Vapour-Liquid Equilibrium of Mixtures

Worked Examples

Module 4

2105603 Advanced Chemical Engineering Thermodynamics

Soorathep Kheawhom

Department of Chemical Engineering
Faculty of Engineering
Chulalongkorn University

Eight problems worked end to end, with the arithmetic shown. Every dataset is a published measurement and carries its DOI. Every numerical answer is reproduced by `toolkit/verify_worked_examples.py`, which runs the same hand chain printed here — rounded intermediates and all — against the `vlekit` library at full precision and compares the two. 102 comparisons; none outside tolerance. Where the last quoted digit differs because the hand chain carried a rounded number, the text says so instead of adjusting the figure.

How to use this

These problems are the calculations of Module 4, in the order the module builds them. They are not a summary of the lectures and they will not substitute for the deck; every one of them assumes you have seen the argument the slide makes and want to know what the argument costs in arithmetic.

Three things are deliberate.

One system runs through most of them. Chloroform (1) + 2-butanone (2) at 303.15 K is the only isothermal P-x-y set in the module, so it is the only one on which Gibbs-Duhem appears in its clean form with no enthalpy term. Problems 1–4, 6 and 8 all use it, and the parameters fitted in Problem 3 are the parameters used in Problems 4, 6 and 8. That continuity is the point: a real reduction is one chain, not seven unrelated exercises. Problem 5 uses acetone + chloroform because it is short enough to integrate by hand, and Problem 7 uses ethanol + water because that is where the constant-α shortcut fails loudly.

The arithmetic is printed, not summarised. Every intermediate that a calculator would display is written down. If a quantity is not in the text you should not need it. Where a step is a loop, the loop is tabulated iteration by iteration rather than described as "iterate to convergence".

Each problem ends by naming the misconception it targets. That paragraph is the reason the problem is in the set. The number is the vehicle.

Conventions

T in K, P in kPa, compositions as mole fractions of component 1 throughout. The gas constant is $R = 8.314\,463\ \mathrm{J/(mol\,K)}$. Antoine correlations, critical constants, liquid molar volumes and UNIFAC group assignments are the ones in `vlekit.data.LIBRARY`, and every one of them is checked in that module against a normal boiling point and a tabulated vapour pressure — which is the only reason a hard-coded table deserves any trust at all.

Results are quoted to five significant figures. That is well beyond the accuracy of any of the models and well beyond the precision of the measurements; it is printed anyway, because a reader who cannot reproduce the fifth figure cannot tell a transcription error from a physical disagreement. The uncertainty of each answer is a separate question, and Problem 2 is about exactly that.

Reproducing the numbers

```
cd vle/toolkit
python3 verify_worked_examples.py        # the comparison table
python3 verify_worked_examples.py -v     # every intermediate as well
```

The script imports `vlekit` and nothing else. It builds each answer twice: once along the rounded chain printed in this document, and once at double precision through the library. The table it prints is the evidence that this document is arithmetic rather than assertion.

Notation

Every symbol Module 4 uses, what it means, where it is defined, and its units. This module carries the heaviest notation load in the course, and most of the confusion in it is not physical: it is four different fugacity-like quantities wearing four different accents.

The four quantities with hats and superscripts

These are the ones that get mixed up. Read the column headed *whose property* first; it is what actually distinguishes them.

symbol	whose property	meaning	units
f_i	pure component i	fugacity of i as a pure substance at the T and P of the mixture	kPa
$\hat{f}_i$	component i *in* a mixture	fugacity of i in solution; equality across phases is the equilibrium condition	kPa
φ_i	pure component i	f_i/P, from an equation of state for the pure substance	—
$\hat{\varphi}_i$	component i in a mixture	$\hat{f}_i/(y_i P)$, from an equation of state applied to the mixture	—
φ_i^{sat}	pure saturated vapour of i	φ_i evaluated at T and $P_i^{\mathrm{sat}}(T)$; the reference the liquid fugacity is measured against	—

φ_i^{sat} does *not* cancel against $\hat{\varphi}_i$, except for a pure component. Problem 1 shows both, and shows that they are nearly equal and opposite — which is why applying one without the other is worse than applying neither.

The full list

symbol	meaning	defined	units
Measured, and derived directly from measurement			
T	temperature	measured	K
P	total pressure	measured	kPa
x_i	mole fraction of i in the liquid	measured	—
y_i	mole fraction of i in the vapour	measured	—
P_i^{sat}	vapour pressure of pure i	Antoine, §4.3	kPa
$\sigma(\ln\gamma_i)$	propagated standard uncertainty of $\ln\gamma_i$	Eq. (18), §4.3	—
$u(\cdot)$	the uncertainty a source paper reports	the paper	as the quantity
The liquid: activity and excess properties			
γ_i	activity coefficient of i, Lewis-Randall reference, so $\gamma_i \to 1$ as $x_i \to 1$	Eq. (3), §4.1	—
γ_i^{∞}	$\lim_{x_i\to 0}\gamma_i$, the infinite-dilution value	§4.2	—
G^{E}	excess Gibbs energy of the liquid	§4.5	J/mol
G^{E}/RT	the dimensionless form every model is written in; $G^{\mathrm{E}}/RT = \sum_i x_i \ln\gamma_i$	§4.5	—

symbol	meaning	defined	units
Q	reduced excess energy $G^{\mathrm{E}}/(RTx_1x_2)$; $Q(0) = \ln\gamma_1^\infty$, $Q(1) = \ln\gamma_2^\infty$	§4.5	—
H^{E}	excess enthalpy; the term that survives in isobaric Gibbs-Duhem	§4.4	J/mol
Δg_{mix}	Gibbs energy of mixing; its convexity, not the equilibrium equations, decides how many phases there are	§4.6	J/mol
The vapour, and the bridge between the phases			
B_{ij}	second virial coefficient; B_{11}, B_{22} pure, B_{12} cross	§4.3	$\mathrm{cm}^3/\mathrm{mol}$
δ_{12}	$2B_{12} - B_{11} - B_{22}$, the only mixing quantity in the two-term virial $\hat{\varphi}$	Eq. (7)	$\mathrm{cm}^3/\mathrm{mol}$
v_i^L	liquid molar volume of pure i, for the Poynting factor	§4.3	$\mathrm{cm}^3/\mathrm{mol}$
Φ_i	the three vapour-phase factors as one group, $\Phi_i = \dfrac{\hat{\varphi}_i}{\varphi_i^{\mathrm{sat}}} \exp\left[-\dfrac{v_i^L(P - P_i^{\mathrm{sat}})}{RT}\right]$; $\Phi_i = 1$ is modified Raoult	Eq. (6), §4.3	—
Model parameters — each meaningless without its model			
A_{12}, A_{21}	three-suffix Margules parameters; $A_{12} = \ln\gamma_1^\infty$, $A_{21} = \ln\gamma_2^\infty$	Eq. (29), §4.5	—
A, B	van Laar parameters; also equal to $\ln\gamma_1^\infty$ and $\ln\gamma_2^\infty$, and the curve is still not the same one	Eq. (66), §4.5	—
$\Lambda_{12}, \Lambda_{21}$	Wilson parameters; ratios of Boltzmann factors, hence strictly positive, hence Wilson can never split a liquid	§4.5	—
τ_{ij}	NRTL (and UNIQUAC) dimensionless interaction energy, $\tau_{ij} = \Delta u_{ij}/RT$	§4.5	—
α_{ij}	NRTL non-randomness parameter, held at 0.3 throughout this course	§4.5	—
G_{ij}	NRTL group $\exp(-\alpha_{ij}\tau_{ij})$	§4.5	—
Δu_{ij}	the interaction energy itself; fit *this*, not τ_{ij}, if you need another temperature	§4.5	J/mol
r_i, q_i	UNIQUAC volume and surface parameters, here summed from UNIFAC group R and Q values	§4.5	—
Ψ_{mn}	UNIFAC group interaction, $\Psi_{mn} = \exp(-a_{mn}/T)$	§4.5	—
Γ_k	UNIFAC residual activity coefficient of group k	§4.5	—
Using the model			
K_i	$y_i/x_i = \gamma_i P_i^{\mathrm{sat}}/P$; a property of the mixture at a state, not of the component	Eq. (68), §4.6	—
α_{12}	relative volatility K_1/K_2; $\alpha_{12} = 1$ *is* the azeotrope	Eq. (68), §4.2	—
$\tilde{\alpha}$	the geometric mean $\sqrt{\alpha_D \alpha_B}$ that Fenske uses in place of $\alpha(x)$	Eq. (72)	—
N_{min}	minimum stages at total reflux, counting the reboiler	Eq. (72)	stages
V	vapour fraction in a flash; the Rachford-Rice unknown	§4.6	—

symbol	meaning	defined	units
z_i	feed mole fraction	§4.6	—
Testing the data			
D	area-test statistic, $100\lvert A^+ - A^-\rvert/(A^+ + A^-)$	Eq. (54), §4.4	%
J	Herington's isobaric allowance $150\,\Delta T/T_{\min}$	§4.4	%
δ	direct-test residual $\ln(\gamma_1/\gamma_2)_{\text{exp}} - \ln(\gamma_1/\gamma_2)_{\text{fit}}$	§4.4	—
δy_1	point-test residual, measured y_1 minus y_1 implied by a Barker fit to P-x	§4.4	—
The equation-of-state route			
a, b	Peng-Robinson attraction and co-volume; a_{mix}, b_{mix} their mixture values	§4.7	Pa m^6/mol^2, m^3/mol
κ	the α-function coefficient, $\kappa = 0.37464 + 1.54226\omega - 0.26992\omega^2$ (written m in `vlekit`)	Eq. (76)	—
$\mathcal{A}, \mathcal{B}$	dimensionless groups $aP/(RT)^2$ and bP/RT	Eq. (81)	—
Z	compressibility factor, the root of the cubic	§4.7	—
k_{ij}	binary interaction parameter — classical, multiplying $\sqrt{a_i a_j}$; *not* the Wong-Sandler k_{ij}, which multiplies the whole second-virial cross coefficient	§4.7	—
C	the infinite-pressure constant of the cubic; $C = \ln(\sqrt{2}-1)/\sqrt{2} = -0.62323$ for Peng-Robinson, $-\ln 2$ for SRK	Eq. (87)	—
G^{E}_{∞}	the excess Gibbs energy at the infinite-pressure reference state that Huron-Vidal matches	§4.7	J/mol
D_{HV}	the group $a_{\text{mix}}/(b_{\text{mix}}RT)$ that both G^{E} mixing rules set	Eq. (87)	—
ω	acentric factor	Module 1	—
T_c, P_c	critical constants	Module 1	K, kPa

Two conventions worth writing on the inside of your eyelids

A parameter has no meaning without its model. $A_{12} = 1.7528$ is a Margules number and a van Laar number and they draw different curves. Worse, r and q summed from UNIFAC groups are not the r and q of the original UNIQUAC tables, so τ values fitted with one convention cannot be used with the other. Problem 8 makes the same point for Huron-Vidal: its NRTL parameters live at a different reference state from a published low-pressure set, and the numbers are not interchangeable.

Component 1 is whichever component the paper says it is. The ethanol + water set used in Problem 7 is archived with the substances listed as (water, ethanol) and the mole fraction tabulated for the *other* one. Physics settles it — the mixture boiling at 351.35 K at $x_1 = 0.972$ can only be ethanol-rich — and the check takes ten seconds. Run it on every machine-readable table you use.

Problem 1 An activity coefficient from one measured point

Problem

Row 10 of the chloroform (1) + 2-butanone (2) table of Clara, Gómez Marigliano and Solimo, measured isothermally at 303.15 K (`doi:10.1021/je060150a`), reads

$$x_1 = 0.5070, \qquad y_1 = 0.6730, \qquad P = 17.69 \text{ kPa}.$$

Evaluate the Antoine correlations, then compute γ_1 and γ_2 under modified Raoult's law. Then redo the calculation with the full set of vapour-phase corrections — $\hat{\varphi}_i$ from the two-term virial equation, φ_i^{sat} and the Poynting factor — and say where the two answers differ.

