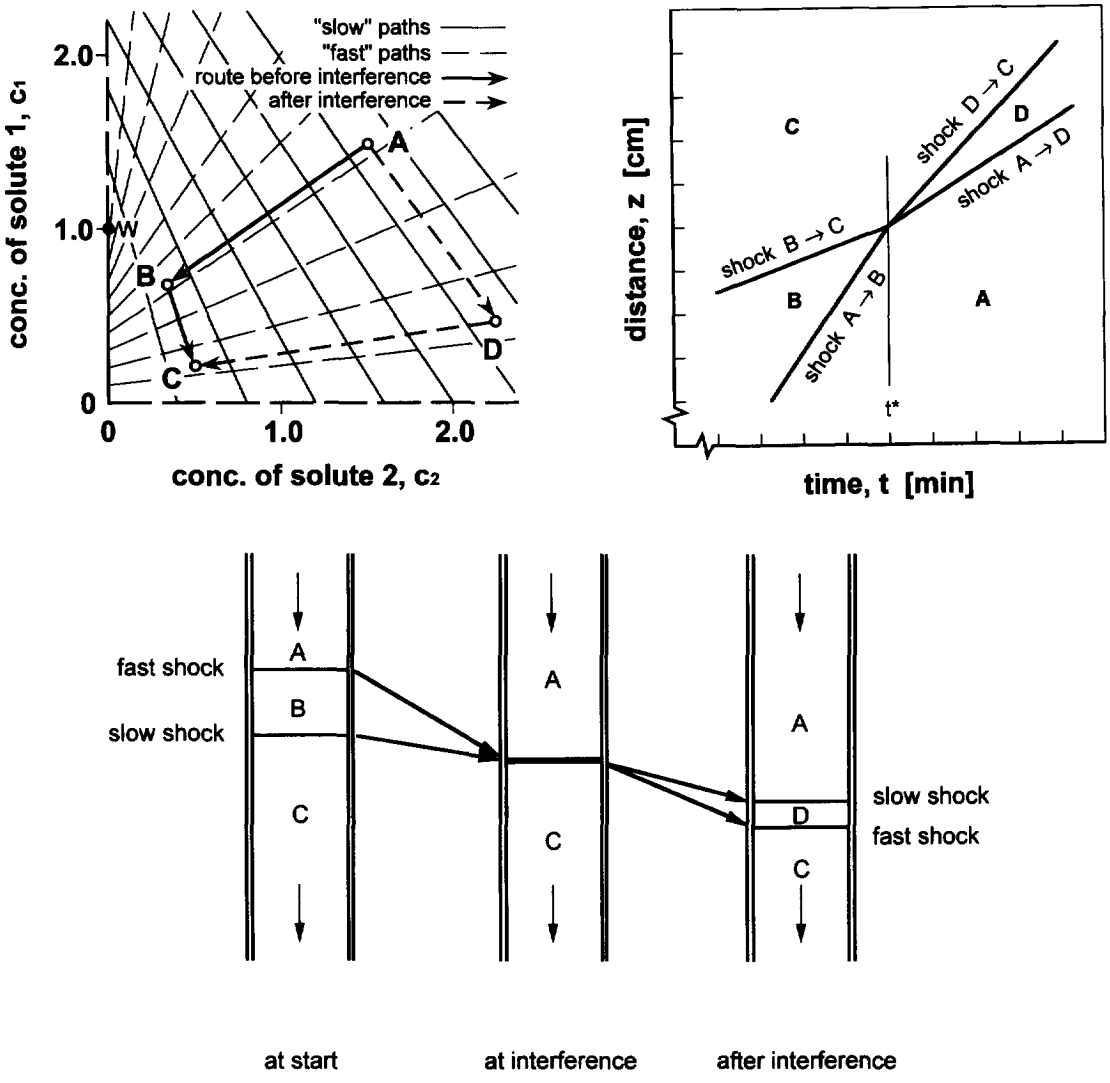

Fig. 13. Interference of two shocks: composition routes before and after interference (top left), distance–time diagram (top right), and resulting column behavior (bottom). t^* = time of interference.

of fluid A into a column initially containing fluid C. After resolution, a 'slow' wave A→D is upstream of a 'fast' wave D→C, as seen in the distance–time diagram. Both new waves are shocks, and a new plateau D grows between them as they pull apart. This example illustrates how the path grid can be used to identify the results of wave interference without further calculation.

The route construction in Fig. 13 yields quantitative information on the composition variations across the two new waves and on the composition of the new plateau between them, but not on the point in space and time at which the interference occurs. The easiest way to identify that point is with the distance–time diagram. Given the initial positions z_1 and z_2 of the shocks A→B and B→C, respectively, at time t^0 and their velocities, their trajectories in the distance–time field can be plotted (the shock velocity $v_{\Delta c}$ equals the trajectory slope dz/dt in the z,t plane). The intersection of the trajectories, at z^* and t^*, is the point at which the shocks interfere.

Interference of a shock with a diffuse wave is more complex. An example is shown in Fig. 14. Interference occurs along a distance–time curve which corresponds to a shock 'cutting through' a diffuse wave of the other family (a 'fast' shock

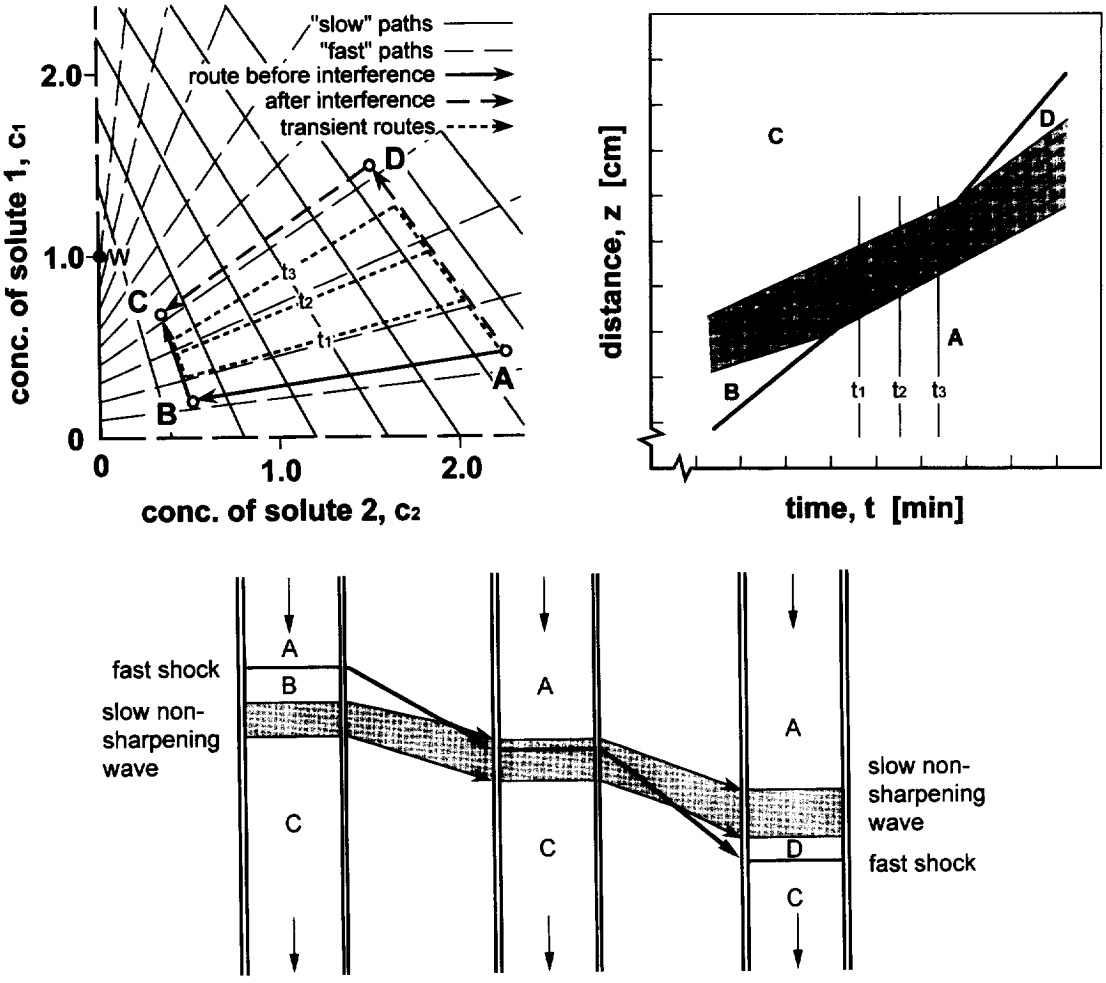

Fig. 14. Interference of 'fast' shock with 'slow' nonsharpening wave: composition profile routes at various times (top left), distance–time diagram (top right), and resulting column behavior (bottom) in two-component system with equilibrium as in Fig. 6. Included are transient routes at times t_1, t_2, and t_3 during interference.

through a 'slow' diffuse wave in the example chosen here). At any moment during interference, the shock meets the coherence condition II.3 although the compositions on its upstream and downstream sides change continuously. During interference, the shock has a portion of the original 'slow' diffuse wave on one side, a portion of the new on the other. As time progresses, the new portion increases at the expense of the old one until the latter has disappeared entirely and interference is complete. In the composition space, the route of the shock shifts gradually and continuously from A→B to D→C. As in interference of two shocks, use of the path grid immediately yields the new coherent route A→D→C from the old route A→B→C. Quantitative identification of the curve along which interference occurs, and thus of the starting points of the new composition and shock trajectories, requires a relatively simple integration along that curve [45].

Interference of two diffuse waves, shown in Fig. 15, is still more complex. Here, a finite distance–time region of noncoherence is generated within which the route changes continuously and does not follow any path, and which therefore contains no composition trajectories. Again, the new route A→D→C can be obtained from the original route A→B→C with use of only the path grid. That is, the composition variations across the new waves and the composition of the plateau in between can be predicted from the grid alone. However, quantitative calculation of the behavior within the region of noncoherence and of the exact locations of the borders of the latter, and thus of the starting points of the new coherent composition trajectories, requires a more lengthy numerical integration over two variables: distance and time.

At least for Langmuir and Langmuir-like systems, an important general rule can be stated for all interferences of waves:

- The sharpening behavior of waves is preserved upon interference.

That is, both the 'slow' and the 'fast' wave remain self-sharpening or nonsharpening, as the case may be. This is true although they change their routes and their positions relative to one another.

The reader interested in further detail and quantitative treatment is referred to Section 3.VI and Appendix II of Ref. [2] and to work of Rhee et al. [32,33].

[In systems with variance 3 or higher, interference of two waves of different families produces two new waves of the same families as the original ones if the path grid is 'orthogonalizable' (see parenthetical comment in Section 6). In the composition space, the entire interference then remains confined to a common surface of mutually intersecting paths of these two families (see Fig. 7, left). If the grid is not orthogonalizable, waves of other families are also produced.]

9. Non-Riemann systems

In a non-Riemann system the route is in general no longer unique, but varies with time or position in the column. Moreover, profile and history routes no longer necessarily coincide with one another. Nevertheless, route construction provides an excellent key to conceptual understanding of the complexities of response behavior under other than Riemann-type conditions and a powerful tool for its quantitative prediction.

