

CHREV 195

## ESTIMATION OF DEAD TIME AND CALCULATION OF KOVATS INDICES

R J SMITH, J K HAKEN\* and M S WAINWRIGHT

*Departments of Industrial Chemistry and Polymer Science, University of New South Wales, P O Box 1, Kensington, N.S.W. 2033 (Australia)*

(Received April 9th, 1985)

### CONTENTS

1. Introduction	95
2. Kovats indices	96
3. Column dead time	97
3.1. Theoretical	97
3.2. Indirect measurement	98
3.3. Direct measurement	98
3.4. Graphical	100
3.5. Classical	101
3.6. Statistical	106
3.7. Iterative	110
3.8. Others	115
4. Linearity of the plot of log adjusted retention time vs carbon number	119
4.1. The experimental evidence	119
4.2. The theoretical evidence	121
5. Comparison of methods	123
6. Conclusions and recommendations	125
7. Summary	125
References	126

### 1 INTRODUCTION

Since the experimental demonstration of gas-liquid chromatography in the early 1950s<sup>1</sup>, the technique has achieved remarkable success. Since, that time, the growth in applications of the procedure has been phenomenal and when reliable, low-cost, easy-to-use gas chromatographs became available in the mid-sixties, a proliferation of new columns, detectors and applications elevated the procedure into the most important and widely used technique in analytical chemistry today.

However, because the raw data (the chromatogram) depends on both the experimental conditions and the equipment used, it soon became apparent that a uniform system of data presentation was necessary to give the accuracy required for the comparison of data between laboratories. This requirement is even greater today with the increasing complexity of separations and the introduction of capillary columns.

Such a system should be as independent as possible of the operating conditions as well as of the experimental equipment. Unfortunately the property directly measured by the analyst, the total retention time,  $t_R$ , is the sum of two factors. One, the

gas hold-up time (or dead time),  $t_M$ , is dependent on the system flow-rate as well as on the void volume of the experimental apparatus. The other, the adjusted retention time,  $t'_R$ , is a characteristic of the separation process

$$t'_R = t_R - t_M \quad (1)$$

and is independent of the equipment used.

Kaiser has stressed the importance of the adjusted retention time in several articles (in particular see refs. 2-4), while Parcher and Johnson<sup>5</sup>, as well as Lin and Parcher<sup>6</sup>, have shown its importance when determining thermodynamic properties such as Henry's law constants.

Although independent of the equipment used, adjusted retention times are still dependent on several variables such as column temperature, flow-rate, pressure drop and liquid phase. A variety of systems have therefore been introduced, which depend only on the column temperature and stationary phase used<sup>7,8</sup>. The most useful of these for qualitative analysis is undoubtedly the retention index system introduced by Kovats<sup>9</sup>.

## 2. KOVATS INDICES

In 1958, Kovats<sup>9</sup> introduced a system to overcome the problem of the uniform reporting of retention data. His system used the  $n$ -alkanes as a series of standards and expressed the retention of other substances relative to them using

$$I = 200 \cdot \frac{\log V_R^0(\text{substance}) - \log V_R^0(nP_z)}{\log V_R^0(nP_{z+2}) - \log V_R^0(nP_z)} + 100 \cdot z \quad (2)$$

where  $V_R^0(nP_z) \leq V_R^0(\text{substance}) \leq V_R^0(nP_{z+2})$ ,  $V_R^0$  = retention volume and  $nP_z$  =  $n$ -paraffin with carbon number  $z$ . This equation gives the retention index of an even-numbered normal paraffin as 100 times its carbon number. Only even-numbered  $n$ -paraffins were used because it was believed (incorrectly) that there would be an oscillation in the properties of successive members of the  $n$ -paraffin series. However, experimental results later indicated that this was not the case and thus in 1964 Kovats redefined the index as

$$I = 100 \cdot \frac{\log V_R^0(\text{substance}) - \log V_R^0(nP_z)}{\log V_R^0(nP_{z+1}) - \log V_R^0(nP_z)} + 100 \cdot z \quad (3)$$

where again  $V_R^0(nP_z) \leq V_R^0(\text{substance}) \leq V_R^0(nP_{z+1})$ , as related by Ettre<sup>10</sup>. In this report Ettre suggested that retention volumes could be replaced by adjusted retention times or corresponding distances on the chromatogram. Kovats<sup>11</sup> confirmed this in a later paper which introduced the more widely known formula

$$I_{\text{substance}}^{\text{stat. phase}}(T) = 100 \cdot \frac{\log t'_{Ri} - \log t'_{Rz}}{\log t'_{R(z+1)} - \log t'_{Rz}} + 100 \cdot z \quad (4)$$

where  $t'_{Rz}$  = adjusted retention time of a homologue with carbon number  $z$  and  $t'_{Ri}$  = adjusted retention time of a substance  $i$ .

It has subsequently been accepted that, under isothermal conditions, retention indices,  $I$ , can be calculated by interpolation within the linear relationship\*

$$\ln t'_{Rz} = bI + c \quad (I = 100 \cdot z) \quad (5)$$

where  $b$  and  $c$  are constants ( $\log t'_{Rz} = bI + c$  is used by some authors) and where the Kovats retention indices for  $n$ -alkanes are defined as  $100z$  for each liquid phase at all temperatures.

The retention index system of Kovats has now been widely accepted not only for comparing retention data, but also in the characterisation of stationary phases<sup>13</sup>. Mazor *et al.*<sup>14</sup> have shown that Kovats retention indices are independent of the carrier flow-rate. Dahlmann *et al.*<sup>15</sup> have shown that while the index is dependent to some degree on both the stationary phase loading and the solid support used, this can be reduced to a minimum by an appropriate choice of the solid support, chemical treatment of the support and the avoidance of very low stationary phase loadings. Finally, the temperature dependence of the index has been discussed by many authors including Wehrli and Kovats<sup>16</sup> who showed that a temperature dependence of less than one unit per degree applies in most instances. Novák *et al.*<sup>17</sup>, amongst others, have reported a relationship between the specific retention volume, carbon number and temperature. A very extensive review of the Kovats retention index system published in 1983 by Budahegyi *et al.*<sup>18</sup> discussed these as well as many other aspects, citing almost 1400 references.

Since the introduction of the retention index system, a large number of papers have analysed its reproducibility. Among others, Takács and Králik<sup>19</sup> introduced the "Erdey" equation, a theoretical equation giving the error of determination of retention indices. Other authors discussed individual contributions to the error and have suggested methods to minimise such errors<sup>4,8,20-24</sup>. The paper by Schomburg and Dielmann<sup>8</sup>, which is a particularly useful discussion on sources of error, suggests an interlaboratory reproducibility of about one unit.

Except a few theoretical methods of calculating retention indices<sup>25,26</sup>, it can be seen from eqns. 1, 4 and 5 that a knowledge of the adjusted retention times and thus the column dead time is essential.

### 3 COLUMN DEAD TIME

There has been increasing discussion in recent years regarding the merits of the various methods for determining dead time. In a number of papers several of these methods have been compared in an attempt to find a simple, accurate method<sup>18,27-32</sup>. Here we update and extend the review carried out by two of us in an earlier publication<sup>32</sup>.

The different techniques can broadly be classified into eight categories.

#### 3.1. Theoretical

Evans and Smith<sup>33</sup> suggested in 1962 that the static column interstitial volume could be calculated from

\* We have used the nomenclature recommended by ASTM E 355<sup>12</sup> where possible

$$d_0 = \frac{2.54 SV}{60 F} \left[ 1 - \frac{B_p}{D_p} - \frac{PB_p}{D_L(100 - P)} \right] \quad (6)$$

where  $d_0$  = distance (cm) from the injection point to the dead volume point on the chart,  $S$  = the recorder chart speed (in./h),  $V$  = volume of the column packed,  $F$  = carrier gas flow-rate (cm<sup>3</sup>/min) at the average column pressure,  $B_p$  = bulk density of the support,  $D_p$  = true density of the support,  $P$  = percentage (w/w) of stationary phase and  $D_L$  = density of stationary phase. This is subtracted from the observed retention distances for the  $n$ -alkanes and the adjusted retention distances are plotted against carbon number on log-linear graph paper. A further correction,  $\delta a$ , to account for the difference between static and dynamic column dead volumes is then applied

$$\delta a = \frac{2.303 \cdot \Sigma \delta \log R}{\Sigma 1/R_{\text{ext.}}} \quad (7)$$

where  $\delta \log R = \log R - \log R_{\text{ext}}$ ,  $R$  = retention distance obtained from the assumed retention volume and  $R_{\text{ext}}$  = value obtained from extrapolation on the linear part of the graph. The resulting adjusted retention distances are then replotted and generally give a straight line. Evans and Smith reported that occasionally a second correction (apparently trial and error) to linearise the "log plot" may be necessary.

### 3.2. Indirect measurement

Kaiser<sup>27</sup> suggested a method which involved the measurement of the total gas volume of the system using an azotometer and the calculation of the hold-up time,  $t_M$  (min), from that total gas volume,  $V_{GC}$  (cm<sup>3</sup>). He described the experimental system in detail and used eqns. 8 and 9 to calculate  $t_M$

$$V_{GC} = V \frac{273.2 p_M}{(273.2 + T_M) 760} + V_D \quad (8)$$

$$t_M = \frac{V_{GC}(273.2 + T_M) 2 [(p_i/p_0)^2 + (p_i/p_0) + 1]}{273.2 F p_M \cdot 3 [(p_i/p_0) + 1]} \quad (9)$$

where  $V$  = azotometer reading (cm<sup>3</sup>),  $p_M$  = gas pressure (Torr) in the azotometer and in the room,  $T_M$  = gas temperature (°C) in the azotometer and in the room,  $V_D$  = volume (cm<sup>3</sup>) of apparatus not registered during purging with CO<sub>2</sub>,  $p_i$  = pressure (Torr) in dosing system,  $p_0$  = pressure (Torr) in the detector and  $F$  = carrier gas flow-rate (cm<sup>3</sup>/min) at  $T_M$  and  $p_M$ . However, as Kaiser pointed out, this procedure is time-consuming, liable to error and requires much experience. Also its application with modern equipment may require extensive modifications.

### 3.3. Direct measurement

In an ideal system the dead time is considered to be the time an infinitesimal

amount of non-adsorbed gas takes to pass through the chromatographic system under identical conditions to those for the sample being analysed. James and Martin<sup>1</sup> designated the air peak as the dead time and measured the retention times of substances being analysed as the distances from this peak. Since then controversy has developed over the question of the best substance to use for dead-time measurement. The problem is especially acute with flame ionisation detectors which do not normally respond to air or inert gases. Thus methane has been suggested as a substitute when using such detectors<sup>34</sup>

Many authors have supported the use of air and/or methane as an indication of dead time, especially in comparison to some of the early mathematical methods of calculating this parameter<sup>5,8,22,23,35-38</sup>. In addition several unique methods have been developed for use with flame ionisation detectors. Hilmi<sup>39</sup> outlined a method in which the carrier gas is presaturated with a low volatility organic solvent, thus producing a negative air peak with a flame ionisation detector. By measuring the retention time of the air peak and making allowance for the vapour pressure of the solvent, he determined the column dead time. The method, which suffers from considerable experimental difficulties, requires large injections of air (of the order of 1 cm<sup>3</sup>) and has not found wide acceptance.