Auxiliary data, and where it came from

None of this is in the paper. All of it is in `vlekit.data.LIBRARY`, and sourcing it is the part of data reduction that actually takes the time.

	chloroform (1)	2-butanone (2)	source
Antoine A (mm Hg, t/°C)	6.95465	7.06356	Yaws
Antoine B	1170.966	1261.34	
Antoine C	226.232	221.969	
T_c / K	536.4	536.8	Poling et al.
P_c / kPa	5472	4207	
ω	0.222	0.323	
v^L / cm^3/mol	80.7	90.2	

Step 1 — the two vapour pressures

The correlation is published as $\log_{10}(P^{\text{sat}}/\text{mm Hg}) = A - B/(t/°\text{C} + C)$, and $T = 303.15$ K is $t = 30.00$ °C. Do the unit conversion at the end, once, where you can see it:

$$\log_{10} P_1^{\text{sat}} = 6.95465 - \frac{1170.966}{30.00 + 226.232} = 6.95465 - 4.56994 = 2.38471$$

$$P_1^{\text{sat}} = 242.499 \text{ mm Hg} \times \frac{101.325}{760} = \boxed{32.3305 \text{ kPa}} \tag{1}$$

$$\log_{10} P_2^{\text{sat}} = 7.06356 - \frac{1261.34}{30.00 + 221.969} = 7.06356 - 5.00593 = 2.05763$$

$$P_2^{\text{sat}} = 114.191 \text{ mm Hg} \times \frac{101.325}{760} = \boxed{15.2242 \text{ kPa}} \tag{2}$$

The free calibration nobody runs. The paper measured both pure end points: $P = 32.19$ kPa at $x_1 = 1$ and $P = 15.74$ kPa at $x_1 = 0$. Against those, Eq. (1) is $+0.44$ % on chloroform and -3.28 % on 2-butanone. Hold on to the second number: a 3.3 % error in P_2^{sat} goes straight into every γ_2 below, and it is larger than every vapour-phase correction on this page put together.

Step 2 — modified Raoult's law

With $\Phi_i = 1$,

$$\gamma_i = \frac{y_i \, \Phi_i \, P}{x_i \, P_i^{\text{sat}}(T)} \tag{3}$$

so

$$\gamma_1 = \frac{0.6730 \times 17.69}{0.5070 \times 32.3305} = \frac{11.9054}{16.3916} = \boxed{0.72631} \tag{4}$$

$$\gamma_2 = \frac{0.3270 \times 17.69}{0.4930 \times 15.2242} = \frac{5.78463}{7.50553} = \boxed{0.77072} \tag{5}$$

Both below one: chloroform's acidic C–H hydrogen-bonds to the 2-butanone carbonyl, mixing is energetically favourable, and this is the negative-deviation system the deck calls Type 4.

Step 3 — the three vapour-phase factors, one at a time

Φ_i is not one correction. It is three, and they do not have the same sign.

$$\Phi_i = \frac{\hat{\varphi}_i}{\varphi_i^{\text{sat}}} \exp\left[-\frac{v_i^L\,(P - P_i^{\text{sat}})}{RT}\right] \tag{6}$$

Second virial coefficients. From the Pitzer-Curl correlation with the Tsonopoulos temperature functions and $k_{12} = 0$, at 303.15 K:

$$B_{11} = -1177.77, \quad B_{22} = -1753.16, \quad B_{12} = -1443.10 \ \text{cm}^3/\text{mol}$$

$$\delta_{12} = 2B_{12} - B_{11} - B_{22} = -2886.20 + 1177.77 + 1753.16 = +44.73 \ \text{cm}^3/\text{mol}$$

δ_{12} is small and positive, which is the whole of the mixing effect in this route: the cross coefficient sits almost exactly at the mean of the two pure ones.

$\hat{\varphi}_i$ in the mixture. With $RT = 8.314463 \times 303.15 = 2520.53\,\text{J/mol}$ and P in Pa,

$$\ln\hat{\varphi}_1 = \frac{P}{RT}\left[B_{11} + y_2^2\,\delta_{12}\right], \qquad \ln\hat{\varphi}_2 = \frac{P}{RT}\left[B_{22} + y_1^2\,\delta_{12}\right] \tag{7}$$

$$\begin{aligned}\ln\hat{\varphi}_1 &= \frac{17690}{2520.53}\left[-1177.77 + (0.3270)^2(44.73)\right] \times 10^{-6} \\ &= 7.01836 \times (-1172.99 \times 10^{-6}) = -0.00823 \ \Rightarrow \ \hat{\varphi}_1 = 0.991801 \end{aligned} \tag{8}$$

$$\begin{aligned}\ln\hat{\varphi}_2 &= 7.01836 \times \left[-1753.16 + (0.6730)^2(44.73)\right] \times 10^{-6} \\ &= 7.01836 \times (-1732.90 \times 10^{-6}) = -0.01216 \ \Rightarrow \ \hat{\varphi}_2 = 0.987912 \end{aligned} \tag{9}$$

φ_i^{sat}, the pure saturated vapour. Same expression with the mixture gone and P replaced by P_i^{sat}:

$$\ln\varphi_1^{\text{sat}} = \frac{32330.5 \times (-1177.77 \times 10^{-6})}{2520.53} = -0.01511 \ \Rightarrow \ \varphi_1^{\text{sat}} = 0.985007 \tag{10}$$

$$\ln\varphi_2^{\text{sat}} = \frac{15224.2 \times (-1753.16 \times 10^{-6})}{2520.53} = -0.01059 \ \Rightarrow \ \varphi_2^{\text{sat}} = 0.989467 \tag{11}$$

Note the sizes. $\ln\hat{\varphi}_1 = -0.00823$ and $\ln\varphi_1^{\text{sat}} = -0.01511$ are the same kind of number with the same sign, because they are the same physics evaluated at two pressures. They are going to cancel, and they are meant to.

Poynting. $P - P_1^{\text{sat}}$ is negative here and $P - P_2^{\text{sat}}$ positive, so the two factors sit on opposite sides of one:

$$\text{PF}_1 = \exp\left[\frac{80.7 \times 10^{-6}\,(17.69 - 32.3305) \times 10^3}{2520.53}\right] = e^{-0.000469} = 0.999531 \tag{12}$$

$$\text{PF}_2 = \exp\left[\frac{90.2 \times 10^{-6}\,(17.69 - 15.2242) \times 10^3}{2520.53}\right] = e^{+0.0000882} = 1.00009 \tag{13}$$

Five parts in ten thousand, and one part in ten thousand. Small — but knowing it is small is not the same as omitting it silently.

Step 4 — assemble Φ_i and correct γ

$$\Phi_1 = \frac{0.991801}{0.985007 \times 0.999531} = \frac{0.991801}{0.984545} = 1.007370 \tag{14}$$

$$\Phi_2 = \frac{0.987912}{0.989467 \times 1.00009} = \frac{0.987912}{0.989556} = 0.998339 \tag{15}$$

and since Φ_i multiplies Eq. (3) directly,

$$\gamma_1 = 0.72631 \times 1.007370 = \boxed{0.73166} \tag{16}$$

$$\gamma_2 = 0.77072 \times 0.998339 = \boxed{0.76944} \tag{17}$$

Answer, and where the difference sits

	modified Raoult	γ-φ, all three factors	shift
γ_1	0.72631	0.73166	+0.00535
γ_2	0.77072	0.76944	−0.00128
$\ln \gamma_1$	−0.31978	−0.31244	+0.00734
$\ln \gamma_2$	−0.26043	−0.26209	−0.00166

The difference is in the third decimal of γ, exactly as advertised. What matters is how it got there. For component 1,

$$\ln \Phi_1 = \underbrace{-0.00823}_{\ln \hat{\varphi}_1} \underbrace{+\,0.01511}_{-\ln \varphi_1^{\mathrm{sat}}} \underbrace{+\,0.00047}_{\text{Poynting}} = +0.00735$$

and the net is smaller than either of the first two terms on its own. (The table above reads +0.00734 rather than +0.00735 because it takes the logarithm of a γ already rounded to five figures. That is the size of the rounding, and it is worth seeing once.) Apply $\hat{\varphi}_i$ and forget φ_i^{sat} and you make an error *larger* than the one you set out to remove, and in the wrong direction.

Against `vlekit`. `VLEData.gamma('ideal')` returns 0.726317 and 0.770724; the hand chain gives 0.72631 and 0.77072. `gamma('virial')` returns 0.731670 and 0.769445 against 0.73166 and 0.76944. The individual factors agree to six figures: $\hat{\varphi}_1 = 0.991801$, $\varphi_1^{\mathrm{sat}} = 0.985007$, $\mathrm{PF}_1 = 0.999531$. The last digit of γ_1 differs by one because the hand chain divided by P_1^{sat} already rounded to 32.3305; carrying the unrounded 32.330215 recovers 0.726317 exactly.

What this was really testing. That a reported γ is not a measurement. It is four measured numbers and two assumptions, and this problem makes you write both assumptions down. The misconception it targets is the belief that the vapour-phase correction is a refinement you can bolt on later, one factor at a time, as time allows. It is not: $\hat{\varphi}_i$ and φ_i^{sat} are one physical quantity evaluated at two pressures, and separating them turns a 0.007 correction into a 0.015 error. The second misconception is quieter and more expensive. Every number in Steps 2 to 4 is worth less than the 3.3 % discrepancy found in Step 1 between the paper's own pure 2-butanone point and the stored Antoine constants. The hard part of reducing published data is never the thermodynamics — it is the auxiliary data the paper does not supply and you must source and check yourself.

Problem 2 What that γ is worth: propagating the measurement error

Problem

The source paper for the point used in Problem 1 reports, as assessed by NIST TRC, $u(x_1) = 0.001$, $u(y_1)$ up to 0.021, $u(P)$ up to 0.65 kPa and $u(T) = 0.01$ K. Propagate these into $\ln\gamma_1$ and $\ln\gamma_2$ at that point. Then answer the question the deck says is the only one that matters: is the vapour-phase correction computed in Problem 1 large enough to bother with on *this* dataset?

The propagation formula, and where it comes from

Take the logarithm of Eq. (3) with $\Phi_1 = 1$:

$$\ln\gamma_1 = \ln y_1 + \ln P - \ln x_1 - \ln P_1^{\mathrm{sat}}(T)$$

Each term is a logarithm, so each partial derivative is a reciprocal, and the four errors are independent:

$$\sigma^2(\ln\gamma_1) = \left(\frac{\sigma_y}{y_1}\right)^2 + \left(\frac{\sigma_P}{P}\right)^2 + \left(\frac{\sigma_x}{x_1}\right)^2 + \left(\frac{\mathrm{d}\ln P_1^{\mathrm{sat}}}{\mathrm{d}T}\,\sigma_T\right)^2 \tag{18}$$

For component 2 the same expression holds with $y_1 \to 1 - y_1$, $x_1 \to 1 - x_1$ and $P_1^{\mathrm{sat}} \to P_2^{\mathrm{sat}}$. The temperature term needs the derivative of the Antoine correlation. In the $\log_{10}$ form,

$$\frac{\mathrm{d}\ln P^{\mathrm{sat}}}{\mathrm{d}T} = \frac{\ln 10 \cdot B}{(T + C)^2} \tag{19}$$

where B and C are in the K form of the correlation, so $C_1 = 226.232 - 273.15 = -46.918$ and $C_2 = 221.969 - 273.15 = -51.181$.