As a first example, Fig. 16 shows successive profile routes and concentration–distance–time behavior for elution of a pulse of a two-component mixture under conditions of concentration and volume overload [30,32,34,46–50]. Until waves begin to interfere, this case can be viewed as a combination of frontal analysis and elution from a saturated column: The start of mixture injection generates a frontal analysis pattern, the switch to injection of eluting solvent or carrier gas generates a pattern as in elution from a uniformly saturated column. Introduction of the mixture of solutes 1 and 2 (composition F), at $t = 0$, into the column free of sorbates (composition I) produces two shocks F→A and A→I and a plateau of pure solute 2 (composition A) growing in between. Start of elution with pure solvent or carrier gas (composition I), at time t', generates two nonsharpening waves I→B and B→F with a plateau B of pure solute 1 growing in between. As the distance–time diagram shows, the first interference to occur is that of the 'fast' wave B→F generated at $t = t'$ with the 'slow' shock F→A generated at $t = 0$. It produces two new waves B→I

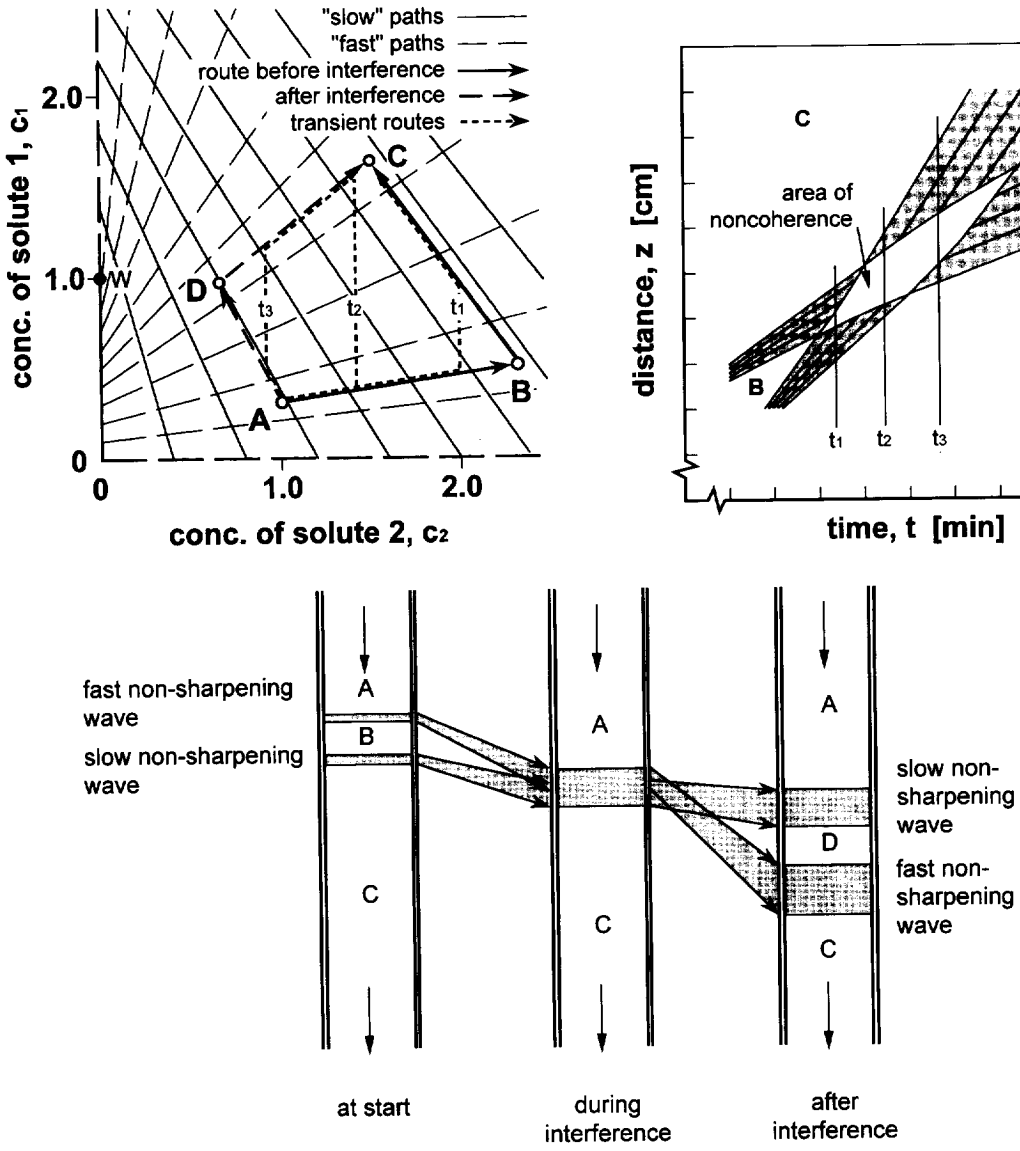

Fig. 15. Interference of two nonsharpening waves: composition profile routes at various times (top left), distance–time diagram (top right), and resulting column behavior (bottom) in two-component system with equilibrium as in Fig. 6. Included are transient routes at times t_1, t_2, and t_3 during interference.

('slow' shock) and I→A ('fast' and nonsharpening wave) and effects resolution into two pure-component pulses with shock fronts and nonsharpening rears. Involving a diffuse wave, this interference requires a finite time, during which the route of the shock gradually shifts from F→A toward B→I. It may or may not be complete by the time the leading edge of the new 'fast' wave I→A catches up with the old 'slow' shock A→I. In the case shown in Fig. 16 it is (see distance–time diagram). Accordingly, the pulse of component 2 still has its flat top at the instant t^* of resolution, and the route at that time is I→B→I→A→I. As time progresses, both pulses keep flattening out, even granted the assumptions of ideal chromatography (see the discussion of single-component overload elution in Section 12 of Part I

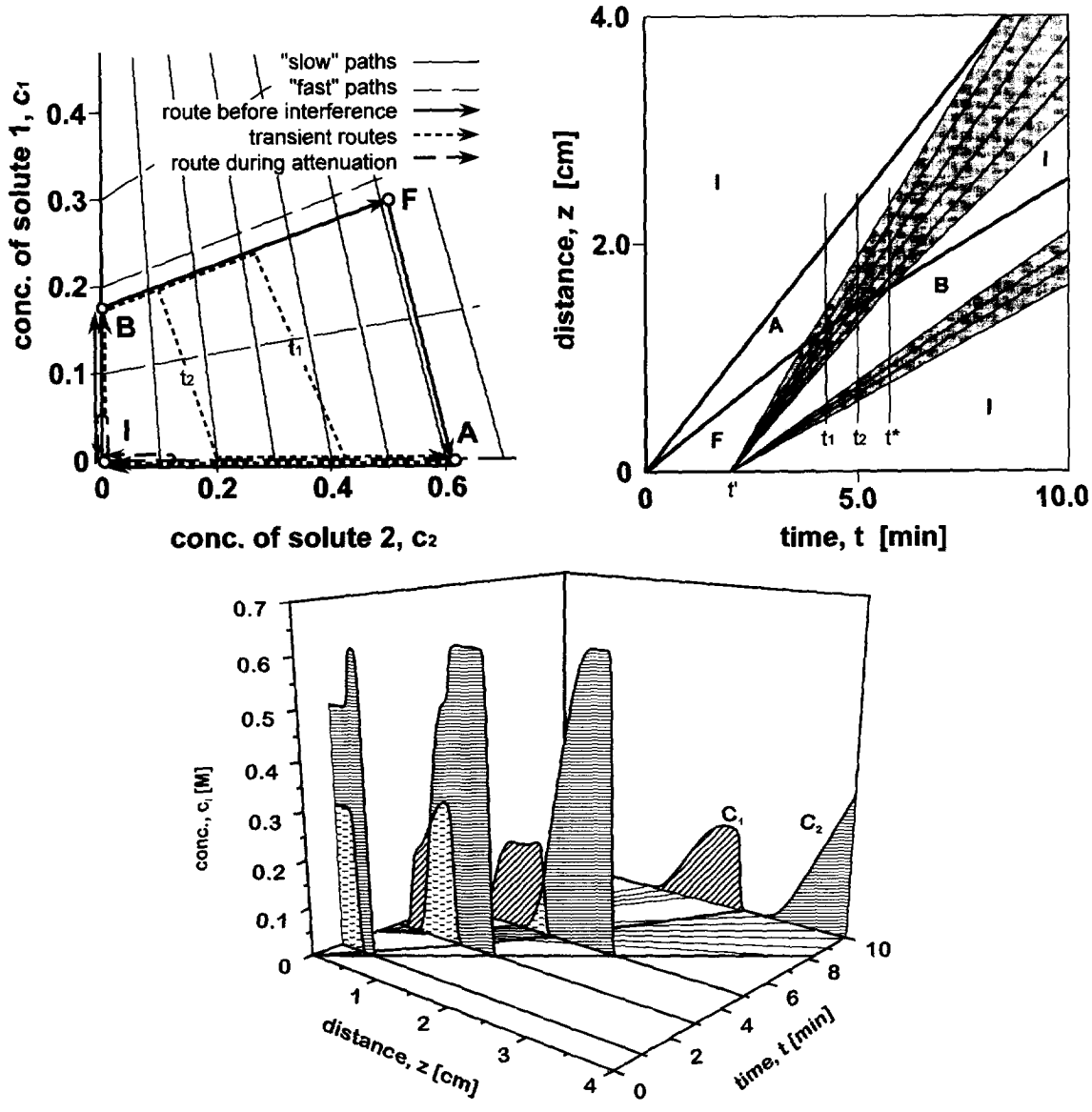

Fig. 16. Elution under overload conditions: profile routes (top left), distance–time diagram (top right), and successive composition profiles (bottom) for elution of two-component mixture F with solvent or carrier gas. Transient routes during interference shown at times t_1 and t_2; t^* = resolution time. Parameter values $a_1 = 6.0$, $a_2 = 3.0$ (both in cm^3/g), $b_1 = 1.0$, $b_2 = 0.5$ (both in cm^3/mmol) in Langmuir equation $q_j = a_j c_j/(1 + \sum_i b_i c_i)$; $v^0 = 3.1831$ cm/min; $\rho/\epsilon = 2.5$ g/cm^3; mixture $c_1 = 0.3$ M, $c_2 = 0.5$ M injected during $0 \le t \le 2.0$ min; t^* = time of resolution.

[1]; in terms of physics, the reason is that the velocity of a shock is an average of the eigenvelocities of the concentrations on the two sides of the shock, so that the velocity of the shock front of a pulse necessarily is lower than the eigenvelocity of the pulse maximum).

At a smaller separation factor $\alpha_{12} \equiv a_1/a_2$ or higher solute concentrations, the flat top A of component 2 is apt to disappear before resolution. The route at the moment t^* of resolution would then be I→B→I→X→I, where X is a point between A and I on the c_2 axis.

This example demonstrates the ease with which the path grid and distance–time diagram allow the evolution of the pattern and the newly arising compositions to be predicted, given the compositions of the column at start and of the entering streams. How this can be done in a quantitative manner will be shown in the next section.

For comparison, Fig. 17 shows an experimentally observed effluent composition history of an elution under overload conditions and with a column too short for resolution [49].

The pulses in Figs. 16 and 17 have sharp fronts and diffuse rears. Such behavior is found in systems with competitive sorption equilibria, including Langmuir and Langmuir-like systems, and is caused by negative isotherm curvatures (see Section 12 in Part I for effect of isotherm curvature on peak shape). Other types of isotherms give different peak shapes. For example, the sorption isotherms of electrolytes on ion exchangers in so-called ion exclusion processes have opposite curvature [51,52] and produce pulses with diffuse fronts and sharp rears. This is true in multicomponent as well as single-component systems.