Riedmann<sup>40</sup> developed a method in which the flow-rate of hydrogen to a flame ionisation detector is reduced to 3-5 cm<sup>3</sup>/min. He found that the injection of a hydrogen sample of about 1 cm<sup>3</sup> and a splitting ratio of 1:100 produced a recorder-pen deflection and could be used as a measure of the system's dead time. Cramers *et al.*<sup>20</sup> described a similar method. In addition, Riedmann described a method suitable for Hewlett-Packard Series 5700 gas chromatographs (or any other chromatograph with a flame ionisation detector in which the ionisation current varies with the air flow). With the detector operating within the recommended air and hydrogen ranges, an injection of  $\geq 1$  mm<sup>3</sup> of air will generate a detectable peak which can then be used as a measure of the column dead time.

Guberska<sup>36,37</sup> investigated the use of methane to measure column dead time at a series of temperatures and compared the results with the method developed by Hansen and Andresen<sup>41</sup>. Guberska found that although the retention time of methane was greater than the dead time obtained by the method of Hansen and Andresen, the standard deviation did not change with temperature, while the standard deviation of the dead time increases with an increase in temperature. He therefore proposed that the retention time of methane, adjusted according to

$$t_M = t_{\text{CH}_4} - (t_{\text{CH}_4}A)/163 \quad (10)$$

where  $A$  = amount (% w/w) of stationary phase in the column, should be used as a measurement of the column dead time.

However, the use of air or methane not only creates experimental difficulties but evidence of a net retention has been presented by many authors. In a series of papers, experimenters from the University of New South Wales have shown that both air and methane show significant retention under the experimental conditions used, with methane showing retention on a wide range of column packings<sup>30,32,42-47</sup>. Garcia Dominguez *et al.*<sup>48</sup> also reported that methane was retained by normal liquid phases at temperatures as high as 180°C. Kaiser<sup>27</sup> warned against the use of several

gases including air and methane. Ettre<sup>13</sup> reported errors when methane was used below 70°C, while Becerra *et al.*<sup>49</sup> reported that methane shows significant retention.

Further, if one accepts the evidence that the use of any homologous series provides the same estimate of dead time<sup>7,50,51</sup>, then one is left with the uncomfortable conclusion that the retention of any C<sub>1</sub> compound should be a good estimate of the column dead time.

The use of other inert gases has been investigated by several authors. Riedo *et al.*<sup>52</sup> presented extensive data on a series of permanent gases and light hydrocarbons at temperatures between 30 and 230°C on a tailor-made C<sub>87</sub> hydrocarbon stationary phase. Neon was found to have the smallest gross retention volume. Nitrogen, hydrogen and argon had the next smallest, with all other substances having much larger gross retention volumes (note that as helium was the carrier gas it could not be tested). Parcher and Johnson<sup>5</sup> compared the dead time estimated using four inert gases with that calculated using C<sub>1</sub>–C<sub>5</sub> hydrocarbons. Their study confirmed that neon had the lowest retention time (again helium was not tested) and that methane had a small but measureable net retention time. Finally, a study by Ezrets and Vigerdgausz<sup>23</sup> confirmed that methane has a small net retention compared to an inert gas (helium and hydrogen), but concluded that it was smaller than that introduced by a mathematical method such as that of Peterson and Hirsch<sup>53</sup>.

We believe that the explanation of the wide divergence of opinion on the use of air and especially methane as a dead-time indicator lies in three areas, the experimental conditions, the accuracy of retention-time measurements and the use to which the retention data are put. The retention of methane in particular is greatly affected by experimental conditions and is larger at low temperatures, in gas–solid chromatography, and at high pressures. The accuracy of measuring retention times becomes of importance under conditions of low methane retention and/or small gross retention times. Finally, the use of methane under conditions of low retention for the calculation of Kovats indices above 400 is unlikely to introduce large errors. On the other hand, its use at low temperatures on porous polymers may introduce significant errors for all but very strongly retained substances. A final problem lies in the comparison of classical methods of dead-time calculation using methane. As will be pointed out later, such methods can also lead to very significant errors.

Therefore methane should only be used with extreme caution as an indication of the column dead time and the results should be verified by another method whenever the experimental conditions are changed.

From the above discussion and the fact that flame ionisation detectors do not respond to permanent gases such as neon, it is clear that there is a need for a mathematical method of dead-time calculation.

### 3.4. Graphical

This method involves plotting the logarithm of the adjusted retention times of the *n*-alkanes against their carbon numbers and linearisation of this plot by trial-and-error variation of the dead time. First reported by Evans and Smith<sup>33</sup> as well as Castello and Parodi<sup>54</sup>, it is tedious and inaccurate and was soon superseded by various mathematical treatments of the data.

## 3.5. Classical

These methods use a limited number of homologues and generally require the solution of a small number (often one) of simple equations. The most widely used of the methods are described below.

(a) The method of Peterson and Hirsch<sup>53</sup> requires three evenly spaced homologues. It is based on the assignment of an arbitrary carrier-gas front (column dead time),  $x_0$ , as shown in Fig. 1 which uses three evenly spaced peaks to illustrate the method. All distances are then measured from this point which differs from the true dead time by  $\delta$ . If  $x_0$  corresponds to the true column dead time then  $\delta = 0$ . In general, however, this is not the case and the distance to a given peak,  $t'_{Rz}$ , is not measured from the true column dead time but from  $x_0$ :

$$t'_{Rz} = t^*_{Rz} + \delta \quad (11)$$

The linear relationship between retention time and carbon number can now be expressed as

$$\log (t^*_{Rz} + \delta) \propto z \quad (12)$$

and when three evenly spaced peaks are used

$$\frac{t^*_{Rz+i} + \delta}{t^*_{Rz} + \delta} = \frac{t^*_{Rz+2i} + \delta}{t^*_{Rz+i} + \delta} \quad (13)$$

or

$$\delta = \frac{t^*_{Rz+i}{}^2 - t^*_{Rz}t^*_{Rz+2i}}{t^*_{Rz+2i} + t^*_{Rz} - 2t^*_{Rz+i}} \quad (14)$$

By measuring from any arbitrary point,  $x_0$ , to the peaks of three evenly spaced homologues, eqn. 13 can be solved for  $\delta$  which is positive if  $t_M$  precedes  $x_0$  and negative if  $t_M$  follows  $x_0$ . The column dead time can then be determined by measuring a distance  $\delta$  from the arbitrary point  $x_0$ .

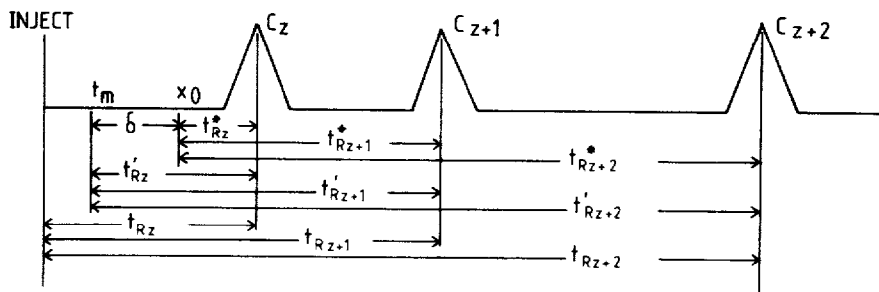


Fig 1 Schematic chromatogram  $t_R$  = Unadjusted retention time,  $t^*_R$  = retention time relative to  $x_0$ ;  $t'_R$  = adjusted retention time

Peterson and Hirsch suggested a further simplification. By designation  $x_0$  as the peak of the second homologue,  $t_{Rz+1}^* = 0$  and the distance measured between the peaks of homologues 2 and 1 multiplied by the distance between homologues 2 and 3, divided by their sum ( $t_{Rz}^*$  is negative in this formulation), is the distance from homologue 2 to  $t_M$ .

(b) Gold<sup>55</sup> improved upon the method of Peterson and Hirsch by developing one which did not require the three homologues to be equally spaced. The method is based on the fact that  $\log(t_{Rz} - t_M)$  is proportional to the carbon number,  $z$ .

$$z = m \log(t_{Rz} - t_M) + k \quad (15)$$

where  $m$  and  $k$  are constants. Letting the difference between carbon numbers be  $z_2 - z_1 = \Delta z_{1,2}$ , then

$$\Delta z_{1,2} = m [\log(t_{Rz_2} - t_M) - \log(t_{Rz_1} - t_M)] \quad (16)$$

and

$$m = \frac{\Delta z_{1,2}}{\log[(t_{Rz_2} - t_M)/(t_{Rz_1} - t_M)]} \quad (17)$$

Rewriting in terms of  $\Delta z_{1,3}$  and solving for  $t_M$  gives:

$$t_M = \frac{t_{Rz_3} - 10^{(\Delta z_{1,3}/m)} \cdot t_{Rz_1}}{1 - 10^{(\Delta z_{1,3}/m)}} \quad (18)$$

Eqs. 17 and 18 are simultaneous and can be solved by the method of successive approximations.

Gold also suggested a graphical procedure in which eqn. 17 is rewritten in terms of  $\Delta z_{1,3}$ , giving two equations for  $m$ . The lines for these two equations are plotted by assuming values for  $t_M$  and calculating the resulting values of  $m$  with each of the two equations. These sets of points are plotted and their common solution is the point at which the two lines intersect.

Both these methods are tedious and have little to recommend them.

(c) Kaiser<sup>56</sup> developed a program for the HP-65 programmable calculator which calculated the dead time by the method of Gold. In the same paper, another program for the HP-65 was presented which calculated the constants in the  $\log t'_R$  vs.  $I$  relationship once the dead time was known. This program was subsequently reprinted with a correction by Ebel and Kaiser<sup>57</sup>. Although this eliminated the tedium from Gold's method, it did not become widely used. This was probably due to the lack of availability of programmable calculators in 1974 and the fact that the advantages of the method over contemporary methods was relatively small.

(d) Hafferkamp<sup>58</sup> as well as Hansen and Andresen<sup>41</sup> developed a method which used three equally spaced homologues as proposed by Petersen and Hirsch. The advantage of their method was that it eliminated the use of an artificial dead time and used a single calculation to determine the column dead time. Their equation can be expressed as follows:



$$t_M = \frac{t_{Rz+i}^2 - t_{Rz}t_{Rz+2i}}{2t_{Rz+2i} - t_{Rz} - t_{Rz+i}} \quad (19)$$

Ettre<sup>31</sup> pointed out that this equation is in fact only a rearrangement of the simplified equations derived by Peterson and Hirsch<sup>53</sup> (last paragraph in section 3.5a). He then proceeded to generalise this approach to the situation where the homologues are separated by unequal distances. Thus  $t_{Rz_3} - t_{Rz_2} = n(t_{Rz_2} - t_{Rz_1})$ . Unfortunately this leads to the complication of solving higher-degree equations. For  $n = 2$  (the only case quoted) a quadratic must be solved.

(e) In 1979 a method with some similarities to that of Gold and partially based on a method developed by Ševčík and Löwentap<sup>29</sup> was described by Al-Thamir *et al.*<sup>59</sup>. It begins with the linear relationship between the adjusted retention time and the carbon number and shows that this leads to a constant ratio of successive differences in retention times

$$\frac{\Delta t_{Rz+1,z+2}}{\Delta t_{Rz,z+1}} = e^b \quad (20)$$

where  $b$  = the slope of the plot of eqn. 5. Hence:

$$b = \ln (\Delta t_{Rz+1,z+2} / t_{Rz,z+1}) \quad (21)$$

Having evaluated  $b$ ,  $c$  is calculated as:

$$c = \ln \left( \frac{\Delta t_{Rz,z+1}}{e^b - 1} \right)^{-bz} \quad (22)$$

This is eqn. 4 of Al-Thamir *et al.*<sup>59</sup> rewritten in a more convenient form. Then  $t_M$  is calculated from:

$$t_M = t_{Rz} - e^{(bz+c)} \quad (23)$$

Now  $t_M$  can be calculated for all homologues employed and a consistent result shows its validity.

