# Surfactant Science
## Principles & Practice
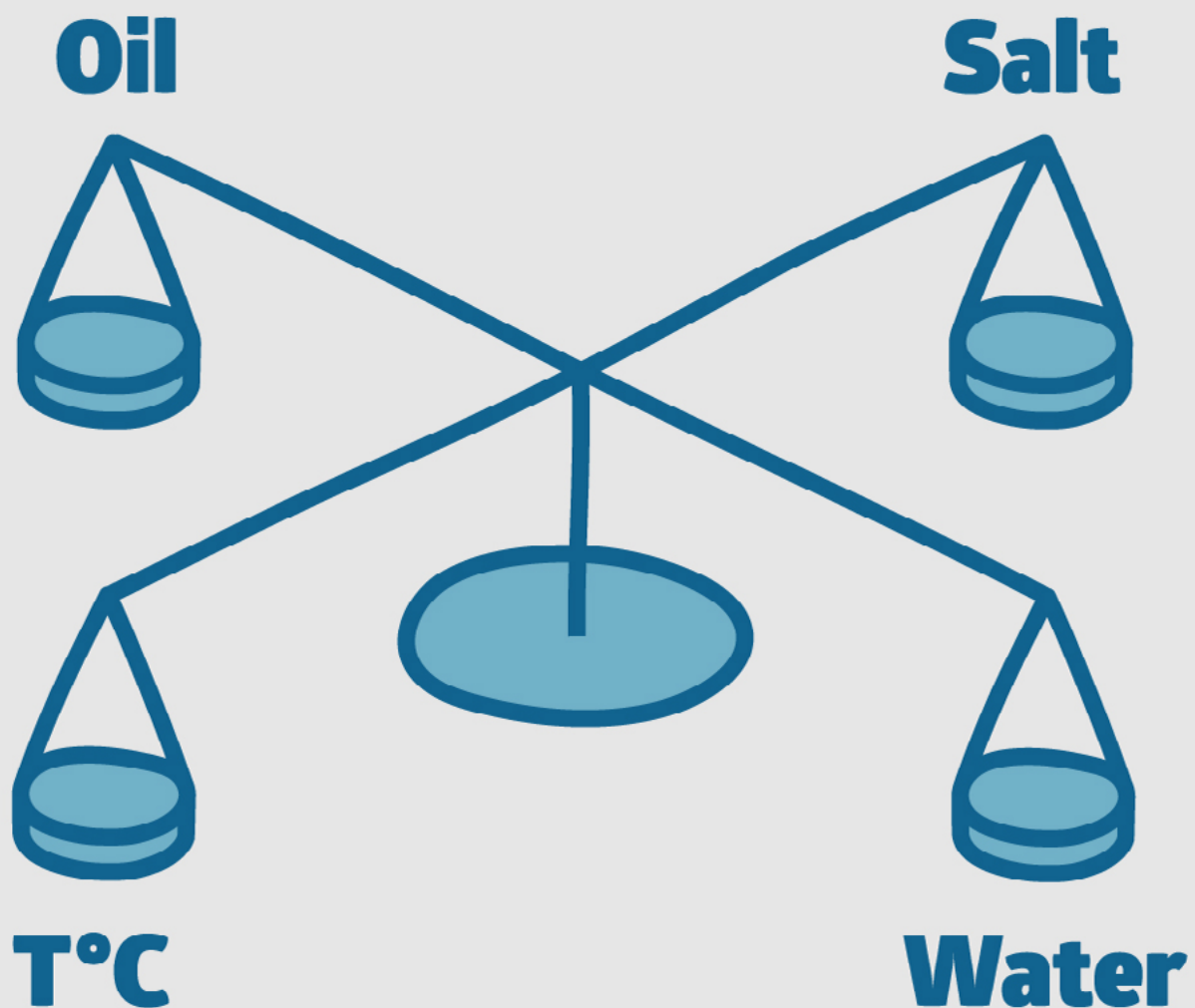
Oil

Salt

T°C

Water

Prof **Steven Abbott**

# Surfactant Science:
# Principles and Practice

Prof Steven Abbott

*Steven Abbott TCNF Ltd, Ipswich, UK*
*and Visiting Professor, University of Leeds, UK*

steven@stevenabbott.co.uk
www.stevenabbott.co.uk

Version history:

| | | |
|---|---|---|
| 1.0.0 | 20 September 2015 | Initial release |
| 1.0.1 | 23 September 2015 | Addition of surfactant solubilizers |
| 1.0.2 | 26 January 2016 | Minor mods |
| 1.0.3 | 07 March 2016 | Update to new website and Fast Scans |
| 1.0.4 | 03 October 2018 | Update on NAC calculations plus minor mods |
| 1.0.5 | 28 November 2019 | Update with the new Acosta Polar Oil approach |
| 1.0.6 | 20 January 2019 | Update on foam making |
| 1.0.7 | 09 December 2024 | Yield stress to stop emulsion creaming |

# Contents

**Preface**

The target audience for this book is a younger version of myself. I have included all those things I wish I had known years ago. Had I had this book, my life of formulating with surfactants would have been so much more productive!

This book, therefore, contains everything that I know about surfactant science that can be useful to someone (such as myself) who formulates with surfactants. So it is a good idea to make it clear right from the start that although there are many good answers to many of our formulation questions (and these answers generally are *not* the ones you read about) there are lots of surfactant questions for which surfactant science currently has no satisfactory answers.

As we will see, I am happy to blame the surfactant community for having wasted decades by *not* working with the best available surfactant theory, and therefore not having the time or the clarity to investigate the extra questions for which formulators really wanted answers. More specifically I lay the blame on those in academia and industry who perpetuated the myth that armed with critical micelle concentration, CMC, and with hydrophilic lipophilic balance, HLB, one could formulate wisely. As countless formulators have found, these, the only numbers that are generally mentioned in the surfactant literature, are largely useless for most purposes. Those who formulate exclusively with ethoxylates will take some exception to this. Because ethoxylates can be formulated with any combination of alkane and ethoxylate chain lengths there have been endless studies and correlations from which a formulator can derive some useful answers armed only with the CMC and HLB values. Ethoxylates, however, are a good part of the larger problem because their somewhat rational behaviour has been used as a cover to pretend that the rest of surfactant science can use ethoxylate-based techniques. This is simply untrue. Outside the world of ethoxylates, HLB is largely meaningless and valueless. And few of us can take long lists of CMC values and reach deep insights into our formulation problems.

The good news is that once you exit the world of HLB there are a set of scientific principles (especially HLD) on which to build a much more solid base for formulations. In some areas the formulations can be entirely rational. In other areas we at least know which bits of science are lacking so the formulator can tell how far to push the science before relying on whatever other resources lie to hand.

I always found surfactant science hard, and regularly skipped over stuff I thought was useless or irrelevant. My aim in this book is to give you the science you really *must* know, and give it in the easiest-possible manner. That means a book/app hybrid. Every key equation in the book is linked to an app that brings the equation to life. Because this is an eBook-only publication, readers simply have to click on the link in whatever eBook Reader they are using (e.g. iBooks, Kindle, Acrobat pdf) and the app will open in their default web browser, ready

for immediate exploration. The inputs are invariably via sliders. Those on PCs and Macs might prefer inputs via text boxes. Those on tablets and phones would hate text inputs – they simply aren't practical when (as is often the case) there are multiple inputs to be provided.

What topics should be included? I am still amazed by the way that "useless" knowledge from one subject area turns out to be useful in another area. I have therefore included chapters on any area where I happened to find some useful surfactant science. Even if you have no interest in, say, foams or emulsion polymerisation there is still a lot to learn that might well apply to your own area of interest. Although, for example, it might be obvious to foam scientists that scientifically a foam is very much the same as an emulsion, it certainly wasn't obvious to me. So ideas in foams carry over to emulsions and vice versa. Ideas that seem to apply only to making fancy microemulsions turn out to be surprisingly useful for those designing laundry detergents – a fact that very much surprised me.

On the other hand there seems little point in going over familiar ground such as lists of surfactant types. If you don't already know that there are anionic, cationic and non-ionic surfactants then this isn't the book for you. Instead you can read all about them in Wikipedia.

## Units

First, I won't be using any US units. It is shocking that these still exist in the scientific literature in the 21st century.

Second, because logarithms also cause confusion, in this book log() means $\log_{10}$ and ln() is natural logarithm.

I *should* be consistent in using either dyne/cm or mN/m but there are times when one is more natural than the other so I have not been rigorous in enforcing a single standard.

## My heroes

It will be impossible to disguise my bias in favour of certain surfactancy schools. Everyone is happy to acknowledge Griffin and his noble attempt to bring order to the (then) chaos of ethoxylates via HLB. Rosen provided a wonderful blend of pragmatism and relevant theory that suits my purposes very well[1]. Salager's early formulation of what is now called HLD should have changed things for the better in the late 70s but for reasons unknown to me it was only in more recent years that HLD has started to have an impact on the surfactant community.

---

1    Throughout this book references to "Rosen" mean his excellent book Milton J. Rosen, Joy T. Kunjappu, Surfactants and Interfacial Phenomena, 4th Edition, Wiley, 2012

Those who built on Salager's work such as Aubry and Sabatini have advanced the cause of rational formulation enormously and Acosta's NAC extension of HLD is behind much of the science and many of the apps in this book. The Denkov and Saint Jalmes groups have been my main source of inspiration on foam science and the Sofia surfactant school in general have furnished me with apps in many areas. Although there are few references directly to the Strey group, the influence of that formidable team are everywhere throughout the modern surfactant world and I am grateful to the Stubenrauch group who have built on that tradition. One Strey paper, in particular, changed my life because it suddenly made clear to me what HLD theory was all about – even though Strey was not writing in HLD terms.

Like most people, I find phase diagrams, and in particular, ternary phase diagrams, very hard to understand. I therefore decided to write a single app to help myself and others to understand them. A casual mention of this over lunch started a fruitful collaboration with Seth Lindberg of Procter & Gamble who not only is a phase diagram expert but is also a brilliant graphic artist. Together we wrote the various Phase Diagram Explorers and the graphics/icons used in those apps were all created by Seth. He has kindly provided more illustrations for this book. While they are not individually acknowledged, you will have no trouble distinguishing between those I created and those created by Seth.

You will find that I am rather critical of surfactant suppliers for failing to provide us with key useful data on their surfactants. If they gave us Cc (a term from HLD theory) we would already be well on the way to making a rational choice, and instead of (just) CMC they should provide the much more important $\Gamma_m$ and K (or a) values as discussed later. One notable exception has been Sasol and, in particular, Dr Charles Hammond who then worked for Sasol. Their provision of key HLD data in the field of extended oil recovery (EOR) has been a good example to other suppliers and has, in part, earned Dr Hammond the 2015 Samuel Rosen Award from AOCS.

Although I would normally offer specific thanks to named individuals I don't wish to imply that they necessarily agree with my view of their work. In all cases I have thanked them personally and when I have benefitted from their specific input to an app that help is acknowledged in the relevant place.

I do wish to point out, however, that I am hugely grateful that so many experts in surfactant science were so generous in their assistance.

The book is deliberately "reference lite". Although I have been through many hundreds of papers in writing this book, relatively few seem to me to be important enough to merit a reference. Most of us don't have the time (or access to the journals) to follow up on references, so why include lots of them? Generally, if a paper is referenced it means that I think you will get a lot out of

reading it. So another way of identifying my heroes is by checking the names of the lead authors on the cited papers.

**Emulsion naming**

It is unfortunate that the word "microemulsion" was used for the thermodynamically stable emulsions that are in the 10-30nm size range and should have been called "nanoemulsions". It is equally unfortunate that the word "nanoemulsion" now exists to describe conventional (thermodynamically unstable) emulsions that are in the "nano" size range, where that word can mean "sub-micron" or "genuinely nano" (i.e. sub 100nm) or just "nano is exciting so let's call it nano". Then there is the name "miniemulsion" which often means nanoemulsions. The doubly-regrettable name "PIT-emulsion" is confusing because it can mean "a normal emulsion prepared via the PIT (Phase Inversion Temperature) method", as is common in the cosmetics industry, or can mean a nanoemulsion produced using the same technique but with more energy and attention to detail. It is a doubly-regrettable name because, as we will see, there is nothing special about a "PIT" emulsion – there are other Phase Inversion Formulation (PIF) techniques that can be used to achieve exactly the same thing.

And what do we call an emulsion? We cannot just use the bare word because all of the above are emulsions. "Standard emulsion" or "Classic emulsion" might be used, but in the end I am forced to use "macroemulsion". If the context makes it unambiguous then the "macro" can be dropped.

Some attempts have been made to define emulsions according to colour. Macroemulsions are generally white because the drop sizes are large enough to cause scattering. Microemulsions can be transparent, but can also have a bluish scattering tinge to them. Small-particle nanoemulsions can vary from transparent to bluish, but if "nano" means "sub-micron" then "nanoemulsions" have plenty of scattering. The problem here is that a clever formulator can match the refractive index of the oil and water phases so that even a macroemulsion will show little scattering, and a nanoemulsions made with a big refractive index difference will be strongly scattering.

Adding further to the confusion is that someone will say "My emulsion has drops of size X" without saying if X is radius or diameter. Even if X is stated explicitly as a diameter, it is frequently unspecified whether this is number average, area average or volume average. Depending on the size distribution, if someone states a number average and someone else thinks they mean volume average then they might be thinking about entirely different emulsions. For those not familiar with this issue, an app is later provided.

A reader has argued that a micro'emulsion' isn't any kind of emulsion but instead is a solution. But then we start to argue what a "solution" is and that is a topic for my next book.

In the end we just have to accept that the situation is a mess, accept that microemulsions are nano-sized and thermodynamically stable, that nanoemulsions are small emulsions that are only kinetically stable, that some people call nanoemulsions miniemulsions, and that macroemulsions are thermodynamically unstable and generally are white because of scattering.