The next examples extend the construction to displacement development of a mixture F1 of solutes 2 and 3 by a development agent F2 consisting of solute 1 of higher affinity. The column is assumed to be long enough for attainment of the final pattern. In displacement development, discussed in Section 13 of Part I, a large amount of mixture is introduced into a solute-free column and is then displaced with an agent of higher affinity for the sorbent. If totally effective, displacement eventually results in a pattern with pure-component bands of uniform concentrations, in the sequence of affinities, separated by shocks, and all advancing at the same velocity. Under real conditions, some overlap remains because the shock layers have a finite, though small thickness. Granted the assumptions of ideal chromatography, the shocks are concentration discontinuities and resolution thus becomes complete. Reference to 'complete' resolution in the following is to be understood in this way.

Fig. 18 shows a classic case of total displacement. Introduction of the mixture F1 generates the profile route F1→A→I with two shocks F1→A ('slow') and A→I ('fast') and a plateau A of pure solute 3 growing in between. This is a frontal analysis pattern. Subsequent introduction of the development agent F2, started after 2 min, generates three new waves and lets the route become F2→D→B→F1→A→I, with D being a plateau of pure solute 2, and B a mixture of solutes 2 and 3 different from F1. The new, upstream portion F2→D→B→F1 is as would develop under Riemann conditions with F2 entering a column initially containing F1 uniformly over its entire length. In Fig. 18 the concentration of the development agent is high enough for all three waves in that portion of the pattern to be shocks. In time, the fastest of these shocks, B→F1, catches up with the 'slow' shock F1→A generated earlier. Wave interference at $t = t^*$ produces two new coherent shocks B→C ('slow') and C→A ('fast'). The 'slow' shock is now upstream of the 'fast' one. The new plateau C that grows between the two contains pure solute 3 at higher concentration than in A. Since the eigenvelocity decreases monotonically from D to C and from C to I, the 'slow' shocks D→B and B→C eventually merge into a single shock D→C, and so do the 'fast' shocks C→A and A→I into a single shock C→I. Only the distance–time diagram can tell which merger occurs first. After both mergers, the new and final route is F2→D→C→I. All its waves are shocks, and with disappearance of the plateau B at time t^* the separation of solutes 2 and 3 is complete. Note that solute 1 is present only upstream of the slowest wave F2→D, a shock; this allows the entire route construction to be conducted in the c_2, c_3 plane once the position of point D has been established.

For comparison, Fig. 19 shows the observed effluent history of a separation achieved under conditions of total displacement, as in Fig. 18 [53].

In Fig. 20 the mixture F1 is the same as in Fig. 18, but the concentration F2 of the development agent is chosen lower so that the new 'fast' wave B→F1 is

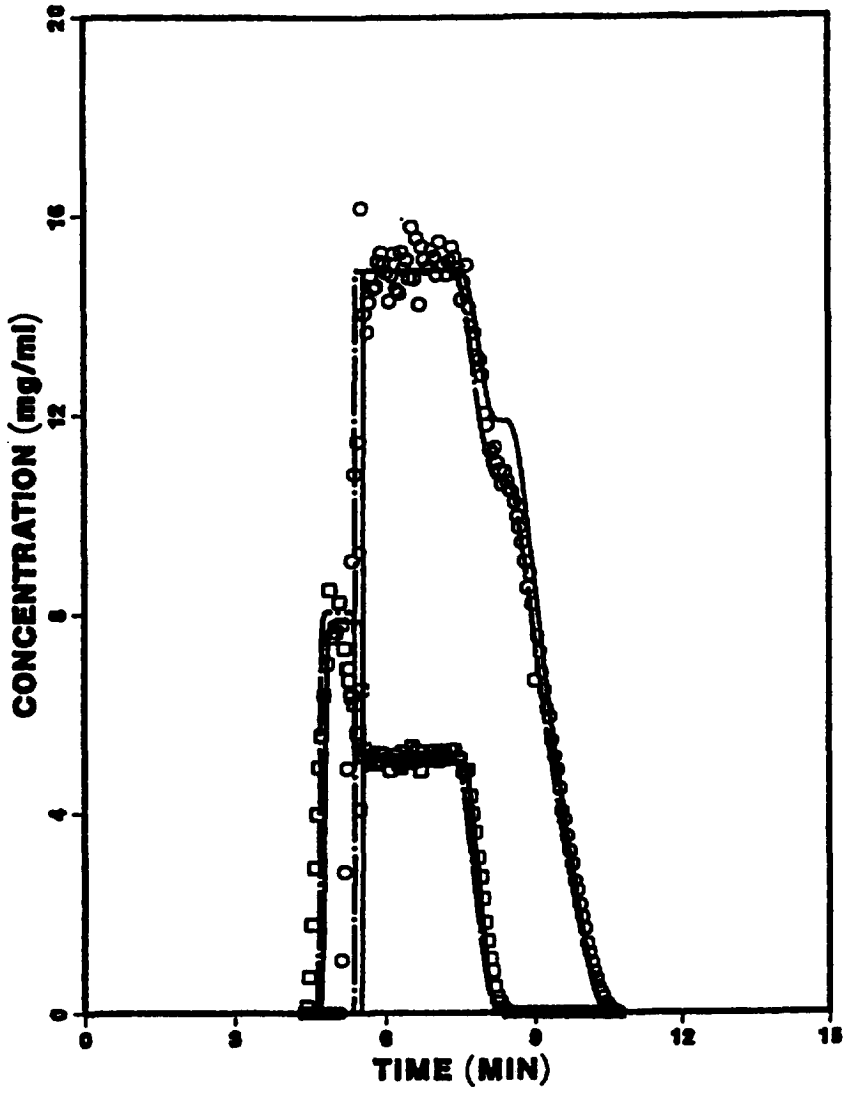

Fig. 17. Observed effluent composition history of elution of a two-component mixture under conditions of concentration and volume overload in column intentionally selected too short for complete resolution. Separation of 2-phenyl-ethanol (□) from 3-phenyl-propanol (○) on ODS silica; curves calculated with model including mass-transfer resistance (from Katti et al., [49] reproduced with permission of AIChE J.).

nonsharpening rather than a shock. Interference of the 'fast' nonsharpening wave B→F1 with the 'slow' shock F1→A moves the route of the latter gradually over toward B→C. This interference may or may not be complete before the plateau A disappears. If it is, the route at the time of its completion becomes F2→D→B→C→A→I. If it is not, the route at that time becomes F2→D→B→C→X→I, with X being a point on the c_2 axis between C and A. The segment C→A→I (or C→X→I) is a pulse of pure component 2 with shock front and nonsharpening rear. This pulse attenuates and eventually disappears completely, leaving only a shock C→I. That happens because the shock velocity of its front is always lower than the eigenvelocity of the maximum concentration, even when the latter approaches C. As in the

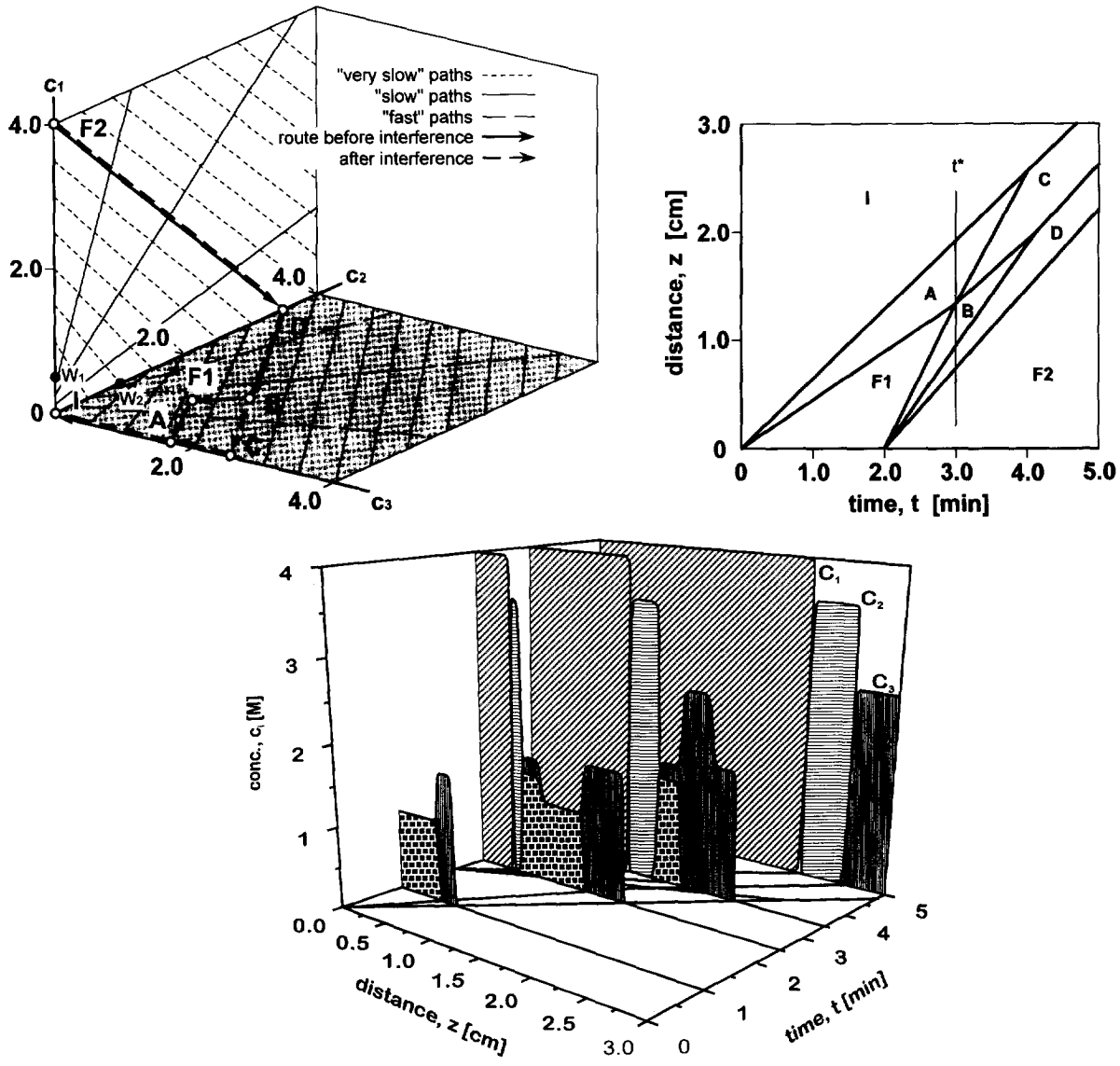

Fig. 18. Complete displacement development: profile routes (top left), distance–time diagram (top right), and successive composition profiles (bottom) for displacement of two-component mixture of solutes 2 and 3 (F1) by solute 1 (F2) at high concentration. Routes shown on path grids of c_1, c_2 and c_2, c_3 planes (latter shaded). t^* = resolution time. Parameter values $a_1 = 12.0$, $a_2 = 6.0$, $a_3 = 3.0$ (all in cm^3/g), $b_1 = 2.0$, $b_2 = 1.0$, $b_3 = 0.5$ (all in cm^3/mmol) in Langmuir equation $q_j = a_j c_j/(1 + \sum_i b_i c_i)$; $v^0 = 3.1831$ cm/min; $\rho/\epsilon = 2.5$ g/cm^3; mixture $c_2 = 1.0$ M, $c_3 = 1.0$ M injected during $0 \le t \le 2.0$ min.; developer $c_1 = 4.0$ M injected during $t \ge 2$ min.

previous example, the shocks D→B and B→C merge at some time, completing resolution. This may happen before or after the pulse C→A→I (or C→X→I) has degenerated into the shock C→I. In either case, the final route is F2→D→C→I, all waves being shocks, as in the previous case. Note, however, that resolution may well be complete earlier, and if so, that the pattern at that time is more favorable than the final one because component 2 then is still concentrated in a narrower band (nonuniform concentration with maximum, corresponding to point X, at its tail end). Such a situation