The four terms, component 1

$$\frac{\sigma_y}{y_1} = \frac{0.021}{0.6730} = 0.031204 \tag{20}$$

$$\frac{\sigma_P}{P} = \frac{0.65}{17.69} = 0.036744 \tag{21}$$

$$\frac{\sigma_x}{x_1} = \frac{0.001}{0.5070} = 0.0019724 \tag{22}$$

$$\frac{\mathrm{d}\ln P_1^{\mathrm{sat}}}{\mathrm{d}T} = \frac{2.302585 \times 1170.966}{(303.15 - 46.918)^2} = \frac{2696.25}{65654.8} = 0.041067\ 1/\mathrm{K}$$

$$\frac{\mathrm{d}\ln P_1^{\mathrm{sat}}}{\mathrm{d}T}\sigma_T = 0.041067 \times 0.01 = 0.00041067 \tag{23}$$

Squaring and adding,

$$\begin{aligned}\sigma^2(\ln\gamma_1) &= 9.7369 \times 10^{-4} + 1.3501 \times 10^{-3} + 3.8904 \times 10^{-6} + 1.6863 \times 10^{-7} \\ &= 2.3279 \times 10^{-3} \\ \sigma(\ln\gamma_1) &= \boxed{0.04825}\end{aligned} \tag{24}$$

The four terms, component 2

$$\frac{\sigma_y}{1-y_1} = \frac{0.021}{0.3270} = 0.064220 \tag{25}$$

$$\frac{\sigma_x}{1-x_1} = \frac{0.001}{0.4930} = 0.0020284$$

$$\frac{\mathrm{d}\ln P_2^{\mathrm{sat}}}{\mathrm{d}T}\sigma_T = \frac{2.302585 \times 1261.34}{(303.15-51.181)^2} \times 0.01 = \frac{2904.34}{63488.4} \times 0.01 = 0.00045746$$

$$\sigma^2(\ln\gamma_2) = 4.1242 \times 10^{-3} + 1.3501 \times 10^{-3} + 4.1144 \times 10^{-6} + 2.09 \times 10^{-7} = 5.4787 \times 10^{-3}$$

$$\sigma(\ln\gamma_2) = \boxed{0.07402} \tag{26}$$

Reading the ranking

term	share of $\sigma^2(\ln\gamma_1)$	share of $\sigma^2(\ln\gamma_2)$	
vapour composition y	41.8 %	75.3 %	the hardest measurement
pressure P	58.0 %	24.6 %	at 17.69 kPa, 0.65 is 3.7 %
liquid composition x	0.17 %	0.075 %	negligible *here*
temperature T	0.007 %	0.004 %	negligible *here*

Two of those four verdicts are conditional and the conditions are worth stating. The composition term is negligible at $x_1 = 0.507$ and only there: at $x_1 = 0.02$ it becomes $\sigma_x / x_1 = 0.05$, five per cent, which is larger than everything else on this table put together — and $x_1 = 0.02$ is exactly where γ^∞ is read off. The temperature term is negligible because this experiment is isothermal and controlled to 0.01 K. On an isobaric set with $u(T) = 0.1$ K the same derivative gives 0.0041 in $\ln\gamma$, comparable to the composition term, and the ranking changes.

Is the vapour correction worth applying?

Problem 1 gave a net correction of $|\Delta\ln\gamma_1| = 0.00734$. The rule the deck insists on is not "correct above one bar"; it is to compare the correction with the uncertainty of the thing being corrected, on your own data:

$$\frac{|\Delta\ln\gamma_1|}{\sigma(\ln\gamma_1)} = \frac{0.00734}{0.04825} = \boxed{0.152} \tag{27}$$

The correction is one seventh of the noise. On this dataset you may neglect the vapour-phase corrections — **and you say so, with that number**, rather than not mentioning that you considered them.

Run the same comparison across the module's five published sets, using the largest correction on each set against the median σ on that set:

system	P / kPa	max $\lvert\Delta\ln\gamma_1\rvert$	med $\sigma(\ln\gamma_1)$	ratio
benzene + toluene	101.3	0.0329	0.0130	2.53
acetone + chloroform	101.3	0.0108	0.0064	1.70
methanol + water	66.5	0.0282	0.0185	1.53
ethanol + water	101.3	0.0214	0.0215	1.00
chloroform + 2-butanone	20.4	0.0086	0.0443	0.19

At atmospheric pressure the correction is comparable to or larger than the noise and may not be neglected. At 20 kPa it is a fifth of it. Note that the ordering is set as much by how carefully the paper was written as by the pressure: the chloroform set is last not because its correction

is uniquely small but because its reported $u(y_1) = 0.021$ and $u(P) = 0.65\,\mathrm{kPa}$ are uniquely generous.

Against `vlekit`. `VLEData.gamma_sigma('ideal')` returns 0.0482477 and 0.0740180 against the hand values 0.04825 and 0.07402. The derivative $\mathrm{d}\ln P_1^{\mathrm{sat}}/\mathrm{d}T$ from `Component.dPsat_dT` is $0.041\,067\,0\,/\mathrm{K}$ against 0.041067. The ratio is 0.15220 from the library and 0.152 by hand, the difference being that the hand chain used $\Delta\ln\gamma_1$ already rounded to 0.00734.

What this was really testing. That "negligible" is not a property of a correction. It is a *relation* between a correction and an uncertainty, and it therefore has to be recomputed for every dataset. The same $\hat{\varphi}$ that is a fifth of the noise on a careless 1970s set is decisive on a careful modern one, and no rule of thumb about pressure will tell you which case you are in. The second target is the belief that quoting more digits is a form of care. A thermometer read to 0.1 K and a vapour sample good to 0.02 cannot support four figures in γ; the fourth figure is arithmetic, not measurement, and Problems 3 to 7 all inherit whatever error was made here.

Problem 3 Fitting a two-parameter model from two measured points

Problem

Using the same chloroform (1) + 2-butanone (2) set at 303.15 K (`doi:10.1021/je060150a`), fit the two parameters of the three-suffix Margules model to just *two* of the twenty-two points, by hand. Then use the fitted model to predict γ_1 and γ_2 at a third point — the one worked in Problem 1 — and compare with the measurement. Finally compare the two-point parameters with a Barker fit to all twenty-two.

Take the two points to be

$$\begin{aligned} &\text{(a)}\quad x_1 = 0.2520,\quad y_1 = 0.2810,\quad P = 15.35\ \text{kPa} \\ &\text{(b)}\quad x_1 = 0.7570,\quad y_1 = 0.9290,\quad P = 24.66\ \text{kPa} \end{aligned}$$

The model

The three-suffix (two-parameter) Margules form is

$$\frac{G^{\mathrm{E}}}{RT} = x_1 x_2 \left(A_{21} x_1 + A_{12} x_2\right) \tag{28}$$

whose partial molar quantities are

$$\ln\gamma_1 = x_2^2\left[A_{12} + 2(A_{21} - A_{12})x_1\right], \qquad \ln\gamma_2 = x_1^2\left[A_{21} + 2(A_{12} - A_{21})x_2\right] \tag{29}$$

Setting $x_1 \to 0$ in the first and $x_1 \to 1$ in the second gives $A_{12} = \ln\gamma_1^{\infty}$ and $A_{21} = \ln\gamma_2^{\infty}$. The parameters *are* the infinite-dilution values. That is convenient and it is also a warning: two numbers determined in the middle of the range are being asked to carry the two limits the data never reach.

Step 1 — G^{E}/RT at the two points

Apply Eq. (3) at each point exactly as in Problem 1, with $P_1^{\mathrm{sat}} = 32.3305$ kPa, $P_2^{\mathrm{sat}} = 15.2242$ kPa and $\Phi_i = 1$:

$$\text{(a)}\quad \gamma_1 = \frac{0.2810 \times 15.35}{0.2520 \times 32.3305} = 0.52942, \qquad \gamma_2 = \frac{0.7190 \times 15.35}{0.7480 \times 15.2242} = 0.96917 \tag{30}$$

$$\text{(b)}\quad \gamma_1 = \frac{0.9290 \times 24.66}{0.7570 \times 32.3305} = 0.93605, \qquad \gamma_2 = \frac{0.0710 \times 24.66}{0.2430 \times 15.2242} = 0.47327 \tag{31}$$

Then $G^{\mathrm{E}}/RT = x_1\ln\gamma_1 + x_2\ln\gamma_2$:

$$\left.\frac{G^{\mathrm{E}}}{RT}\right|_{\mathrm{a}} = 0.2520(-0.63597) + 0.7480(-0.03132) = -0.16027 - 0.02342 = -0.18369 \tag{32}$$

$$\left.\frac{G^{\mathrm{E}}}{RT}\right|_{\mathrm{b}} = 0.7570(-0.06609) + 0.2430(-0.74809) = -0.05003 - 0.18179 = -0.23181 \tag{33}$$

Step 2 — two equations, two unknowns

Divide Eq. (28) by x_1x_2 so that what is left is linear in the parameters. Write $r \equiv G^{\mathrm{E}}/(RTx_1x_2)$, which is the reduced quantity Q of the notation card:

$$r_{\mathrm{a}} = \frac{-0.18369}{0.2520 \times 0.7480} = \frac{-0.18369}{0.188496} = -0.97450 \tag{34}$$

$$r_{\mathrm{b}} = \frac{-0.23181}{0.7570 \times 0.2430} = \frac{-0.23181}{0.183951} = -1.26017 \tag{35}$$

so that

$$\begin{cases} 0.2520\,A_{21} + 0.7480\,A_{12} = -0.97450 \\ 0.7570\,A_{21} + 0.2430\,A_{12} = -1.26017 \end{cases} \tag{36}$$

By Cramer's rule, with

$$\Delta = (0.2520)(0.2430) - (0.7570)(0.7480) = 0.061236 - 0.566236 = -0.505000$$

$$A_{21} = \frac{(-0.97450)(0.2430) - (-1.26017)(0.7480)}{-0.505000} = \frac{-0.236804 + 0.942607}{-0.505000} = \boxed{-1.39763} \tag{37}$$

$$A_{12} = \frac{(-1.26017)(0.2520) - (-0.97450)(0.7570)}{-0.505000} = \frac{-0.317563 + 0.737697}{-0.505000} = \boxed{-0.83195} \tag{38}$$

Both negative, as they must be for a system whose G^{E} is negative throughout.

Step 3 — predict the third point

At $x_1 = 0.5070$, with $A_{21} - A_{12} = -1.39763 + 0.83195 = -0.56568$,

$$\begin{aligned} \ln\gamma_1 &= (0.4930)^2\,[-0.83195 + 2(-0.56568)(0.5070)] \\ &= 0.243049\,(-0.83195 - 0.573600) = 0.243049 \times (-1.405550) = -0.34162 \\ \gamma_1 &= \boxed{0.71062} \end{aligned} \tag{39}$$

$$\begin{aligned} \ln\gamma_2 &= (0.5070)^2\,[-1.39763 + 2(+0.56568)(0.4930)] \\ &= 0.257049\,(-1.39763 + 0.557760) = 0.257049 \times (-0.839870) = -0.21589 \\ \gamma_2 &= \boxed{0.80582} \end{aligned} \tag{40}$$

Answer, against the measurement

at $x_1 = 0.5070$	measured (Problem 1)	predicted	error
γ_1	0.72631	0.71062	−2.16 %
γ_2	0.77072	0.80582	+4.55 %
G^{E}/RT	−0.29052	−0.27964	3.7 % too small in magnitude

Two per cent on one and four and a half on the other, from two points chosen sensibly and arithmetic done correctly. That is what a two-point fit is worth.

What all twenty-two points say

A Barker fit — minimise $\sum_k (P_k^{\mathrm{calc}} - P_k^{\mathrm{exp}})^2$ over all twenty-two points, with y_1 never entering the objective — gives

$$A_{12} = -0.9122, \qquad A_{21} = -1.2988, \qquad \mathrm{rms}\,\delta P = 0.3101\ \mathrm{kPa} \tag{41}$$

against the two-point $(-0.83195, -1.39763)$. Now compare what the two parameter sets imply about the ends, which is where a two-parameter model does its real extrapolating:

	two points	all 22, Barker	difference
$A_{12} = \ln\gamma_1^{\infty}$	−0.83195	−0.9122	8.8 %
$A_{21} = \ln\gamma_2^{\infty}$	−1.39763	−1.2988	7.6 %
γ_1^{∞}	0.43520	0.40164	8.4 %
γ_2^{∞}	0.24718	0.27285	9.4 %

Eight to nine per cent in the extrapolated limits, from parameters that agree to a few per cent in the middle. And note the direction: A_{12} moved up and A_{21} moved down — in *opposite* directions. That is the parameter correlation the deck's bootstrap figure is about — the two numbers are not separately determined by this experiment, only a combination of them is, and quoting two independent standard errors for them describes a rectangle the fit never occupied.

Against `vlekit`. Solving the same 2×2 system in double precision gives $A_{12} = -0.8318648$ and $A_{21} = -1.3975914$ against the hand values -0.83195 and -1.39763 — a disagreement of 1×10^{-4}, and it is worth knowing where it comes from. The two vapour pressures were rounded to six figures before dividing, r_{a} and r_{b} to five before the determinant was formed, and $\Delta = -0.505000$ is itself a difference of two numbers ten times larger. Rounding survives that sequence at the fourth decimal and no further. The predicted γ from the exact parameters are 0.710627 and 0.805843 against 0.71062 and 0.80582. `regress.barker_fit(d,'margules')` returns $(-0.9121743, -1.2988434)$ with rms $\delta P = 0.310\,099\,5$ kPa.