## Interactions

As scientists we should not be afraid of formulae. But most of us are. This book tries to persuade you to understand formulae by bringing them to life as apps. Although in principle the apps can work on the page of a modern eBook, my efforts to make this happen have failed. Instead, for each relevant idea and formula there is a link to click on that takes you via your default browser straight to the app. You can thus swap between text and app easily, allowing you interactively to understand the science. The apps are all HTML5/CSS3/ Javascript which means they are safe, are acceptable within corporate firewalls and, to the best of my ability, run on all platforms: smartphones; tablets; computers; iOS; Android; Mac; PC. I have written all the apps myself and the act of writing them was a great way for me to better understand the science, or, to put it another way, to overcome my own fear of formulae.

The book is interactive in another sense. As this book is entirely free (no hidden adverts, you don't end up on a marketing email list) and electronic, I can update it any time I like. This means that the edition you have might differ from the current edition. No problem – just download the current version.

The advantage of all this is that I can respond to critical comments from the readers. I thrive on dialogue and although I like being right, being proven wrong is more exciting because then I'm learning new stuff. So don't hesitate to comment on faults or omissions. If I don't have your favourite topic, or if I deal with it too lightly, just let me know, point out the sorts of things I should be including and I'll give it a go. If you think I am wrong, just say so and give me some reasons or academic references to change what I have written.

Similarly, if there are bugs or issues in any of the apps, let me know and I will be delighted to fix the problems.

The layout of the book was entrusted to the expert hands of Sean Cooper of Three&me who compressed about a week of layout work into a few hours of genius.

Finally, in terms of the apps I need to warmly thank Mark Abbott of Fingertip Scientific Ltd for dragging my apps into the world of Responsive Design so that they work across all platforms and for solving so many of the maddening problems that make cross-platform functionality so hard to get right.

**Update for v1.0.3**

I have been astonished by the reception given to the eBook and the large number of downloads. Especially gratifying has been the number of emails expressing relief and delight that someone at last has stood up and accused HLB of not being fit for purpose.

So far I have not had to retract or revise anything, though I made a few delicate adjustments in response to some sophisticated feedback.

Excitingly, evidence is growing that the formulation community is starting to seriously challenge the suppliers to give them the Cc and EACN data they need for rational formulation.

Thank you to everyone for their feedback. Feel free to send me yours!

## Abbreviations and Symbols

A      Head area

l      Tail length

CMC    Critical Micelle Concentration

$\Gamma_m$    Surface Excess, i.e. Surface Concentration in mol/m²

K, a   The air/water partition coefficient of a surfactant (K) or its inverse (a=1/K)

HLB    Hydrophilic-Lipophilic Balance

HLD    Hydrophilic-Lipophilic Difference

NAC    Net Average Curvature

CPP    Critical Packing Parameter

PP     Packing Parameter

EO     Ethylene Oxide, i.e. $-CH_2CH_2O-$ "ethoxylate" groups

PO     Propylene Oxide, i.e. $-CH_2CHMeO-$ "propoxylate" groups

O/W    Oil in Water emulsion

W/O    Water in Oil emulsion

PIT    Phase Inversion Temperature

PIF    Phase Inversion Formulation

APG    Alkyl Poly Glycoside surfactants – alkyl substituted sugars

kT     Boltzmann's Constant times Absolute Temperature – thermal fluctuation energy

RT     Universal Gas Constant times Absolute Temperature – bulk thermal energy

# 1    Some Basics

| Lost in surfactant space? | Too little useful information? | Sufactancy is a balance |
|---|---|---|
| **HLD=S-k.EACN-αΔT+Cc** | **w/o ~ o/w**<br>**HLD > 0   HLD = 0   HLD < 0** | **Net Average Curvature Predictive Power** |
| Key concepts explained | Ternary diagrams explained | Foam Science Too |
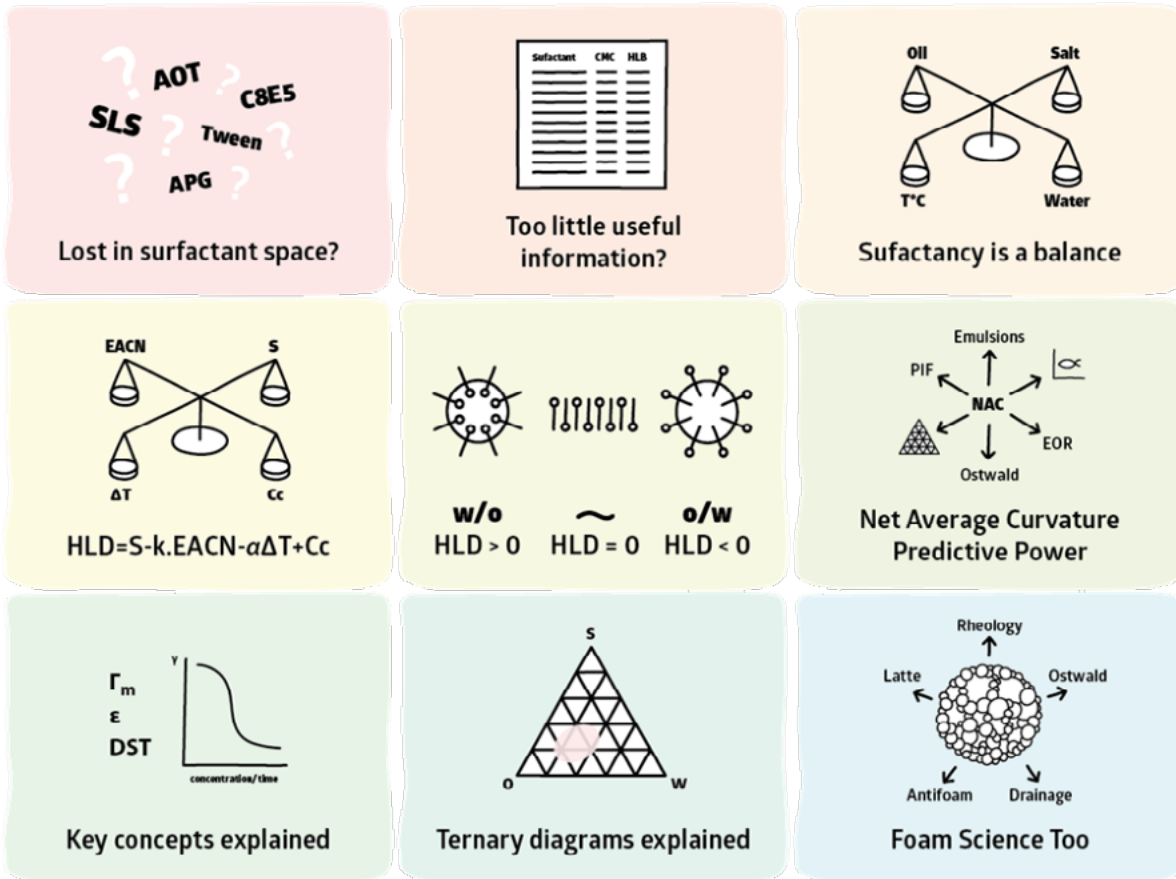
## 1.1    A summary in 9 images

The whole of the book can be described in 9 images.

- Many of us find ourselves lost in surfactant space. There are so many surfactants to choose from, so how does one choose rationally?

- Any "Handbook" of surfactants will tend to list 1000's of surfactants and might provide CMC and HLB values with general terms such as "a good emulsifier" which are usually unhelpful.

- The fact is that surfactancy is always a balance between oil, water, salinity and surfactant. Something like HLB is worse than useless because it tries to describe the balance of a system as a number for the surfactant. This glaring defect is so obvious that HLB should have imploded decades ago. But it hasn't.

- Fortunately the more sophisticated HLD system does a good job of describing the balance as the key equation contains all four components: oil, salinity, temperature and surfactant.

- The numbers from HLD tell you immediately if the system is in the oil in water (O/W), water in oil (W/O) or an intermediate state where many interesting things happen.

- When you start to use HLD-NAC you add the power of understanding the curvature which is so important in so much of surfactancy.

- The plot of surface tension with surfactant concentration which gives us the Critical Micelle Concentration, CMC, which is surprisingly unhelpful, actually gives us values such as $\Gamma_m$ and a which are much more useful but largely ignored. When you also have a feel for elasticity and dynamic surface tension effects then things start to make a lot more sense.

- Because surfactant space is complicated we often need to resort to phase diagrams. I was never able to grasp ternary diagrams so I was delighted to work with a phase diagram expert to create a suite of phase diagram apps that bring all the key concepts to life.

- Finally, much of foam science can be better understood via apps covering the key aspects.

Now you know what the book is about we can get started.

## 1.2    What's not included

There are plenty of surfactant books with descriptions of the different types of surfactants followed by long tables of values such as CMC and, perhaps, HLB. The assumption here is that you already know that the most-used surfactants are anionics such as SLS or nonionics such as C8EO6, and that cationics can have some unpleasant effects when they contact human skin or get swallowed. In addition to such information being readily available elsewhere (the surfactants entry on Wikipedia is as good as many books in terms of describing the different surfactants), the astonishing fact is that for most of us the information is nearly useless. Few formulators have derived much value by comparing CMC values of various ethoxylates or anionics. If the CMC of a surfactant is $10^{-5}$m, that means, for a MWt of 400, 4mg/l, yet we will typically be using 10g/l, i.e. 1% in a formulation. The CMC of typical anionics like SLS are 2-3 orders of magnitude higher than typical ethoxylates, yet that doesn't mean that anionics are 2-3 orders of magnitude less used. In any event, CMC is a measurement that relates only to surfactant-water interactions and for many formulators that is of no interest because our interest is in oil-surfactant-water interactions. Even "advanced" ideas such as CPP (Critical Packing Parameter) are at the same time wildly wrong and largely useless.

The reason for writing this book is that the surfactants community has focussed too much on the wrong things and too little on the things that we can use to create better surfactant-based formulations. When I needed to formulate a simple microemulsion for a cosmetics application I was astonished to find so little useful information to help me formulate. That's what led, eventually, to this book.

So the deal with the reader is that any surfactant science described here comes with a description of why the science will be useful. No formulator will need all the science described here, though it is generally a good idea to have at least a passing knowledge of all the main topics. But all the science will be of practical use to some formulators in the surfactant community. I know this because I have directly used, or known those who have used, every bit of science included in the book. However, if you really want a summary of surfactant types I have one at https://www.stevenabbott.co.uk/practical-surfactants/Surfactant-Types.php and if you want a list of CMCs I have one at https://www.stevenabbott.co.uk/practical-surfactants/CMC-Values.php.

## 1.3     CMC, Γm, A, K, E. Who cares?

As we all know, if one starts with the surface tension of pure water and slowly adds drops of surfactant, the surface tension decreases till it reaches a minimum value. The concentration of surfactant where that minimum is attained is the point where extra surfactant ends up as micelles in the water rather than doing any further reduction in surface tension. We know that this value depends on surfactant chain length, and that for ionics it is a higher value because the charged heads tend to repel each other, and that added salts decrease the CMC because the reduce the head-to-head repulsion. Discussions on these matters tend to fill many pages, despite the fact that they are largely irrelevant. Although CMC is not all that useful or very profound, especially for those of us (the majority) who are more interested in oil/water interfaces than in air/water interfaces, the process of measuring CMC gives us another set of values, $\Gamma_m$, A, K and E which are often of great value.

The real interest to formulators comes from the curves of surface tension with concentration. These curves are called "isotherms" and can be fitted to various equations. What is important to those of us who are interested in the air/water interface (e.g. for foams) is that these curves provide some important insights into what the surfactant is doing at the air/water interface. Even for those interested in oil/water interfaces the basic numbers are useful. Different surfactant systems might show different types of curves that might be of interest to academics. For our purposes we can assume that the curve is a Langmuir-Szyszkowski isotherm. This says that if the starting surface tension is $\gamma_0$ then at a concentration of surfactant c, and with two constants K (the absorption coefficient) and $\Gamma_m$ (limiting surface concentration) the surface tension $\gamma$ is given by:

$$\gamma = \gamma_0 - 2.3RT\Gamma_m \log(1 + Kc) \qquad \text{1-1}$$

Others prefer to work in terms of the inverse absorption coefficient, a, which is 1/K – there is an approximately 50:50 split in the literature. K has units of l/mol and a is the concentration (mol/l) at which half the surface is covered by surfactant.

As mentioned in the Preface, the convention is that log() means $\log_{10}$ and ln is the natural logarithm. Isotherms follow natural logarithms and the 2.3 in the equation is from the conversion to $\log_{10}$.

To disentangle this equation and to get at the key constants we need to do a bit of work. The first thing is to note that $\gamma_0$-$\gamma$ is called the *surface pressure*, Π. Why do we need a term to describe the reduction in surface tension? Because it is making an important point. The surfactant isn't magically reducing surface tension simply by waggling some hydrophobic tails in the air. Those tails, packed increasingly tightly are exerting a pressure. It is this pressure that is reducing the surface tension. As soon as we introduce a word such as pressure we start to imply some effort. Those tails don't want to be near each other – they get in each other's way. And each tail comes with a head, and those heads might also not like being together. Indeed they exert an anti-pressure that would *increase* surface tension. The surfactant molecules are only at the interface because on balance the tails' interaction with water outweigh, up to a limit, the surface pressure.

The "concentration" at the surface, expressed as mol/m² (rather than the conventional solubility in solution which is in mol/m³) is, for reasons explained later, called "surface excess", and is given the symbol Γ. For each surfactant there is a maximum value of this surface excess, $\Gamma_m$, though the "m" is for monolayer, not maximum. Surprisingly, this value is not reached at the CMC but long before – again pointing out the CMC is not special. In general (see Rosen) it is reached when the surface pressure is 20 mN/m, i.e. when the surface tension is 52 mN/m.