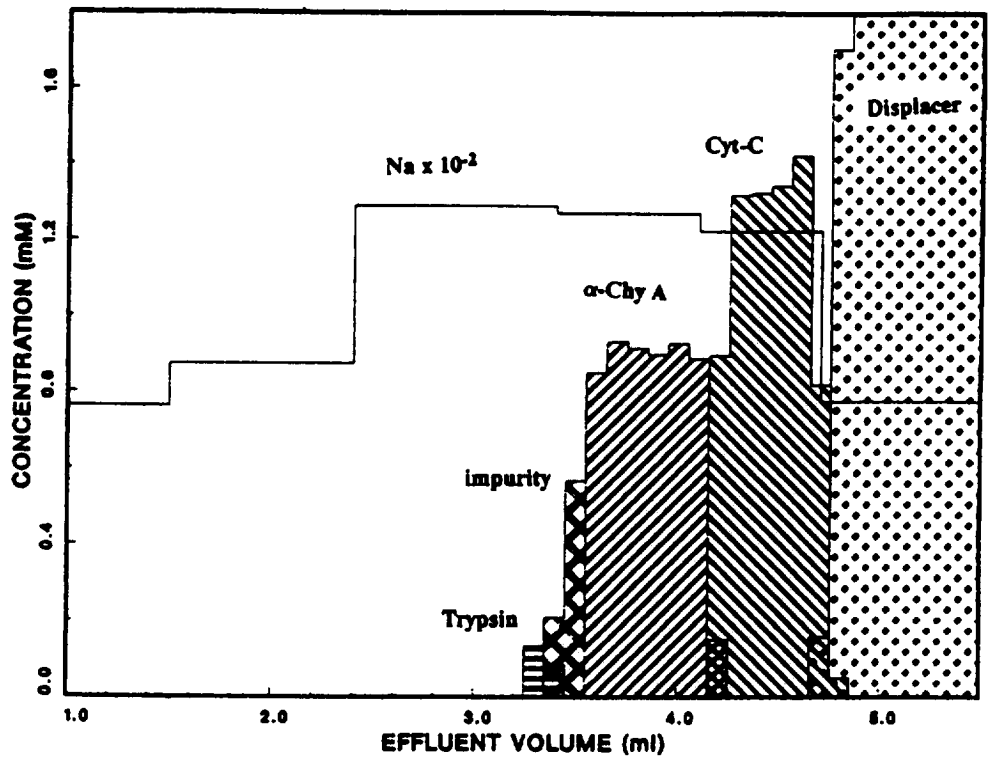

Fig. 19. Observed effluent composition history of protein separation (α-chymotrypsinogen A, cytochrome c, two impurities) on strong-acid cation exchanger by development with dextran polyelectrolyte displacer at concentration high enough for total displacement (from Jayaraman et al., [53] reproduced with permission of J. Chromatogr.).

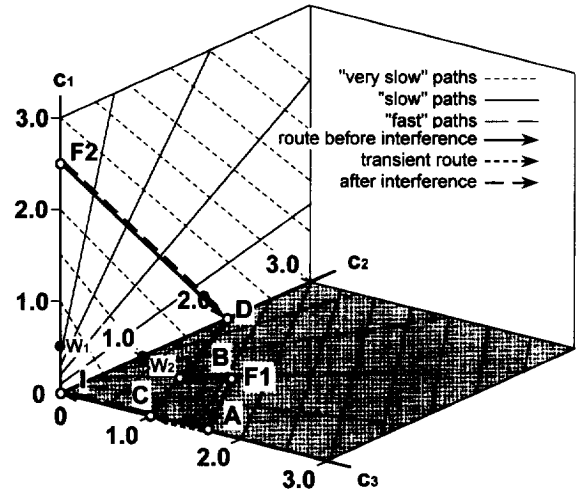

Fig. 20. Displacement development with resolution before attainment of final pattern: profile routes for displacement as in Fig. 18 but with lower concentration of developer, $c_1 = 2.5$ M. Routes shown on paths grids of c_1, c_2 and c_2, c_3 planes (latter shaded).

has been discussed in more detail by Gadam et al. [54].

For comparison, Fig. 21 shows the observed and simulated effluent history of a separation achieved under conditions corresponding to those in Fig. 20. (The simulation is based on numerical integration of the differential mass-conservation equations without recourse to wave theory.)

Figs. 22 and 23 illustrate the chromatographic response with the same mixture F1, but with developer concentrations too low to achieve the normal displacement pattern. In Fig. 22 the developer concentration is such that the 'very slow' path originating from F1 leads to a point D below the so-called watershed point W_2 on the c_2 axis. Development now produces a route F2→D→B→F1→A→I with B on that axis. Interference of the 'fast' nonsharpening wave B→F1 with the 'slow' shock F1→A requires a finite time and, at $t = t^*$, gives resolution. This interference may or may not be complete before the

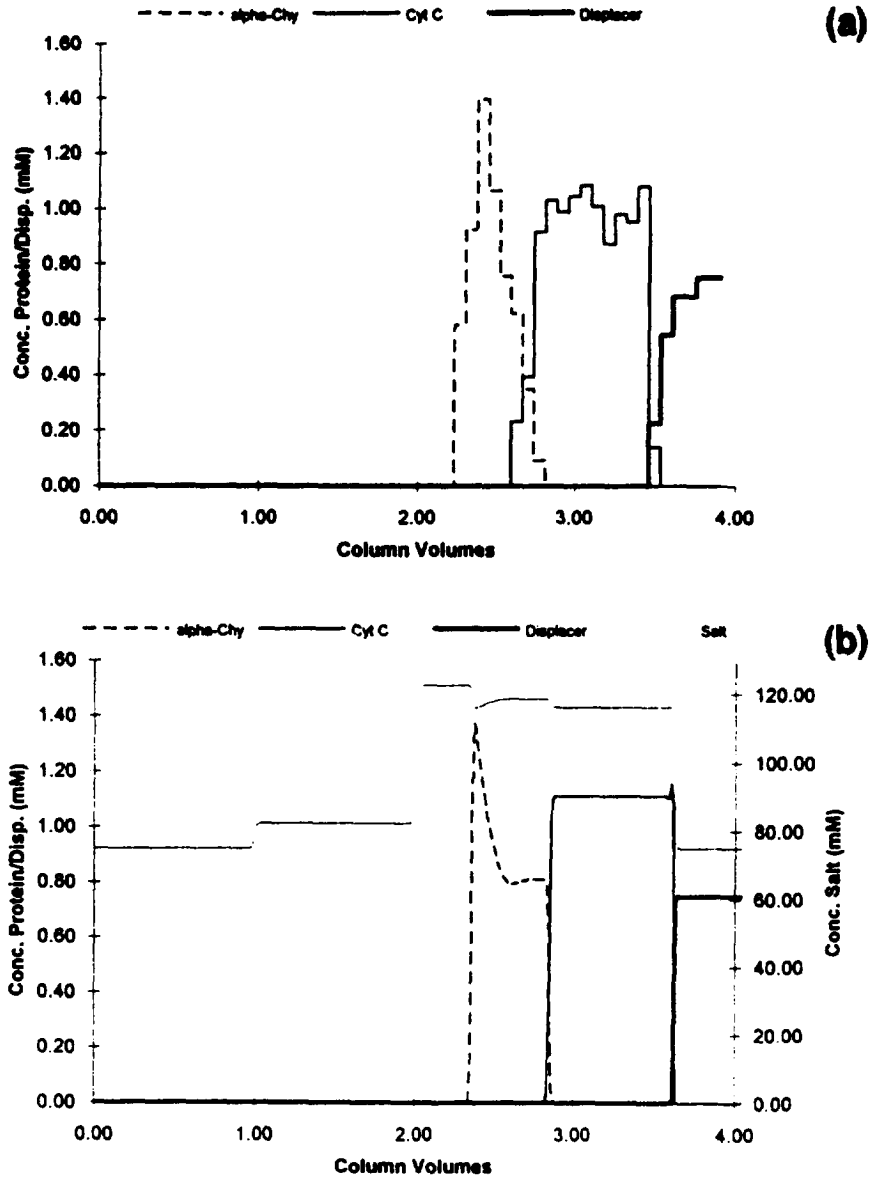

Fig. 21. Observed (top) and simulated (bottom) effluent histories of protein separation (α-chymotrypsinogen, cytochrome c) on strong-acid cation exchanger by development with phosphate-buffered DEAE-dextran displacer of pH 6 under conditions giving complete resolution before attainment of final pattern (from Gadam et al., [54] reproduced with permission of AIChE J.).

leading edge of the new 'fast' nonsharpening wave catches up with the also 'fast' shock A→I and makes the flat top A disappear. In Fig. 22 it is not, and the route upon resolution is F2→D→B→I→X→I, where X is a point on the c_3 axis between A and I. The portion I→X→I (or I→A→I) of the pattern is an attenuating pulse with shock front and nonsharpening rear, as in elution under overload conditions. The shocks D→B and B→I eventually merge into a single shock B→I. The overall effect of the lower concentration of the developer is that the latter still displaces solute 2, but is too slow to displace solute 3. As a result, that solute 'runs away' from the displacement pattern in form of an overload peak.