A comparison of the method with that of Hansen and Andresen<sup>41</sup> (eqn. 19) shows that although the method involves more calculation, it provides no advantage over Hansen and Andresen's method.

Ettre<sup>31</sup> pointed out that the method is equally valid for homologues which differ by any distance,  $i$ , and thus is not restricted to consecutive homologues. Thus the method can be used for homologues with retention times  $t_{Rz}$ ,  $t_{Rz+i}$  and  $t_{Rz+2i}$ . In addition, he showed how the method could be generalised even further to the case where  $t_{Rz_3} - t_{Rz_2} = n(t_{Rz_2} - t_{Rz_1})$ . Unfortunately this leads to the need to solve higher order equations ( $n = 2$  leads to a quadratic) and thus the method becomes very difficult for all but the simplest extensions.

However, Ettre<sup>31</sup> criticised the approach because of the "danger of 'linearizing' the  $t_{Ri}$  vs.  $c_i$  relationship even where it is not valid". He also pointed out that "individual gas hold-up time values are obtained for the individual homologues which,

of course, is absurd since one can only have one single gas hold-up time for one chromatogram".

(f) In a recent paper Tóth and Zala<sup>60</sup> extended the method of Al-Thamir *et al.*<sup>59</sup> by deriving an expression for  $b$  which takes into account all possible ratios of the differences in retention times of neighbouring members of an homologous series. Expressions were also presented for  $c$  and  $t_M$  which take into consideration the retention values of all homologues of a consecutive series.

According to Al-Thamir *et al.*<sup>59</sup> (eqn 21) the ratio between two neighbouring retention-time differences can be expressed as

$$b = \ln (\Delta_z/\Delta_{z-1}) \quad (24)$$

where  $\Delta_z$  and  $\Delta_{z-1}$  are consecutive retention differences and  $z > 1$ . Thus  $b$  can be more generally expressed by not necessarily using consecutive differences, giving

$$b_{z/z-n} = \frac{1}{n} \cdot \ln (\Delta_z/\Delta_{z-n}) \quad (1 \leq n \leq z - 1) \quad (25)$$

where  $b_{z/z-n} = b$  value obtained using the ratio of the  $z$ th and  $(z - n)$ th retention differences.

For an homologous series with  $N$  members,  $N - 1$   $\Delta$  values are obtained. By combining these differences in pairs, the total number of pairs can be expressed as:

$$\frac{(N - 1)!}{2[(N - 1) - 2]!} \quad (26)$$

Therefore  $N - 2$  values of  $b$  can be calculated from an homologous series having  $N$  members ( $N - 1$   $\Delta$  values), using only neighbouring  $\Delta$  values (in this case  $n = 1$ ):

$$\sum_{z=2}^{N-1} b_{z/z-1} = \ln \left( \frac{\Delta_2 \Delta_3 \dots \dots \Delta_{N-2} \Delta_{N-1}}{\Delta_1 \Delta_2 \dots \dots \Delta_{N-3} \Delta_{N-2}} \right) \quad (27)$$

Also  $N - 3$  values are obtained using the ratio of those  $b$  values which follow each other by the omission of one  $\Delta$  (in this case  $n = 2$ ). The sum can be expressed as:

$$\sum_{z=3}^{N-1} b_{z/z-2} = \frac{1}{2} \ln \left( \frac{\Delta_3 \Delta_4 \dots \dots \Delta_{N-2} \Delta_{N-1}}{\Delta_1 \Delta_2 \dots \dots \Delta_{N-4} \Delta_{N-3}} \right) \quad (28)$$

By increasing the value of  $n$  the ratio of the first and last  $b$  value is obtained:

$$b_{z/z-(N-2)} = \frac{1}{N - 2} \cdot \ln (\Delta_{N-1}/\Delta_1) \quad (29)$$

Now the sum of all  $b$  values expressed by all possible  $\Delta$  values is given by

$$\Sigma b_{z/z-n} = \ln \prod_{z=1}^L (\Delta_{N-z}/\Delta_z)^{1/z+1/(z+1)+\dots+1/(N-z-1)} \tag{30}$$

where  $L = (N - 1)/2$  if  $N$  is odd and  $L = (N/2) - 1$  if  $N$  is even.

Taking into account the total number of pairs of  $\Delta$  values (eqn. 26), the mean value of  $b$  can be expressed as:

$$\bar{b} = \frac{2[(N - 1) - 2]!}{(N - 1)!} \cdot \ln \prod_{z=1}^L (\Delta_{N-1}/\Delta_z)^{1/z+1/(z+1)+\dots+1/(N-z-1)} \tag{31}$$

After calculating the mean value of  $b$ , the values of  $c$  and  $t_M$  can be obtained from:

$$c = \ln \left( \frac{t_{Rz+n} - t_{Rz}}{\exp[b(z + n)] - \exp[bz]} \right) \quad (z + n > z) \tag{32}$$

$$t_M = t_{Rz} - \exp(bz + c) \tag{33}$$

Now these equations can be used to express the mean values of  $c$  and  $t_M$  as follows:

$$\bar{c} = \ln \left[ \frac{(\Delta_1 \Delta_2 \dots \Delta_{N-1})^{1/(N-1)}}{(\exp \bar{b}) - 1} \right] - \frac{\bar{b}(2z_1 + N - 2)}{2} \tag{34}$$

$$\bar{t}_M = \frac{\Sigma t_{Rz}}{N} - \frac{\exp[z_1 \bar{b} + c][\exp(\bar{b}N) - 1]}{N[\exp(\bar{b}) - 1]} \tag{35}$$

Tóth and Zala<sup>60</sup> went on to show that their method gave similar results to both the statistical method of Grobler and Balizs<sup>61</sup> and their earlier iterative method (see section 3.7e).

(g) Garcia Dominguez *et al.*<sup>48</sup> described a method (“approximate series method”) using any three homologues to calculate approximate retention times (and therefore the dead time). Two homologues are used to determine the approximate slope from:

$$b \approx [\log(t_{Rz_3}/t_{Rz_2})]/(z_3 - z_2) \tag{36}$$

A first approximation to the adjusted retention time of the first homologue can then be calculated from:

$$t'_{Rz_1} = (t_{Rz_2} - t_{Rz_1})/\{10^{[(z_2 - z_1) \log(t_{Rz_3}/t_{Rz_2})/(z_3 - z_2)]} - 1\} \tag{37}$$

The accuracy of  $t'_{Rz_1}$  will increase as the difference  $z_3 - z_1$  increases and for lines with higher slopes. For increased accuracy the process can be repeated to obtain a better estimate of  $t'_{Rz_1}$ . If three estimates are obtained, a better approximation can be made using

$$t'_{Rz_1} = t_{Rz_1} + d_1/(1 - d_2/d_1) \quad (38)$$

where  $t'_{Rz_1}$  is the first approximation of  $t'_{Rz_1}$  and  $d_1, d_2$  are the differences between the first and second and the second and third approximations of  $t'_{Rz_1}$ .

The dead time can then be calculated by subtracting the value found using eqn. 38 from the experimentally measured total retention time. However, the method shows no significant improvement over other classical methods and involves more effort in calculation than most. A second approximate method for use with only two homologues ("slope method"), presented in the same paper, is even less accurate and thus has little to recommend it.

(h) Recently Tóth and Zala<sup>62</sup> introduced a method based on three non-successive  $n$ -alkanes. They developed a series of equations for calculating the slope,  $b$ , given different combinations of known and unknown  $n$ -alkanes within a group of either four or five consecutive  $n$ -alkanes. The dead time,  $t_M$ , as well as  $c$  are then calculated using simple formulae. Unfortunately a different equation is required for the calculation of  $b$  in each case. In addition, no indication was given as to whether the method can be extended to the more general case where the three homologues are separated by random distances, *e.g.*, homologues with carbon numbers  $z, z + x$  and  $z + y$  where  $x$  and  $y$  are random integers. The method is more complex than other classical methods and therefore does not offer any advantage over other less complex methods.

### 3.6. Statistical

The previous classical methods used a limited number of homologues and thus their accuracy was poor (see section 5 for a more detailed discussion). In addition some used graphical or numerical trial and error which decreased their usefulness. As the use of programmable calculators and microcomputers increased it became clear that far more sophisticated methods were not only feasible, but were a necessity if consistently reliable results were to be obtained.

(a) Grobler and Balizs<sup>61</sup> developed a method which relied on the use of two linear regressions. The derivation, which is fully described in a previous paper by two of us<sup>32</sup> (not in Grobler and Balizs' paper), starts with eqn. 5. From this the following relationship between the difference in retention times and the carbon number can be derived

$$\ln(t_{Rz+1} - t_{Rz}) = \ln A + bz \quad (39)$$

where  $A = e^c(e^b - 1)$ , and therefore a linear regression on eqn. 39 of  $\log(t_{Rz+1} - t_{Rz})$  against  $z$  will give  $b$  as the slope. Therefore:

$$b = \frac{(n-1) \sum_{i=z}^{z+n-1} i \cdot \ln [t_{R(i+1)} - t_{Ri}] - \sum_{i=z}^{z+n-1} i \cdot \sum_{i=z}^{z+n-1} \ln [t_{R(i+1)} - t_{Ri}]}{(n-1) \sum_{i=z}^{z+n-1} i^2 - \left( \sum_{i=z}^{z+n-1} i \right)^2} \quad (40)$$

Now

$$t_{Rz} = t_M + Aq^z \tag{41}$$

where  $A = e^c$  and  $q = e^b$ . Therefore a further linear regression can be carried out on eqn. 41 of  $\ln t_{Rz}$  against  $z$ . This gives

$$t_M = \frac{\sum_{i=z}^{z+n} q^i \cdot \sum_{i=z}^{z+n} t_{Ri} q^i - \sum_{i=z}^{z+n} q^{2i} \cdot \sum_{i=z}^{z+n} t_{Ri}}{\left(\sum_{i=z}^{z+n-1} q^i\right)^2 - n \sum_{i=z}^{z+n-1} q^{2i}} \tag{42}$$

and

$$c = \sum_{i=z}^{z+n} \ln t'_{Ri} - \left(b \sum_{i=z}^{z+n} i\right) / n \tag{43}$$

Eqns. 40, 42 and 43 allow the retention index of any compound to be calculated from:

$$I = 100 (\ln t'_R - c) / b \tag{44}$$

(b) This method has been extended by Van Tulder *et al.*<sup>63</sup>. In addition they carried out an extensive evaluation of the method and produced a series of criteria to allow the best selection of homologues to be used in the two linear regressions.

They proposed an extension in which homologues with equidistant carbon numbers were used instead of consecutive homologues. Thus the method starts by assuming a set of  $n$  consecutive homologues, starting with homologue  $z^0$  as follows\*:

$$(t_{Rz}, t_{Rz+1}, \dots, t_{Rz+n-1})$$

or

$$(t_{Ri} \text{ where } i = z \text{ to } z + n - 1) \tag{45}$$

Now a subset of  $n^1$  homologues are selected with carbon numbers at equal distance,  $m$ , from each other, starting with  $z^1$ . This subset can be characterised by:

$$[t_{Rj} \text{ where } j = z^1 \text{ to } z^1 + m(n^1 - 1)] \tag{46}$$

Therefore eqn. 5 can be rewritten as

$$\ln (t_{Rj} - t_M) = b(z^0 + j) + c \tag{47}$$

---

\* Note that in the paper of Van Tulder *et al.*<sup>63</sup>  $i$  and  $j$  are the sequence numbers of the homologues (starting with 1), while in our review they are the true homologue numbers

where  $z^0$  represents the carbon number of the basic group and  $j$  is the homologue number ( $z^0$  will be set to 0 for the remainder of this review).