What this was really testing. That a fit is not an answer, it is an answer plus a statement of what it rests on. Two well-chosen points and twenty-two points give parameters that agree to a few per cent where the data are and disagree by nine per cent where they are not — and the second number is the one a stripper design would use. The misconception under attack is that more data buys accuracy uniformly. It does not: it buys accuracy *where the data are*, and the infinite-dilution limits are outside the measured range no matter how many interior points you have. The second target is subtler. Because A_{12} and A_{21} move in opposite directions between the two fits, a report that quotes them with independent error bars is describing a region the regression never visited. Uncertainty in a two-parameter fit is a correlated cloud, not two numbers, and that is why the deck asks for a confidence region rather than two standard errors.

Problem 4 Classifying a system, and locating its azeotrope

Problem

(a) Chloroform (1) + 2-butanone (2) at 303.15 K. Using the fitted parameters of Eq. (41) and the two vapour pressures, decide without touching any interior data point whether the system has an azeotrope and of which kind. Then locate it, and compare with the measurement.

(b) Methanol (1) + water (2) at 66.52 kPa (`doi:10.1021/je300613v`) has $\gamma_1^\infty = 2.413$ and $\gamma_2^\infty = 1.596$ from an NRTL fitted by Barker — a large positive deviation. Show that it nevertheless has no azeotrope, and say what makes the difference.

The criterion

At an azeotrope $y_1 = x_1$, so $y_2 = x_2$, and the two equilibrium relations divide to give

$$y_1 = x_1 \iff \gamma_1 P_1^{\mathrm{sat}} = \gamma_2 P_2^{\mathrm{sat}} \iff \alpha_{12} \equiv \frac{K_1}{K_2} = \frac{\gamma_1 P_1^{\mathrm{sat}}}{\gamma_2 P_2^{\mathrm{sat}}} = 1 \tag{42}$$

Now take the two limits. As $x_1 \to 0$, $\gamma_1 \to \gamma_1^\infty$ and $\gamma_2 \to 1$; as $x_1 \to 1$ it is the other way round. So

$$\alpha_{12}|_{x_1\to 0} = \gamma_1^\infty \frac{P_1^{\mathrm{sat}}}{P_2^{\mathrm{sat}}}, \qquad \alpha_{12}|_{x_1\to 1} = \frac{1}{\gamma_2^\infty}\frac{P_1^{\mathrm{sat}}}{P_2^{\mathrm{sat}}} \tag{43}$$

If α_{12} is above 1 at one end and below 1 at the other, a continuous $\alpha_{12}(x_1)$ has to cross, and the crossing is the azeotrope. Only two infinite-dilution values and a vapour-pressure ratio are needed: *no interior data at all.*

This is sufficient, not necessary. A G^{E} with three or more parameters can make γ_1/γ_2 non-monotonic and place two azeotropes with both ends on the same side. Assuming monotonicity is an assumption, and it is being made here.

(a) Chloroform + 2-butanone

The ratio. From Problem 1, $P_1^{\mathrm{sat}}/P_2^{\mathrm{sat}} = 32.3305/15.2242 = 2.12362$.

The two limits. From Eq. (41), $\gamma_1^\infty = e^{-0.9122} = 0.40164$ and $\gamma_2^\infty = e^{-1.2988} = 0.27285$, so

$$\alpha_{12}|_{x_1\to 0} = 0.40164 \times 2.12362 = \boxed{0.8529} \tag{44}$$

$$\alpha_{12}|_{x_1\to 1} = \frac{2.12362}{0.27285} = \boxed{7.783} \tag{45}$$

Below one at the dilute-chloroform end and well above one at the other, so the curve crosses: **there is an azeotrope**. It straddles *from below*, which is the negative-deviation case, so P has an interior *minimum* and therefore T an interior *maximum*: a **maximum-boiling** azeotrope. Equivalently: $\gamma_i < 1$ puts $P = x_1\gamma_1 P_1^{\mathrm{sat}} + x_2\gamma_2 P_2^{\mathrm{sat}}$ *below* the Raoult straight line everywhere inside, and a continuous curve that starts and ends on a line and stays below it in between has an interior minimum.

Locating it. Substituting Eq. (29) into Eq. (42), the Margules $\ln(\gamma_1/\gamma_2)$ collapses to a quadratic — expand and the cubic terms cancel:

$$\ln\frac{\gamma_1}{\gamma_2} = 3(A_{12}-A_{21})\,x_1^2 + (2A_{21}-4A_{12})\,x_1 + A_{12} = \ln\frac{P_2^{\mathrm{sat}}}{P_1^{\mathrm{sat}}} \tag{46}$$

The right-hand side is $\ln(1/2.12362) = -0.75311$, and the coefficients are

$$\begin{aligned} a &= 3(-0.9122 + 1.2988) = 3(0.3866) = 1.15980 \\ b &= 2(-1.2988) - 4(-0.9122) = -2.59760 + 3.64880 = 1.05120 \\ c &= -0.9122 - (-0.75311) = -0.15909 \end{aligned} \tag{47}$$

so

$$\begin{aligned} x_1 &= \frac{-1.05120 + \sqrt{(1.05120)^2 - 4(1.15980)(-0.15909)}}{2(1.15980)} \\ &= \frac{-1.05120 + \sqrt{1.105021 + 0.738050}}{2.31960} = \frac{-1.05120 + 1.357598}{2.31960} \\ &= \frac{0.306398}{2.31960} = \boxed{0.13209} \end{aligned} \tag{48}$$

(The other root, -1.0381, is outside $(0,1)$ and is discarded.) The pressure there follows from Eq. (29) at $x_1 = 0.13209$, which gives $\gamma_1 = 0.46577$ and $\gamma_2 = 0.98911$, hence

$$P_{\mathrm{az}} = 0.13209(0.46577)(32.3305) + 0.86791(0.98911)(15.2242) = \boxed{15.058\ \mathrm{kPa}} \tag{49}$$

Against the measurement. In the table, $y_1 - x_1$ changes sign between $(0.1420,\ 0.1190)$ and $(0.2520,\ 0.2810)$. Linear interpolation through zero:

$$x_{\mathrm{az}} = 0.1420 + (0.2520 - 0.1420)\,\frac{0.0230}{0.0230 + 0.0290} = 0.1907 \tag{50}$$

and the paper's own value, obtained independently by the authors, is 0.199 at 15.3 kPa.

	x_{az}	P_{az} / kPa
measured, interpolated from the table	0.1907	15.33–15.35
the paper's own reduction	0.199	15.3
Margules-2 fitted by Barker to P-x	0.13209	15.058

The model misses the azeotropic composition by 0.058 — thirty per cent — while reproducing all twenty-two pressures to rms $\delta P = 0.31$ kPa, which is half the paper's own $u(P)$. That is not a contradiction and it is not a bad fit. It is the Gibbs-Konovalov theorem being unhelpful: $\mathrm{d}P/\mathrm{d}x_1 = 0$ at the azeotrope, so the P-x curve is *flat* there and the pressures carry almost no information about where the stationary point sits. Barker's method minimises exactly those pressures. If the azeotropic composition is the number your design needs, it must appear in the objective function; otherwise you are fitting a variable that is by construction insensitive to it.

(b) Methanol + water: a big deviation and no azeotrope

This set is isobaric, so the two ends sit at different temperatures — pure water boils at 361.51 K at 66.52 kPa and pure methanol at 327.45 K — and the vapour-pressure ratio must be evaluated at each end separately:

$$\left.\frac{P_1^{\mathrm{sat}}}{P_2^{\mathrm{sat}}}\right|_{361.51\,\mathrm{K}} = \frac{240.20}{65.914} = 3.6441, \qquad \left.\frac{P_1^{\mathrm{sat}}}{P_2^{\mathrm{sat}}}\right|_{327.45\,\mathrm{K}} = \frac{66.354}{15.299} = 4.3371 \tag{51}$$

$$\alpha_{12}|_{x_1 \to 0} = 2.413 \times 3.6441 = \boxed{8.793} \tag{52}$$

$$\alpha_{12}|_{x_1 \to 1} = \frac{4.3371}{1.596} = \boxed{2.717} \tag{53}$$

Both ends are above one, so α_{12} never approaches unity and there is **no azeotrope** — confirmed by the raw data, in which $y_1 > x_1$ at every one of the sixteen points.

The point. Methanol + water has a genuine, substantial positive deviation: γ_1 reaches 2.17 in the table. Chloroform + 2-butanone has $\gamma_1^{\infty} = 0.40$, a comparable departure the other way.

One azeotropes and one does not, and the deviation is not what decides it. What decides it is whether γ_1^∞ is large enough to overcome the vapour-pressure ratio — and for methanol + water that ratio is already 3.6 to 4.3, so a γ_1^∞ of 2.4 never gets close. **Type 2 and Type 3 differ by the vapour pressures, not by the size of the deviation.**

Against `vlekit`. Solving Eq. (46) for the exact fitted parameters with `brentq` gives $x_{az} = 0.1320752$ against the hand quadratic's 0.13209; the bubble pressure there is 15.058 26 kPa against 15.058. The measured interpolation is 0.190654 against 0.1907. The two end volatilities are 0.852956 and 7.78322 against 0.8529 and 7.783. For methanol + water, `barker_fit(d,'nrtl')` gives $\gamma_1^\infty = 2.41346$ and $\gamma_2^\infty = 1.59574$, and the end volatilities 8.79506 and 2.71803 against the hand 8.793 and 2.717.

What this was really testing. Two things, and the second is the expensive one. First, that an azeotrope is not a category a mixture belongs to but a root of one equation, and that the root can be located — or ruled out — from two infinite-dilution values and a vapour-pressure ratio, with no interior data and no curve-fitting by eye. The misconception it kills is that Type 2 is a weak Type 3, that a big enough positive deviation always produces an azeotrope. Methanol + water is the counterexample and the vapour-pressure ratio is the reason.
Second, and this is the one that damages designs: a fit can be excellent in the variable it was fitted to and thirty per cent wrong in the quantity you actually need. The flatness of $P(x_1)$ at the azeotrope is not a numerical nuisance; it is a theorem, and it guarantees that pressure data are nearly blind to azeotropic composition. Any regression report that quotes rms δP and then states an azeotrope is quoting a residual that does not constrain the claim.

Problem 5 The area test, by trapezoid, on a short real dataset

Problem

Acetone (1) + chloroform (2), isobaric at 101.3 kPa, nine points, from Gao, Zhao, Jia and Zhang (`doi:10.3390/app8091519`, Table A1, open access). Compute $\ln(\gamma_1/\gamma_2)$ at every point, integrate it across the composition range by the trapezoidal rule, form the area-test statistic D, and say what the result does and does not establish.

Where the test comes from

The Gibbs-Duhem equation at constant T and P,

$$x_1 \frac{\mathrm{d}\ln\gamma_1}{\mathrm{d}x_1} + x_2 \frac{\mathrm{d}\ln\gamma_2}{\mathrm{d}x_1} = 0$$

integrates once, using $G^{\mathrm{E}}/RT = x_1\ln\gamma_1 + x_2\ln\gamma_2$ and the fact that G^{E} vanishes at both pure ends, to

$$\int_0^1 \ln\frac{\gamma_1}{\gamma_2}\,\mathrm{d}x_1 = 0, \qquad D = 100\,\frac{|A^+ - A^-|}{A^+ + A^-} \tag{54}$$

where A^+ and A^- are the areas the curve encloses above and below the axis. Data satisfy this only within their errors, and D — the imbalance as a percentage of the total area — is the statistic. The conventional threshold is $D < 10$.

Step 1 — reduce every point

The set is isobaric, so P_i^{sat} must be evaluated at each row's own temperature. Acetone and chloroform Antoine constants come from `LIBRARY`; $\Phi_i = 1$ throughout, and the columns are $\gamma_i = y_i P/(x_i P_i^{\mathrm{sat}})$.

x_1	y_1	T/K	P_1^{sat}/kPa	P_2^{sat}/kPa	γ_1	γ_2	$\ln(\gamma_1/\gamma_2)$
0.102	0.067	335.7	125.241	105.854	0.5313	0.9943	−0.6267
0.202	0.160	336.9	130.222	110.024	0.6162	0.9692	−0.4529
0.301	0.280	337.6	133.199	112.516	0.7075	0.9274	−0.2707
0.396	0.401	337.6	133.199	112.516	0.7701	0.8929	−0.1479
0.494	0.544	336.9	130.222	110.024	0.8566	0.8297	+0.0319
0.587	0.660	335.9	126.061	106.541	0.9035	0.7828	+0.1435
0.679	0.765	334.5	120.412	101.810	0.9478	0.7284	+0.2633
0.786	0.866	332.5	112.691	95.340	0.9904	0.6653	+0.3979
0.893	0.945	331.0	107.161	90.703	1.0004	0.5741	+0.5554

As a spot check on the middle row: $\gamma_1 = (0.544)(101.3)/[(0.494)(130.222)] = 55.107/64.330 = 0.8566$, and $\gamma_2 = (0.456)(101.3)/[(0.506)(110.024)] = 46.193/55.672 = 0.8297$, giving $\ln(0.8566/0.8297) = \ln(1.03242) = +0.0319$.