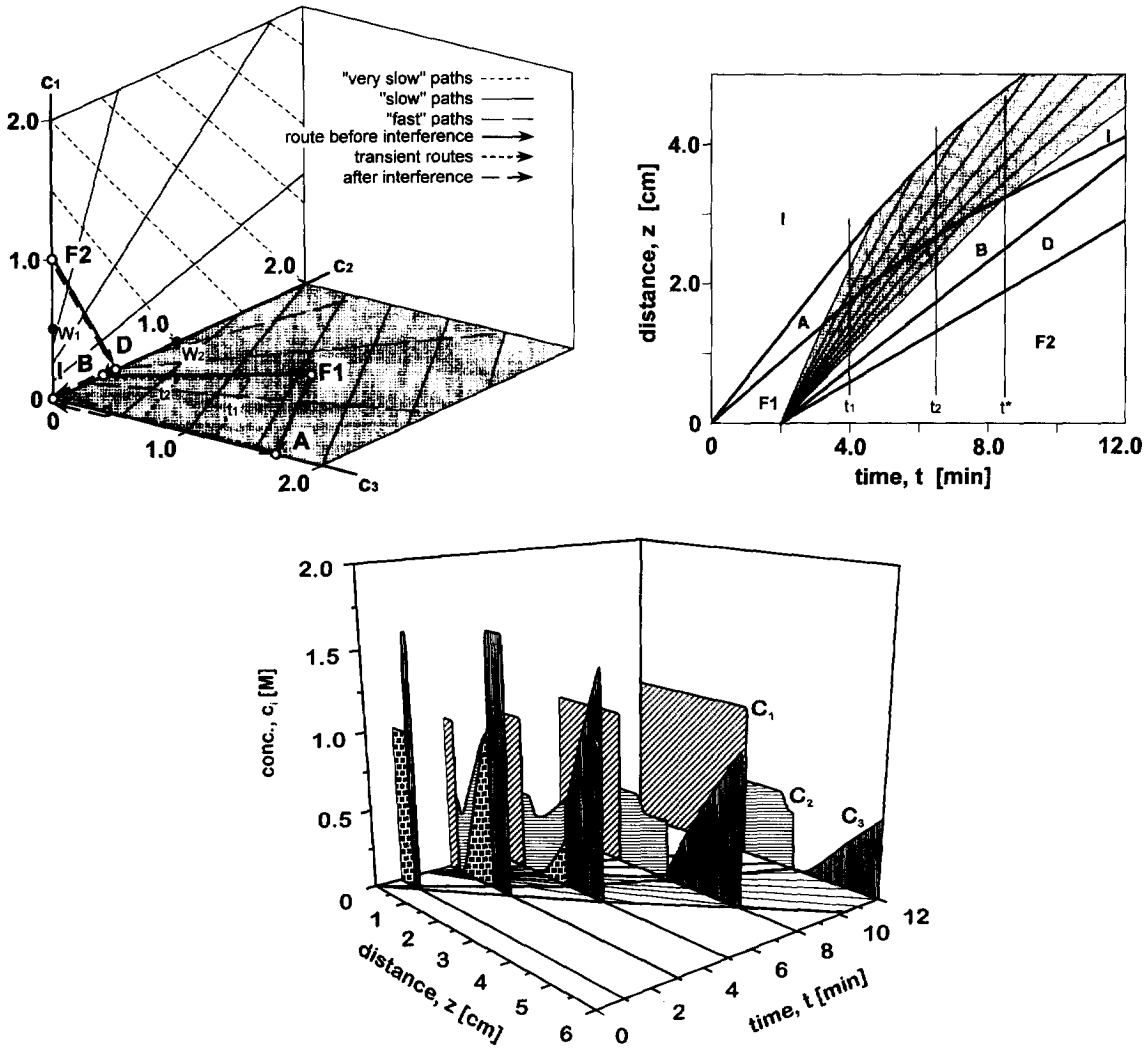

Fig. 22. Incomplete displacement development: profile routes (top left), distance–time diagram (top right), and successive composition profiles (bottom) for displacement as in Fig. 18 but with developer at $c_1 = 1.0$ *M*, too dilute to displace solute 3. Routes shown on paths grids of c_1, c_2 and c_2, c_3 planes (latter shaded). Transient routes during interference shown at times t_1 and t_2; t^* =resolution time.

Fig. 24 shows an experimental effluent history observed under such conditions [55].

In Fig. 23 the developer concentration is even lower, namely, below the watershed point W_1 on the c_1 axis. Here, development produces a route F2→I→B→F1→A→I. The front of the developer now is so slow that both components of the mixture 'run away' from it. No displacement occurs. Rather, the mixture F1 is eluted in exactly the same way as a pulse under overload conditions, as though only pure solvent or carrier gas were introduced into the column instead of a displacement agent. The behavior is the same as in Fig. 16, except for the difference in composition of the original mixture.

The situation in Figs. 22 and 23 corresponds to what has been alluded to in the discussion of final displacement patterns in Section 13 of Part I: The single-component isotherms are not intersected by the straight line connecting the q_i, c_i point of the development agent with the origin, $q_i = 0$, $c_i = 0$ ('Tiselius line'). This can be seen in Fig. 25, in which the single-component isotherms of the three

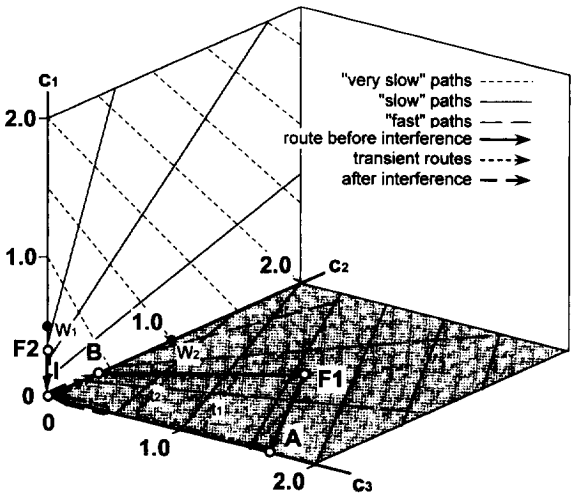

Fig. 23. Ineffective displacement development: profile routes for separation as in Fig. 18 but with developer at $c_1 = 0.3$ *M*, too dilute to displace either solute. Routes shown on paths grids of c_1, $c_2$and c_2, c_3 planes (latter shaded); transient routes during interference at times t_1 and t_2.

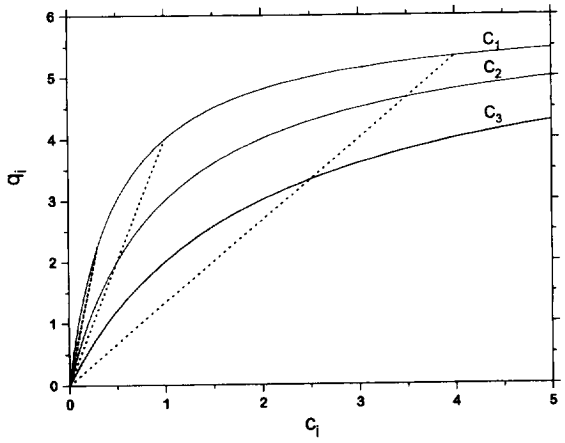

Fig. 25. Single-component isotherms of solutes as in Figs. 18, 20, 22, and 23, with Tiselius lines for developer at 4.0, 1.0, and 0.3 *M*.

solutes are shown with the Tiselius lines for the operations in Figs. 18, 22 and 23 (see also Fig. 16 in Part I).

Note that a knowledge of the path grid allows the

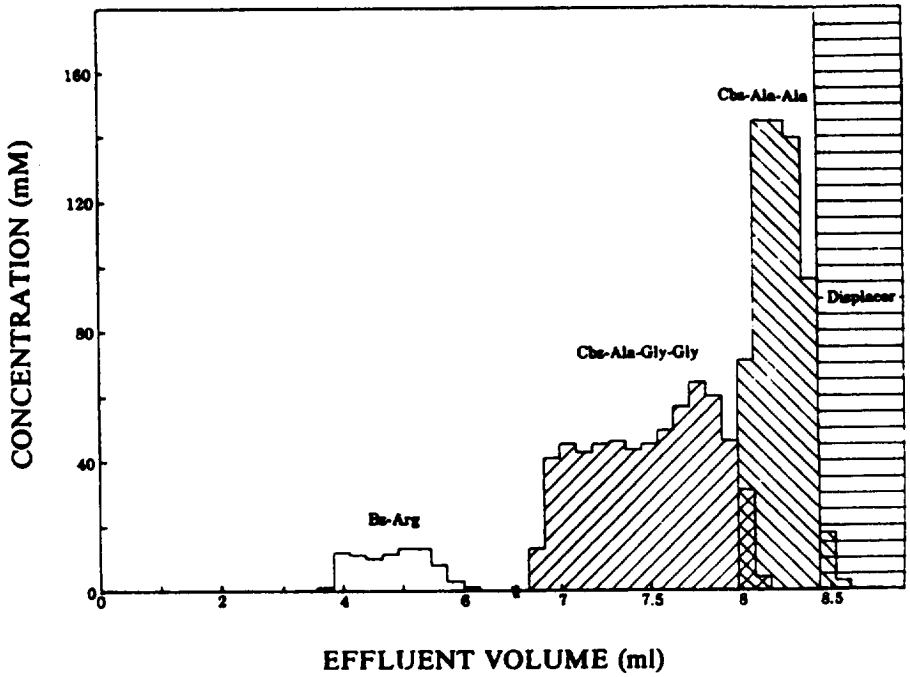

Fig. 24. Observed effluent composition history of peptide separation with developer concentration too low to displace solute of lowest affinity. Column: Zorbax ODS; carrier: 40% methanol in 50 m*M* phosphate buffer pH 2.2; displacer: 30 mg/ml BEE. (From Subramanian et al., [55] reproduced with permission of J. Chromatogr.)

type of development behavior to be predicted without further calculation: Normal behavior results if the path from D across the c_2, c_3 plane cuts through compositions with higher concentrations than those of the mixture; a normal final pattern but earlier resolution with a still nonuniform band of solute 3 results if that path cuts through lower concentrations than those of the mixture, but D itself is higher up on the c_2 axis than the watershed point W_2; solute 3 'runs away' if D is below W_2 on that axis; and no displacement of either solute occurs if the point F2 of the displacer is below the watershed point W_1 on the c_1 axis.

Regardless of the number of solutes, the final pattern of a normal displacement development can be found much more easily by the Tiselius construction [10] shown in Fig. 16 in Part I and Fig. 25. However, route construction is needed to identify the compositions of the transient plateaus, as may be necessary if the separation of two solutes not adjacent in the final pattern is of interest. Also, route construction in combination with the distance–time diagram provides information on the column lengths and development times needed to separate any pair of solutes. Equations for calculation of these requirements for mixtures with arbitrary number of components have been developed for ion exchange with constant separation factors [56] and adsorption with Langmuir isotherms [57,58] based on the mathematics to be shown in Part III of this series. Note also that separation may be complete before the final pattern exclusively with flat-top bands has been attained, and that in such cases the pattern at the moment of complete resolution is more favorable than the final one given by the Tiselius construction (see Figs. 20 and 21) [54].

Ion-exchange displacement development [56,59] is simpler: No counterion can 'run away' from the train of others because pure solvent produces no movement of counterions in the direction of flow. The pattern generated thus always is of the type shown in Fig. 18, except the concentrations in all bands of the final pattern are the same as that of the displacing counterion.

The distance–time diagrams of non-Riemann systems may contain finite distance–time regions of noncoherence, produced by interference of diffuse waves (see Fig. 15), by gradual variation of the composition of the entering solution, or by an initial composition profile with gradual variation (see Fig. 26). Any composition within such a noncoherent area exists only at one point in distance and time, and profile and history routes necessarily involve different sequences of such transient compositions. The distance–time diagrams in Figs. 15 and 26 suggest that noncoherence extending over finite distance–time regions can be viewed as a superposition of coherent waves that have not yet separated

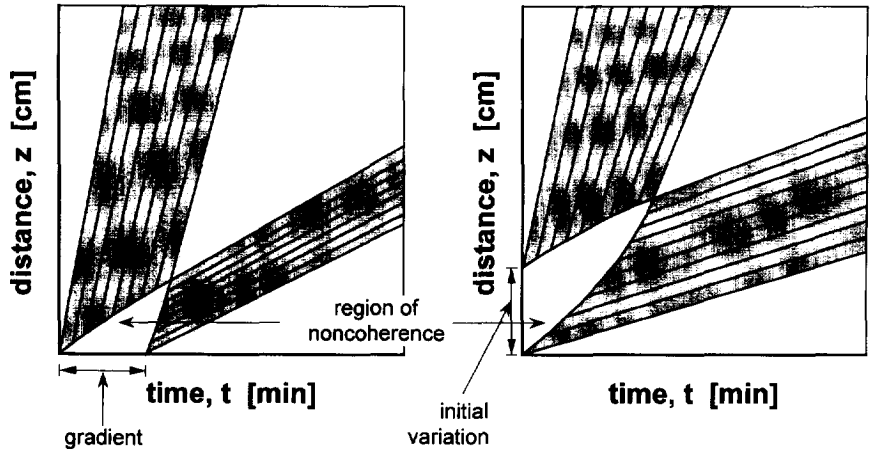

Fig. 26. Finite distance–time regions of noncoherence produced by gradual variation of composition of entering fluid (left) and nonuniform initial composition of column (right), for system with variance 2 producing two nonsharpening waves (schematic).

from one another. This idea helps in forming a qualitative picture of events, although it is of no use in quantitative calculations because the superposition is not additive.

In so-called stratified beds with sorbent layers of different equilibrium properties, each layer has its own, different path grid. Accordingly, when a coherent wave enters a new layer, it becomes noncoherent because its route does not coincide with a path of the new grid. The wave thus breaks up into a set of new coherent waves. If the original wave was diffuse, noncoherence extends over a finite distance–time region because the noncoherent composition variation at the entry to the layer extends over a finite time.