What about the K in the isotherm equation? This is an absorption coefficient – the preference for the surfactant to be at the interface rather than in the water. The higher K, the lower the CMC, because the surface will reach saturation faster. Because CMC is a concept that is easier to relate to than K (though we will return to K as it is very useful), and because we know the equilibrium value of the surface tension, $\gamma_c$, it is possible to define some values via the knowledge that $\Gamma_m$ is attained at 52 mN/m, i.e. 0.052N/m:

$$\Gamma_m = \frac{1}{2.3RT} \frac{0.052 - \gamma_c}{\log(CMC) - \log\left(\dfrac{CMC}{10}\right)} \quad \text{1-2}$$

$$x = \frac{0.072 - \gamma_c}{2.3RT\Gamma_m} \quad \text{1-3}$$

$$a = \frac{CMC}{10^x - 1} \quad \text{1-4}$$

From these we can now plot the isotherm of γ versus concentration using the re-statement of the Langmuir-Szyszkowski isotherm in terms of a (as discussed, a=1/K) and c:

$$\gamma = 0.072 - 2.3RT\Gamma_m \log\left(1 + \frac{c}{a}\right) \qquad \text{1-5}$$

Finally we can calculate the important Area per Surfactant Molecule, A, using the simple fact that one mole of surfactant contains Avogadro's number (6e23) of molecules:

$$A = \frac{1}{6e23\Gamma_m} \qquad \text{1-6}$$

All this has been done to introduce the first app, though for reasons described below it is slightly different from what has just been described:



**App 1-1** https://www.stevenabbott.co.uk/practical-surfactants/CMC.php
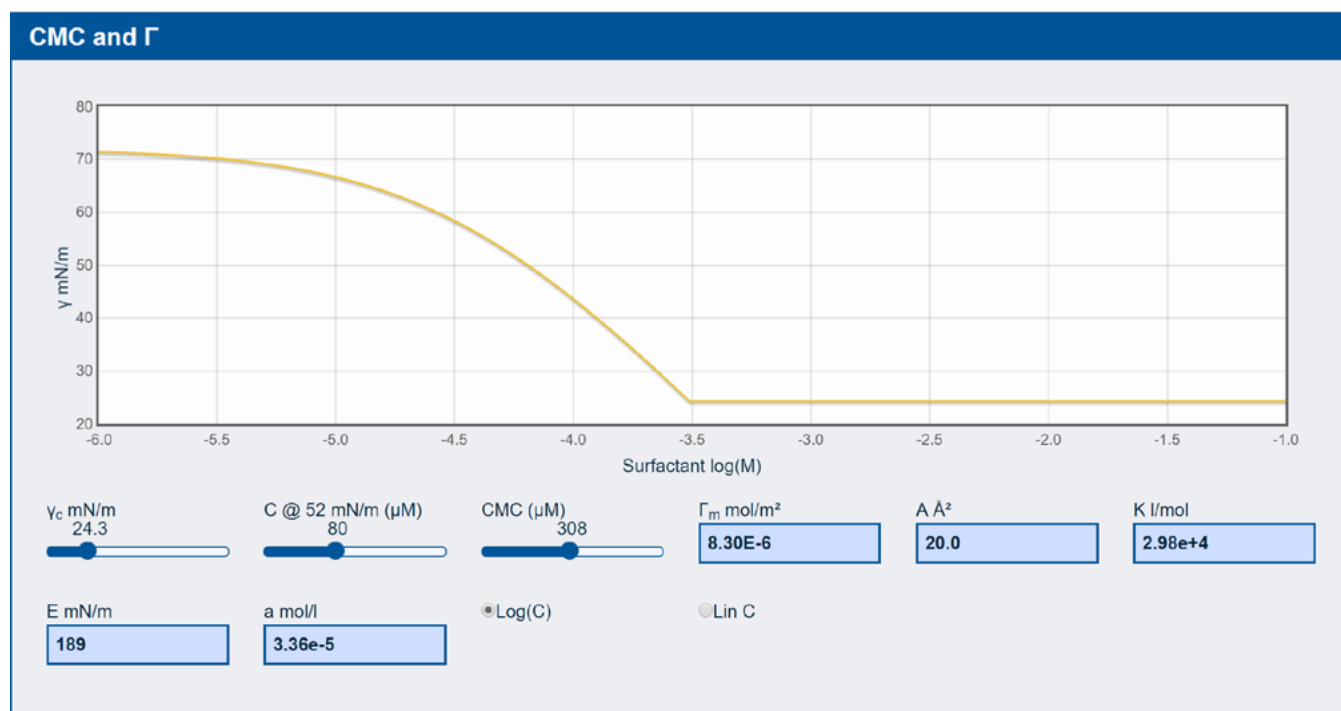
The first thing to do is to recognise that, like the other apps in this book, all you have to do is click on the link and it is there in your default browser ready for you to play with. The second thing to do is to start playing with it, getting some idea of what happens when input values are changed. For example, you might not know what Log(C) and Lin C mean (though a tooltip appears to explain), so just click on them and it will become obvious. The third thing is to go between the app and the theory to make sense of both.

In this case the conclusions are very simple:

- You can ignore most of the equations – they were just there to reach the desired destination.

- The CMC has no effect (within the approximations behind the app) on either $\Gamma_m$ or A.

- For a very low equilibrium surface tension, a very tightly packed surfactant with a very small A is required.

- K is also calculated – a high CMC and low $\gamma_c$ are equivalent to a (comparatively) low K. Its inverse, a = 1/K, is also shown because the surfactants world is almost equally divided between those who use K and those who use a.

- The elasticity, E, is discussed later.

The app is a simplification. In reality the experimental curve is fitted to the appropriate isotherm and the parameters derived from the fitted curve. Nonetheless it is a rather robust simplification, though it seems mysterious that $\Gamma_m$ is reached when the surface pressure has risen to 20 mN/m. If the maximum surface excess is reached long before the CMC, what is happening as more surfactant is added? The answer is that the surfactant appears as rafts of close-packed areas and added surfactant merely increases the area of those rafts. Simple calculations from Rosen show that when the surface pressure has risen by 20 mN/m the surface is 84-99.9% saturated and the remaining packing out of the surface is a relatively minor effect. So the CMC is of no significance in terms of surfactant packing (which is why it has no effect on $\Gamma_m$ and A). Instead, CMC governs the area of those rafts, by controlling the partition coefficient not between water and surface (which is what many people think) but between monomers and micelles. To a first approximation, for a given length of ethoxylate chain, simple ethoxylates have the same $\Gamma_m$ and A values. As the length of the alkyl chain increases the CMC goes down because the rafts get larger at lower concentrations of the surfactant. CMC, therefore, is about (micelle) partition and not packing, while $\gamma_c$ is about packing at the interface and not about partition. As we shall shortly see, the fact that the portioning to the interface (K) mirrors CMC does *not* imply that CMC is a surface phenomenon.

## CMC at 52 mN/m

An email from surfactant expert Sanja Natali alerted me to an improved way to get more realistic values. The key equation used CMC/10 as an automatic reference point. It is better to include a user-based parameter CMC at 52 mN/m to control the curve. The significance of 52 is that it is how $\Gamma_m$ is defined within the curve. The key benefit is that Area values are larger and more realistic.

All this has been a way to alert you that CMC is, in many ways, far less interesting than $\Gamma_m$ $\gamma_c$ and K. Because CMC is so easy to measure and looks so profound it has taken on a far greater symbolic importance than it should.

CMC's importance can further be demoted by comparing the conventional log plot with the unfamiliar linear plot (select the Lin option). The CMC in each case is 1000µM:
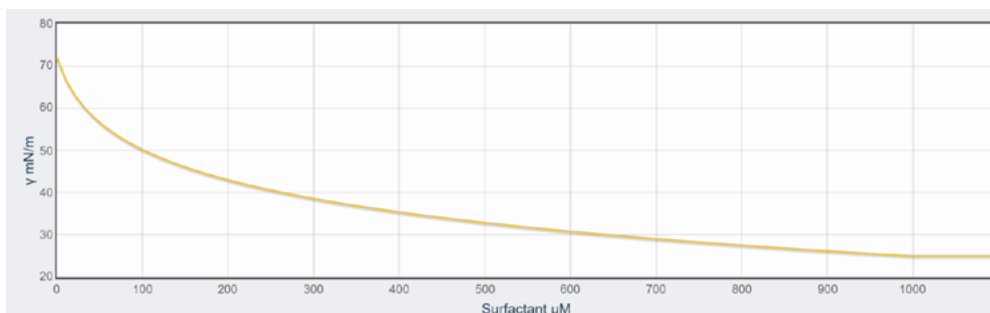


**Figure 1-1 The CMC looks less important in the linear plot of the same data**

Instead, the hitherto mysterious importance of the 20 mN/m drop now takes on more significance. Most of the "CMC" action takes place near the start (as mentioned above, the surface is >80% saturated at this stage) and the CMC itself is merely the tail end of the process. This is further emphasised by going back to the definition of "a" in the isotherm which is the concentration where half the surface is covered with surfactant you can check using the mouse on the graph that this typically takes place when the surface tension has fallen from 72 to 66 mN/m.

The formation of micelles is also no big deal. Micelles are a statistical construct, fleeting assemblies of molecules. The clear images from countless portrayals of micelles give them a solidity they do not deserve. When we come to discuss dynamic surface tension we will find that to everyone's surprise, micelles are of no significance. Remember, too, that in general any surfactant in a micelle is a waste of surfactant. Most of the time we want surfactants to be at an interface, not in their own little world. The exception is when we want micelles to help carry other water-insoluble materials, an important yet still narrow application area.

Finally we need to get rid of the idea that the CMC kicks in when the surface no longer absorbs additional surfactant. This is clearly wrong – the CMC is a bulk, internal, thermodynamic phenomenon which can be measured (in theory) within a large spherical vessel with a surface that attracts no surfactant (i.e. the interfacial tension is already 0). The measurement can be via one of the many other properties such as conductivity, self-diffusion coefficient or turbidity so CMC knows nothing of surfaces. The equilibrium coverage at the surface depends on the separate partition coefficient of monomers into micelles, with a

smaller dependence on the number of monomers within a micelle, the micelle aggregation number. This all makes sense following the approach of Durbut.[2]

When there is free surface there are two possible equilibria. The first is the adsorption at the surface, given by the Langmuir equation where m is the concentration of free surfactant molecules, the monomers, and A (in this section) is the concentration of absorption sites. Although the $K_{ads}$ discussed here is essentially the same as K discussed above, it is given the subscript to distinguish it from $K_M$ which is the micellar equilibrium:

$$K_{ads} = \frac{(m - A)}{m.A} \qquad 1\text{-}7$$

The second equilibrium is the one that controls CMC and which does not involve the surface. It is that between surfactant monomers, m and micelles M (containing n monomers per micelle) with an equilibrium constant $K_M$ given by:

$$K_M = \frac{M}{m^n} \qquad 1\text{-}8$$

For simplicity, $K_M$ is kept as 1 (so 1 needs to be multiplied by the CMC in whichever units you choose). With the constraint that the total surfactant concentration, S, is given by:

$$S = m + nK_M m^n \qquad 1\text{-}9$$

it is then possible to see what happens as $K_{ads}$ and n change.



**App 1-2** https://www.stevenabbott.co.uk/practical-surfactants/CMCSat.php

---

2    Patrick Durbut, Surface Activity, Chapter 3 in Handbook of Detergents, Part A,Surfactant Science Series 82, ed. Guy Broze, Marcel Decker, 1999
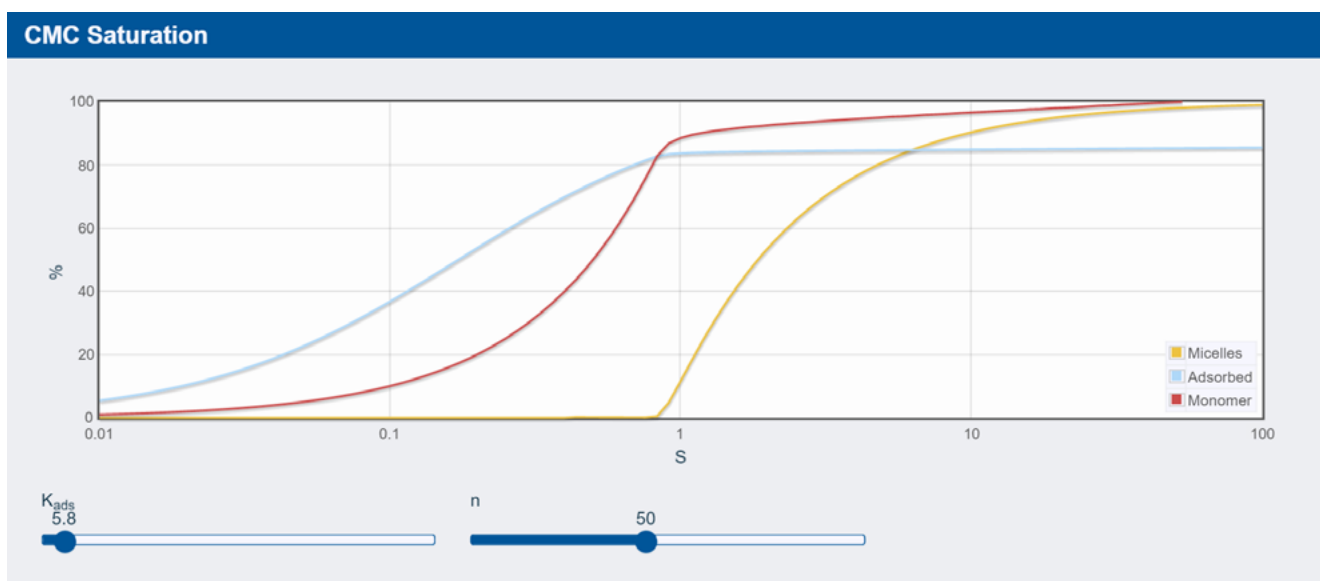
The effect of n, the micelle aggregation number, is to sharpen the transitions because the effect on the overall monomer concentration of 100 monomers all disappearing into one micelle are stronger than the effects of 25 monomers disappearing into a micelle. However, the effects of n are modest once n>20.