Step 2 — the trapezoid, interval by interval

Each interval contributes $h_k\,\frac{1}{2}(f_k + f_{k+1})$ with $h_k = x_{1,k+1} - x_{1,k}$ and $f = \ln(\gamma_1/\gamma_2)$. The second numeric column is the same thing with $|f|$, which is the denominator of Eq. (54).

interval	h_k	$\frac{1}{2}(f_k + f_{k+1})$	contribution to $\int f$	contribution to $\int \lvert f \rvert$
$0.102 \to 0.202$	0.100	-0.53980	-0.053980	0.053980
$0.202 \to 0.301$	0.099	-0.36180	-0.035818	0.035818
$0.301 \to 0.396$	0.095	-0.20930	-0.019884	0.019884
$0.396 \to 0.494$	0.098	-0.05800	-0.005684	0.008810
$0.494 \to 0.587$	0.093	$+0.08770$	$+0.008156$	0.008156
$0.587 \to 0.679$	0.092	$+0.20340$	$+0.018713$	0.018713
$0.679 \to 0.786$	0.107	$+0.33060$	$+0.035374$	0.035374
$0.786 \to 0.893$	0.107	$+0.47665$	$+0.051002$	0.051002
sum			$\mathbf{-0.00212}$	**0.23174**

Only the fourth interval differs between the two columns, because it is the one that straddles zero.

Step 3 — the statistic

$$D = 100 \times \frac{0.00212}{0.23174} = \boxed{0.915\,\%} \tag{55}$$

against a threshold of 10. **This dataset passes the area test, and comfortably.**

The two lobes, and a piece of bookkeeping worth seeing once. Eq. (54) is usually stated in terms of A^+ and A^-. Where does the curve cross? Linear interpolation on the fourth interval:

$$x_{\text{cross}} = 0.396 + 0.098 \times \frac{0.1479}{0.1479 + 0.0319} = 0.4766$$

so that interval contributes $\frac{1}{2}(0.4766 - 0.396)(-0.1479) = -0.005961$ to A^- and $\frac{1}{2}(0.494 - 0.4766)(0.0319) = +0.000277$ to A^+. Adding these to the unambiguous intervals,

$$A^- = 0.11564, \qquad A^+ = 0.11352, \qquad A^+ - A^- = -0.00212$$

which reproduces the net area, as it must. But $A^+ + A^- = 0.22917$, not 0.23174. The two disagree by 1.1 % because the trapezoid drawn through $|f|$ on the straddling interval is not the same shape as the two triangles the sign change actually makes. It changes D from 0.915 to 0.926, and it changes nothing at all about the verdict — but if your number differs from someone else's in the second decimal, this is why. `vlekit` uses $\int |f|\,\mathrm{d}x_1$, which is the version quoted here.

Step 4 — what the number does not establish

Three qualifications, and each of them is the point of the exercise.

The integral was taken over 0.102 to 0.893, not 0 to 1. The pure end points were never measured. Everything outside that window is invisible to this calculation. `vlekit`'s default fits a Redlich-Kister polynomial to $\ln(\gamma_1/\gamma_2)$ and integrates *that* from 0 to 1, which gives $D = 1.439$ rather than 0.915: a 57 % change in the statistic from a choice about the ends that no threshold mentions.

The set is isobaric, and the isothermal form was used anyway. T runs from 331.0 to 337.6 K, so the true Gibbs-Duhem equation carries $-(H^{\mathrm{E}}/RT^2)(\mathrm{d}T/\mathrm{d}x_1)$ on its right-hand side and this integral does not. Everyone does this, because H^{E} is almost never available; Herington's correction is the conventional patch, $J = 150\,\Delta T/T_{\min} = 150(6.6)/331.0 = 2.99$, and $|D - J| = 2.07 < 10$ passes as well. (`vlekit` reports 1.55 for the same test, because its Herington routine feeds it the Redlich-Kister $D = 1.439$ rather than the trapezoid 0.915. Two

defensible conventions, one threshold, and the verdict does not change — but the number does, so say which you used.) Herington is a correlation, not a derivation. It is reported because reviewers ask for it.

An integral cannot see an error that changes sign. This is structural and no threshold repairs it. Perturbing the vapour compositions by δy_1 shifts the integrand by $\delta \ln(\gamma_1/\gamma_2) = \delta y_1/[y_1(1-y_1)]$; choose $\delta y_1 = \varepsilon\, y_1(1-y_1)(2x_1-1)$ and the shift is $\varepsilon(2x_1-1)$, odd about $x_1 = \frac{1}{2}$, contributing *exactly nothing* to $\int_0^1$. At $\varepsilon = 0.4$ the dataset is grossly wrong and the area test still reports $D \approx 1$. Only the *even* Redlich-Kister coefficients of $\ln(\gamma_1/\gamma_2)$ can make a set fail here.

So the honest sentence is: this dataset passes a necessary condition. On the same nine points the point test gives mean $|\delta y_1| = 0.0059$ (index 3) and the direct test rms $\delta = 0.0414$ (index 2), and *those* are the results worth quoting, because nothing in them is integrated.

Against `vlekit`. `consistency.area_test(d, extrapolate=False)` returns $D = 0.9184$ against the hand 0.9148. The disagreement is 0.4 % of the statistic and it is instructive: the net area, -0.00212, is a sum of eight numbers each up to twenty-five times larger, so the fourth decimal of the tabulated $\ln(\gamma_1/\gamma_2)$ is the last digit that survives into D. Carrying the unrounded integrand gives -0.0021282, hence 0.9184. The denominator, by contrast, agrees to five figures (0.23174 by hand against 0.231730 unrounded) because nothing cancels in it. `area_test(d)` with the default Redlich-Kister extrapolation returns 1.4386.

What this was really testing. That passing a consistency test is not a certificate. Three separate things were assumed to get $D = 0.914$ — that the integrand can be extrapolated to ends that were never measured, that the excess-enthalpy term can be dropped on a set whose temperature moves 6.6 K, and that no error in the data happens to be antisymmetric about the equimolar point — and only the third is even in principle testable by this statistic. The misconception targeted is that the area test is a weak version of the point test, fixable with a tighter threshold. It is not weak, it is *blind*: no threshold on an integral can detect a function that integrates to zero. The other half of the lesson is arithmetic rather than thermodynamic. A statistic built as the difference of two nearly equal areas inherits every rounding error in the table, which is why the same dataset can be quoted as 0.91, 0.92 or 1.44 by three careful people who each made a different defensible choice. Say which one you made.

Problem 6 BUBL P to convergence, and the DEW T nest

Problem

Chloroform (1) + 2-butanone (2) with the Margules-2 parameters of Eq. (41).

(a) Compute the bubble pressure and the vapour it produces at $x_1 = 0.500$, $T = 303.15\,\mathrm{K}$, first with an ideal vapour and then with the two-term virial equation, carrying the second calculation to convergence and showing every iteration.

(b) Set up DEW T at $y_1 = 0.600$, $P = 20.00\,\mathrm{kPa}$ with the iteration structure written out, and take two outer steps.

The four calculations, and where the unknowns sit

All four rest on $y_i P = x_i \gamma_i P_i^{\mathrm{sat}}$. What changes is which two of P, T, x, y you were given:

	given	equation	difficulty
BUBL P	x, T	$P = \sum_i x_i \gamma_i P_i^{\mathrm{sat}}$	none: explicit
DEW P	y, T	$1/P = \sum_i y_i/(\gamma_i P_i^{\mathrm{sat}})$	γ_i needs x, which is the answer
BUBL T	x, P	$\sum_i x_i \gamma_i P_i^{\mathrm{sat}}(T) = P$	P_i^{sat} needs T, which is the answer
DEW T	y, P	$\sum_i y_i/[\gamma_i P_i^{\mathrm{sat}}(T)] = 1/P$	**both**, and they nest

The bubble forms are sums, the dew forms sums of reciprocals; everything about which one needs a loop follows from that and from where the unknown sits. And there is a third difficulty, switched on in part (a): turn on the vapour corrections and $\hat{\varphi}_i$ becomes a function of y and P, so even BUBL P stops being explicit.

(a) BUBL P, ideal vapour — one line

At $x_1 = 0.5$, Eq. (29) simplifies pleasantly: the bracket for γ_1 becomes $A_{12} + (A_{21} - A_{12}) = A_{21}$, and vice versa, so

$$\ln\gamma_1 = (0.5)^2 A_{21} = 0.25(-1.2988) = -0.32470 \;\Rightarrow\; \gamma_1 = 0.72274 \tag{56}$$

$$\ln\gamma_2 = (0.5)^2 A_{12} = 0.25(-0.9122) = -0.22805 \;\Rightarrow\; \gamma_2 = 0.79608 \tag{57}$$

$$\begin{aligned} P &= 0.5(0.72274)(32.3305) + 0.5(0.79608)(15.2242) \\ &= 11.68327 + 6.05984 = \boxed{17.7431\ \mathrm{kPa}} \end{aligned} \tag{58}$$

$$y_1 = \frac{11.68327}{17.7431} = \boxed{0.65847} \tag{59}$$

No iteration at all. Note that the mixture boils at a *lower* pressure than either pure component would suggest a mole-fraction average to be ($0.5 \times 32.33 + 0.5 \times 15.22 = 23.78\,\mathrm{kPa}$): 25 % below the Raoult line, which is what a strong negative deviation looks like.

(a) BUBL P with the virial vapour — successive substitution

Now the full γ-φ statement, $y_i \hat{\varphi}_i P = x_i \gamma_i P_i^{\mathrm{sat}} \varphi_i^{\mathrm{sat}} \mathrm{PF}_i$, must be solved for P and y_1 together. Write the composition-independent left side of each species as a fugacity,

$$f_i \equiv x_i \gamma_i P_i^{\mathrm{sat}} \varphi_i^{\mathrm{sat}}\, \mathrm{PF}_i(P), \qquad P = \frac{f_1}{\hat{\varphi}_1(y_1, P)} + \frac{f_2}{\hat{\varphi}_2(y_1, P)}, \qquad y_1 = \frac{f_1/\hat{\varphi}_1}{P} \tag{60}$$

and iterate from the ideal answer. $\varphi_1^{\mathrm{sat}} = 0.985007$ and $\varphi_2^{\mathrm{sat}} = 0.989467$ were computed in Problem 1 and do not change, because they depend on T alone.

iter	$\hat{\varphi}_1$	$\hat{\varphi}_2$	f_1/kPa	f_2/kPa	P/kPa	y_1
start	—	—	—	—	17.7431	0.65847
1	0.991780	0.987869	11.5027	5.99655	17.6682	0.65644
2	0.991815	0.987920	11.5027	5.99655	17.6675	0.65644
3	0.991815	0.987920	11.5027	5.99655	17.6675	0.65644

Take iteration 1 apart, since the rest is the same arithmetic:

$$\begin{aligned}
\mathrm{PF}_1(17.7431) &= \exp\left[\frac{80.7\times10^{-6}(17.7431-32.3305)\times10^{3}}{2520.53}\right] = 0.999533\\
f_1 &= 0.5(0.72274)(32.3305)(0.985007)(0.999533) = 11.5027\\
f_2 &= 0.5(0.79608)(15.2242)(0.989467)(1.000090) = 5.99655\\
\hat{\varphi}_1 &= \exp\left[\frac{17743.1}{2520.53}(-1177.77+(0.34153)^2\,44.73)10^{-6}\right] = 0.991780\\
P &= \frac{11.5027}{0.991780}+\frac{5.99655}{0.987869} = 11.5981+6.07018 = 17.6682\ \text{kPa}\\
y_1 &= \frac{11.5981}{17.6682} = 0.65644
\end{aligned} \tag{61}$$

Converged to six figures in **two iterations**:

$$P = \boxed{17.6675\ \text{kPa}}, \qquad y_1 = \boxed{0.65644} \tag{62}$$

Was it worth it? The vapour corrections moved P by 0.076 kPa, which is 0.43 %, against a reported $u(P) = 0.65$ kPa on this dataset — a ratio of 0.12. Same verdict as Problem 2, reached from the other direction. The reason to do the calculation is to be able to say that, not to improve the answer.