[If the operation involves two or more interferences of waves of different families, one of them may occur earlier in time than another, but at a longer distance from the column inlet. If so, the transient coherent history and profile routes between the two interferences run along different composition paths. For successive histories (composition as function of time), the route switch caused by the earlier interference comes first; for successive profiles (composition as a function of distance), the switch caused by the later interference comes first because it occurs nearer the inlet.]

10. Quantitative construction of column response

The composition route, in graphical or mathematical form, gives complete quantitative information about the sequence of compositions in the column, but does not indicate when and where the respective compositions exist. The distance–time diagram supplies that missing information. The combination of the two provides a complete, compact, quantitative description of the column response. From that description, concentration profiles and histories can be generated as may be desired. In essence: The route diagram provides the concentrations and shocks that will occur. For a concentration profile at time t', the distance–time diagram gives the distances z of selected concentrations or shocks as the intersections of their trajectories with the line $t = t'$; similarly, for a concentration history at $z = z'$, it provides the times t as the intersections of the respective trajectories with the line $z = z'$.

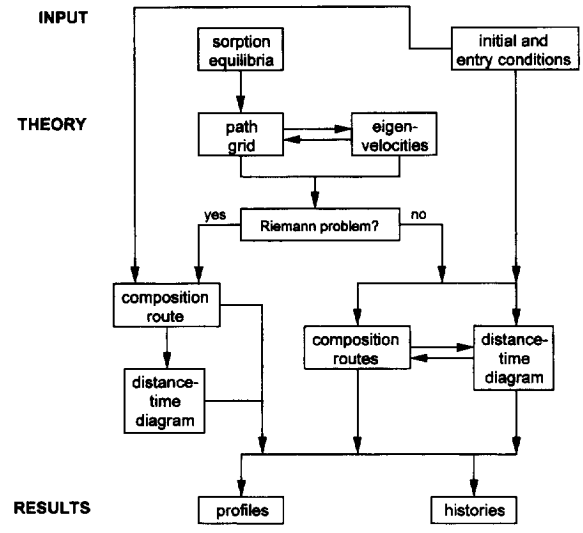

Fig. 27. Flow sheet for quantitative construction of column response.

Fig. 27 shows a flow sheet for such purposes. The first step is to construct the path grid, either as a graph or in abstract mathematical form. If this is done step by step by repeated solution of the eigenvalue problem, as described in Section 6, eigenvelocities have been generated in the process and should be noted. If a short-cut method has been used, eigenvelocities are calculated at a number of grid points. The procedure from here on depends on whether this is or is not a Riemann problem.

10.1. Riemann problems

If this is a Riemann problem, its composition route is established with the sequence rule given in Section 7. Next, its distance–time diagram is constructed with the relevant eigenvelocities of compositions within any nonsharpening waves, and shock velocities of any shocks (the velocity dz/dt is the slope of the respective trajectory in the z,t plane); this is done with Eq. I.6 for shocks and with Eq. II.4 for the eigenvelocities of the respective compositions and paths for nonsharpening waves.

Fig. 28 illustrates a convenient procedure of constructing the concentration history of a solute i at a given location z' with use of the distance–time diagram (z' may or may not be the column exit). A horizontal line corresponding to $z = z'$ is entered into the distance–time diagram. To construct the history, the times t at which trajectories intersect the $z = z'$ line are transposed to the abscissa of the history diagram below. For any composition trajectory, the time t of intersection is the time at which the concentration c_i of that composition has reached location z'; the respective c_i value, read off the route diagram or retrieved from the set $\{c_1, \ldots, c_n\}$ that was used to calculate the eigenvelocity, is plotted at t in that diagram. Similarly, for any shock trajectory, the time t of its intersection with the line $z = z'$ is the time at which the shock has reached the location z'; the upstream and downstream concentrations of i at the shock are given as the two c_i values used in Eq. I.6 and are plotted at t in the history diagram. In this way the complete history $c_i(t)$ at z' is compiled point by point.

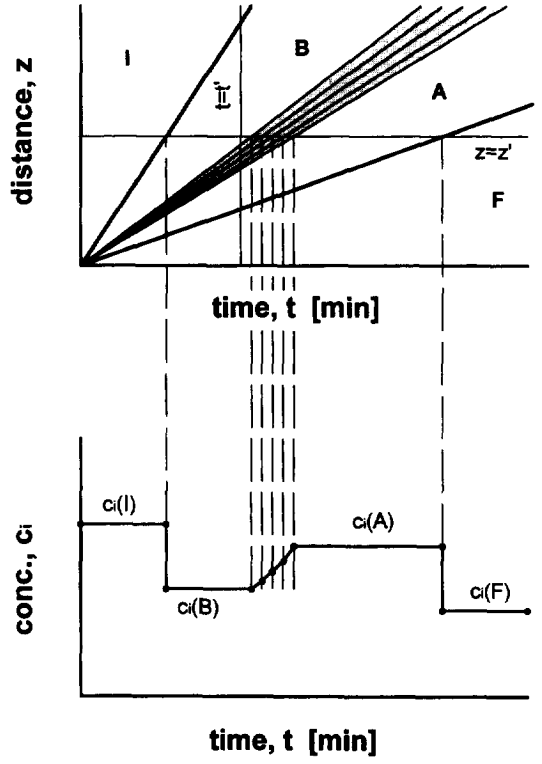

Fig. 28. Use of distance–time diagram for quantitative construction of concentration histories, for trivariant Riemann system with two shocks and one nonsharpening wave in history route I→B→A→F. History at location z' compiled from trajectory intersections with line $z = z'$.

The construction of the concentration profile of a solute i at given time t' is analogous. Here, a vertical line corresponding to $t = t'$ is entered into the distance time diagram, intersection distances z of trajectories with that line are transposed to the abscissa of the $c_i(z)$ profile diagram being constructed, and the respective values of c_i are obtained for plotting as above. (If, for convenience, direct graphical transposition as in Fig. 28 is used, the profile diagram appears to the right or left of the distance–time diagram and is tilted 90°.)

10.2. Non-Riemann problems

The basic procedure is the same for non-Riemann problems, with one exception. The profile or history route in general varies with time or distance, respectively, and this calls for a construction in which the distance–time diagram and the route diagram are used in tandem.

The case shown in Fig. 18 may serve as example. Known are the compositions I, F1, and F2 and the time span $t = 0$ to $t = t'$ over which the introduction of F1 extends. The path grid is constructed, and from it the intermediate compositions A, B, C, and D are established as explained in Section 9. For construction of the distance–time diagram, composition I is entered along the z axis, and compositions F1 and F2 along the t axis between 0 and t' and beyond t', respectively. The route diagram before interferences consists of two quasi-Riemann portions F2→D→B→F1 and F1→A→I, all waves being shocks. Their velocities are calculated with Eq. I.6. This allows their trajectories, originating from the points ($z = 0$, $t = 0$) and ($z = 0$, $t = t'$), to be plotted in the distance–time diagram (the shock velocity gives the trajectory slope in the z,t plane). The intersection of the B→F1 and F1→A trajectories is the point of the first interference and the origin of

new shock trajectories B→C and C→A. Composition C is found in the route diagram as the intersection of the 'fast' path through A with the 'slow' path through B. With knowledge of A, B, and C the velocities of the new shocks are obtained from Eq. I.6 and their trajectories are plotted. This procedure is repeated for the subsequent mergers of D→B with B→C and of C→A with A→I. The distance–time diagram is now complete and can be used for construction of concentration profiles or histories as described earlier.

11. Prediction without calculation

To this day, wave theory has not been very popular among practicing chromatographers. The principal reason appears to be the expectation that experience in higher mathematics would be needed and complex concepts and techniques would have to be mastered. That is true to some extent for accurate quantitative constructions. However, the basic ideas are very simple: development toward coherence as a 'stable' state, and the coherence condition of equal wave velocities of all solutes within a coherent wave. Once these ideas are accepted, they can often be used to for qualitative predictions of column response with little or no further calculation or graphical construction. This will be illustrated with an example, of pH variations in an ion-exchange column during stepwise elution with a buffer.

Suppose stepwise elution with a sodium-acetate buffer is to be used for a separation of trace amounts of amphoteric substrates such as amino acids or peptides on a strong-base anion-exchange column. Of interest are the pH variations in the column induced by step changes in the composition (total concentration and pH) of the entering buffer.

The wave equation

$$v_{c_i} = \frac{v^0}{1 + (\rho/\epsilon)\mathrm{d}q_i/\mathrm{d}c_i} \qquad \text{(I.4)}$$

is valid for the co-ion Na^+ as well as for the counterions OH^- and OAc^-. Na^+ is the only cation, so its concentration equals the total electrolyte concentration. As a co-ion, it is largely excluded from the anion exchanger by the Donnan effect, so that its concentration q_{Na} on the ion exchanger is and remains close to zero. Accordingly, $\mathrm{d}q_{Na}/\mathrm{d}c_{Na}$ remains very small in any wave that involves a significant variation of its solution concentration c_{Na}, and thus of total concentration. This being the case, Eq. I.4 shows that such a wave travels essentially at the velocity of mobile-phase flow:

$$v_c \cong v^0 \quad \text{for wave with variation of } c_{Na} \qquad \text{(II.5)}$$

If such a 'salinity wave' is coherent, Eq. II.5 applies to the counterions also. Eq. I.4 then requires $\mathrm{d}q_{OAc}$ and $\mathrm{d}q_{OH}$ to remain very small. That is, the salinity wave cannot entail any significant exchange of OAc^- for OH^-.

Where any uptake or release of OAc^- by the ion exchanger in exchange for OH^- occurs, $\mathrm{d}q_i$ and therefore $\mathrm{d}q_i/\mathrm{d}c_i$ are significant for these ions. Here, application of Eq. I.4 shows that any wave involving such exchange is retarded relative to mobile-phase flow:

$$v_c < v^0 \quad \text{for wave with variation of } q_{OAc} \text{ and } q_{OH} \qquad \text{(II.6)}$$

If such an 'ion-exchange wave' is coherent, condition II.6 applies to Na^+ also. Eq. I.4 then is seen to require $\mathrm{d}q_{Na}/\mathrm{d}c_{Na}$ to be significant; because Donnan exclusion keeps $\mathrm{d}q_{Na}$ very small, it follows that $\mathrm{d}c_{Na}$ must also be very small: The wave cannot involve a significant change of total concentration.[4]

· Under Riemann conditions the coherent pattern must have two waves: a fast 'salinity wave' involving the variation of the total concentration but no significant variation of the state of the ion exchanger, and a slower 'ion-exchange wave' involving the exchange of counterions but no variation of total concentration.

[4] If Donnan exclusion is ideal so that $\mathrm{d}q_{Na} = 0$ (or $\Delta q_{Na} = 0$), the coherence condition requiring $\mathrm{d}q_{Na}/\mathrm{d}c_{Na} > 0$ (or $\Delta q_{Na}/\Delta c_{Na} > 0$) in that wave is satisfied with $\mathrm{d}c_{Na} = 0$ (or $\Delta c_{Na} = 0$), so that the derivative (or finite-difference ratio) becomes indefinite.

The salinity wave is indifferent (i.e., neither self-sharpening nor nonsharpening); the ion-exchange wave is either self-sharpening or nonsharpening, depending on whether the concentration of the preferred ion (OAc^- for most strong-base anion exchangers) decreases or increases across the wave in the direction of flow. In the plateau between the two waves, the ion exchanger is still in its initial state, but the total concentration of the solution has already changed to the new value (see Fig. 29, left). Since OAc^- and OH^- ions have the same valence, their separation factor can be assumed to vary little with total concentration. To the extent that this is true, the OH^-:OAc^- ratio in the solution between the two waves is about the same as downstream of the fast wave, although the total concentration is different.