The effect of $K_{ads}$ is more interesting in this context. With low $K_{ads}$ the supply of monomer is so limited above CMC that the surface only reaches, say, 50% saturation. To those looking at a plot of surface tension versus concentration, there is no hint that the minimum surface tension is not at all associated with the surface being filled with surfactant monomers. With high $K_{ads}$ the saturation of the surface is closer to 100%. So saturation is governed *not* by CMC but by $K_{ads}$. As mentioned earlier, typical surfactants (which tend to be reasonably good as the bad ones are generally not studied) tend to saturate at anywhere from 84 to 99.9% coverage. All this is another way of saying that our focus should not be on CMC but $\Gamma m$, $\gamma c$ and K or a.

Having put CMC into context, it needs to be pointed out that $\Gamma_m$ itself is not all that exciting. Over the classic range of surfactants it varies (as Rosen is keen to point out) from $1-5 \times 10^{-6}$ mol/m² and A is in the 30-40Å² range. They become interesting when their values are exceptional. Polymeric surfactants, for example, are a symbol of potency because their CMC values can be phenomenally low – which makes them sound like super-surfactants. In fact in many ways they are super-useless. CMC is not directly related to $\gamma_c$, and many polymeric surfactants have high $\gamma_c$ values so do not do a great job in reducing surface tension. From the simple theory this means that A is in the 80+ range – there isn't (not surprisingly) a tight packing of polymeric surfactants at the interface. As we will see, their kinetics are amazingly slow which means that they are useless for many applications and especially useful in laundry detergents because any soil that is first cleared by conventional surfactants and then becomes engulfed (slowly) by the polymeric surfactants is unlikely to get re-deposited on the laundry.

This means that the most interesting factor is K and it becomes especially interesting in the context of dynamic surface tension, DST. There are two problems with K. The first is its units. $5.10^4$ l/mol can also be shown as 50 m³/mol or $5.10^7$ cm³/mol. The second is that its inverse, a =1/K is often used instead and can vary between units of mol/cm³ to mol/m³ via mol/l or simply M for Molar or μM for micromolar. Comparing surfactant datasheets and academic papers is made exceedingly difficult by the need to swap between different ways of expressing the same idea.

There is some other important basic science about surfactant films that must wait while we dispose of a few other much-quoted items that happen also to be unimportant. We also need to note a much more interesting, but much less well-known number, CμC, the critical microemulsion concentration which will also be discussed later.

## 1.4    Why CPP is mostly useless

The idea of micelles exerts a curious hold over how surfactants are discussed. One such idea is that of the Critical Packing Parameter, CPP, which can provide some modest guidance into how high concentrations of surfactants behave in water. This perfectly reasonable insight has somehow become hijacked as an attempt to explain how to optimise emulsions. It should be obvious that a parameter which is based only on considerations of water/surfactant systems has no necessary link to water/surfactant/oil systems. And, indeed, CPP has no such link. As we will see, PP (Packing Parameters) at oil/water interfaces are of huge importance, though it turns out that the ideas behind PP can be better expressed in the complementary methodology of HLD theory which is discussed in the next chapter.

CPP is a dimensionless number that is a ratio of V, the volume of the tail of the surfactant and the product of the head area A and the tail length l. So

$$CPP = \frac{V}{A.l} \qquad 1\text{-}10$$

The app brings the formula to life. As you slide the controls to change V, A and l, the graphic gives a general view of the balance of the three parameters and an idea of how the curvature of an assembly of surfactant might change. This view has a CPP<0.33 and the surfactant will form a classic micelle shape:



App 1-3  https://www.stevenabbott.co.uk/practical-surfactants/CPP.php
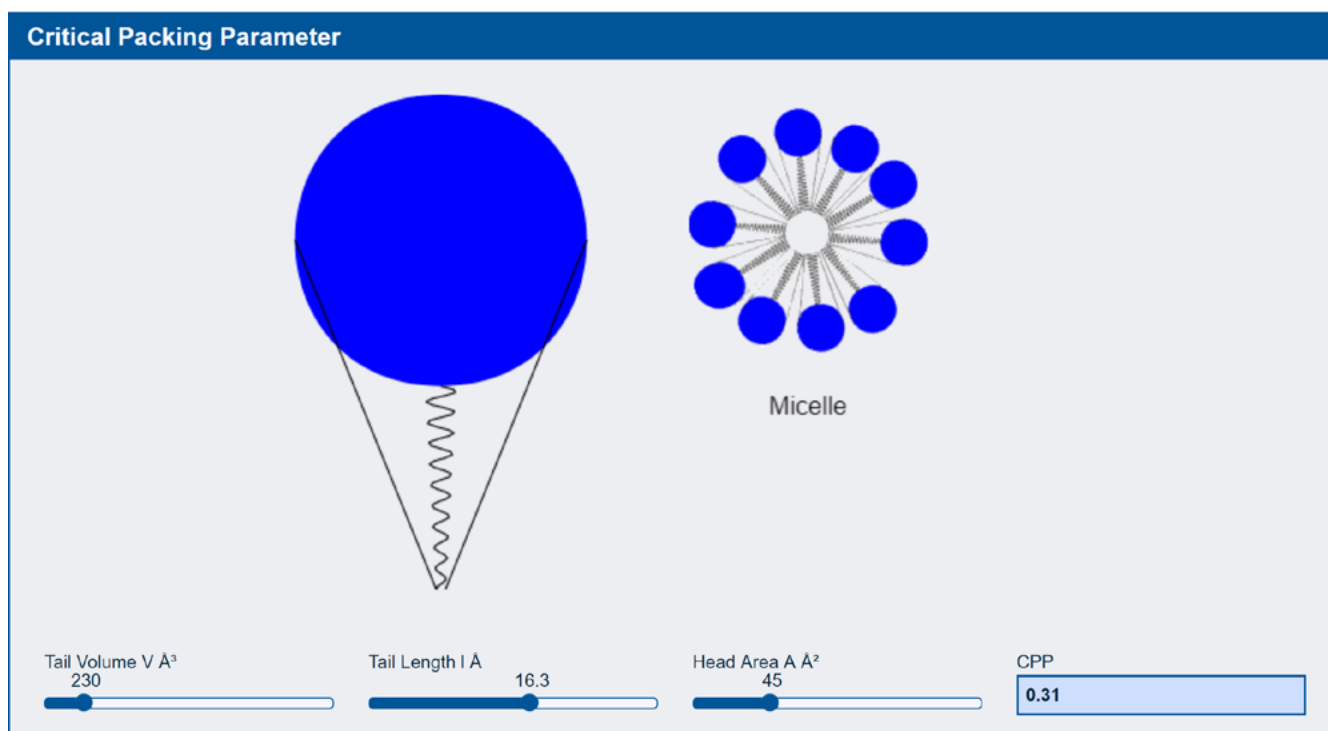
The figure shows what happens as CPP is steadily raised. The natural curvature first encourages hexagonal, then lamellar and then there is a flip into inverted phases with the head inside and the tail outside.
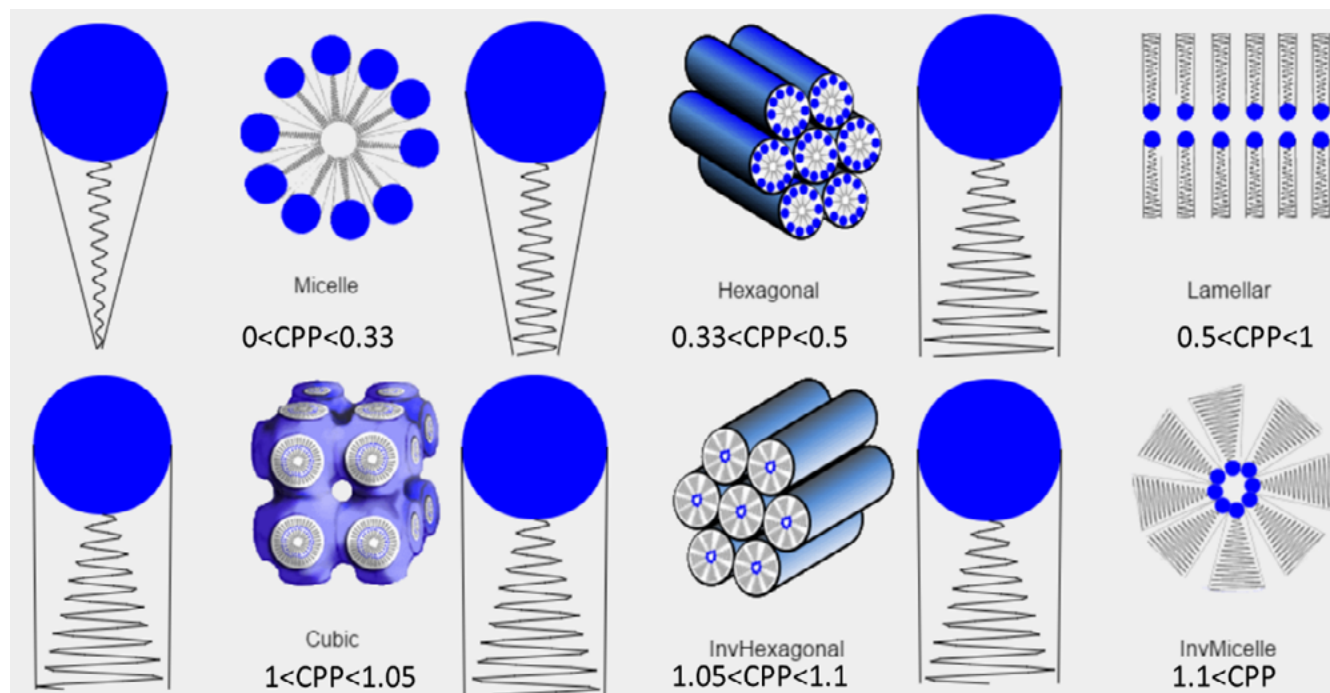


**Figure 1-2 The standard story behind CPP.**

For those who have a special interest in concentrated phases of surfactants in water (lyotropic liquid crystals), the CPP story is of some relevance. Surfactants with very large heads tend to form micelles and those with very large tails tend to form inverted micelles, and those with a balanced head and tail tend to form lamellae. Unfortunately this generality is of little help because the same surfactant at different concentrations will form different phases across this spectra. CPP is about tendencies, not final outcomes.

A nice story about concentration dependence can be produced if another form of packing is introduced. As soon as micelles start to bump into each other they can start to override the intrinsic curvature of CPP. The much-cited paper from the Tiddy group at Unilever[3] paints a beautiful picture of how the systems might evolve with concentration.

3    D. John Mitchell, Gordonj. T. Tiddy et al, Phase Behaviour of Polyoxyethylene Surfactants with Water, J. Chem. SOC., Faraday Trans. I , 1983, 79, 975-1000
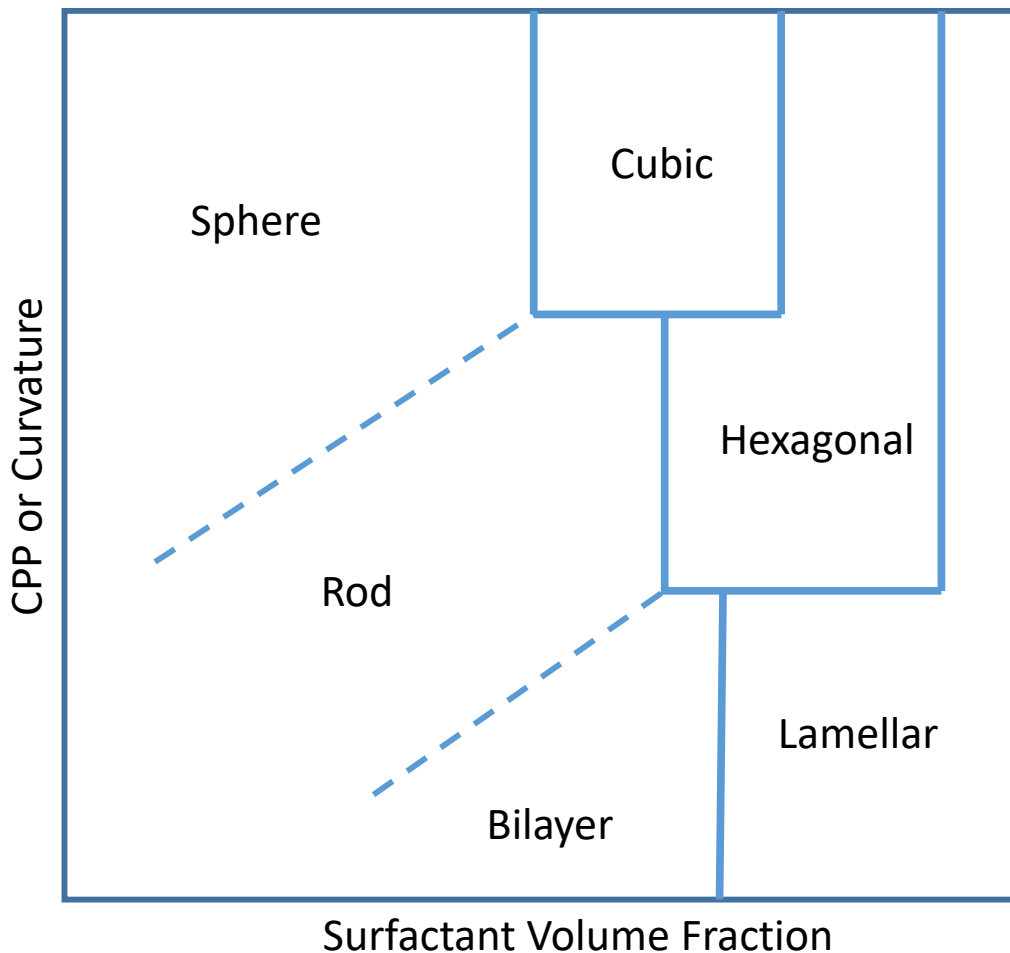
**Figure 1-3 How CPP copes with packing through increased volume fraction.**

It is a nice image. The main thing wrong with it is that the real world does not conform well to it, as shown by their analysis of many phases across $C_xEO_Y$. The Discussion part of the paper starts: "In this section we review the agreement (or lack of it) between the theoretical ideas and the experimental phase diagrams." There is indeed some agreement, which is encouraging, along with plenty of oddities which cannot readily be explained. The general idea is that increased volume fraction creates compression so that spheres naturally start to pack into cubic structures, rods naturally start to pack into hexagonal structures and bilayers pack into lamellae. Missing from this simple story are phases such as bicontinuous cubic which can form from many precursor shapes and can be transitional from, say, hexagonal to lamellar.