(b) DEW T — the structure before the arithmetic

Two unknowns are coupled through two different functions, so there are two loops and they nest:

outer: choose T
  evaluate $P_1^{\mathrm{sat}}(T)$, $P_2^{\mathrm{sat}}(T)$ — Antoine at the trial T
  inner: start from $x_1 = y_1$
    evaluate $\gamma_1(x_1)$, $\gamma_2(x_1)$
    $P^{\mathrm{calc}} = \left[\dfrac{y_1}{\gamma_1 P_1^{\mathrm{sat}}}+\dfrac{y_2}{\gamma_2 P_2^{\mathrm{sat}}}\right]^{-1}$
    $x_1 \leftarrow \dfrac{y_1 P^{\mathrm{calc}}}{\gamma_1 P_1^{\mathrm{sat}}}$; repeat until x_1 stops moving
  compare P^{calc} with the specified P; adjust T and repeat

The cost is *multiplicative*, not additive: the inner loop runs to convergence at every outer trial temperature. It also explains why a loose inner tolerance is worse than a slow one — it makes the outer function stop being a smooth function of T, and the outer root-find then converges to whatever the noise allows.

Outer step 1: $T = 305.00$ K. Antoine gives $P_1^{\mathrm{sat}} = 34.8632$ and $P_2^{\mathrm{sat}} = 16.5583$ kPa. Starting the inner loop at $x_1 = y_1 = 0.600$:

inner	γ_1	γ_2	P^{calc}/kPa	new x_1
1	0.80237	0.70032	17.8751	0.38341
2	0.63158	0.88618	18.3456	0.49990
3	0.72266	0.79618	18.4650	0.43975
4	0.67499	0.84588	18.4995	0.47168
5	0.70018	0.82031	18.5088	0.45494
6	0.68693	0.83395	18.5114	0.46378
⋮				
42	0.69151	0.82929	**18.5124**	**0.46074**

Iteration 1 in full:

$$P^{\text{calc}} = \left[\frac{0.600}{(0.80237)(34.8632)} + \frac{0.400}{(0.70032)(16.5583)}\right]^{-1} = \left[\frac{0.600}{27.9731} + \frac{0.400}{11.5961}\right]^{-1}$$

$$= [0.0214492 + 0.0344945]^{-1} = \frac{1}{0.0559436} = 17.8751 \text{ kPa}$$

$$x_1 = \frac{0.600 \times 17.8751}{27.9731} = 0.38341$$

Look at the x_1 column: 0.383, 0.500, 0.440, 0.472, 0.455, 0.464… Successive substitution here *oscillates*, halving its error each pass, and needs 42 iterations to settle at 10^{-12}. That is not a pathology; it is what a fixed-point map with a negative derivative near $-\frac{1}{2}$ does, and it is the practical reason the deck reports DEW T as costing two orders of magnitude more work than BUBL P. Damping — accepting $\frac{1}{2}(x^{\text{old}} + x^{\text{new}})$ instead of x^{new} — removes the oscillation and most of the cost.

Outer step 2: $T = 310.00$ K. $P_1^{\text{sat}} = 42.5205$, $P_2^{\text{sat}} = 20.6547$ kPa; the inner loop runs 43 iterations to $P^{\text{calc}} = 22.8532$ kPa, $x_1 = 0.46440$.

Closing the outer loop. P^{calc} is smooth and monotone in T, so a secant step is enough:

$$T^{(2)} = 305.00 + (20.00 - 18.5124)\,\frac{310.00 - 305.00}{22.8532 - 18.5124} = 305.00 + \frac{1.4876 \times 5}{4.3408} = \boxed{306.713 \text{ K}} \tag{63}$$

which returns $P^{\text{calc}} = 19.9164$ kPa — within 0.4 % of the specification after two outer steps. Two more secant steps give

$$T_{\text{dew}} = 306.812 \text{ K}, \qquad x_1 = 0.4621 \tag{64}$$

The check that costs nothing

At the same P and the same specified composition, the bubble temperature must lie *below* the dew temperature — the bubble point is the first bubble of vapour, the dew point the last drop of liquid, and the two-phase region is between them. BUBL T at $x_1 = 0.600$, $P = 20.00$ kPa gives 303.354 K against DEW T's 306.812 K. Ordered correctly.

Run the inequality before trusting any new routine. It catches swapped components, swapped phases and reciprocal errors in one line, and unlike a comparison against somebody else's published number it costs nothing and depends on nothing.

> **Against** `vlekit`. `regress.bubble_P` with `vapour='ideal'` returns 17.742 958 kPa and $y_1 = 0.658465$ against the hand 17.7431 and 0.65847; with `vapour='virial'`, 17.667 327 kPa and 0.656436 against 17.6675 and 0.65644. Both hand values are high by one or two units in the sixth figure, inherited from the vapour pressures rounded at Eq. (1). `flash.dew_T` returns 306.811 90 K

and $x_1 = 0.462081$ against 306.812 and 0.4621; `regress.bubble_T` returns 303.354 15 K against 303.354. The inner-loop pressures at the two trial temperatures are 18.512448 and 22.853173 against the tabulated 18.5124 and 22.8532.

What this was really testing. That these are not four algorithms but one equation with four different things unknown, and that the cost of each is decided by *which unknown appears inside which function* — not by the solver, not by the physics, and not by how non-ideal the mixture is. Nothing was made harder between (a) and (b): same system, same model, same tolerance. Moving the specification from the liquid to the vapour and from P to T turned an explicit one-liner into two nested loops and about a hundred function evaluations, and that ratio is where a flowsheet actually fails.

The second target is the belief that "iterate until converged" is a complete instruction. The inner loop in (b) oscillates, and a student who stops it at five iterations gets $x_1 = 0.455$ instead of 0.461 and an outer function that is no longer smooth. The third is quieter: it is worth noticing that 0.600 means different things in (a) and (b). A liquid at $x_1 = 0.6$ and a vapour at $y_1 = 0.6$ are different states, so the answers are not a comparable pair — only the ordering check is.

Problem 7 K-values, relative volatility, Fenske, and what constant α costs

Problem

Ethanol (1) + water (2) at 101.3 kPa, 21 published points (`doi:10.1021/je2008704`). A van Laar model fitted to those points by Barker's method gives

$$A = 1.76673, \qquad B = 0.92108 \tag{65}$$

Compute K_1, K_2 and α_{12} at $x_1 = 0.02$ and at $x_1 = 0.84$; obtain $N_{\min}$ from Fenske for that separation; then step the stages off one at a time with the real $\alpha(x)$ and report the error the shortcut makes.

The model, and the reason for choosing van Laar here

$$\ln\gamma_1 = A\left(\frac{Bx_2}{Ax_1 + Bx_2}\right)^2, \qquad \ln\gamma_2 = B\left(\frac{Ax_1}{Ax_1 + Bx_2}\right)^2 \tag{66}$$

As with Margules, $A = \ln\gamma_1^\infty$ and $B = \ln\gamma_2^\infty$, so Eq. (65) says $\gamma_1^\infty = 5.852$ and $\gamma_2^\infty = 2.512$. On these 21 points van Laar reaches rms $\delta T = 0.045$ K and rms $\delta y_1 = 0.0040$, against a reported $u(y_1) = 0.015$: it is inside the measurement. Margules-2, which is symmetric, manages only 0.226 K and 0.0068 on the same points — ethanol + water is a strongly skewed system and the symmetric chord cannot follow it.

Step 1 — the two states

This is an isobaric problem, so each composition sits at its own bubble temperature. Solving $\sum_i x_i\gamma_i P_i^{\mathrm{sat}}(T) = 101.3$ kPa (a BUBL T, the outer loop of Problem 6 without an inner one) gives $T = 367.808$ K at $x_1 = 0.02$ and $T = 351.305$ K at $x_1 = 0.84$.

At the bottoms, $x_1 = 0.02$.

$$\begin{aligned} Ax_1 + Bx_2 &= 1.76673(0.02) + 0.92108(0.98) = 0.035335 + 0.902658 = 0.937993 \\ \ln\gamma_1 &= 1.76673\left(\frac{0.902658}{0.937993}\right)^2 = 1.76673(0.962330)^2 = 1.76673(0.926079) = 1.63614 \\ \gamma_1 &= 5.1351 \\ \ln\gamma_2 &= 0.92108\left(\frac{0.035335}{0.937993}\right)^2 = 0.92108(0.037670)^2 = 0.0013070 \;\Rightarrow\; \gamma_2 = 1.0013 \end{aligned} \tag{67}$$

At 367.808 K, Antoine gives $P_1^{\mathrm{sat}} = 188.500$ and $P_2^{\mathrm{sat}} = 83.5032$ kPa, so

$$K_i = \frac{y_i}{x_i} = \frac{\gamma_i P_i^{\mathrm{sat}}(T)}{P}, \qquad \alpha_{12} = \frac{K_1}{K_2} = \frac{\gamma_1 P_1^{\mathrm{sat}}}{\gamma_2 P_2^{\mathrm{sat}}} \tag{68}$$

$$\begin{aligned} K_1 &= \frac{5.1351 \times 188.500}{101.3} = \frac{967.97}{101.3} = 9.5554, \qquad K_2 = \frac{1.0013 \times 83.5032}{101.3} = 0.82539 \\ \alpha_{12} &= \frac{9.5554}{0.82539} = \boxed{11.577} \end{aligned} \tag{69}$$

At the distillate, $x_1 = 0.84$.

$$\begin{aligned} Ax_1 + Bx_2 &= 1.76673(0.84) + 0.92108(0.16) = 1.484053 + 0.147373 = 1.631426 \\ \ln\gamma_1 &= 1.76673(0.090334)^2 = 0.014417 \;\Rightarrow\; \gamma_1 = 1.0145 \\ \ln\gamma_2 &= 0.92108(0.909664)^2 = 0.762164 \;\Rightarrow\; \gamma_2 = 2.1429 \end{aligned} \tag{70}$$

At 351.305 K, $P_1^{\text{sat}} = 101.159$ and $P_2^{\text{sat}} = 44.0183$ kPa:

$$K_1 = \frac{1.0145 \times 101.159}{101.3} = 1.0131, \qquad K_2 = \frac{2.1429 \times 44.0183}{101.3} = 0.93116$$
$$\alpha_{12} = \frac{1.0131}{0.93116} = \boxed{1.088} \tag{71}$$

What just happened. α_{12} fell from 11.6 to 1.09 across the column — a factor of **10.6** — while the vapour-pressure ratio $P_1^{\text{sat}}/P_2^{\text{sat}}$ moved only from 2.257 to 2.298. Every bit of that collapse is γ_1/γ_2, which runs from 5.13 down to 0.473. K is not a component property; it is a property of the mixture at a state, which is why a De Priester chart works for light hydrocarbons and is useless here.

Step 2 — Fenske

$$N_{\min} = \frac{\ln\left[\dfrac{x_D}{1-x_D} \cdot \dfrac{1-x_B}{x_B}\right]}{\ln \tilde{\alpha}}, \qquad \tilde{\alpha} = \sqrt{\alpha_D\, \alpha_B} \tag{72}$$

with four assumptions attached: total reflux, α constant over the column and taken as the geometric mean of the two ends, equilibrium stages, and one binary pair. $N_{\min}$ counts the reboiler.

$$\tilde{\alpha} = \sqrt{1.088 \times 11.577} = \sqrt{12.5958} = 3.5491$$
$$\frac{x_D}{1-x_D} \cdot \frac{1-x_B}{x_B} = \frac{0.84}{0.16} \cdot \frac{0.98}{0.02} = 5.25 \times 49 = 257.25$$
$$N_{\min} = \frac{\ln 257.25}{\ln 3.5491} = \frac{5.55005}{1.26669} = \boxed{4.382 \text{ stages}} \tag{73}$$

Step 3 — stepping the stages with the real $\alpha(x)$

At total reflux the passing streams have $y_{n+1} = x_n$, so no extra assumption is needed: the same van Laar that produced α produces the stage count. Start with the vapour leaving the top, $y = x_D = 0.84$; find the liquid in equilibrium with it (a DEW T, Problem 6b); set the next vapour equal to that liquid; repeat until $x \le x_B$.

stage n	y_n	x_n	T/K	α at x_n
1	0.84000	0.82585	351.332	1.107
2	0.82585	0.80706	351.374	1.134
3	0.80706	0.78099	351.444	1.173
4	0.78099	0.74245	351.571	1.237
5	0.74245	0.67962	351.836	1.359
6	0.67962	0.55738	352.524	1.685
7	0.55738	0.24086	355.325	3.969
8	0.24086	0.02798	366.281	11.02
9	0.02798	0.00222	372.407	12.95

$x_B = 0.02$ falls between stages 8 and 9. Interpolating logarithmically,

$$N = 8 + \frac{\ln 0.02798 - \ln 0.02}{\ln 0.02798 - \ln 0.00222} = 8 + \frac{0.33576}{2.53398} = \boxed{8.13 \text{ stages}} \tag{74}$$

Answer

	stages	
Fenske at $\tilde{\alpha} = 3.5491$	4.38	the shortcut
stepped with $\alpha(x)$ from the same model	8.13	the same physics, no averaging
shortfall	**3.75**	Fenske is short by 46 %

And the error is in the dangerous direction: it builds a column that cannot make the specification. The reason is visible in the α column of the stepping table. Fenske is exact when $\ln \alpha$ is linear in the stage number. Here α stays below 1.7 for six of the nine stages, then leaps to 11.0 in two; the geometric mean 3.55 describes no stage in the column at all. Two-thirds of the stages are being spent in a region where the true α is a third of the mean or less, and the shortcut has no way to know.