If the step change in composition of the entering buffer involves a variation of total electrolyte concentration, the pH in the solution between the response waves may be higher, or lower, than it was in both the initial and entering fluids at any time. For example, suppose the step change is from a sodium-acetate buffer of 0.1 *M* concentration and pH 9 to one of 0.01 *M* and pH 8.5. As was shown above, between the two waves the OH^-:OAc^- ratio is about the same as downstream of the fast wave, where the pH is 9 and the total concentration is higher by a factor 10. Accordingly, between the waves the OH^- concentration is about one order of magnitude lower than at pH 9, that is, close to 8 and thereby significantly lower than in the entering fluid both before and after the step change. That a composition variation of an entering buffer can produce such a pH excursion is counter-intuitive, but not unusual [60,61] (for an example, see Fig. 29, right).

12. Summary

Non-linear wave theory provides effective tools for qualitative insight into and quantitative treatment of multicomponent chromatographic systems at concentrations high enough for solutes to affect one another's behavior. A key concept is that of coherence: An arbitrary composition variation entering the column is in general noncoherent and is resolved into a set of coherent waves that travel at different velocities. The condition for coherence is that all solutes have the same wave velocity at any point in

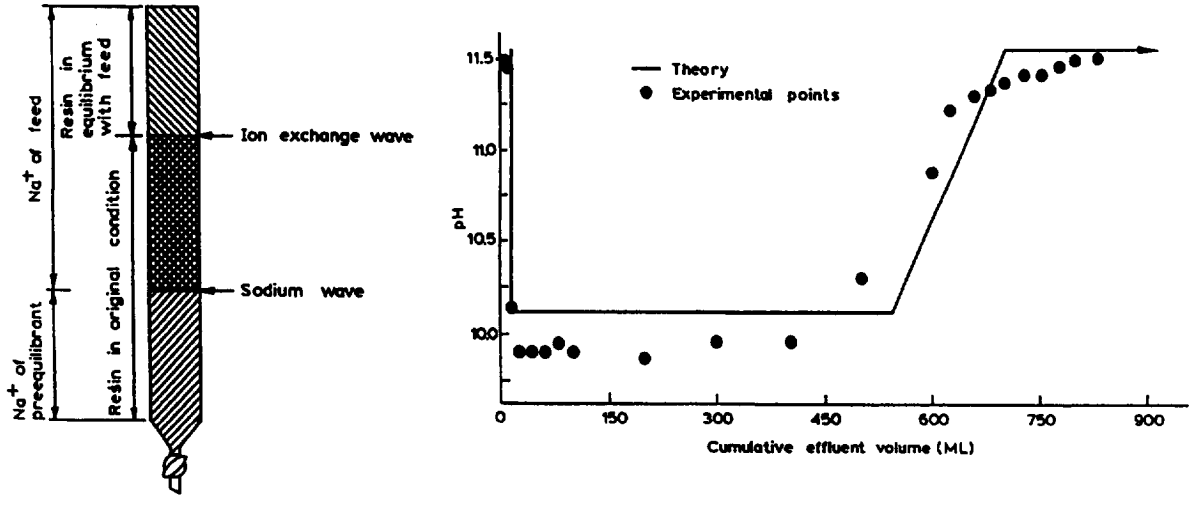

Fig. 29. Column response in a buffer system: salinity wave (sodium wave) and ion exchange wave in anion exchange column in response to step change in composition of entering sodium acetate buffer. Left: positions of waves in column. Right: observed and theoretical effluent pH history; column IRA-400, buffer variation from 0.495 *M* and pH 11.5 to 0.015 *M* and pH 11.5. (From Helfferich and Bennett [61], reproduced with permission of React. Polymers).

distance and time within the respective wave. With this condition, composition variations compatible with coherence can be mapped as so-called composition paths in the composition space (with concentrations as the coordinates). The path grid so obtained depends only on the equilibrium isotherm, not on the initial and entry conditions of the respective operation. Accordingly, once constructed for a particular system, the path grid can be used to predict behavior under any initial and entry conditions, much as one establishes the route of a car trip on a road map. The combination of such routes in the composition space with a distance–time diagram provides a complete, compact description from which concentration profiles and histories can be generated as desired.

Typical applications include frontal analysis, elution from a uniformly saturated column, elution of a pulse under overload conditions, displacement development, and elution with a buffer. In each case, the typical features familiar from experimental results are seen to arise from properties of the path grid. Route construction in the latter allows the compositions of transient plateaus to be established without further calculation, and constructions with route and distance–time diagrams in tandem can provide information on time and column length requirements for specified separations.

13. Glossary of symbols and terms

a_i coefficient in Langmuir isotherm equation ($cm^3\ g^{-1}$)

b_i coefficient in Langmuir isotherm equation ($mmol^{-1}\ cm^3$)

c_i concentration of solute i in moving phase (per unit volume of moving phase) ($mmol\ cm^{-3}$)

k_i $\equiv q_i/c_i$, distribution coefficient ($cm^3\ g^{-1}$)

q_i sorbent loading with solute i: amount of i in sorbent (averaged over bead) per unit mass of sorbate-free sorbent ($mmol\ g^{-1}$)

t time (s)

v^0 linear velocity of moving-phase flow ($cm\ s^{-1}$)

v_c linear eigenvelocity of composition $\{c_1, \ldots, c_n\}$ in coherent wave ($cm\ s^{-1}$)

v_{c_i} linear wave velocity of concentration c_i ($cm\ s^{-1}$)

$v_{\Delta c_i}$ linear wave velocity of shock Δc_i ($cm\ s^{-1}$)

z column position (linear distance from inlet end of packing) (cm)

Δ finite difference across shock (operator)

ϵ fractional volume of moving phase in column (dimensionless)

ρ bulk density of sorbent: mass of sorbate-free sorbent per unit volume of column ($g\ cm^{-3}$)

Solutes are numbered $1,2,\ldots,n$ in the sequence of decreasing affinity for the sorbent.

Compositions are denoted as follows: I=initial state of column; F, F1, F2=compositions introduced during operation; A, B, C,...=plateau compositions arising in column during operation; X=variable (non-plateau) composition arising in column during operation.

In distance–time diagrams, shock trajectories are shown as heavy lines; coherent nonsharpening waves are shown shaded and with trajectories of compositions as thin lines.

Definitions or explanations of frequently used technical terms are given in Table 1.

Acknowledgments

We are indebted to Lloyd R. Snyder, Editor, for his encouragement and for arranging for the publication of this series.

Appendix A

Outline of mathematics of coherence

This appendix provides an overview of the relevant mathematics of ideal multicomponent theory. It is not needed for general understanding, and is intended only for orientation. The reader interested in carrying out quantitative calculations is referred to

Table 1
Glossary of frequently used terms

Coherence	State in which coexisting concentrations advance jointly, at same velocity (state which waves in column strive to attain) (Section 4)
Competitive equilibrium	Equilibrium in which competition for sorption sites depresses uptake by sorbent (Section 6)
Composition	Ensemble of concentrations of all solutes
Composition path	Curve in composition space along which coherence condition is met (Section 6)
Composition route	Curve in composition space corresponding to sequence of compositions in column at given time or distance (Section 7)
Composition space	Coordinate space with concentrations as coordinates (Section 6)
Eigenvelocity	Natural velocity of composition in coherent wave (Section 5)
Front (of x)	Wave or shock with species x present only on its upstream side
History	Concentration or composition as function of time at given location
Langmuir system	System with Langmuir sorption isotherm, $q_j = a_j c_j / [1 + \sum_i (b_i c_i)]$
Langmuir-like system	System with competitive sorption equilibrium and no selectivity reversals (Section 6)
Natural velocity	Velocity at which a concentration or composition would advance under conditions of ideal chromatography (Section I.2 in Part I)
Nonsharpening wave	Wave with natural tendency to spread (Section I.4 in Part I and Section 3 in the present paper)
Particle velocity	Average velocity of molecules of x in direction of flow (Section I.2)
Path	See composition path
Path grid	Grid of composition paths in composition space (Section 6)
Plateau	Zone of uniform composition extending over finite distance or time
Profile	Concentration or composition as function of distance at given time
Riemann problem	System with uniform initial composition of medium and constant composition of entering fluid (Section 4)
Route	See composition route
selectivity	preference of sorbent for one solute over another
Selectivity reversal	Change in preference of sorbent for one solute over another
Self-sharpening wave	Wave with natural tendency to sharpen (Section I.4 and Section 3)
Shock	Wave that has remained (or has sharpened into) ideal discontinuity (Section I.4 and Section 3)
Trajectory	Curve traced by composition or shock in distance–time plane (Section 2)
Wave	Variation of dependent variables (concentrations) with distance or time (Sections I.1 and Section 3)
Wave velocity	Velocity of a concentration or composition in direction of flow (Sections I.2 and Section 3)

textbooks and reviews as guides to original literature that might be of help [2,14,21,33,62].

The basis of the mathematical treatment are the wave equation of a solute j

$$v_{c_j} = \frac{v^0}{1 + (\rho/\epsilon)(\partial q_j/\partial c_j)_z} \tag{I.3}$$

and the differential coherence condition of equal wave velocities

$$(\partial q_i/\partial c_i)_z = \lambda \quad \text{for all } i \tag{II.7}$$

The essential premises are local equilibrium, ideal plug flow, and mass transfer in axial direction by convection only.

Any coherent wave is represented by a curve in the composition space (a path segment), and this curve is the same regardless of whether the wave is viewed as a composition variation with time t or distance z. Accordingly, the coherence condition can be written in terms of a total derivative

$$\mathrm{d}q_i/\mathrm{d}c_i = \lambda \quad \text{for all } i \tag{II.2}$$

(constrained to coherent behavior) instead of the partial derivative as in Eqs. I.3 and II.7 above.

At local equilibrium, the stationary-phase concentration q_j is a function of the moving-phase composition:

$$q_j = q_j(c_1, \ldots, c_n) \tag{II.1}$$

The total differential $\mathrm{d}q_j$ in Eq. II.2 can therefore be expanded:

$$\mathrm{d}q_j = f'_{j1}\mathrm{d}c_1 + f'_{j2}\mathrm{d}c_2 + \cdots + f'_{jn}\mathrm{d}c_n \tag{II.8}$$

where the f'_{jk} are defined as

$$f'_{jk} \equiv (\partial q_j / \partial c_k)_{c_i,\, i \neq k} \qquad \text{(II.9)}$$

With Eq. II.8, the coherence condition II.2 becomes, in matrix form:

$$\begin{bmatrix} f'_{11}-\lambda & f'_{12} & f'_{13} & \cdots & f'_{1n} \\ f'_{21} & f'_{22}-\lambda & f'_{23} & \cdots & f'_{2n} \\ f'_{31} & f'_{32} & f'_{33}-\lambda & \cdots & f'_{3n} \\ & & \cdots & & \\ f'_{n1} & f'_{n2} & f'_{n3} & \cdots & f'_{nn}-\lambda \end{bmatrix} \times \begin{Bmatrix} dc_1 \\ dc_2 \\ dc_3 \\ \cdots \\ dc_n \end{Bmatrix} = 0 \qquad \text{(II.10)}$$

This is a typical eigenvalue problem, with n eigenvalues λ that can be obtained by setting the determinant of the matrix in Eq. II.10 equal to zero (in bivariant systems the eigenvalues are the two roots of a quadratic equation).