In any case, it is not clear that these guidelines are of much help as they are *only* about pure surfactant/water mixtures. It tells us almost nothing about oil/surfactant/water interfaces which are often our key focus of interest. To put it bluntly, you cannot formulate with CPP. It is a beautiful intellectual construct which offers some insight into liquid crystal phases and, at best, some broad trends in emulsion formation. Most formulators, most of the time, should completely ignore CPP because the key information can be obtained by more direct routes that allow rational formulation.

By all means think in terms of Packing Parameters, which will be discussed in detail. Just enjoy the pretty pictures from CPP and move on.

In the early stage of development of HLD theory, which is really about the curvature of the oil/water/surfactant interface, a term was needed for the surfactant-specific part of the theory. It seemed a good idea to call it Cc, Characteristic Curvature with some reference to the sorts of natural curvatures seen in CPP. It was later realised that this is not a helpful idea, but the name Cc has stuck.

## 1.5    Bancroft's "Rule"

An observation by Bancroft in the early 1900's has been enshrined as a rule: "*The phase in which an emulsifier is more soluble constitutes the continuous phase*." So hydrophilic surfactants tend to give O/W emulsions and hydrophobic surfactants tend to give W/O emulsions. Although it is generally true it is often unreliable and of little help in practice because if you want an O/W emulsion and if (as there are) there are 1000's of hydrophilic surfactants to choose from, nothing in the "rule" tells you how to make a more intelligent choice.

There are many ways to explain why Bancroft's rule is generally correct. My favourite is found in an excellent review[4] by Danov and colleagues at U Sofia which is not for the faint-hearted but contains a lot of deep insights. It starts with a 50:50 O:W mix and the idea that when energy is put into the mix it creates (assuming that viscosities are comparable) equal numbers of water drops and oil drops. Suppose that the surfactant is soluble in the water phase. Oil drops will naturally find their surfaces supplied with surfactant and can start to form a classic shell of surfactants with heads pointing outwards. The water drops equally form a classic shell with tails pointing outwards. At this stage there is no obvious difference between these two conditions. All questions of packing parameter etc. are totally irrelevant – both types of drops have a surfactant shell.

Now let the oil drops approach each other. As we will discuss when we get to the section on elasticity, the shells can be considered rather robust. Our instinct is that this robustness might stop the drops from coalescing when they meet and this is, of course, true. But if particles touch then they tend to stick together (flocculate) which then gives them a chance to cream (rise to the surface driven by gravity) or to coalesce at their leisure, so the shell on its own is not enough. What is needed is some mechanism to stop them touching in the first place. What is therefore of importance is that the speed at which (essentially) rigid drops can approach each other is greatly reduced by the need for the intervening liquid to move out of the way. The so-called Taylor velocities for rigid drops are greatly reduced as they get closer. This at least reduces the speed at

4    Krassimir D. Danov, Peter A. Kralchevsky, and Ivan B. Ivanov, *Dynamic Processes in Surfactant Stabilized Emulsions*, Chapter 26 in Encyclopedic Handbook of Emulsion Technology (J. Sjöblom, Ed.), Marcel Dekker, 2001

which they collide, but is not enough on its own. An extra mechanism is required so that the oil drops, on average, collide much less than the water drops.

Let the water drops approach each other. The surfactant shells act as though they aren't there at all! This can be shown theoretically and experimentally. Without going into theory, why is this the case? Suppose the surface of the drop starts to deform. This opens a gap in the surfactant layer, something that should be difficult to do. The difficulty is overcome because there is plenty of surfactant inside the drop (remember, it is the phase in which the surfactant is soluble) so as any gap appears a new surfactant molecule can take its place. The water drop, therefore, has no elasticity, it just conforms to whatever is happening. Taylor theory shows that such deformable drops can approach each other at much higher speed. The oil drops take on the full slowing down from the Taylor effect so collide far less often.

So, according to this way of explaining Bancroft, the reason for getting O/W emulsions for a water-soluble surfactant is that the water drops coalesce almost as soon as they are created and the oil drops coalesce (if at all) much more slowly. The argument is symmetric when the surfactant is oil soluble, so oil drops coalesce (because they have lots of surfactant to fill any gaps caused as a drop deforms) and the water drops survive.

This velocity-based argument has much to recommend it. One striking aspect of the theoretical predictions is that for relatively rigid, small drops, the velocity differences between oil drops and water drops decrease as the drops get smaller. So the "wrong" emulsion is more easily formed with smaller drops. The argument also contains its own counter-argument. Because it is a pure *kinetic* argument it can be overridden by *thermodynamics*. Or, in more common language, if, the balance of solubilities or elasticities shifts then the kinetic balance becomes less significant and other effects, such as curvature, start to dominate.

The argument is further weakened away from 50:50 O:W because the probabilities of oil drops coalescing at low % oil or water drops coalescing at low % water go down for purely statistical reasons so it is easier to create the "wrong" emulsion. Given that we tend to make, say, 20% O/W emulsions the Bancroft "rule" becomes less compelling.

Some variants on these ideas have been neatly captured in another review[5] about Bancroft. At very low concentrations of a surfactant that happens to prefer to be in the oil phase, you get anti-Bancroft O/W emulsions. The definition of "very low" is an interesting one. It means "below the CµC" the critical microemulsion concentration, a value that is typically 2-10x larger than the CMC. Once above the CµC the emulsion follows the "curvature" rules discussed

---

5    Eli Ruckenstein, Microemulsions, Macroemulsions, and the Bancroft Rule, Langmuir 1996, 12, 6351-6353

at length in the HLD chapter. As is often the case, CMC is not important and another factor, CµC is important. This is not surprising because the generally-quoted CMC knows nothing of the oil (though you can measure an oil-dependent CMC if you wish) while the CµC depends strongly on the oil for reasons that HLD theory makes clear. The review also points out that at super-high surfactant concentrations other effects (black films) dominate

As the Danov review emphasises, Bancroft is not a rule – it is one heuristic applied to an immensely complicated process.

There is much more to explore in terms of complications – with the promise that approaches more powerful than Bancroft will emerge.

## 1.6    Why HLB should be banned

Some people might be surprised by the title of this section. Hydrophilic-Lipophilic Balance, HLB, appears as a standard section in just about any presentation or course on surfactants. Nevertheless, it has done huge damage to the surfactants community. The damage is not because it is a particularly bad theory. The damage arose because it never moved on from its own limitations. It got stuck in a time warp and blocked the development of a much more fruitful approach that not only built on the foundations of HLB but has a similar name, HLD.

HLB arose with the development of the then new class of ethoxylate (EO) surfactants. Because they are based on a simple combination of hydrophobic tail and ethoxylate head, and because head and tail can be altered methodically, it was possible to create a systematic array of surfactants, ranging from those that were very hydrophilic (long EO chains, readily soluble in water) to those that were very hydrophobic (long hydrophobic tails, readily soluble in oils). This led Griffin to classify them based on a "neutral" value of 7 (presumably in analogy to the neutral pH 7 of water) with hydrophobic ones at lower values and hydrophilic ones at higher values. In effect it is a more sophisticated version of Bancroft.

At the time, this was hugely liberating. Using the known HLB values of the ethoxylate series it became possible (within limits discussed shortly) to do rational formulations. If a surfactant was too hydrophilic for the given oil, just throw in some more hydrophobic one and vice versa.

And this is when the trouble started. The HLB bandwagon was rolling so those who had no rational formulation scheme of their own attempted to use the prestige of HLB to create a good-enough approach. The central problem was that the HLB numbers were easy to derive for simple ethoxylates. A few, difficult and not-very-well-characterised techniques could pin down a few HLB numbers for ethoxylates and the rest could be generated via a simple formula. Such a formula simply did not apply to different classes of surfactants. To this day

papers are being published that purport to come up with a better scheme for calculating HLB values from surfactant structures. In practice such schemes are largely worthless.

The other problem with HLB is that it is a number given to a surfactant molecule. This is a nonsense. The surfactant is supposed to balance between water and oil, and different oils require a different balance. Hexane is different from hexadecane. So HLB *cannot* be ascribed to a surfactant. It gets worse. Different "waters" (i.e. water with different salt concentrations) behave differently. Worse still, temperature changes the oil/surfactant/water balance. And, finally, different surfactant classes have different temperature effects. As it happens, the ethoxylates get less water soluble as the temperature is increases, anionics get more soluble and "sugar" surfactants such as the APGs happen to have no significant temperature effect.

So even in terms of its own name, HLB, being a value for a surfactant, cannot do what it says it does – it cannot impose a hydrophilic lipophilic balance because the balance is a function of the system, not just of the surfactant. To confuse things even further, attempts have been made to give oils an HLB with the intention that the matching of surfactant and oil HLBs will give the desired outcome – which is generally a perfect emulsion. The famous ICI guide[6] provides a compromise for a given oil: the "Required HLB of O/W emulsions" and no doubt works as long as you stay with conventional ethoxylates, their oils and, bizarrely, room temperature where the data seem to apply. Why bizarrely? Because the ethoxylates attain a hydrophilic-lipophilic balance only at an elevated temperature, the phase inversion temperature. This erroneous approach is very popular in cosmetic emulsions and is taught on all reputable courses. The secret to this comparative success is the secret of most cosmetic formulations: basically they are all the same. So *if* the formulator is using conventional ethoxylates and *if* the formulator is using typical plant oils then the HLB technique is fine. As soon as some attempt at novelty is made with non-conventional surfactants or oils then the whole approach falls apart. It doesn't work for the simple reason that it cannot work. Hydrophilic lipophilic balance is a property of the system, whereas HLB is a property of the surfactant so *must* be wrong when salinity, oil and temperature change. The cosmetic industry is especially keen on using the phase inversion temperature, so the one time they really use a balance is at a temperature for which HLB is not specified.

It is useful to quote a higher authority, Rosen: "It has been pointed out [since 1968] that a single surfactant can produce either an O/W or a W/O emulsion, depending on the temperature at which the emulsion is prepared, the shear rate, or, at high oil concentrations, and depending on the concentration of the oil, O/W

---

6    The HLB System, a time-saving guide to emulsifier selection, ICI Americas, 1984, readily downloadable via a Google search

emulsions can be prepared with certain surfactants over the entire range of HLB numbers from 2 to 17."

As Stubenrauch likes to point out in lectures[7], the HLB of C8E4, C10E5 and C12E6 are all ~12.5 but their CMC values change from 8.6e-3 to 8.0e-5. And of course the solubilizing power of a shorter C8 and E4 chains is less than that of a longer C12 and E6 chains, so irrespective of CMC the C12 is generally a more effective surfactant.

So why, in the 21st century, do books, courses and surfactant suppliers persist with the notion that HLB is of some value? When you talk to surfactant suppliers in private you quickly realise that they regard HLB as valueless. They provide the numbers because customers expect them. When you talk to customers, they say they ask about HLB because they must be important since suppliers provide the information (along with CMC).

It is a system that keeps going despite the fact that it is of little use to anyone, because no one seems to have the courage to say that it should be banned.

The real crime of HLB is that it has stood in the way of a much better system, designed first in the late 1970's by Salager and developed in subsequent decades by Salager and others. That system made one strategic mistake – it called itself SAD (Surfactant Affinity Difference), not an acronym that would ever be popular. In recent years SAD has been regenerated as HLD (Hydrophilic Lipophilic Difference; in fact HLD=SAD/RT). HLD contains all the insights of HLB and solves the problems of finding the right balance with different oils, different salinities and different temperatures. HLD is a property of the *system*, not of the *surfactant*, and that makes all the difference.

## 1.7    Surfactant blends

There is one more issue to be discussed in the context of failed attempts to help surfactant formulators to formulate. It is well known that single surfactants are not much use on their own and that formulators end up with surfactant blends. This is a different issue from the fact that all commercial surfactants are, for cost and raw material issues, mixtures of different chain lengths. There are many "explanations" of why surfactant blends are superior and I have generally found them unconvincing. I have my own explanation: "Single surfactants are generally unsatisfactory because they are in a useless part of surfactant space". It turns out that most of the common surfactants are far too extreme to be of much use in any type of subtle formulation. This means they have to be blended with other surfactants (too extreme in the other direction) before they enter a sensible part of surfactant space. To some extent the various recipes for blending surfactants of different HLB are an acknowledgement of this. In principle one can always

---

7    I am grateful to Prof Stubenrauch for giving me the appropriate slide from her lecture

choose an appropriate ethoxylate from the full range, but manufacturers seem to have settled on relatively few of them to be made in bulk at low cost and these tend to be unsatisfactory on their own. Once we understand HLD and have some objective characteristics of the surfactants that can be fitted in to the requirements of the overall system, the need for blends (if one insists on using bad surfactants) becomes clearer. Two useless surfactants can become useful. It would be better if the default surfactants from suppliers were in a generally more useful range so we didn't have to create blends of surfactants that are too extreme. A supplier who intelligently offered intrinsically useful surfactants with appropriate data would have a significant competitive advantage.

This is not quite the whole story. It is argued that some surfactants might well be useful but are too insoluble to have the desired effect. The presence of other surfactants can solubilise the desirable surfactant, rendering it useful. There are other, more subtle, reasons for fine-tuning surfactants. Each of these will be discussed when relevant.

## 1.8 Elasticity and bending

Most of us can survive without knowing much about elasticity and bending of surfactant monolayers. So maybe you can skip this section. My experience has been that the modest effort needed to understand these effects is worthwhile in the long run. So have a go and see for yourself. In this context, elasticity is simply resistance to stretching the surfactant film along a straight line – just like pulling an elastic band. Bending is necessarily more complicated because usually it involves three dimensions.