And it gets worse. This van Laar puts the azeotrope at $x_1 = 0.9132$ (measured: 0.8946). Push x_D towards it and the true stage count diverges while Fenske, which knows nothing about an azeotrope, returns a comfortable finite number. At $x_D = 0.912$ the shortcut answers about five stages for a separation that needs infinitely many. **Constant α does not fail loudly. It fails by returning a plausible number.**

Against `vlekit`. `regress.barker_fit(d,'vanlaar')` returns $(1.7667256, 0.9210752)$, whose γ^∞ are 5.85166 and 2.51199. `regress.bubble_T` gives 367.808 42 K and 351.305 39 K; the activity coefficients there are 5.135235, 1.001308, 1.014521 and 2.142950 against the hand 5.1351, 1.0013, 1.0145 and 2.1429. K_1 and K_2 are 9.555712 / 0.825394 and 1.013108 / 0.931184 against 9.5554 / 0.82539 and 1.0131 / 0.93116; α is 11.57716 and 1.087978 against 11.577 and 1.088; $\tilde{\alpha} = 3.549040$ and $N_{\min} = 4.38158$ against 3.5491 and 4.382. The stepped count from `flash.dew_T` is 8.13244 against 8.13, and the azeotrope of this model is 0.913151 against 0.9132.

What this was really testing. That K and α are the only place the whole module's machinery becomes visible to a process engineer, and that they are functions of state rather than tabulated component properties. A student who has internalised $K_i = \gamma_i P_i^{\mathrm{sat}} / P$ can say immediately why a De Priester chart is fine for propane and fatal for ethanol: the γ_i that a hydrocarbon chart quietly sets to one is here responsible for a factor of ten in α.
The second target is the geometric mean. Most students expect a shortcut to be conservative, and expect $\tilde{\alpha} = \sqrt{\alpha_D \alpha_B}$ to "average out" a varying α. It does neither. $\tilde{\alpha}$ is exact only if $\ln \alpha$ is linear in stage number, which is a statement about the mixture that nobody checked, and when it is false the shortcut is short — here by nearly a factor of two, in the direction that under-builds the column. The final point is the one the deck ends on: the failure is silent. Fenske returns 5.1 stages for a distillate specification that lies past the azeotrope and therefore needs infinitely many, and nothing in the arithmetic warns you.

Problem 8 One $\hat{\varphi}$ from Peng-Robinson, and a Huron-Vidal a_{mix}

Problem

(a) Compute $\hat{\varphi}_1$ for the vapour of Problem 1 — chloroform (1) + 2-butanone (2), $y_1 = 0.673$, $T = 303.15\,\text{K}$, $P = 17.69\,\text{kPa}$ — from the Peng-Robinson equation with the classical mixing rule and $k_{12} = 0$. Compare with the two-term virial value found in Problem 1.

(b) At $x_1 = 0.500$, compute a_{mix} and b_{mix} from the Huron-Vidal rule using the G^{E} of Eq. (41), and compare with the classical rule at $k_{12} = 0$.

(a) Peng-Robinson, from constants to $\hat{\varphi}_1$

Pure-component terms. With $\kappa = 0.37464 + 1.54226\omega - 0.26992\omega^2$,

$$\begin{aligned}\kappa_1 &= 0.37464 + 1.54226(0.222) - 0.26992(0.222)^2 = 0.703719\\ \kappa_2 &= 0.37464 + 1.54226(0.323) - 0.26992(0.323)^2 = 0.844629\end{aligned} \tag{75}$$

$$\alpha_i(T) = \left[1 + \kappa_i\left(1 - \sqrt{T/T_{c,i}}\right)\right]^2 \tag{76}$$

$$\begin{aligned}T_{r,1} &= 303.15/536.4 = 0.565157, & \sqrt{T_{r,1}} &= 0.751769\\ \alpha_1 &= [1 + 0.703719(0.248231)]^2 = (1.174675)^2 = 1.37988 & &\\ T_{r,2} &= 303.15/536.8 = 0.564736, & \sqrt{T_{r,2}} &= 0.751489\\ \alpha_2 &= [1 + 0.844629(0.248511)]^2 = (1.209900)^2 = 1.46386 & &\end{aligned} \tag{77}$$

With $a_{c,i} = 0.45724\,(RT_{c,i})^2/P_{c,i}$ and $b_i = 0.07780\,RT_{c,i}/P_{c,i}$, working in Pa and m^3/mol:

$$\begin{aligned}RT_{c,1} &= 8.314463(536.4) = 4459.88, & a_{c,1} &= \frac{0.45724(4459.88)^2}{5.472\times10^6} = 1.66205\\ a_1 &= 1.66205 \times 1.37988 = 2.29343, & b_1 &= \frac{0.07780(4459.88)}{5.472\times10^6} = 6.34098\times10^{-5}\\ RT_{c,2} &= 8.314463(536.8) = 4463.20, & a_{c,2} &= \frac{0.45724(4463.20)^2}{4.207\times10^6} = 2.16504\\ a_2 &= 2.16504 \times 1.46386 = 3.16932, & b_2 &= \frac{0.07780(4463.20)}{4.207\times10^6} = 8.25380\times10^{-5}\end{aligned} \tag{78}$$

so $b_1 = 63.41$ and $b_2 = 82.54\,\text{cm}^3/\text{mol}$.

Mixture terms at the vapour composition. With $k_{12} = 0$ the cross term is $a_{12} = \sqrt{a_1 a_2} = \sqrt{2.29343 \times 3.16932} = 2.69604$, and

$$\begin{aligned}a_{\text{mix}} &= y_1^2 a_1 + 2y_1y_2a_{12} + y_2^2a_2\\ &= 0.452929(2.29343) + 0.440142(2.69604) + 0.106929(3.16932)\\ &= 1.038761 + 1.186639 + 0.338891 = 2.56429\end{aligned} \tag{79}$$

$$b_{\text{mix}} = 0.673(6.34098) + 0.327(8.25380) = 6.96647\times10^{-5}\ \text{m}^3/\text{mol} \tag{80}$$

The dimensionless groups. $RT = 2520.53\,\text{J/mol}$, $P = 17\,690\,\text{Pa}$:

$$\mathcal{A} = \frac{a_{\text{mix}}P}{(RT)^2} = \frac{45362.3}{6.35307\times10^6} = 0.00714021, \qquad \mathcal{B} = \frac{b_{\text{mix}}P}{RT} = 0.000488932 \tag{81}$$

Both tiny, as they must be at 0.18 bar: the vapour is nearly ideal, and whatever $\hat{\varphi}_1$ comes out to will be within a per cent of one.

The root. The cubic in Z is

$$Z^3 - (1-\mathcal{B})Z^2 + (\mathcal{A} - 3\mathcal{B}^2 - 2\mathcal{B})Z - (\mathcal{A}\mathcal{B} - \mathcal{B}^2 - \mathcal{B}^3) = 0$$

Newton from $Z = 1$ converges in three steps to the vapour root

$$Z = \boxed{0.993311} \tag{82}$$

(The cubic also has roots at 0.005617 and 0.000583; both lie at or below $\mathcal{B}$-scale volumes and are not liquid states of this mixture at 0.18 bar. The vapour root is unambiguous here, which is exactly why this is a safe place to practise — the root-selection trap belongs to high pressures.)

The fugacity coefficient.

$$\ln\hat{\varphi}_i = \frac{b_i}{b_{\text{mix}}}(Z-1) - \ln(Z-\mathcal{B}) - \frac{\mathcal{A}}{2\sqrt{2}\,\mathcal{B}}\left[\frac{2\sum_j y_j a_{ij}}{a_{\text{mix}}} - \frac{b_i}{b_{\text{mix}}}\right]\ln\frac{Z+(1+\sqrt{2})\mathcal{B}}{Z+(1-\sqrt{2})\mathcal{B}} \tag{83}$$

Term by term for $i = 1$:

$$\begin{aligned}
\frac{b_1}{b_{\text{mix}}} &= \frac{6.34098}{6.96647} = 0.910214 \\
\textstyle\sum_j y_j a_{1j} &= 0.673(2.29343) + 0.327(2.69604) = 2.42508 \\
\frac{2(2.42508)}{2.56429} - 0.910214 &= 1.891425 - 0.910214 = 0.981211 \\
\frac{\mathcal{A}}{2\sqrt{2}\,\mathcal{B}} &= \frac{0.00714021}{2.828427 \times 0.000488932} = 5.16319 \\
\ln\frac{Z+2.414214\mathcal{B}}{Z-0.414214\mathcal{B}} &= \ln\frac{0.994491}{0.993108} = \ln(1.0013925) = 0.0013915
\end{aligned} \tag{84}$$

$$\begin{aligned}
\ln\hat{\varphi}_1 &= 0.910214(-0.006689) - \ln(0.992822) - 5.16319(0.981211)(0.0013915) \\
&= -0.006088 + 0.007204 - 0.007050 = -0.005934 \\
\hat{\varphi}_1 &= \boxed{0.994083}
\end{aligned} \tag{85}$$

The three terms are each about 7×10^{-3} and they cancel to 6×10^{-3}. Compute $\hat{\varphi}$ rather than $\ln\hat{\varphi}$ and you lose that structure; it is one reason flash algorithms are written in the logarithm.

(a) Against the virial route

	$\hat{\varphi}_1$	implied mixture B / cm^3/mol
two-term virial, Pitzer-Curl/Tsonopoulos (Problem 1)	0.991801	−1356.1
Peng-Robinson, $k_{12} = 0$	0.994083	−947.7

The second column is the point. Every cubic has a low-density expansion whose leading coefficient is

$$B(T,x) = b(x) - \frac{a(T,x)}{RT} = 6.96647 \times 10^{-5} - \frac{2.56429}{2520.53} = -9.47697 \times 10^{-4}\ \text{m}^3/\text{mol} \tag{86}$$

and the correlation used in Problem 1 gives $y_1^2 B_{11} + 2y_1 y_2 B_{12} + y_2^2 B_{22} = -1356.1$ cm^3/mol. Peng-Robinson is 30 % short. So the 0.23 % disagreement in $\hat{\varphi}_1$ is not rounding and it is not a bug: it is two different descriptions of the same vapour disagreeing by a third about how non-ideal it is, at a pressure so low that a third of very little is still very little. *Both* answers are inside the noise established in Problem 2. Neither would be at 5 bar.