Now eqn. 39 can be rewritten as

$$\ln (t_{R_{j+m}} - t_{R_j}) = \ln A^1 + bj \quad (48)$$

where  $A^1 = e^c(e^{mb} - 1)$ . Therefore a linear regression on eqn. 48 produces

$$b = \frac{(n^1 - 1) \sum_{j=z^1}^{Z-m} j \cdot \ln [t_{R_{(j+1)}} - t_{R_j}] - \sum_{j=z^1}^{Z-m} j \cdot \sum_{j=z^1}^{Z-m} \ln [t_{R_{(j+1)}} - t_{R_j}]}{(n^1 - 1) \sum_{j=z^1}^{Z-m} j^2 - \left( \sum_{j=z^1}^{Z-m} j \right)^2} \quad (49)$$

where  $Z = z^1 + mn^1$  and thus:

$$t_{R_j} = t_M + A^1 q^j \quad (50)$$

A second linear regression on eqn. 38 gives

$$t_M = \frac{\sum_{j=z^1}^Z q^j \sum_{j=z^1}^Z t_{R_j} q^j - \sum_{j=z^1}^Z q^{2j} \cdot \sum_{j=z^1}^Z t_{R_j}}{\left( \sum_{j=z^1}^{Z-m} q^j \right)^2 - n \sum_{j=z^1}^{Z-m} q^{2j}} \quad (51)$$

and

$$c = \frac{\sum_{j=z^1}^Z \ln t'_{R_j} - b \sum_{j=z^1}^Z j}{n^1} \quad (52)$$

Van Tulder *et al.*<sup>63</sup> carried out an extensive computer simulation to investigate the influence of statistical error on the relative standard deviation of the dead time, R.S.D. ( $t_M$ ). The study indicated that the minimum number of homologues required depended on the slope  $b$ . From a large number of computer simulations, the following empirical equation was developed

$$N_0 = (3/b) + 1 \quad (53)$$

where minimum  $N_0$  is 3 and  $N_0 =$  the optimum number of homologues. In addition, surprisingly it was found that three homologues with a separation distance,  $m$ , of  $1.5/b$  gave the lowest R.S.D. ( $t_M$ ) values. Hence the following procedure to reduce the statistical error in dead-time calculation to a minimum was proposed.

First an initial calculation should be performed with as many consecutive homologues as possible, excluding any lower members which show non-linear behaviour. From this calculation a value for  $b$  is obtained and used to calculate the distance,  $m$ , between homologues to obtain a slope factor  $\geq 1.5$ . Finally the  $t_M$  calculation is repeated with the properly selected homologues.

Although extensive experience in our laboratory has shown that in general this method selects homologues which give a good estimate of dead time, two problems are apparent. The first is the authors' use of  $R^2$  as an indication of linearity. In fact,  $R^2$  is particularly insensitive in this application and is of no use when using only three homologues. A better test of linearity is needed. The second problem, which is discussed in more detail in our recent paper<sup>64</sup>, is that the method only ensures that over a large number of trials it will give the lowest R.S.D. ( $t_M$ ). However there is no guarantee that, given a particular set of retention times, the method will produce an acceptable estimate of the dead time. Our experience has been that in some cases the method can lead to very poor results. The reason is not difficult to understand. While the method provides the range of homologues needed to obtain a good estimate, the use of such a limited number of homologues for the actual dead-time calculation means that errors in the retention times of the chosen homologues can, under certain conditions, be magnified in the final dead time<sup>28,32</sup>. We therefore recommend caution in the use of this section of the method.

(c) Ambrus<sup>76</sup> recently presented a method which determines both the column dead time,  $t_M$ , and the slope in a single linear regression. Where Kovats indices are required,  $c$  can then be calculated either from a second regression or as the mean of a set of  $c$  values calculated using each adjusted retention time.

Starting with the usual assumption of linearity of the log plot, an expression for relative retention can be calculated as follows:

$$\ln(t_{Rz} - t_M) = bz + c \quad (\text{eqn. 5})$$

$$t_{Rz} - t_M = e^{bz+c} = e^{bz}e^c \quad (54)$$

and

$$t_{R(z+1)} - t_M = b(z+1) + c = e^{bz}e^b e^c \quad (55)$$

thus

$$\frac{t_{R(z+1)} - t_M}{t_{Rz} - t_M} = \frac{e^{bz}e^b e^c}{e^{bz}e^c} = e^b \quad (56)$$

which is a constant. Letting the constant be  $q$  gives

$$t_{R(z+1)} - t_M = q t_{Rz} - q t_M \quad (57)$$

and thus

$$t_{R(z+1)} = q t_{Rz} - B \quad (58)$$

where  $B = t_M(q - 1)$ . Therefore a regression of  $t_{R(z+1)}$  against  $t_{Rz}$  will give  $q$  as the slope (from which  $b$  can be calculated as  $\ln q$  and  $B$  as the intercept [from which  $t_M$  equals  $B/(q - 1)$ ]).

For the calculation of Kovats indices,  $c$  can now be calculated from eqn. 5 by a further regression of  $\ln t'_{Rz}$  against  $I$ . Alternatively, a series of values of  $c$  can be obtained by substituting each retention time into eqn. 5, and the mean of these values used as an estimate of  $c$ .

Ambrus' methods<sup>76</sup> can also be extended to homologues with equidistant carbon numbers by the technique of Van Tulder *et al.*<sup>63</sup> (section 3.6b).

### 3.7. Iterative

(a) Guardino *et al.*<sup>24</sup> presented a method in which an iteration is carried out on  $t_M$ , with  $b$  and  $c$  calculated using a least-squares fit. The optimum values of  $t_M$ ,  $b$  and  $c$  are determined by minimising the sum of squares of the differences between the known and calculated retention-index values. A flow chart of the method is shown in Fig. 2, where UPLIM and LOWLIM are the upper and lower limits, respectively of the sum of squares of the deviation, TM is the dead time, INC is the increment in the dead time, IC is the calculated Kovats retention index, SUM is the sum of squares of the deviations, TR is the unadjusted retention times of the homologues, I is the known Kovats retention index ( $100z$  where  $z$  is the carbon number) and PREC is the precision to which the answer is required.

The method starts with an initial estimate of the dead time, which is used to determine adjusted retention times. A linear regression then allows  $b$  and  $c$  to be calculated and thus retention indices can be determined. Subtracting these from the known values gives a sum of differences which is compared to the upper and lower limits. If the estimate of  $t_M$  is below the lower limit, the limits are reduced and the estimate of  $t_M$  is increased. When this estimate increases above the lower limit, it is decreased and the increment is lowered by a factor of 10. The whole process is repeated until the increment is less than the required precision (PREC).

A very important point when using this method (not mentioned in the paper) is that the initial estimate of dead time must be less than the true mathematical dead time or the method will fail. Guardino *et al.* suggested using 98% of the value determined by the method of Haferkamp<sup>58</sup> as the initial estimate for their method. However, due to the occasional large errors that can occur when using any of the classical methods of estimating dead time, there is the possibility that this estimate may lead to an incorrect dead time. The best way of ensuring this does not happen is to monitor the difference in consecutive sums of squares and ensure that a change of sign occurs before the sum of squares becomes larger than the lower limit and that the program does not stop within the first two iterations.

(b) Bellas<sup>65</sup> also reported an iterative method of calculating dead time. The program, written in FORTRAN, makes an initial estimate for the dead time of one eighth of the minimum retention time. A linear regression is then carried out and the sum of squares of the deviations of the indices calculated for the reference compounds is found. The dead time is increased incrementally until a minimum is passed in the sum of squares. The increment is then halved and changed in sign and the fitting continued. Thus, after the first passage the increment adopts successive values of



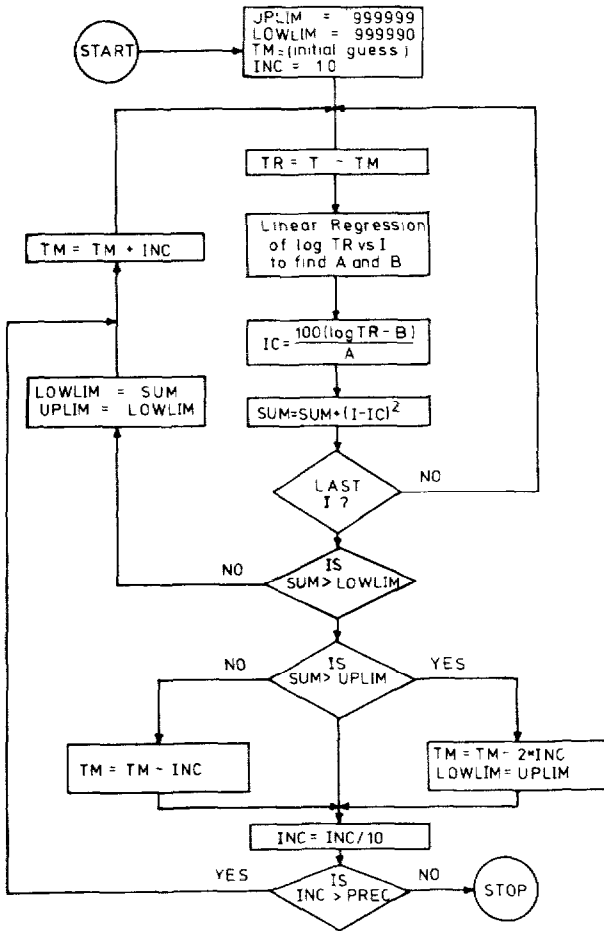


Fig 2. Flow chart for calculation of mathematical dead time by the method of Guardino *et al*<sup>24</sup>

$(-1)^n \cdot n \cdot (\text{initial increment})$ . Twelve is set as the maximum value of  $n$ . In addition, limits are placed on the permissible values for  $t_M$ , with negative values and values greater than the minimum retention time being rejected

Bellas also reported that the program had been extended to calculate the index for data from linear temperature-programmed gas chromatography by fitting a polynomial of degree equal to the number of data points.

Because of the similarity to the method of Guardino *et al.*<sup>24</sup> (the data used by Bellas give virtually identical results when used in the method of Guardino *et al.*), it will be assumed that the methods are equivalent and only the method of Guardino *et al* will be discussed further.

(c) Garcia Dominguez *et al.*<sup>48</sup> reported an iterative method (“exact calculator method”) in a paper which introduced several new methods (section 3.5g). A flow chart of the method is shown in Fig. 3.