We have already done much of the hard work to understand elasticity. From the curve of interfacial tension versus surfactant concentration it was easy to extract $\Gamma_m$ which is the concentration of surfactant at the surface which is measured in moles/cm² rather than the conventional moles/cm³ for bulk concentration. It is often called "surface excess" because it is trying to differentiate between the numbers of molecules at the surface for positive reasons rather than if they were there statistically via the bulk concentration. As the calculation requires a theoretical construct called a Gibbs dividing surface which itself divides the thermodynamic community, we can just get used to thinking of all references to $\Gamma$ as being a common-sense surface concentration.

The curve also required an absorption coefficient K. From these two values it is possible to calculate a maximum elasticity, E, at the maximum possible relevant surfactant concentration which is CMC:

$$E = -RT\Gamma_m K.CMC \qquad \text{1-11}$$

The CMC app (https://www.stevenabbott.co.uk/practical-surfactants/CMC.php) discussed earlier calculates E. This is, unfortunately, a nearly useless theoretical

construct, so do not get too excited about the precise values. First we need to see why the ideas behind it are useful (and why they are being discussed here) then see why the actual value is of not much help. Note that when we discuss foams, the elasticity value calculated by a similar method is of great practical value.

Elasticity (N/m) is the stress (equivalent to surface tension) needed to produce a unit strain – i.e. to double the surface area. The classic definition is that $E=2A\delta\gamma/\delta A$, the change of surface tension $\gamma$ with area A. The more molecules that are packed at the interface (high $\Gamma_m$), with a great desire to be there (high K) backed up by a high concentration (high CMC), the more resistant the interface is to stretching. Because of the interconnection of these variables in the idealised setup used in the app, increasing CMC reduces K by the same amount, so changing CMC has no effect on elasticity (yet another reminder that CMC is not an especially significant parameter). Decreasing the equilibrium surface tension via a "better" surfactant (low $\gamma_C$) decreases $\Gamma_m$ so makes the interface less elastic. Another way to look at this is that if the interfacial energy is very low then it takes very little effort to disturb the surface. Yet another way is to remember that a low $\gamma_C$ implies a large $(\gamma_0-\gamma_C)$ which is a high surface pressure $\Pi$, and a high pressure means that there is a "push" that makes it easier to stretch the interface.

Why is the elasticity calculated from the surface tension plot merely "ideal" and not very useful? Because in the real world, as the interface is stretched, opening up an energy-demanding hole, another surfactant molecule from the bulk can readily pop in to fill the gap. This makes it much easier to stretch the interface. Attempts to measure surface elasticity tend to raise more questions than they answer as the results depend strongly on the speed of measurement. At low speeds the elasticity is close to 0 because the constant supply of fresh molecules means that the surface doesn't even know that it is being stretched. At high speeds it will be close to the theoretical value.

So, curiously, scientific attempts to validate ideas of elasticity tend to give results that are not too useful. Practically, the predicted relatively high elasticities should be of great importance because they can make emulsion particles behave as solid spheres during high-speed processes such as emulsion creation. Indeed, the Danov explanation of the Bancroft rule was based on "solid" particles so these high elasticities may well be real.

The Danov review also covers what happens when drops with low elasticities and therefore high surface mobilities interact. These follow different rules. For many reasons it turns out that 1μm is a significant size for emulsion particles. Below this they tend to be rigid and tend to obey Brownian rules of motion. Above this they tend to be less rigid and also subject to more conventional (gravitational) rules of motion. So the worlds of those who routinely handle 10μm emulsions are rather different from those who routinely handle 500nm

emulsions. There are, as ever, many exceptions to these broad statements. For example, big drops made with large, slow surfactants will be more rigid both because their theoretical elasticity is higher and because the surfactants are much slower at coming to the surface to reduce the elasticity.

All this means that many bad surfactants in terms of common ideas such as having a high interfacial tension (with air) and high CMC can be excellent stabilizers *if they can be brought sufficiently well to the interface*. It can be hard to get bad surfactants to do what you want, but once they do it, they will stick around doing it very well. A large amount of food science relies on this profound truth. Many food surfactants are big, slow and inefficient – and give long-lasting emulsions once those obstacles have been overcome. The limit of this idea can be found with Pickering emulsions which are stabilised by large nanoparticles. If you can make a Pickering emulsion then it is wonderfully stable. The problem is such emulsions are harder to make (to the right size) in the first place.

Bending at first seems simple. Take a planar film and bend it like a sheet of paper. This requires energy and the bending constant k is intuitively straightforward. This can be considered as an intrinsic property of the interfacial film. Now wrap the sheet of paper into some shape such as a cylinder and try to bend it. The resistance depends strongly on the overall curvatures of the sheet of paper in two orthogonal directions, $c_1$ and $c_2$, as well as on the properties of the paper itself. So we end up with a new bending constant, $k_c$, which is then applied to the mean curvature, $H=0.5(c_1+c_2)$. The complication is that this curvature is not enough to describe the whole physics. The system also needs the "splay" or "Gaussian" curvature constant $\bar{k}_c$ which is applied to $K_G=c_1c_2$ (the subscript G is not normally used but we have confusion here with the absorption coefficient K in the CMC formulae). Finally, because surfactants have a natural curvature, $H_0$, from their shape any bending has to be seen as being away from that natural curve – i.e. a straight interface is bent! The surface energy γ for a mean curvature H starting with an unbent energy $\gamma_0$ is then given by:

$$\gamma = \gamma_0 + 2k_c\left(H - H_0\right) + \bar{k}_c K_G \qquad \text{1-12}$$

So we have an equation with two curvatures, $c_1$ and $c_2$ and two constants, $k_c$ and $\bar{k}_c$ as well as the natural curvature $H_0$. If we add Helfrich's idea of torque, τ – force applied over a distance – then we start to get some powerful insights into what is happening at an interface. As Fraaije argues[8], and as is discussed later, it is possible to link together all these ideas and gain powerful insights into what is happening at curved interfaces, especially the curved oil/water interface which is generally more interesting than water/air. The core problem is that it is hard to know what all these constants are; $k_c$ in particular seems to be hard to estimate.

---

8    Johannes G. E. M. Fraaije et al, Method of Moments for Computational Microemulsion Analysis and *Prediction in Tertiary Oil Recovery*, Langmuir 2013, 29, 2136−2151

The standard illustration of these bending parameters tends, unfortunately, to belong to the world of CPP, giving the impression that this is where curvature thinking is important. Now you are alerted to the fact that micelles, as generally understood, are of little importance to much of surfactant formulation, we can show the common diagrams with examples of different curvatures, but you can imagine them in terms of more interesting phenomena more closely associated with oils, solubilisation etc.



Curvature $c = \dfrac{1}{r}$

Mean Curvature $H = \dfrac{1}{2}(c_1 + c_2)$

Gaussian Curvature $K_G = c_1 c_2$

Micellar
$c_1 = c_2$
$H = \dfrac{1}{r}$, $K_G > 0$

Hexagonal
$c_1 = \dfrac{1}{r}$ $c_2 = 0$
$H = \dfrac{1}{2r}$, $K_G = 0$

Lamellar
$c_1 = c_2 = 0$
$H = 0$, $K_G = 0$

Bicontinuous
$c_1 = -c_2$
$H = 0$, $K_G < 0$
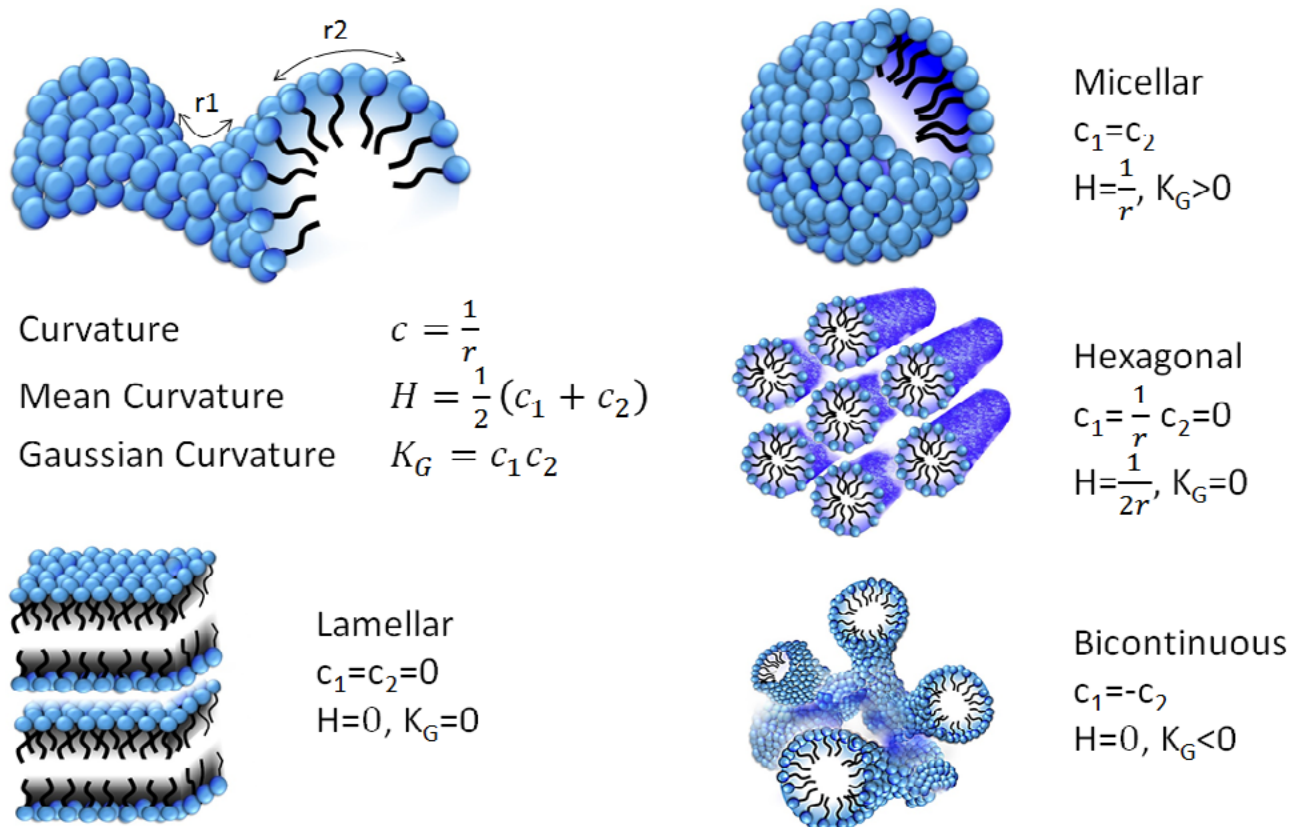
**Figure 1-4 The classic descriptions of curvature.**

At this point it is enough to have pointed out these powerful ideas exist, even if the surfactant community is not yet well placed to take full advantage of them outside the conventional analysis of surfactant/water liquid crystal phases.

## 1.9    Cloud Point and Krafft Point

No book on surfactant science can omit these numbers, even though they aren't all that significant except in a negative sort of way discussed shortly. Neither of them is general because cloud point applies (mostly) to ethoxylates and Krafft point (mostly) to ionics.

The cloud point is the temperature at which a clear solution (typically 1%) of a surfactant turns cloudy when it is heated from room temperature. Ethoxylates are peculiar because the solubility of the ethylene oxide chain decreases with temperature as its helical shape unwinds and reveals more of the hydrophobic

portion of the chain. For a pure ethoxylate the transition from clear to cloudy is quite sudden so the cloud point is well-defined, for a more normal commercial version of the "same" surfactant, the transition to cloudy can take place over a broader temperature range because the different tails and heads in the mix have different cloud points.

It is sometimes said that "you should formulate an emulsion at the cloud point temperature". This is mostly wrong for multiple reasons. When we come to the HLD explanation of surfactancy issues, the idea that ethoxylates change phase at a certain temperature is certainly correct, but this temperature depends on the oil, salinity and other additives such as alcohols. Those who claim that the cloud point is significant are confusing a property of surfactant + water (which is of little interest to most of us) with a property of surfactant + water + oil + additives which is of great interest to us.

For those who are interested in surfactant + water the cloud point is still of little interest as it is defined at one arbitrary concentration. In general the surfactant/water/temperature behaviour is complex and needs to be described with a binary phase diagram. Such diagrams are explored in detail later in the book. It is possible to obtain a cloud point from a binary phase diagram, it is not possible to obtain a binary phase diagram from a cloud point.

The Krafft temperature (often misspelled as Kraft) is a phenomenon relevant to those surfactants (especially ionics) whose solubility increases with temperature. The diagram shows what happens in typical cases.
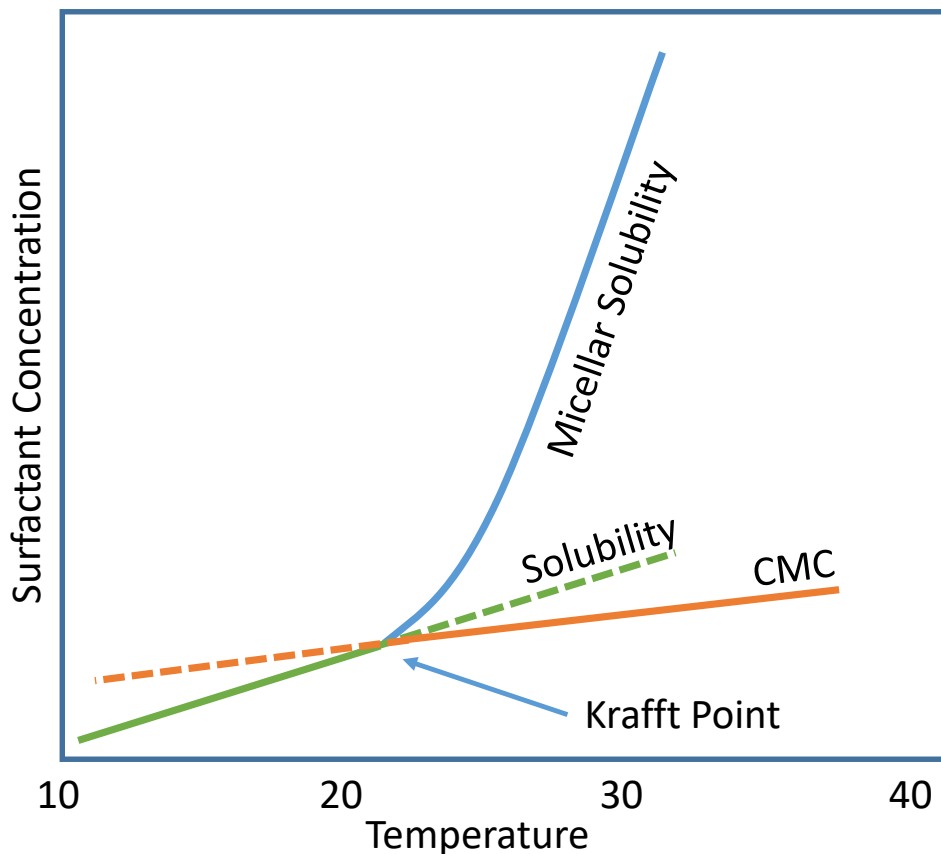
**Figure 1-5 The classic Shinoda diagram explaining the Krafft point. Reality is generally more complex, but the principle is clear.**

At low temperatures the surfactant is a normal molecule with some sort of solubility curve that would tend to increase as shown. But surfactants have CMCs which also change with temperature, as shown. The point where the two curves cross is the Krafft point. What is the significance of this point? As soon as you have micelles then the "solubility" of the surfactant is defined by micelle formation and, roughly speaking, you can have large amounts of micellar solubility because adding more surfactant simply creates more micelles. But if you are at a temperature where the solubility is less than the CMC there are no micelles so you cannot have any form of micellar solubility.