In view of the wave equation I.4, the eigenvelocity in terms of the respective eigenvalue λ is given by

$$v_c = \frac{v^0}{1 + (\rho/\epsilon)\lambda} \qquad \text{(II.4)}$$

This equation, with the n eigenvalues λ, gives the n eigenvelocities that can be assumed by the composition $\{c_1, \ldots, c_n\}$ for which the eigenvalue problem was solved.

Each eigenvalue λ is associated with an eigenvector $\{dc_1, \ldots, dc_n\}$ which indicates the direction of the respective route in the composition space. The eigenvectors can be obtained from Eq. II.10 after the respective value of λ has been substituted.

Note that Langmuir systems can be handled with much simpler mathematics, to be shown in Part III of this series.

Appendix B

Shock routes in nonideal non-Langmuir systems

In non-Langmuir systems with curved differential composition paths, the fact that nonidealities produce shock layers instead of ideal shocks (i.e., concentration discontinuities) has an interesting consequence. The nonideality lets the compositions between the upstream and downstream composition end points become physically realized. The sequence of these compositions in general does not follow a (differential) composition path and so is noncoherent. The local tendency to establish coherence distorts the composition route of the shock layer slightly, usually toward an S shape. The shock can be viewed as a continuous, traveling source of minor waves of other families that run ahead or are left behind. A quantitative treatment requires step-by-step integration of the conservation equations over distance and time and is tedious. Fortunately, the effect is quite minor under most conditions of interest.

Appendix C

Construction of a Langmuir path grid

The path grid of any two-component Langmuir system can be constructed as follows.

All paths are straight lines. The grid has a so-called watershed point (point W in Fig. 30) which divides the c_1 axis into two segments: 'Slow' paths originate from the segment above the watershed

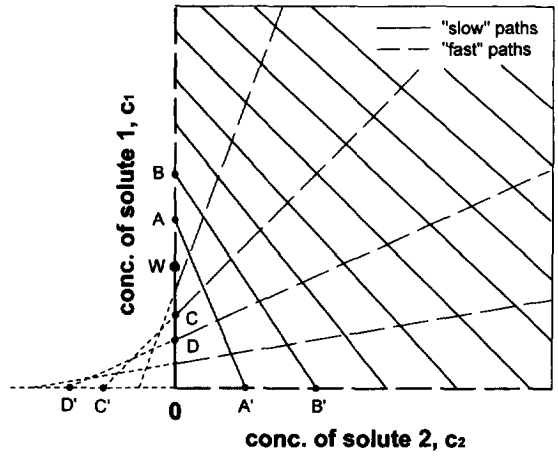

Fig. 30. Construction of Langmuir path grid with rule of equal intercept ratios (for $a_1 b_2 / a_2 b_1 = 0.67$).

point; 'fast' paths from that below. The watershed point is at

$$c_1 = \frac{(a_1/a_2) - 1}{b_1}, \quad c_2 = 0 \qquad \text{(II.11)}$$

After the watershed point has been plotted, all paths can be drawn in with the so-called rule of equal intercept ratios [63]: Paths of the same family that intersect the c_1 axis at equidistant points also intersect the c_2 axis (or its extension to negative c_2 values) at equidistant points. For any two paths of the same family, the ratio of the distance between their intersections with the c_1 axis to that between their intersections with the c_2 axis is given by

$$\text{ratio} = a_1 b_2 / a_2 b_1 \qquad \text{(II.12)}$$

(This is also the ratio of the ultimate capacities a_1/b_1 to a_2/b_2 in the respective single-component systems.) For example, in Fig. 30, the ratio of the distances A–B to A′–B′ (for a 'slow' path) is given by Eq. II.12, and so is that of the distances C–D to C′–D′ (for a 'fast' path). Note that the c_1 axis itself is a 'slow' path between 0 and W, and a 'fast' path above W.

Purists confine Langmuir equilibria to systems in which the ratio a_i/b_i is the same for all solutes (thermodynamics requires this for sorption in a idealized system that, among other facets, implies equal ultimate capacities $q_i = a_i b_i$ for all competing solutes). For such systems the ratio given by Eq. II.12 is unity, regardless of the absolute values of the Langmuir coefficients and separation factors. Accordingly, the grids of such Langmuir systems differ from one another only in the concentration scales on their axes. In contrast, Eq. II.12 is more generally applicable to empirical Langmuir isotherms whose best-fit coefficients are not confined to equal a_i/b_i ratios.

References

[1] F.G. Helfferich and P.W. Carr, J. Chromatogr., 629 (1993) 97; erratum: 633 (1993) 323.

[2] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382).

[3] J.A. Berninger, R.D. Whitley, X. Zhang and N.-H.L. Wang, Computers Chem. Eng., 15 (1991) 749.

[4] N.-H.L. Wang, VERSE (computer simulation package for large-scale adsorption and chromatography processes), Purdue Research Foundation, West Lafayette, IN, 1994. For details of or help with this program, contact N.-H. Linda Wang at Dept. Chem. Eng., Purdue Univ.

[5] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Sec. 5.III.A.

[6] J.J. van Deemter, F.J. Zuiderweg and A. Klinkenberg, Chem. Eng. Sci., 5 (1956) 271.

[7] J.C. Giddings, Dynamics of Chromatography, Dekker, New York, 1956.

[8] C.F. Poole and S.K. Poole, Chromatography Today, Elsevier, Amsterdam, 1991.

[9] L.R. Snyder, in E. Heftmann (Editor), Chromatography, Part A: Fundamentals and Techniques, 5th ed., Elsevier, Amsterdam, 1992, Ch. 1.

[10] A. Tiselius, Arkiv Kemi Mineral. Geol., B14 (22) (1940).

[11] S. Claesson, Arkiv Kemi Mineral. Geol., A23 (1) (1946).

[12] L.G. Sillén, Arkiv Kemi, 2 (1950) 477.

[13] S. Claesson, Discussions Faraday Soc., 7 (1949) 34.

[14] G. Guiochon, S. Golshan-Shirazi and A. Katti, Fundamentals of Preparative and Nonlinear Chromatography, Academic Press, New York, NY, 1994 (for frontal analysis see Chaps. 12 and 15).

[15] A. Jeffrey and T. Taniuti, Non-linear Wave Propagation, Academic Press, New York, 1964.

[16] F.G. Helfferich, Soc. Petrol. Eng. J., 21 (1981) 51.

[17] G.J. Kynch, Trans. Faraday Soc., 48 (1952) 166.

[18] F.G. Helfferich, AIChE J., 35 (1989) 75.

[19] G. Klein, in A.E. Rodrigues and D. Tondeur (Editors), Percolation Processes: Theory and Applications, Sijthoff and Noordhoff, Alphen aan den Rijn, 1981, pp. 363–421.

[20] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Sec. 5.IV.A.4.

[21] D. Tondeur and M. Bailly, Chem. Eng. Process, 22 (1987) 91.

[22] F.G. Helfferich, Ind. Eng. Chem. Fundam., 6 (1967) 362.

[23] F.G. Helfferich, AIChE Symp. Ser. 80 (1984), No. 233, 1.

[24] F.G. Helfferich, in G.E. Keller II and R.T. Yang (Editors), New Directions in Sorption Technology, Butterworths, Boston, 1989, Chapter 1.

[25] F.G. Helfferich, Chem. Eng. Commun., 44 (1986) 275.

[26] G.G. Baylé and A. Klinkenberg, Rec. Trav. Chim., 73 (1954) 1037.

[27] T.-P. Liu, Commun. Math. Phys., 55 (1977) 163.

[28] F.G. Helfferich, presented at NATO Advanced Study Institute on Migration and Fate of Pollutants in Soils and Subsoils, Maratea, Italy, May 1992.

[29] Q. Yu and N.-H.L. Wang, Sep. Purif. Methods, 15 (1986) 127.

[30] E. Glückauf, Proc. Roy. Soc., A186 (1946) 35.

[31] J.I. Coates and E. Glueckauf, J. Am. Chem. Soc., 69 (1947) 1309.
[32] H.-K. Rhee, R. Aris and N.R. Amundson, Phil. Trans. Roy. Soc., A267 (1970) 419.
[33] H.-K. Rhee, R. Aris and N.R. Amundson, First-order Partial Differential Equations, Vol. II, Prentice-Hall, Englewood Cliffs, NJ, 1986, Chaps. 2–4.
[34] S. Golshan-Shirazi and G. Guiochon, J. Phys. Chem., 93 (1989) 4341.
[35] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Sec. 3.IV.C.3.
[36] D. Tondeur, Thesis, Université de Nancy, 1969.
[37] G. Klein, personal communication, 1970.
[38] H. Poppe, J. Chromatogr., 556 (1991) 95.
[39] E.A. Abbott, Flatland, 2nd revised ed., Seeley, London, 1884; reprinted by Buccaneer Books, Catchogue, NY, 1976.
[40] A.I. Liapis and D.W.T. Rippin, Chem. Eng. Sci., 33 (1978) 593.
[41] E. Glueckauf, Discussions Faraday Soc., 7 (1949) 12.
[42] G. Klein, D. Tondeur and T. Vermeulen, Ind. Eng. Chem. Fundam., 6 (1967) 339.
[43] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), p. 304.
[44] J.H. Harwell, F.G. Helfferich and R.S. Schechter, AIChE J., 28 (1982) 448.
[45] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Appendix III.
[46] G. Guiochon and L. Jacob, Chromatogr. Rev., 14 (1971) 77.
[47] J.E. Eble, R.L. Grob, P.E. Antle and L.R. Snyder, J. Chromatogr., 405 (1987) 1.
[48] J.E. Eble, R.L. Grob, P.E. Antle, G.B. Cox and L.R. Snyder, J. Chromatogr., 405 (1987) 31.
[49] A.M. Katti, Z. Ma and G. Guiochon, AIChE J., 36 (1990) 1722.
[50] L.R. Snyder, J.W. Dolan and G.B. Cox, J. Chromatogr., 540 (1991) 21.
[51] W.C. Bauman, R.M. Wheaton and D.W. Simpson, in F.C. Nachod and J. Schubert (Editors), Ion Exchange Technology, Academic Press, New York, NY, 1956, p. 182.
[52] F. Helfferich, Ion Exchange, McGraw-Hill, New York, 1962, Sec. 9-4 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2003414).
[53] G. Jayaraman, S.D. Gadam and S.M. Cramer, J. Chromatogr., 630 (1993) 53.
[54] S.D. Gadam, S.R. Gallant and S.M. Cramer, AIChE J., 41 (1995) 1676.
[55] G. Subramanian, M.W. Phillips and S.M. Cramer, J. Chromatogr., 439 (1988) 341.
[56] F. Helfferich and D.B. James, J. Chromatogr., 46 (1970) 1.
[57] H.-K. Rhee and N.R. Amundson, AIChE J., 28 (1982) 423.
[58] J. Frenz and C. Horvath, AIChE J., 31 (1985) 400.
[59] F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Sec. 4.IV.B.
[60] B.J. Bennett and F.G. Helfferich, in D. Naden and M. Streat (Editors), Ion Exchange Technology, Ellis Horwood, Chichester, 1984, pp. 322–330.
[61] F.G. Helfferich and B.J. Bennett, React. Polym., 3 (1984) 51.
[62] A.M. Katti and G. Guiochon, Adv. Chromatogr., 31 (1992) 1.
[63] G. Klein, personal communication, 1965. See also F. Helfferich and G. Klein, Multicomponent Chromatography, Dekker, New York, 1970 (out of print, available from University Microfilms International, Ann Arbor, MI, catalogue # 2050382), Sec 3.IV.C.2.