The method initially uses the retention times of substances with known Kovats

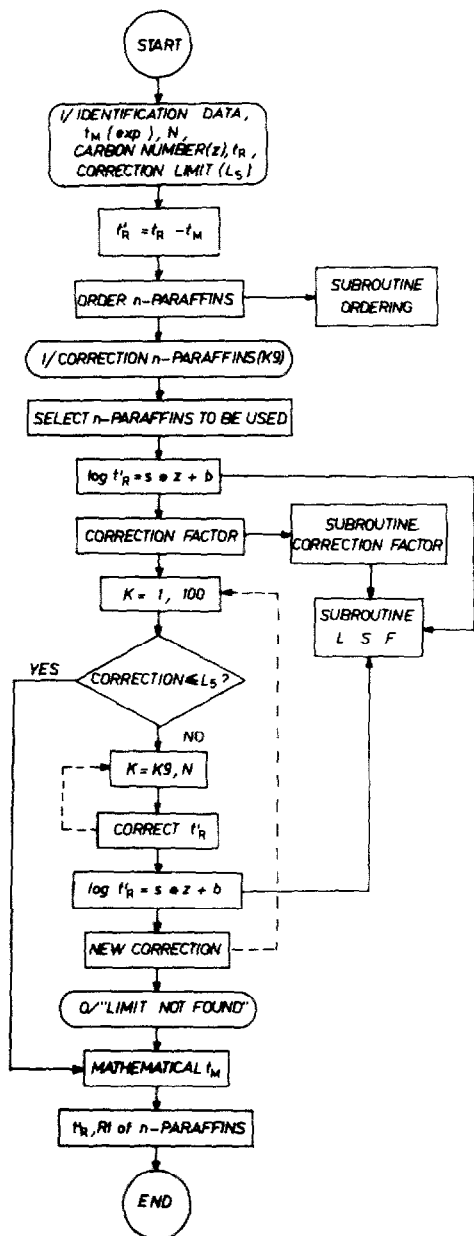


Fig 3. Flow chart for the "exact calculator method" of Garcia Dominguez *et al*<sup>48</sup>

indices (such as the  $n$ -alkane homologues), measured from any point on the chromatogram, in a linear regression. Letting the initial retention time of the alkane with the least retention be  $t_{Rz_0}^1$ , the coefficients determined from the regression can be used to determine a new retention time for that alkane of  $t_{Rz}^1$ . A correction factor is then calculated:

$$d_1 = t_{Rz_1^1} - t_{Rz_0^1} \quad (59)$$

This correction, which can be either positive or negative depending on whether the initial retention time of the first alkane,  $t_{Rz_0^1}$ , is shorter or longer than the true retention, is now applied to all peaks, producing a second set of retention times. These times are subsequently used in another linear regression and a second correction is calculated from the new retention of the alkane which initially had the lowest retention time:

$$d_2 = t_{Rz_2^1} - t_{Rz_1^0} \quad (60)$$

This equation is equivalent to eqn. 3 of Garcia Dominguez *et al.*<sup>48</sup>, except for a correction. The original equation implies that the correction is always the difference between the latest two values of the retention time of the lowest alkane. In fact, for the method to converge, the correction must be the difference between the latest value of the retention time and the originally measured retention time.

After a sufficient number of cycles a final set of adjusted retention values is obtained. A refinement to the program is that the actual correction applied to each set of retention times can be multiplied by a "correction factor" to reduce the number of iterations required by the program. Fig. 4 shows the effect the process has on the linearity of the log plot; the curvature of the log plot is reduced as the adjusted retention times approach their true values.

(d) Tóth and Zala<sup>62</sup> mentioned an iterative method and also introduced a classical method (see section 3.5f). The method calculates the parameters in eqn. 5 by first determining  $t_M$ . On the basis of data pairs  $(t_{Rz}, I)$  obtained by measurements on the  $n$ -alkanes, the coefficient of determination,  $R^2$ , between the quantities  $I$  and  $\ln(t_{Rz} - t_M)$  as a function of  $t_M$  can be expressed as:

$$R^2 = R^2(t_M) = \frac{\left\{ \left[ \sum_{i=z}^{z+n} \ln(t_{Ri} - t_M) I_i \right] - \frac{\sum_{i=z}^{z+n} I_i \cdot \sum_{i=z}^{z+n} \ln(t_{Ri} - t_M)}{n} \right\}^2}{\left\{ \sum_{i=z}^{z+n} I_i^2 - \frac{\left( \sum_{i=z}^{z+n} I_i \right)^2}{n} \right\} \left\{ \sum_{i=z}^{z+n} [\ln(t_{Ri} - t_M)]^2 - \frac{\left[ \sum_{i=z}^{z+n} \ln(t_{Ri} - t_M) \right]^2}{n} \right\}} \quad (61)$$

The dead time can be calculated from the value at which  $R^2(t_M)$  is a maximum. This value of  $t_M$  can be obtained by solving the non-linear equation

$$dR^2/dt_M = 0 \quad (62)$$

by an iterative technique. With a knowledge of  $t_M$ , estimates of  $b$  and  $c$  can be computed by linear regression of  $I$  versus  $\ln(t_{Rz} - t_M)$ .

Although no method was given for the iterative solution of  $dR^2/dt_M = 0$ ,

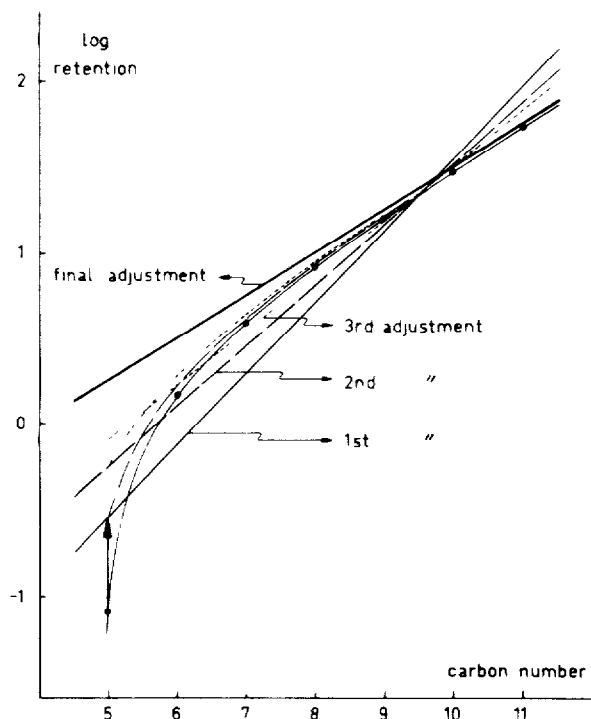


Fig 4 Schematic drawing of the "exact calculator method" of Dominguez *et al.*<sup>48</sup>. The points represent initial retention values. The arrow shows the first correction applied.

several methods of numerical differentiation are given in a book by Dorn and McCracken<sup>66</sup>. It should be noted that numerical differentiation suffers from subtractive cancellation, which can lead to very large errors. Because of the difficulty of overcoming this problem, numerical differentiation is not normally recommended if there are other methods available to solve the problem.

An alternative solution which is not only faster, but also does not involve the problems associated with numerical differentiation is to minimise  $1 - R^2$  using an appropriate iterative technique.

(e) One of the most versatile methods of dead-time estimation is undoubtedly the simultaneous non-linear least-squares estimation of  $t_M$ ,  $b$  and  $c$ <sup>28,30</sup> by the use of numerical minimisation<sup>67</sup>. When using this method it is necessary to define the objective function which is to be minimised. This should be done very carefully because different objective functions will lead to different estimations of  $t_M$ ,  $b$  and  $c$  as reported by us<sup>30</sup>. For the calculation of Kovats indices, the sum of squares of the differences between the known and calculated  $I$  values is a suitable objective function.

However, this method suffers from three problems. First it is very complex to program. Secondly, the execution time is much longer than in non-iterative methods and it can take as long as half an hour to achieve maximum accuracy on a micro-computer. Finally the method may converge to a local minimum rather than the global minimum. This is especially true with complex functions such as higher-degree

polynomials. The method thus requires good initial estimates of  $t_M$ ,  $b$  and  $c$  as well as a critical evaluation of the output of the program.

Therefore the method is not suitable for routine use and is only useful in the evaluation of other methods of calculating dead time.

### 3.8. Others

(a) A method was reported by Ševčík<sup>68</sup> and later extended by Ševčík and Löwentap<sup>29</sup> which does not require calculation of the column dead time. Later<sup>69</sup> the method was used for the classification of stationary phases. That paper showed that  $A$  was independent of the carrier gas flow-rate and pressure as well as a high degree of reproducibility with time. The method, which can be best understood by reference to Fig 5, uses the ratio of the differences in retention times between consecutive  $n$ -alkanes,  $A$ , to calculate adjusted retention times.

Defining the difference between consecutive gross retention times as  $\Delta$  gives:

$$\Delta_z = t_{Rz} - t_{R(z-1)} \quad (63)$$

Now, from Fig. 4, the adjusted retention time of a peak can be expressed as

$$t'_{Rz} = \Delta_1 + \Delta_2 + \Delta_3 + \dots + \Delta_{z-1} + \Delta_z \quad (64)$$

where  $\Delta_1$  is the time difference between the elution of a substance with  $I = 0$  and one with  $I = 100$ , etc.;  $\Delta_1$  actually equals the adjusted retention time of a homologue with a carbon number of one, i.e.,  $\Delta_1 = t'_{R1}$ .

Experimentally, Ševčík and Löwentap<sup>29</sup> found that the ratio of successive time differences in an homologous series was constant. Therefore:

$$\Delta_z = A \Delta_{z-1} \quad (65)$$

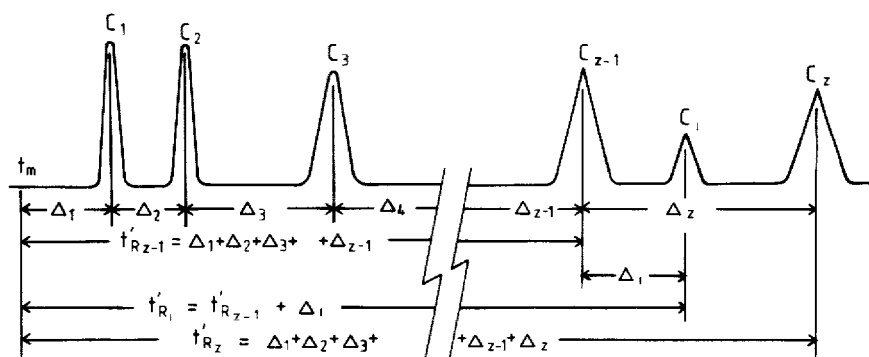


Fig 5. Schematic chromatogram illustrating the method of Ševčík and Löwentap<sup>29</sup>.  $t'_R$  = Adjusted retention time;  $\Delta_z$  = time difference between consecutive homologues;  $\Delta_i$  = time difference between compound  $i$  and the nearest homologue

Thus

$$\begin{aligned} \Delta_{z-1} &= A^{-1}\Delta_z \\ \Delta_{z-2} &= A^{-1}\Delta_{z-1} = A^{-2}\Delta_z \\ \Delta_{z-3} &= A^{-1}\Delta_{z-2} = A^{-3}\Delta_z \\ &\vdots \\ \Delta_2 &= A^{-1}\Delta_3 = A^{2-z}\Delta_z \\ \Delta_1 &= A^{-1}\Delta_2 = A^{1-z}\Delta_z \end{aligned}$$

and eqn. 64 can be written as.

$$\begin{aligned} t'_{Rz} &= \Delta_z + \Delta_{z-1} + \Delta_{z-2} + \dots + \Delta_2 + \Delta_1 \\ &= \Delta_z + A^{-1}\Delta_z + A^{-2}\Delta_z + \dots + A^{2-z}\Delta_z + A^{1-z}\Delta_z \\ &= \Delta_z(1 + A^{-1} + A^{-2} + \dots + A^{2-z} + A^{1-z}) \end{aligned} \quad (66)$$

This is a geometric series and can be summed as follows:

$$t'_{Rz} = \frac{\Delta_z[1 - (1/A)^z]}{1 - (1/A)} = \frac{\Delta_z(A^{z+1} - A)}{A^{z+1} - A^z} \quad (67)$$

Therefore the method allows the calculation of the adjusted retention time of any homologue without reference to the dead time. The adjusted retention of other substances can be calculated by adding the difference in retention times between the substance and any alkane to the adjusted retention time of the alkane as calculated by eqn. 64 or 67. Therefore in Fig. 5

$$t'_{Ri} = t'_{Rz-1} + \Delta_i$$

where  $t'_{Rz-1}$  is calculated by eqn. 64 or 67 and  $\Delta_i$  is the difference in retention times between a substance  $i$  and a standard  $z - 1$ .