To put it crudely, below the Krafft point you have a useless surfactant. And that is about all that needs to be said. Except for an important negative point. As we shall see, a great way to increase the efficiency of a surfactant in terms of emulsification is to extend the tail. So why aren't all (ionic) surfactants, say, C20? The reason is that the solubility of the long-chain surfactants is low, so at room temperature they are below their Krafft point and are generally hopeless.

Once more, this is only partially true. In general we don't care about CMC and micelles because they are a surfactant + water phenomenon. We care more about surfactant + water + oil and the Krafft point is largely irrelevant. As with the cloud point, the real answer lies in phase diagrams, but in this case, ternary phase diagrams with water, surfactant and oil. Nevertheless, it is generally true

that long-tail surfactants have horrible phases in such diagrams which make them effectively unusable for typical emulsion applications.

An alternative to long tails for efficient surfactants is to have long middles. This idea is discussed when we come to extended surfactants.

## 1.10 Dynamic Surface Tension

Elasticity turns out to be of huge importance in terms of how surface tensions change with time during dynamic processes, Dynamic Surface Tension or DST. The classic discussion of DST is concerned only with the issue of how the surface tension changes once some fresh surface has been formed, and the first section explores those ideas in detail. The surface tension starts off as that of water and ends up as $\gamma_C$. Yet DST is also relevant for when surfaces are deformed from $\gamma_C$ via processes that can include compression (reducing $\gamma$ to *less* than $\gamma_C$, an idea that is relatively unfamiliar) as well as expansion (increasing $\gamma$). For those who would like to know why DST during compression might be of interest, perhaps the idea that we would not be able to breathe without that effect is reason enough. The DST variation around equilibrium is discussed in the second section.

### 1.10.1 Classic DST

The classic study of DST follows how the high initial interfacial tension decreases to reach the equilibrium low tension.

There are two cases where classic DST is important. First, because emulsions require low interfacial tensions, it would be ideal if the surfactant appeared instantly at the surface of the newly-formed drop. Unfortunately, at the instant the drop has formed very little of the surfactant has had any reason to position itself at the surface. Only when the surface has appeared is there a thermodynamic driving force and a concentration gradient necessary for the surfactant molecules to move to the interface. Second, when coating or printing onto a substrate it would be highly desirable if the added surfactant gave the desired low surface tension (e.g. to allow spreading) instantly rather than after a delay, by which time other processes might have ruined the coated/printed surface.

The first of those cases involves DST at the water/oil interface, the second is the much more studied case of the water/air interface. Given that water/oil is generally more interesting, why is that not the default system for study? Even the experts admit that DST studies in oil/water are of great difficulty so tend to rely on water/air results to illustrate the salient points. That is the approach adopted here. Even with this simplification, things are sufficiently complicated. The app attempts to handle the complications making it as easy as possible to explore the different aspects of the process:

**App 1-4** https://www.stevenabbott.co.uk/practical-surfactants/DST.php
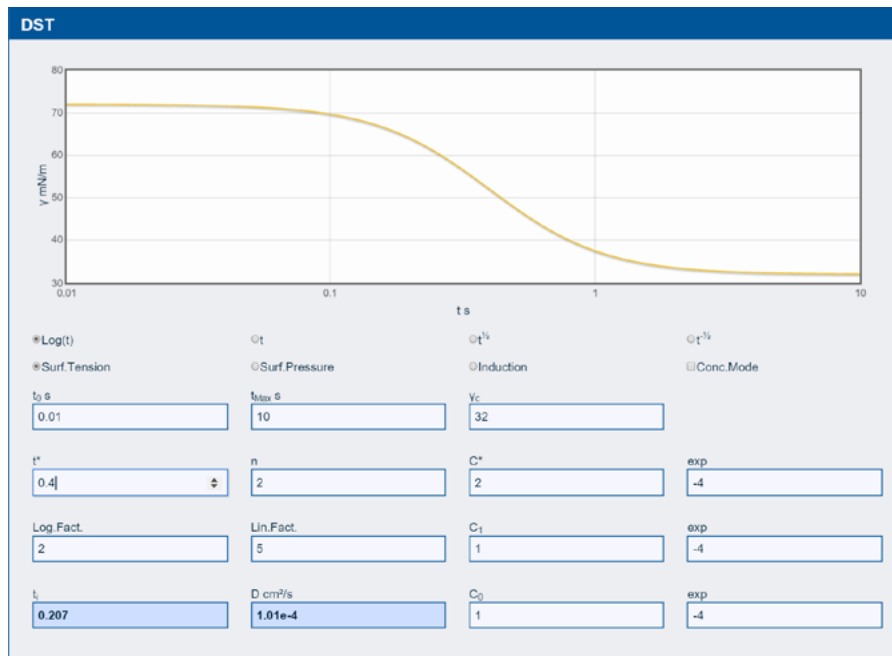
The default view shows a typical academic's DST curve. The plot is between 0.1 and 10 seconds (plotted logarithmically), it starts at 72 mN/m and falls to the equilibrium $\gamma_c$ of 32 mN/m. A careful reading of academic papers shows that there are lots of simpler theories to describe the initial slow fall from 72 to, say, 65 and also the slow equilibrium from, say, 35 to 32 mN/m. These theories are of little practical interest to formulators who are more interested in how quickly the surface tension falls to a low value such as 40. The more interesting regime requires a solution to the Ward-Tordai equation, discussed later, which is amenable only to numerical, rather than algebraic, solutions.

Fortunately, Rosen has produced a formula that describes all such curves with just a few parameters. He admits that the curve has no justification other than the fact that it works very well. As happens so often in the surfactant community, there is little sign of suppliers of surfactants providing "Rosen parameters" (t* and n described below) that would allow us to compare the dynamic properties of two surfactants. Suppliers delight in telling us that they have a "fast" surfactant, while neglecting to give us a standardised set of numbers that would allow us to make sense of how appropriate that surfactant is for our particular needs. An alternative set of "Stebbe" parameters is described below.

The Rosen formula gives the surface tension at time t, $\gamma_t$, starting at $\gamma_0$ and with equilibrium value $\gamma_c$ depending on a characteristic time t* and a power dependency n:

$$\gamma_t = \gamma_c + \frac{\gamma_0 - \gamma_c}{1 + \left(\frac{t}{t^*}\right)^n}$$

1-13

If the formula is shown using the same data but with a linear plot, the academic focus on the start and end slopes of the curve looks less necessary:
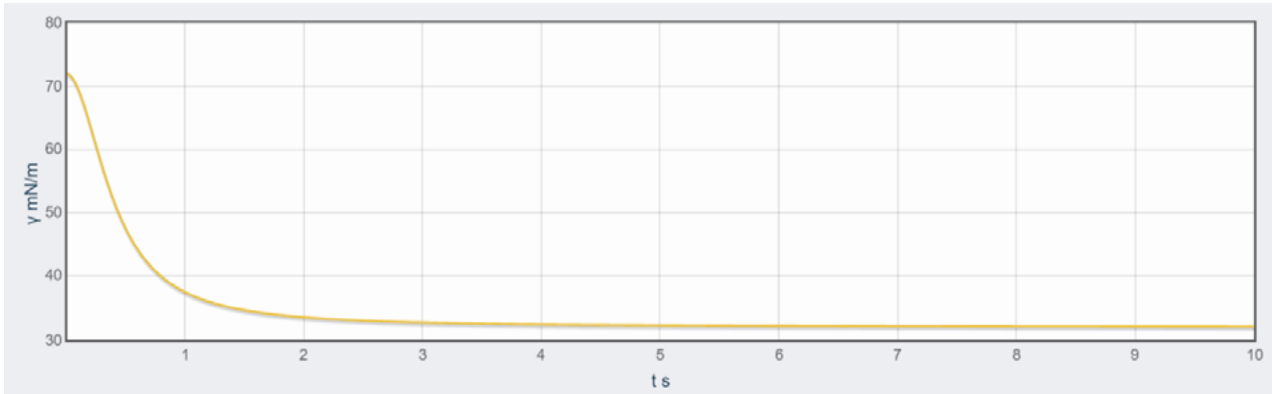


**Figure 1-6 For those of us who simply want a significant and rapid practical reduction in surface tension, the linear plot of the same data looks less interesting but is more relevant.**

For those who just want a fast-acting surfactant the need is for the lowest possible t* and the highest possible n, while providing the required low $\gamma_c$ and high stiffness interface for long-term performance. The many intellectual niceties of the Rosen formula are of little extra value to most of us but are included in the app for those who are interested.

Trade-offs are inevitable. Large and/or rigid surfactants can provide excellent long-term stability whilst being useless in terms of generating emulsions because their DST behaviour is far too slow. Surfactants that can zip quickly to the interface tend to be less robust in the long term.

The key number related to speed is the Diffusion coefficient, D. The app uses standard approximations to estimate D from the two Rosen parameters. As the log plot implies, there tends to be an induction period, $t_i$ which is controlled by D. It is derived from the Rosen parameters as:

$$\ln(t_i) = \ln(t^*) - \frac{1.137}{n} \qquad \text{1-14}$$

Any curve is defined for one concentration of surfactant. More surfactant gives a faster curve. The effect of surfactant concentration C fits into the Rosen approach via t*:

$$\log(t^*) = a\log\left(\frac{\Gamma_m}{C}\right) + b \qquad \text{1-15}$$

$\Gamma_m$ is the surface excess discussed in the first chapter and represents the saturated concentration at the interface. The two constants a and b are fitting parameters. Because a small t* means faster equilibrium a higher C (obviously) is a good idea. An unfortunate contradiction arises from the $\Gamma_m$ term. For low

equilibrium surface tension, $\gamma_c$, and for a nice, robust interface a high $\Gamma_m$ is needed; this automatically increases t*. This makes sense – it takes longer to pack in more surfactant molecules – whilst still being unfortunate.

There is one academic nicety that everyone can appreciate. Standard diagrams of how surfactants work make it look as if surfactant molecules get locked into large micelles above the CMC. It would then seem as though diffusion should slow down markedly once micelles start to form. So it is self-evident that curves of DST plotted against surfactant concentration should show a clear "blip" at the CMC. To general astonishment, no such blip can reliably be found. Micelles, it seems, are such loose statistical constructs that in terms of diffusion to an interface they don't impose a significant barriers.

The app has many more features that the interested reader can explore. Usually anyone who has the equipment to measure DST will have software that can do all the relevant analyses. If the software does not follow the Rosen approach then it should be possible to translate the data from the software via the app.

An alternative approach from Prof Stebe[9] and colleagues (which they explicitly link to the Rosen approach) takes as a starting point the standard equilibrium surface tension versus concentration plot discussed in the CMC section. The title of the paper is encouraging for a formulator: *Which surfactants reduce surface tension faster? A scaling argument for diffusion-controlled adsorption*.

Using the $\Gamma_m$, *a* and $\gamma_C$ values from the Langmuir fit to the data it is possible to predict DST behaviour provided the diffusion coefficient D is known. Measuring D is not easy for most of us so the starting point can be that typical surfactants have D values in a small range, typically $3\text{-}7.10^{-6}$ cm²/s (larger protein surfactants have smaller D values, say $5.10^{-7}$). Because $\Gamma_m$ does not vary much between surfactants, and because $\gamma_C$ is what it is, it looks as though all DSTs are about the same. Fortunately there is one extra variable, the surfactant concentration C, which brings everything to life. The Stebe approach makes sense of all this. The full approach recognises that not all surfactants fit a Langmuir curve. Those surfactants that attract or repel each other require a Frumkin curve and ionics have the complexity of ionic strength so require a Davies curve. These extras are not included in the app as the aim is to point out the elegant simplicity of the methodology which does *not* depend strongly on the subtleties of surface tension curves.