(b) Huron-Vidal

The rule requires the cubic to reproduce a chosen G^{E} in the infinite-pressure limit, where $v \to b$, the excess volume vanishes and the expression closes in a and b alone:

$$\frac{a_{\mathrm{mix}}}{b_{\mathrm{mix}}RT} = \sum_i x_i \frac{a_i}{b_i RT} + \frac{G^{\mathrm{E}}(T,x)}{RT\,C}, \qquad b_{\mathrm{mix}} = \sum_i x_i b_i, \qquad C = \frac{\ln(\sqrt{2}-1)}{\sqrt{2}} = -0.62323 \tag{87}$$

The activity model. At $x_1 = 0.5$, Eq. (28) with the parameters of Eq. (41):

$$\frac{G^{\mathrm{E}}}{RT} = (0.5)(0.5)\left[(-1.2988)(0.5) + (-0.9122)(0.5)\right] = 0.25(-1.10550) = -0.27637 \tag{88}$$

The pure-component sum.

$$\begin{aligned} \frac{a_1}{b_1 RT} &= \frac{2.29343}{(6.34098\times10^{-5})(2520.53)} = \frac{2.29343}{0.159826} = 14.3495 \\ \frac{a_2}{b_2 RT} &= \frac{3.16932}{(8.25380\times10^{-5})(2520.53)} = \frac{3.16932}{0.208040} = 15.2342 \\ \sum_i x_i \frac{a_i}{b_i RT} &= 0.5(14.3495) + 0.5(15.2342) = 14.7919 \end{aligned} \tag{89}$$

Assemble.

$$\begin{aligned} D &= 14.7919 + \frac{-0.27637}{-0.62323} = 14.7919 + 0.443448 = \boxed{15.2353} \\ b_{\mathrm{mix}} &= 0.5(6.34098) + 0.5(8.25380) = 7.29739\times10^{-5}\ \mathrm{m^3/mol} \\ a_{\mathrm{mix}} &= D\,b_{\mathrm{mix}}RT = 15.2353(7.29739\times10^{-5})(2520.53) = \boxed{2.80227} \end{aligned} \tag{90}$$

against the classical rule at the same composition,

$$a^{\mathrm{cl}}_{\mathrm{mix}} = 0.25(2.29343) + 0.5(2.69604) + 0.25(3.16932) = 2.71371 \tag{91}$$

with the same b_{mix}, because Huron-Vidal keeps b linear.

Check the sign, which is where implementations die. C is **negative**. This system has $G^{\mathrm{E}} < 0$, so $G^{\mathrm{E}}/C > 0$ and Huron-Vidal *raises* a/b by 3.3 % relative to the classical rule — a stronger unlike attraction, which is exactly what a negative deviation from Raoult's law means, and exactly the direction a *negative* k_{12} moves it. Several references define C with the opposite sign and then subtract. Both are in print; only one works with the value in Eq. (87). Settle it with a numerical limit, never with a citation.

And what the classical rule was assuming all along. With $k_{12} = 0$ the bracket in Eq. (87) run backwards factorises into a perfect square, so the classical rule *is* a G^{E} rule with a van Laar activity model whose parameters are fixed by the pure components alone:

$$A = 0.0840, \qquad B = 0.1093$$

Both positive — so at $k_{12} = 0$ the classical rule can only produce *positive* deviations. It offers $G^{\mathrm{E}}/RT = +0.024$ at equimolar where this system needs -0.276: the wrong sign, not merely the wrong size. One fitted $k_{12} = -0.072$ buys -0.305 and rescues it, but the *shape* stays a single skewed lobe whose skew is set by $b_1 : b_2 = 63.4 : 82.5$ and by nothing about the mixture. k_{ij} scales the lobe and can flip its sign; it cannot move the peak or change the curvature.

Against `vlekit`. `cubic.PengRobinson.ln_phi` returns $\ln\hat{\varphi}_1 = -0.00593438$ and $\hat{\varphi}_1 = 0.9940832$ against the hand -0.005934 and 0.994083; the root is $Z = 0.99331128$ against 0.993311, and $\mathcal{A}$, $\mathcal{B}$ are 0.00714015 and 0.000488908 against 0.00714021 and 0.000488932. `gemix.HuronVidal` gives $D = 15.235923$, $b_{\mathrm{mix}} = 7.297\,021 \times 10^{-5}\,\mathrm{m^3/mol}$ and $a_{\mathrm{mix}} = 2.802245$ against the hand 15.2353, 7.29739×10^{-5} and 2.80227; the classical a_{mix} is 2.713676 against 2.71371. `gemix.C_PR` is -0.62322524. `vanlaar_equivalent_of_classical` returns $(0.0839552, 0.1092811)$, and `ClassicalGE` gives $G^{\mathrm{E}}/RT = +0.023740$ at $k_{12} = 0$ and -0.305136 at $k_{12} = -0.072$.

What this was really testing. Part (a) targets the belief that agreement between two models is evidence about the fluid. The virial route and the cubic differ by 30 % on the second virial coefficient of the same vapour at the same state, and they still agree on $\hat{\varphi}_1$ to two parts in a thousand — because at 0.18 bar there is almost nothing for them to disagree *about*. Read $\hat{\varphi}$ alone and you would conclude the two descriptions were interchangeable. They are not; the pressure was doing the work.

Part (b) targets the idea that a G^{E} mixing rule is a more sophisticated alternative to the classical one. It is not an alternative: at $k_{12} = 0$ the classical rule already *is* a G^{E} rule, with one particular and rather rigid activity model — a van Laar whose two parameters are determined by the pure-component a_i and b_i and cannot be adjusted at all. Recognising that turns k_{ij} from a free knob into a scale factor on a lobe of fixed shape, and it explains in one sentence why a single k_{ij} can rescue a moderately non-ideal binary and cannot rescue an alcohol + water one. The sign check is the other half: $C < 0$ is a fact about the cubic, not a convention, and getting it backwards turns every positive-deviation system into a negative-deviation one while raising no error at all.

Appendix A The data, and where every number came from

Four published binaries are used. None of them was copied from a textbook table; each was transcribed mechanically from the publisher's record or from the NIST TRC ThermoML digitisation of it, and each is checked against its own pure-component end points before use. The full provenance — including the transcription route, the typo and smoothness checks, and the consistency verdicts, failures included — is in `vlekit/datasets.py`.

A.1 Chloroform (1) + 2-butanone (2), isothermal, 303.15 K

Used in Problems 1, 2, 3, 4, 6 and 8. Twenty-two points; the only isothermal P-x-y set in the module, and therefore the only one on which Gibbs-Duhem appears with no enthalpy term.

Clara, R. A.; Gómez Marigliano, A. C.; Solimo, H. N. *Density, Viscosity, Vapor-Liquid Equilibrium, Excess Molar Volume, Viscosity Deviation, and Their Correlations for the Chloroform + 2-Butanone Binary System.* J. Chem. Eng. Data **2006**, 51(4), 1473–1478. `doi:10.1021/je060150a`

Uncertainties (NIST TRC, combined expanded, 95 %): $u(x_1) = 0.001$, $u(y_1) = 0.001$–0.021, $u(P) = 0.39$–0.65 kPa, $u(T) = 0.01$ K. *End-point check:* the table's pure values are 32.19 and 15.74 kPa; the stored Antoine sets give 32.33 (+0.44 %) and 15.22 (−3.28 %). *Consistency:* area 3.18 pass, point 0.0071 pass (index 3), direct 0.0757 fail (index 4), Kojima 1181 fail badly — the last because the set contains both pure end points, where one γ is $0/0$.

A.2 Acetone (1) + chloroform (2), isobaric, 101.3 kPa

Used in Problem 5. Nine points, none of them pure. Open access, so a student can read the source table directly.

Gao, Q.; Zhao, Z.; Jia, P.; Zhang, X. *Effect of Ionic Liquids on the Isobaric Vapor-Liquid Equilibrium Behavior of Acetone-Chloroform.* Applied Sciences **2018**, 8(9), 1519. `doi:10.3390/app8091519` (Table A1, the ionic-liquid-free binary; CC-BY)

Uncertainties as stated by the authors: $u(x_1) = u(y_1) = 0.002$, $u(T) = 0.1$ K; pressure read on a mercury manometer and corrected to 101.3 kPa. *Consistency:* five tests out of five pass — the only clean sweep in the module, helped considerably by the absence of pure end points, which leaves the Kojima test nothing to trip on. *Caveat:* nine points is a thin dataset, and it is here for recognisability and for the trapezoid arithmetic, not because it is the best Type 4 data available.

A.3 Ethanol (1) + water (2), isobaric, 101.3 kPa

Used in Problem 7. Twenty-one points, x_1 from 0.018 to 0.972.

Kamihama, N.; Matsuda, H.; Kurihara, K.; Tochigi, K.; Oba, S. *Isobaric Vapor-Liquid Equilibria for Ethanol + Water + Ethylene Glycol and Its Constituent Three Binary Systems.* J. Chem. Eng. Data **2012**, 57(2), 339–344. `doi:10.1021/je2008704`

Uncertainties: $u(x_1) = 0.001$, $u(y_1) = 0.015$, $u(P) = 0.1$ kPa, $u(T) = 0.1$ K. *Azeotrope:* $y_1 - x_1$ crosses zero at $x_{az} = 0.8946$, against an accepted literature value of 0.894. *Component numbering — a real trap.* The ThermoML record lists the substances as (water, ethanol) and tabulates the mole fraction of the *other* one. Physics settles it: the mixture boils at 368.18 K at $x_1 = 0.018$ and 351.35 K at $x_1 = 0.972$, and pure water and pure ethanol boil at 373.15 and 351.4 K, so the tabulated component is ethanol. Run this check on every machine-readable table. *Consistency:* four of five pass; the direct test misses by 6 %.

A.4 Methanol (1) + water (2), isobaric, 66.52 kPa

Used in Problem 4(b). Sixteen points.

Yang, C.; Sun, F.; Ma, S.; Yin, X.; Zeng, H. *Organic Salt Effect on Vapor-Liquid Equilibrium of the Methanol + Water System at Subatmospheric Pressure.* J. Chem. Eng. Data **2012**, 57(10), 2696–2701. `doi:10.1021/je300613v` (the salt-free binary)

End-point check: pure methanol $-0.25\,\%$, pure water $-0.91\,\%$ against the stored Antoine sets — both acceptable. The companion 37.5 kPa set from the same paper is $-4.33\,\%$ on pure methanol and fails the point test; it is not used here, and the reason is worth knowing. *Consistency:* Herington and the point test pass; the area, direct and Kojima tests fail marginally.

A.5 Auxiliary data

Antoine constants, critical constants, acentric factors and liquid molar volumes are the entries in `vlekit.data.LIBRARY`: Smith, Van Ness & Abbott Appendix B in the $\ln(\mathrm{kPa})$ form for most species, the Yaws $\log_{10}(\mathrm{mm\,Hg})$ set for chloroform and 2-butanone (because the SVA coefficients miss their normal boiling points by several kelvin), and Poling, Prausnitz & O'Connell for T_c, P_c and ω. Every entry is checked against a normal boiling point and a 298.15 K vapour pressure in the toolkit's test suite. Second virial coefficients use the Pitzer-Curl correlation with the Tsonopoulos temperature functions and $k_{12} = 0$; that correlation is for non-polar and slightly polar species, so on associating vapours it understates the deviation, and the $\hat{\varphi}$ of Problem 1 is itself uncertain by more than the Poynting rung it is being compared against.

Appendix B Verification

Every printed answer is produced twice — once along the rounded chain this document shows, once by `vlekit` at double precision — and the two are compared by `toolkit/verify_worked_examples.py`. The current run makes **102 comparisons and none falls outside tolerance**. The eight largest relative disagreements are all rounding, and all of them are already discussed in the problem that produced them:

problem	quantity	relative disagreement
5	$A^+ - A^-$, the net area	3.9×10^{-3}
5	D, trapezoid over the data	3.5×10^{-3}
5	net area $\int \ln(\gamma_1/\gamma_2)\,\mathrm{d}x_1$	3.4×10^{-3}
2	correction-to-noise ratio	1.3×10^{-3}
3	error in the predicted γ_2	9.6×10^{-4}
1	net $\Delta \ln \gamma_1$	5.8×10^{-4}
3	error in the predicted γ_1	5.6×10^{-4}
8	van Laar A implied by the classical rule	5.3×10^{-4}

(The script's own printed list shows the area-test D twice, because it is compared both with a direct trapezoid and with `vlekit`'s own routine; the duplicate is removed above.)

Three of the top four are in Problem 5, and for one reason: the area-test statistic is a difference between two nearly equal areas, so it loses about two significant figures to cancellation before it is even formed. Every other quantity in this document agrees with the library to better than one part in a thousand, and most to one part in 10^5.

Running it.

```
cd vle/toolkit
python3 verify_worked_examples.py        # the 102-row comparison table
python3 verify_worked_examples.py -v     # every intermediate as well
```

The script imports only `vlekit` and `numpy/scipy`. It takes a few seconds, most of it in the stage-stepping of Problem 7, which runs nine nested dew-point calculations.

What the check does and does not establish. It establishes that the arithmetic printed here is the arithmetic the library performs, and that a student who follows every line with a calculator will land where the toolkit lands. It establishes nothing whatever about whether the models are right, whether the data are right, or whether either is applicable to a problem you have. That is the subject of the module, and it is not something a test suite can settle.