The method can also be used for extrapolation by noting that eqn. 65 means that the dependence of the logarithm of the difference in times,  $\ln \Delta_z$ , versus  $z$  is linear. Thus:

$$\ln \Delta_z = dz + e \quad (68)$$

where  $d$  and  $e$  are constants.

Ševčík and Löwentap<sup>29</sup> also showed that their method leads to the conclusion that relative retention is dependent on the carbon number. Eqn. 64 can be expressed as:

$$\begin{aligned} t'_{Rz+1} &= \Delta_{z+1} + \Delta_z + \Delta_{z-1} + \dots + \Delta_3 + \Delta_2 + \Delta_1 \\ &= \Delta_{z+1} + \frac{\Delta_z(A^{z+1} - A)}{A^{z+1} - A^z} \end{aligned} \quad (69)$$

Combining eqn. 69 with eqn. 65 gives:

$$\begin{aligned} t'_{Rz+1} &= A\Delta_z + \frac{\Delta_z(A^{z+1} - A)}{A^{z+1} - A^z} \\ &= \frac{\Delta_z[A(A^{z+1} - A^z) + (A^{z+1} - A)]}{A^{z+1} - A^z} \\ &= \frac{\Delta_z(A^{z+2} - A)}{A^{z+1} - A^z} \end{aligned} \quad (70)$$

Dividing eqn. 70 by eqn. 67 gives:

$$\begin{aligned} \frac{t'_{Rz+1}}{t'_{Rz}} &= \frac{\Delta_z(A^{z+2} - A)}{A^{z+1} - A^z} \cdot \frac{1}{\Delta_z} \cdot \frac{A^{z+1} - A^z}{A^{z+1} - A} \\ &= \frac{A^{z+2} - A}{A^{z+1} - A} \end{aligned} \quad (72)$$

Rearranging eqn. 72 gives:

$$\frac{t'_{Rz+1}}{t'_{Rz}} = \frac{A - (1/A^z)}{1 - (1/A^z)} \quad (73)$$

From this expression it can be seen that as  $z \rightarrow \infty$  then  $t'_{Rz+1}/t'_{Rz} \rightarrow A$ . Thus the method predicts that relative retention will only be constant at high values of  $z$ .

Unfortunately a similar analysis of eqn. 5 gives a very different result

$$\frac{t'_{Rz+1}}{t'_{Rz}} = A \quad (74)$$

where  $A = e^b$  (see eqn. 56). This discrepancy leads to one of two difficulties with the method. Because the method has no independent procedure for calculating Kovats indices of substances other than the homologues, it is left without a consistent procedure for calculating Kovats indices, as the usual formula (eqn. 5) is no longer valid and their relationship (eqn. 68) is only suitable for calculating adjusted retention times. It is not suitable for the reverse process of calculating  $z$  numbers (and thus  $I$  values) from the adjusted retention time of some unknown substance.

The second problem arises because the method leads to different dead times for each homologue. Although the authors assume that this is due to errors in reported retention times, their own data show a disturbing variation in dead time between the measured data (high  $z$ ) and the extrapolated data (low  $z$ ) as shown in Table 1. Ettre<sup>31</sup> has discussed this particular problem with specific reference to the method of Al-Thamir *et al.*<sup>59</sup>. He does, however, point out that the problem exists in all methods which calculate adjusted retention times without first estimating a gas hold-up time.

Finally it is worth pointing out that workers from the University of New South Wales have found that the method is a very sensitive indicator of linearity and have published a series of papers<sup>43-47</sup> in which it is used to test the linearity of a range of homologues on a wide selection of columns.

TABLE I  
DATA FROM ŠEVČÍK AND LÖWENTAP<sup>20</sup> SHOWING VARIATION IN DEAD TIME

$z$	Measured $t_{Rz}$ (mm)	$\Delta_z$ (mm)	$A$	$\Delta_z$ (mm) (calc)	$t'_{Rz}$	$r = \frac{t'_{Rz+1}}{t'_{Rz}}$	$t_M^*$ (mm)
20	2598	704		703.56	2564.46	1.3781	33.5
19	1894	510	1.3804	510.86	1860.90	1.3784	33.1
18	1384	371	1.3747	370.94	1350.04	1.3789	34.0
17	1013	269	1.3792	269.34	979.10	1.3795	33.9
16	744	196	1.3724	195.57	709.76	1.3804	34.2
15	548	142	1.3803	142.00	514.19	1.3815	33.8
14	406	103	1.3786	103.11	372.19	1.3832	33.8
13	303	77	1.3377	74.87	269.08	1.3856	33.9
12	226	56	1.3750	54.36	194.21	1.3888	31.8
11	170	40	1.4000	39.47	139.85	1.3931	30.2
10	130	29	1.3793	28.66	100.38	1.3996	29.6
9	101			20.81	71.72	1.4090	29.3
8	(80.19)**			15.11	50.91	1.4222	29.3
7	(65.08)			10.97	35.80	1.4414	29.3
6	(54.11)			7.97	24.83	1.4745	29.3
5	(46.14)			5.78	16.85	1.5212	29.3
4	(40.36)			4.20	11.07	1.6114	29.3
3	(36.16)			3.05	6.87	1.7984	29.3
2	(33.11)			2.21	3.82	2.3727	29.3
1	(30.90)			1.61			29.3
0	(29.29)						

\* Estimated column dead time,  $t_M = t_{Rz} - t'_{Rz}$ .

\*\* Figures in brackets are calculated gross retention times,  $t_{Rz} = t_{Rz+1} - \Delta_z$ .

(b) All methods discussed so far have relied on the accepted linear relationship between  $\ln(t_{Rz} - t_M)$  and  $I$ , even though its linearity for homologues with low carbon numbers has been questioned (see section 4). However, a new method based on a cubic relationship between  $\log t'_{Rz}$  and  $I$  has been reported by Heeg *et al.*<sup>70</sup>. The assumption of non-linearity was based on reports in the literature of experimental results<sup>74</sup> as well as on thermodynamic considerations<sup>71</sup>.

Heeg *et al.*<sup>70</sup> compared several traditional methods of calculating retention indices with methods based on a quadratic and a cubic fit:

$$\log t'_{Rz} = az^3 + bz^2 + cz + d \quad (75)$$

For Kovats index calculation, eqn. 75 can be more conveniently expressed in the form:

$$I = a'(\log t'_{Rz})^3 + b'(\log t'_{Rz})^2 + c'(\log t'_{Rz}) + d' \quad (76)$$

This comparison led to the conclusion that a quadratic fit was inadequate and that a cubic fit (eqns. 75 and 76) was necessary to obtain good accuracy.

However, the method presented a problem because the authors were not able to manipulate the cubic equation in such a way that would allow it to be solved for



$t_M$ . Thus they recommended that  $t_M$  be first calculated using two linear regressions as outlined by Grobler and Balizs<sup>61</sup>. Having calculated  $t_M$ , the constants  $a$ ,  $b$ ,  $c$  and  $d$  can be found by using one of several methods of fitting data by a cubic equation.

All methods designed to fit cubic and higher-degree polynomials are based on matrix manipulations. Such methods are relatively complex and suffer from round-off and subtractive cancellation errors. It is therefore essential that the maximum number of significant figures be carried throughout the calculations. We have found that the method of Gaussian elimination with partial pivotal condensation<sup>66</sup> using double precision arithmetic (with 16 digits accuracy) gives acceptable results.

A final problem with this method is its use of a dead-time value calculated by a linear method when the data are actually being fitted by a cubic polynomial. This is mathematically undesirable and likely to produce a less than optimum fit. A way of overcoming this problem is to modify one of the iterative methods to allow a cubic fit instead of a linear fit. Such an approach sacrifices speed in an attempt to improve the fit.

#### 4 LINEARITY OF THE PLOT OF LOG ADJUSTED RETENTION TIME VS CARBON NUMBER

Almost all methods of calculating Kovats indices are based on eqn. 5 which states that the relationship between the log of the adjusted retention time and carbon number is linear. Although some authors have accepted that this relationship is linear over the entire range of carbon numbers<sup>72,73</sup>, an increasing amount of evidence has accumulated that the relationship is non-linear for low carbon numbers. This evidence is both experimental and theoretical

##### 4.1. The experimental evidence

In a large number of papers experimental evidence has been presented as to the non-linearity of the log relationship for homologues with low carbon numbers. The most extensive evidence has been presented by Haken, Wainwright and Srisukh from the University of New South Wales<sup>42-47</sup>. Although their first paper<sup>42</sup> (which is more concerned with mathematical dead time *versus* methane retention than the question of linearity) present evidence for the linearity of this relationship for C<sub>1</sub>-C<sub>9</sub> alkanes on both SE-30 and OV-25, a close examination of the results in conjunction with results presented in a later paper<sup>43</sup> strongly indicates that the apparent linearity was a result of the relative high temperatures used in the study.

First the results on OV-25 show a calculated methane retention time which is consistently 2-3 sec less than the measured retention time, thus indicating a slight non-linearity. The retention times, which were calculated using the method of Grobler and Balizs<sup>61</sup>, are confirmed in Table 2, where the retention time of methane has been recalculated for five sets of conditions using the method developed by Ševčík and Löwentap<sup>29</sup>. Although the discrepancy is small it translates into Kovats indices which range from approximately 180 to 250. Secondly, although the SE-30 data show linearity at temperatures of 120-140°C, a later paper<sup>43</sup> shows non-linearity for C<sub>1</sub>-C<sub>4</sub> alkanes on both SE-30 and OV-25 at 30°C.

The other papers by these workers provide evidence for non-linearity of the relationship for lower members of *n*-alkanes, *n*-alcohols, *n*-aldehydes, acetates and

TABLE 2

DATA FROM REF. 42 RECALCULATED USING THE METHOD OF ŠEVČÍK AND LÖWEN-TAP<sup>29</sup>

Carbon number	Retention time* $\Delta$ (sec)	A	Retention time* $\Delta$ (sec)	A	Retention time* $\Delta$ (sec)	A
SE-30	(120°C, 30 ml/min)		(130°C, 40 ml/min)		(140°C, 45 ml/min)	
1	116.0 (116.5)** (2.45)		96.2 (96.8)** (1.53)		80.7 (81.3)** (1.35)	
2	(118.95) (4.33)	(1.77)	(98.30) (2.70)	(1.76)	(82.66) (2.28)	(1.69)
3	(123.28) (7.66)	(1.77)	(101.20) (4.75)	(1.76)	(84.94) (3.85)	(1.69)
4	(130.94) (13.56)	(1.77)	(105.95) (8.35)	(1.76)	(88.79) (6.51)	(1.69)
5	144.5 24.0	(1.77)	114.3 14.7	(1.76)	95.3 11.0	(1.69)
6	168.5 43.40	1.81	129.8 27.5	1.87	106.3 18.9	1.72
7	211.9 76.1	1.75	157.3 46.4	1.69	125.2 31.5	1.69
8	288.0 134.1	1.76	203.7 79.5	1.71	156.7 52.0	1.65
9	422.1		283.2		208.7	
OV-25	(120°C, 45 ml/min)		(130°C, 45 ml/min)			
1	197.2 (194.9)** (2.03)		191.1 (189.4)** (1.72)			
2	(196.94) (3.53)	(1.74)	(191.12) (2.93)	(1.70)		
3	(200.47) (6.14)	(1.74)	(194.05) (4.98)	(1.70)		
4	(206.61) (10.69)	(1.74)	(199.03) (8.47)	(1.70)		
5	217.3 18.6	(1.74)	207.5 14.4	(1.70)		
6	235.9 31.9	1.72	221.9 25.8	1.79		
7	267.8 55.7	1.75	247.7 42.1	1.63		
8	323.5 96.9	1.74	289.8 71.0	1.69		
9	420.4		360.8			

\* Measured retention times (sec) are the average values from two or three determinations

\*\* Calculated values are shown in brackets

methyl ketones on a series of porous polymers as well as for *n*-alkanes on porous silica.

Heeg *et al.*<sup>70</sup> found that the log relationship was non-linear and developed a method based on a cubic fit to determine retention indices. In addition, Heldt and Köser<sup>51</sup> presented evidence in their 1980 paper that the relationship is non-linear.

Parcher and Johnson<sup>5</sup> as well as Lin and Parcher<sup>6</sup> presented evidence for the non-linear behaviour of methane, and suggested that an effective carbon number of 0.5 should be assigned to it in order to linearise the log relationship. This approach has been discussed by Wainwright and Haken<sup>4,7</sup> who pointed out several problems, including their findings that the effective carbon number of methane varied with the column packing, that it was closer to 1.0 than 0.5 and that other *n*-alkanes showed non-linear behaviour. There is also a fundamental problem with changing the carbon number of methane. This relates to the fact that the retention index system is based on the definition that the retention index of an *n*-alkane is 100 times its carbon number. The setting of the effective carbon number of methane to anything other than 1 is theoretically undesirable. Even if the practice led to accurate dead times, it would lead to difficulties in the definition and calculation of retention indices.

Riedo *et al.*<sup>52</sup> carried out an extensive study during an 8-month period. About

1500 retention values for *n*-alkanes from pentane to pentadecane were determined on two columns at temperatures from 30 to 250°C. The results were used to calculate thermodynamic properties which were then plotted, together with the results for the light hydrocarbons (C<sub>1</sub>–C<sub>4</sub>), as a function of the carbon number of the *n*-alkane. They concluded that “this plot is not linear. Therefore, the determination of the starting point of the chromatogram is not justified by linearization of the logarithms of the retention data of *n*-alkanes”.

#### 4.2. The theoretical evidence

Rohrschneider<sup>74</sup> was one of the first to discuss the linearity of the log plot in relation to basic thermodynamic parameters. He reported that the slope of the *n*-alkane curves depended on the ratio of the vapour pressure, *p*, and activity coefficient,  $\gamma$ :

$$b = \log \frac{t'_{R_{z+1}}}{t'_{R_z}} = \log \frac{p_z}{p_{z+1}} + \log \frac{\gamma_z}{\gamma_{z+1}} \quad (77)$$

By calculating this value for C<sub>2</sub> through C<sub>10</sub> he was able to show that *n*-alkane curves show a definite deviation from linearity caused above all by the non-linear vapour-pressure dependence.

Sojak *et al.*<sup>75</sup> proposed a relationship between the retention index, boiling point and the activity coefficient of a substance. Starting with the same equation as did Rohrschneider<sup>74</sup> (eqn. 77), they extended the relationship to boiling points using a relation derived by Purnell for a compound at a given temperature:

$$\log p^0 = k_1 + k_2 T_b \quad (78)$$

where *k*<sub>1</sub> and *k*<sub>2</sub> are constants. Thus:

$$\log (p_2^0/p_1^0) = k_2(T_{b,2} - T_{b,1}) \quad (79)$$

Combining eqn. 79 with eqn. 77 gives:

$$\log \frac{t'_{R_2}}{t'_{R_1}} = -k_2(T_{b,2} - T_{b,1}) - \frac{\log \gamma_2}{\gamma_1} \quad (80)$$

Although eqns. 78–80 are only approximations and are not valid for substances capable of hydrogen bonding, the general derivation supports Rohrschneider's work in this area.

Bach *et al.*<sup>71</sup> discussed, using basic thermodynamics, the dependence of retention times on boiling points. By using an equation derived by Klages, in which the square of the boiling temperature of an organic substance is represented as an additive quality of the molecule, they established an analytical relationship between retention times and carbon numbers. Simplifying approximations enabled the relationship to be linearised and thus the Kovats system could be introduced. The range of linearity was found to commence approximately at pentane.

Therefore, the theoretical evidence supports the experimental observations of the non-linearity of the log plot at low carbon numbers. The reason that this non-linearity has not been uniformly recognised is a result of several factors. First the non-linearity is minimised for homologues with large carbon numbers, at high temperatures and on non-polar stationary phases. Secondly, it can be obscured (if not large) by the mathematical treatment of the data. This problem has been discussed by both Rohrschneider<sup>74</sup> and Ettre<sup>31</sup>.

The data in Table 3 show the problems that can arise. The retention data in column 1 were generated assuming a dead time of 100 sec and adjusted retention times conforming to eqn. 79:

$$\ln t'_{Rz} = 0.5z + 0.3 \quad (81)$$

Column 2 shows the inaccuracies introduced purely by rounding off the data to the nearest 0.1 sec. It should be noted that already an error of up to 1.3 units in Kovats retention indices has resulted, even though the coefficient of determination,  $R^2$ , is 0.99998. Column 3 adds a non-linearity for  $C_1$ – $C_3$ ; their effective carbon numbers are shown in brackets. The unexpected finding is that even though the curvature is significant, the accuracy of fit as shown by both  $R^2$  and the individual index values is almost identical to that in column 2. Column 4 introduces an even larger non-

TABLE 3

HYPOTHETICAL COMPARISON OF THE EFFECTS OF A NON-LINEAR BEHAVIOUR OF THE LOWER ALKANES WITH SMALL RANDOM DEVIATIONS

$z$	$t_{R_1}$ (correct)	$t_{R_2}$ (rounded off)	$t_{R_3}$ (non-linear)	$t_{R_4}$ (non-linear)	$t_{R_5}$ ( $t_{R_2}$ + deviat)	$t_{R_6}$ ( $t_{R_4}$ + deviat)
1	102.23	102.2	101.9 [0.7]*	101.7 [0.5]*	102.1	101.6
2	103.67	103.7	103.4 [1.8]*	103.2 [1.7]*	103.8	103.3
3	106.05	106.1	105.9 [2.9]*	105.5 [2.8]*	106.2	105.6
4	109.97	110.0	110.0	109.5 [3.9]*	110.0	109.5
5	116.44	116.4	116.4	116.4	116.3	116.3
6	127.11	127.1	127.1	127.1	127.0	127.0
7	144.70	144.7	144.7	144.7	144.7	144.7
$t_M^{**}$	100.0	99.9	99.5	99.5	99.6	99.2
$b^{**}$	0.500	0.495	0.488	0.502	0.480	0.486
$c^{**}$	0.297	0.336	0.396	0.306	0.435	0.408
$R^{2**}$	1.00000	0.99998	0.99998	0.99992	0.99967	0.99973
<i>Kovats retention indices</i>						
1	100.0	99.3	100.3	99.7	96.8	97.3
2	200.0	201.3	199.1	201.9	206.4	207.0
3	300.0	300.6	300.2	297.4	301.5	298.3
4	400.0	399.4	401.5	398.7	396.8	396.0
5	500.0	498.7	498.9	503.0	495.9	500.1
6	600.0	599.9	599.4	600.6	599.3	600.0
7	700.0	700.8	700.5	698.8	703.4	701.3

\* Effective carbon number

\*\* Calculated by the method of Guardino *et al*<sup>24</sup>

linearity into the  $C_1$ – $C_4$  data. This time the fit is not as good with a maximum error in the Kovats indices of 3. However,  $R^2$  is still a respectable 0.99992. An important feature of this analysis is that it is impossible to assign the less than perfect fit to its true cause, the non-linearity. The variation in Kovats indices appears to be completely random. This problem is emphasised by the data in columns 5 and 6. These columns have had an identical deviation of  $\pm 0.1$  sec added both to the “correct” data and the most non-linear data. The result is that the data from both columns give fits of almost equal accuracy, showing that the non-linearity has been “eliminated” by mathematical manipulation.

Therefore this analysis is in agreement with those authors<sup>28,30,31,74</sup> who have argued both that random deviations can introduce significant errors into the calculation of mathematical dead time and that the actual treatment of the data can obscure a non-linearity. In particular, Ettre<sup>31</sup> was critical of the use of the full range of homologues for this very reason, suggesting that only the linear part of the curve should be used. However, as pointed out in our recent paper<sup>64</sup>, this is not really a problem when calculating Kovats indices where the real requirement is accurate and reproducible index values. The exclusion of some of the data from the analysis is likely to lead to significant errors in those regions which have been excluded.

On the other hand, the problem of random deviations is a real one and thus a higher accuracy than has generally be recognised is required to obtain reproducible results. Retention times should be measured to at least the nearest 0.1 sec unless very long times are involved, in which case a lower accuracy may be acceptable.

## 5 COMPARISON OF METHODS

Several papers<sup>18,27,28,30,32</sup> have previously reviewed the calculation of dead time in general, while a larger number of papers have compared two or three specific methods (see for example refs. 31, 36–38). In this section we will summarise the conclusions that have been drawn in previous sections of this paper as well as presenting results from our recent paper<sup>64</sup> in which we carried out an extensive comparison of several statistical and iterative methods.

The use of theoretical and graphical methods as well as methods involving indirect measurement was discussed in section 3. As these are considered complex, time-consuming and of very limited usefulness, they will not be discussed further. Direct-measurement techniques were extensively covered in section 3.3. In general such methods are either experimentally difficult or likely to introduce significant errors into dead-time and retention-index calculations. The use of methane in particular was shown to lead to large errors in many cases and should be discouraged. The only method which was shown to lead to accurate results under a wide variety of conditions was the use of neon or possibly helium together with the appropriate treatment of the data. In fact, the use of multiple analyses using neon is, in our opinion, the most reliable and accurate method of determining system dead time and is the method of choice when determining absolute data such as Henry's law values or other thermodynamic data. However, due to the difficulty of using this method (gas chromatography–mass spectrometry has been recommended for the determination), a mathematical determination is more useful for the vast majority of analyses, where the final data are comparative, such as in the calculation of Kovats indices.

Of the various mathematical methods, several authors have been critical of the use of the classical methods, especially when compared to the use of methane. In particular, the results of Guberska<sup>36,37</sup>, Sharples and Vernon<sup>38</sup> as well as our own results reported in an earlier paper<sup>30</sup> show that large errors can result from the use of such methods. This problem is especially acute when inaccurate timing methods are used, as reported by Sharples and Vernon. We therefore believe that the inaccuracies inherent in these methods, together with the availability of inexpensive programmable calculators and notebook-size computers, render them of limited use other than to provide an initial estimate for a more accurate iterative method.

Of the statistical and iterative methods, the method of Ševčík and Löwentap<sup>29</sup> was discussed in section 3.8a. Because of the problems discussed in that section, including Ettore's<sup>31</sup> comments on a similar method, we believe that the method offers no advantage over other methods. However, it should be noted that several workers in our laboratories<sup>42-47</sup> have reported that the method is a sensitive indicator of linearity. This also means that it is a good indicator of the degree of scatter of the data and therefore holds merit as an initial test of the accuracy of a new set of retention data.

This leaves the following statistical and iterative methods.

(1) The method of Grobler and Balizs<sup>61</sup> as extended by the technique of Van Tulder *et al.*<sup>63</sup>

(2) The method van Guardino *et al.*<sup>24</sup>

(3) The "exact calculator method" of Garcia Dominguez *et al.*<sup>48</sup>

(4) The method of Ambrus<sup>76</sup>, also extended by the technique of Van Tulder *et al.*<sup>63</sup>

(5) The iterative method mentioned by Tóth and Zala<sup>62</sup>

(6) The flexible Simplex method<sup>67</sup>

(7) The method of Heeg *et al.*<sup>70</sup>

These seven methods, together with two modified approaches allowing the optimisation of  $t_M$  while simultaneous fitting a cubic or fifth-degree polynomial to the data, have been compared in our recent paper<sup>64</sup>. That paper showed that the linear methods (methods 1-6 above) gave similar estimates of the mathematical dead time in most cases. In fact, the methods of Tóth and Zala<sup>62</sup>, Guardino *et al.*<sup>24</sup> and flexible Simplex<sup>67</sup> gave identical results in all cases.

The paper also showed that for the calculation of Kovats indices the method of Guardino *et al.*<sup>24</sup> was not only the fastest, but the most accurate. This conclusion, while differing from those expressed in earlier papers, results from the recently recognised instability inherent in methods such as those of Ambrus<sup>76</sup> or Grobler and Balizs<sup>61</sup>. These methods, while providing an acceptable estimate of the equation parameters in the majority of cases, can lead to large errors in other cases. This means that all homologue combinations must be checked to reduce the likelihood of introducing significant errors into the calculation. However, this procedure increases both the program's complexity and its running time.

With regard to the polynomial methods, the paper concluded that in general they offer only a small improvement over the linear methods. Therefore, since these methods involve greater complexity, including the need to carry a large number of significant figures throughout the calculations, and can be difficult to interpret, they

should only be used when extreme accuracy is required or where highly variable data are involved. In this case, an improvement in the experimental technique is more likely to increase accuracy than the use of a more complex method.

In addition these methods do not give an estimate of the mathematical dead time. The parameter ( $t_M$ ), which in the linear methods is accepted as an estimate of the system dead time, shows no relation to the dead time in polynomial methods and thus should be treated in the same manner as the other equation parameters, with the single exception that its value must be less than the retention time of the homologue with the lowest retention time.

Finally the paper investigated the problems associated with extrapolation, particularly with the polynomial methods. It concluded that this technique should be avoided where good accuracy is required.

## 6 CONCLUSIONS AND RECOMMENDATIONS

In choosing a method to analyse the raw data from a chromatographic analysis, several factors should be taken into account. These include the accuracy of the original data, the required accuracy of the final result and the purpose to which the data are to be put.

When the resultant data are to be used to calculate some property of the column or stationary phase such as Henry's law constants, the system dead time should, where possible, be determined experimentally using an inert gas such as neon. On the other hand, where the data are to be used in a comparative method, such as the calculation of Kovats indices, a mathematical method is more appropriate.

With the availability of inexpensive programmable calculators, the method of choice is an iterative one such as that of Guardino *et al.* This method, while more time-consuming than the classical methods, overcomes the problems with accuracy inherent not only in the classical methods, but to a lesser degree in the statistical methods.

For maximum accuracy a polynomial method can be employed. However, not only must double precision be used throughout these methods, but all results should be critically evaluated before being accepted. In addition such methods cannot be used for the estimation of the system dead time.

The reproducibility of Kovats indices is dependent not only on the mathematical method used to manipulate the raw data, but also on the original data. The combination of a well designed experimental technique, an accurate timing mechanism (nearest tenth of a second) and an appropriate mathematical method will give an interlaboratory reproducibility of one unit for larger values of Kovats indices and two units for indices below approximately 400 (ref. 64).

## 7. SUMMARY

To increase the usefulness of gas chromatography, a method of data presentation is required that removes interlaboratory variations. All such methods require the determination of the system dead time, with the most useful being the retention index system of Kovats. This paper reviews the many theoretical, experimental and mathematical methods of estimating system dead time and evaluates their usefulness,

not only in the estimation of dead time, but also in the subsequent calculation of Kovats indices. Recommendations are made as to the appropriate method to be used, and the expected accuracy in retention index values is estimated.

## REFERENCES

- 1 A T James and A. J P. Martin, *Biochem J*, 50 (1952) 679
- 2 R. Kaiser, *Chromatographia*, 2 (1969) 125.
- 3 R. Kaiser, *Chromatographia*, 3 (1970) 127.
- 4 R. Kaiser, *Chromatographia*, 3 (1970) 383
- 5 J F Parcher and D M Johnson, *J. Chromatogr Sci.*, 18 (1980) 267.
- 6 P J. Lin and J F Parcher, *J Chromatogr. Sci.*, 20 (1982) 33
- 7 H. Van Den Dool and P Dec Kratz, *J. Chromatogr.*, 11 (1963) 463
- 8 G. Schomburg and G Dielmann, *J. Chromatogr Sci.*, 11 (1973) 151
- 9 E Kovats, *Helv Chim Acta*, 41 (1958) 1915
- 10 L S Ettre, *Anal Chem*, 36 (8) (1964) 31A
- 11 E Kovats, *Chimia*, 22 (1968) 459.
- 12 *1984 Annual Book of ASTM Standards*, Vol. 14 01, E 355-77 (Reapproved 1983)
- 13 L. S. Ettre, *Chromatographia*, 6 (1973) 489
- 14 L. Mazor, E B Molnar, L Kaplar and J Takacs, *J Gas Chromatogr*, 6 (1968) 401
- 15 G. Dahlmann, H. J K Köser and H. H. Oelert, *J. Chromatogr*, 171 (1979) 398
- 16 A Wehrli and E Kovats, *Helv. Chim Acta*, 42 (1959) 2709
- 17 J. Novák, J Vejrosta, M Roth and J Janák, *J Chromatogr*, 199 (1980) 209.
- 18 M V Budahegyi, E R Lombosi, T S Lombosi, S Y Mészáros, Sz Nyiredy, G Tarján, I Timár and J. M. Takács, *J Chromatogr*, 271 (1983) 213.
- 19 J Takács and D Králik, *J. Chromatogr*, 50 (1970) 379
- 20 C A. Cramers, J. A. Rijks, V. Pacáková and I Ribeiro de Andrade, *J. Chromatogr*, 51 (1970) 13
- 21 M. B. Evans and J F Smith, *J Chromatogr.*, 6 (1961) 293
- 22 B Peetre, *Chromatographia*, 6 (1973) 257
- 23 V. A Ezrets and M S Vigdergauz, *Chromatographia*, 9 (1976) 205
- 24 X Guardino, J Albaigés, G. Firpo, R Rodríguez-Viñals and M Gassiot, *J Chromatogr*, 118 (1976) 13.
- 25 R. Vanheertum, *J Chromatogr Sci.*, 13 (1975) 150
- 26 G Tarjan, *J Chromatogr Sci.*, 14 (1976) 309
- 27 R Kaiser, *Chromatographia*, 2 (1969) 215
- 28 J K. Haken, M S Wainwright and R J Smith, *J Chromatogr*, 133 (1977) 1
- 29 J Ševčík and M. S. H. Löwentap, *J. Chromatogr.*, 147 (1978) 75.
- 30 R J. Smith, J. K. Haken and M. S. Wainwright, *J Chromatogr.*, 147 (1978) 65
- 31 L. S. Ettre, *Chromatographia*, 13 (1980) 73
- 32 M. S. Wainwright and J K Haken, *J Chromatogr*, 184 (1980) 1
- 33 M. B Evans and J. F Smith, *J. Chromatogr*, 9 (1962) 147
- 34 W O. McReynolds, *J Chromatogr Sci*, 8 (1970) 685
- 35 M. B. Evans, *Chromatographia*, 6 (1973) 301
- 36 J Guberska, *Chem Anal (Warsaw)*, 18 (1973) 1059
- 37 J Guberska, *Chem. Anal. (Warsaw)*, 19 (1974) 161
- 38 W. E Sharples and F. Vernon, *J. Chromatogr*, 161 (1978) 83.
- 39 A K. Hilmi, *J Chromatogr*, 17 (1965) 407.
- 40 M. Riedmann, *Chromatographia*, 7 (1974) 59
- 41 H L Hansen and K Andresen, *J Chromatogr*, 34 (1968) 246
- 42 M S Wainwright, J. K Haken and D Srisukh, *J. Chromatogr*, 179 (1979) 168
- 43 M S Wainwright, J. K. Haken and D. Srisukh, *J. Chromatogr*, 188 (1980) 246
- 44 M S. Wainwright, J. K Haken and D Srisukh, *J. Chromatogr*, 198 (1980) 156
- 45 M. S Wainwright, J K Haken and D. Srisukh, *J. Chromatogr.*, 236 (1982) 1
- 46 J K Haken, M S Wainwright and D Srisukh, *J. Chromatogr*, 257 (1983) 107.
- 47 M S. Wainwright and J. K Haken, *J. Chromatogr.*, 256 (1983) 193.
- 48 J A Garcia Dominguez, J Garcia Munoz, E Fernandez Sanchez and M. J Molera, *J Chromatogr Sci.*, 15 (1977) 520



- 49 M R Becerra, E Fernandez Sanchez, J A Garcia Dominguez, J Garcia Munoz and M J Molera, *J Chromatogr. Sci.*, 20 (1982) 363.
- 50 J R Ashes, S. C Mills and J. K Haken, *J. Chromatogr.*, 166 (1978) 391
- 51 U Heldt and H J K Köser, *J Chromatogr.*, 192 (1980) 107
- 52 F. Riedo, D Fritz, G Tarján and E sz Kováts, *J Chromatogr.*, 126 (1976) 63
- 53 M L Peterson and J Hirsch, *J. Lipid Res.*, 1 (1959) 132
- 54 G Castello and P. Parodi, *Chromatographia*, 4 (1971) 147.
- 55 H J Gold, *Anal. Chem.*, 34 (1962) 174
- 56 R E Kaiser, *Chromatographia*, 7 (1974) 251
- 57 S Ebel and R E Kaiser, *Chromatographia*, 7 (1974) 696.
- 58 M Haferkamp, in R Kaiser (Editor), *Chromatographie in der Gasphase*, Teil II, Bibliographisches Institut, Mannheim, 1966, p. 93.
- 59 W K Al-Thamir, J H Purnell, C A Wellington and R. J. Laub, *J Chromatogr.*, 173 (1979) 388.
- 60 A Tóth and E Zala, *J Chromatogr.*, 298 (1984) 381.
- 61 A Grobler and G Balizs, *J. Chromatogr. Sci.*, 12 (1974) 57
- 62 A Tóth and E Zala, *J Chromatogr.*, 284 (1984) 53
- 63 P J. M van Tulder, J P. Franke and R. A. de Zeeuw, presented at the VIIth International Symposium on Column Liquid Chromatography, Baden-Baden, May 3–6, 1983.
- 64 R J. Smith, J K Haken, M. S. Wainwright and B. G. Madden, *J Chromatogr.*, 328 (1985) 11
- 65 T. E Bellas, *Chromatographia*, 8 (1975) 38
- 66 W S Dorn and D. D McCracken, *Numerical Methods with Fortran IV Case Studies*, Wiley, New York, 1972
- 67 J A. Nelder and R Mead, *Comput. J.*, 7 (1965) 308
- 68 J Ševčík, *J Chromatogr.*, 135 (1977) 183.
- 69 J Ševčík and M S. H. Löwentap, *J Chromatogr.*, 217 (1981) 139.
- 70 F J. Heeg, R. Zinburg, H. J Neu and K. Ballschmiter, *Chromatographia*, 12 (1979) 451.
- 71 R W. Bach, E Dotsch, H. A Friedrichs and L Marx, *Chromatographia*, 4 (1971) 459.
- 72 J F Smith, *Nature (London)*, 193 (1962) 679
- 73 H. Groenendijk and A. W C van Kemenade, *Chromatographia*, 1 (1968) 472
- 74 L. Rohrschneider, *Chromatographia*, 2 (1969) 437
- 75 L. Sojak, J Krupčík and J Ríjks, *Chromatographia*, 7 (1974) 26
- 76 L. Ambrus, *J. Chromatogr.*, 294 (1984) 328