The first thing to do is to plot the surface tension curve from the Langmuir parameters. Unlike the CMC app, these are inputs, not outputs. Next, the DST curve can be calculated via the Ward and Tordai equation. The details of this need not detain us other than to repeat that there is no algebraic solution to the

9    James K. Ferri, Kathleen J. Stebe, Which surfactants reduce surface tension faster? A scaling argument for diffusion-controlled adsorption, Advances in Colloid and Interface Science 85, 2000, 61-97

equation which itself contains an integral, so a numerical solution is required. The app relies on the clear methodology[10] provided by Prof Stevenson and colleagues for solving the equation – something we should be grateful for because the published "standard" solutions to Ward-Tordai are incomprehensible to most of us. We now have everything that is needed to optimise for DST.
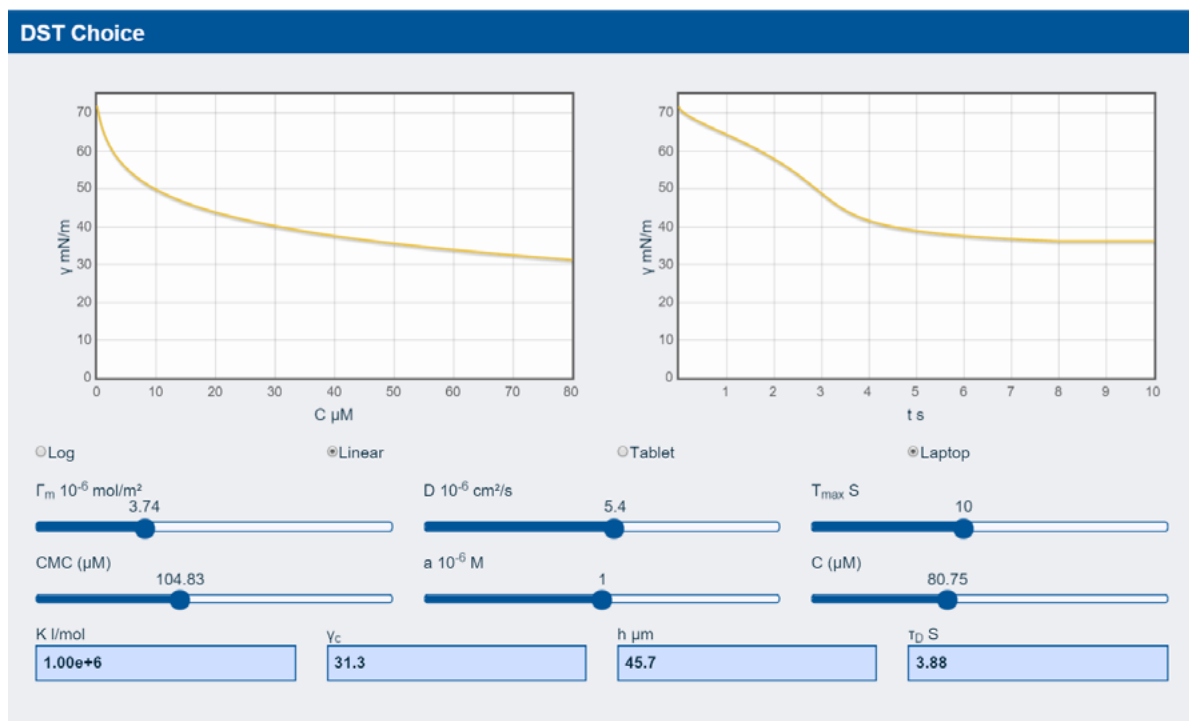
Moving the sliders about is one way of doing it. The Stebe approach is to focus on two parameters: h and $\tau_D$. The h parameter is the "adsorption depth" which describes how far down into the bulk solution it is necessary to go to find all the surfactant required to cover the surface with its current concentration $\Gamma$. A large h is clearly a bad thing for rapid response because molecules have to diffuse over a longer distance to reach the surface. The definition of h at a given surfactant concentration C is simple and can be in terms of the current $\Gamma$ or in terms of the constant $\Gamma_m$ and the parameter a:

$$h = \frac{\Gamma}{C} = \frac{\Gamma_m/a}{1 + C/a} \qquad \text{1-16}$$

The $\tau_D$ parameter is a time-scale which tells you, as the authors say, "within a factor of 10" when the surface tension will approach the desired $\gamma_C$. This is actually all you need and you can neglect everything else about the app!

---

10   Xueliang Li, et. al., A simple numerical solution to the Ward–Tordai equation for the adsorption of non-ionic surfactants, Computers and Chemical Engineering 34 (2010) 146–153

$$\tau_D = \frac{h^2}{D} \quad \text{1-17}$$

The $\tau_D$ equation says that for a fast DST equilibrium all that is needed is a small h and a large D. Because D doesn't change much, that leaves h. $\Gamma_m$ values do not change much so this leaves *a* and C. The *a* value varies strongly between surfactants so can be used to tune DST. It is also related to the limiting surface tension so is not an independent variable. The real secret is to throw in a lot of surfactant (high C) to get fast results. This sounds rather crude.

The Stebe approach is especially relevant for those who have the need for a very rapid reduction in surface tension when fresh surface is created during a coating or printing process. The fact is that many of the more successful fast-acting surfactants used in formulations are not-very-good surfactants as they might require, say, 0.5% surfactant to get a result. Many "good" surfactants at 0.5% are grossly over CMC so their *effective* concentration will be close to CMC and won't deliver the required benefits – though the absence of correlation between DST and CMC mentioned earlier rather dilutes this point.

Good surfactants might also have problems from interactions with other components in the formulation. For example, some polymers *love* to soak up good surfactants which wrap around them – doing strange things to the formulation and to the DST behaviour. Some "bad" surfactants are not especially attracted to polymers in a formulation so cause no such formulation problems and can be added at relatively high concentrations (though still "small" in overall formulation terms). If these bad surfactants are also volatile then they do their DST job at the in the early instants of the process before disappearing. This takes us into the domain of simply adding alcohols to a formulation – which is what is often done. These are very bad surfactants, but are at such high concentrations that the DST response is instantaneous. Their downside is the high concentrations and flammability.

There is no single right answer. In some cases, super-efficient low CMC good surfactants are required, in some cases alcohols are fine, and in many cases, the right "bad" surfactant turns out to be best for the job. It is, for example, frequently noted that some acetylenic surfactants are peculiarly good as practical additives for good DST performance. This has often produced speculation about what is so special about acetylenics. The speculation is unnecessary. It just turns out that the popular ones are in a sweet spot of "bad" surfactants that are beneficial at, say, 0.5% and some are sufficiently volatile that they disappear when their excellent job is done.

The "within a factor of 10" accuracy for $\tau_D$ doesn't sound all that good, but is academic modesty – it is more generally a factor of 2. Mostly you can just compare $\tau_D$ values and get a rather good idea of what to expect in reality. Even if

$\tau_D$ were perfectly accurate you could not expect too much accuracy from all this; the Langmuir curve is an approximation, there is frequently some extra barrier at the interface that makes a kinetic D too simple, and the "same" surfactant can vary strongly from batch to batch because relatively small changes in the components can affect different parts of the DST curve.

A further intellectual nicety is that measurements of DST are often done with bubbles and the spherical interface can impose an extra kinetic barrier – so such DST measurements are not directly applicable to planar surfaces without some intelligent corrections. Hopefully anyone with such a measurement device will also have the software to make the appropriate corrections.

### 1.10.2 DST around equilibrium

Imagine a drop that has reached a stable surface tension equilibrium after a few seconds. Now compress or expand the drop. This might happen in an experimental apparatus (with a syringe controlling the drop volume), on the surface of the lung as you breathe in and out, or in an emulsion drop subject to various hydrodynamic forces.

We already have enough information to know what will happen to the surface tension. If the change in surface area is made slowly then there will be no change in surface tension – fresh surfactant can easily fill in any gaps during expansion or exit the interface during compression. If the change is very fast then the surface tension will decrease on compression and increase on expansion.

There are a number of ways to model what happens during such processes. Some focus on the diffusional aspects – related to the previous discussion on classic DST. Some focus on the isotherm kinetics. Others assume that the limiting step is not just diffusion but absorption too – with barriers both to the arrival of new surfactants and to their departure. Of course these approaches are inter-related, they each come back to the basic science of what it means to have an excess or shortage of surfactant at the interface and they all have time-dependency built-in to their rate-limiting steps. To allow the reader to explore these phenomena while using a language consistent with the rest of the book, my choice was the CRM (Compression Relaxation Model)[11],[12] from a team at U. Toronto. This was developed specifically for handling the fascinating behaviour of lung surfactants which give apparently unexceptional surface tension behaviour at equilibrium (25 mN/m) but can go down to a mere 5 mN/m on compression. The model happens to better fit the dynamic behaviour of

11   Sameh M.I. Saad, A.Wilhelm Neumann, Edgar J. Acosta, A dynamic compression–relaxation model for lung surfactants, Colloids and Surfaces A: Physicochem. Eng. Aspects 354 (2010) 34–44

12   Sameh M.I. Saad et al, Effect of surfactant concentration, compression ratio and compression rate on the surface activity and dynamic properties of a lung surfactant, Biochimica et Biophysica Acta 1818 (2012) 103–116

lung surfactants than the other models they tested, though it is possible that those models are better for other cases. In any event, the focus is the general behaviour that any reasonable model will reproduce – the details are for the experts in their given field.

The basic CRM formula tells us that the surface tension, γ changes with time, t, according to two surface tension effects, $\gamma_1$ and $\gamma_2$, with the proviso that it can never go below $\gamma_{min}$.

$$\frac{d\gamma}{dt} = \frac{d\gamma_1}{dt} + \frac{d\gamma_2}{dt} \quad \text{if γ>=γmin} \quad \text{1-18}$$

$$\frac{d\gamma}{dt} = 0 \quad \text{if γ<γmin} \quad \text{1-19}$$

The $\gamma_1$ term depends on the absorption and relaxation coefficients, $k_a$ and $k_r$ which have units of 1/s. It also depends on whether the current γ is greater than or less than the equilibrium value $\gamma_c$.

$$\frac{d\gamma_1}{dt} = k_a \left( \gamma_{eq} - \gamma \right) \quad \text{if γ>=γc} \quad \text{1-20}$$

$$\frac{d\gamma_1}{dt} = k_r \left( \gamma_{eq} - \gamma \right) \quad \text{if γ<γc} \quad \text{1-21}$$

The $\gamma_2$ term depends on the elasticity of the system. In the papers two separate values are used for absorption and relaxation, though experimentally these seem to be very similar so for simplicity just one value, ε, is used in the app:

$$\frac{d\gamma_2}{dt} = \varepsilon \left( \frac{1}{A} \cdot \frac{dA}{dt} \right) \quad \text{1-22}$$

which is the classic definition of elasticity – the change of surface tension with area expanded to include the change of area with time.

Why might there be different absorption and relaxation coefficients? For simple surfactants the answer is that the differences might not be large. For many other surfactants there can be large differences, especially for those surfactants that are essentially insoluble in the aqueous phase, such as the lung surfactants, many crude oil surfactants (e.g. the asphaltenes), classical biological surfactants such as the phospholipids used in liposomes and vesicles and many of the surfactants/emulsifiers used in food science.

**App 1-6** https://www.stevenabbott.co.uk/practical-surfactants/CRM.php
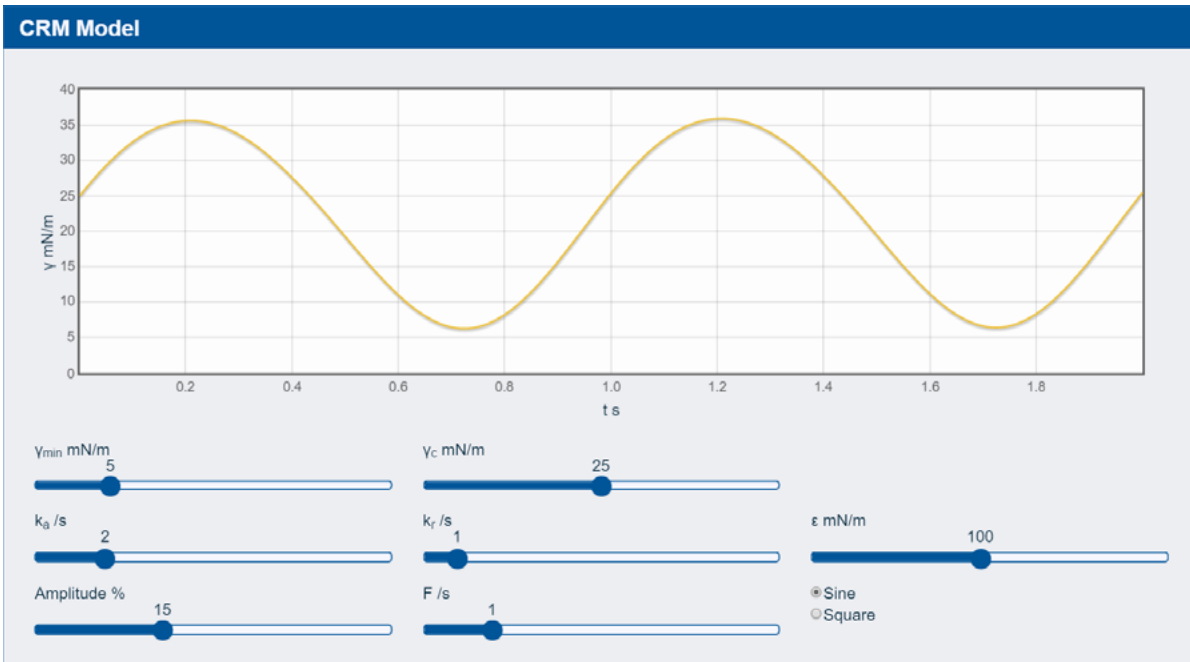
The app can use a basic sine wave oscillation of varying frequency and amplitude (in terms of % compression/expansion of the surface). Over much of parameter space the response is a sine wave. Changing the relative values of $k_a$ and $k_r$ simply change whether the sine wave is symmetrical around $\gamma_c$. Not surprisingly a low $k_r$ biases the curve below $\gamma_c$ and a low $k_a$ biases it above $\gamma_c$. A larger amplitude just makes the sine response larger, with a cut-off at $\gamma_{min}$. Above a certain level, frequency makes no difference to the response curve. Once it falls below the values of k then the response becomes smaller. Very slow changes to a surface give no change in surface tension because surfactants can always migrate to accommodate the change.

Two cycles are shown. The first cycle is starting from an equilibrium state so may not be the same as the second. As the frequency is changed, the time axis changes so that there are always exactly two cycles in view.

The behaviour is visually more interesting when a square wave is used. Irrespective of the chosen frequency, a square wave always has a high-frequency component and the responses can vary strongly as the input parameters are changed.
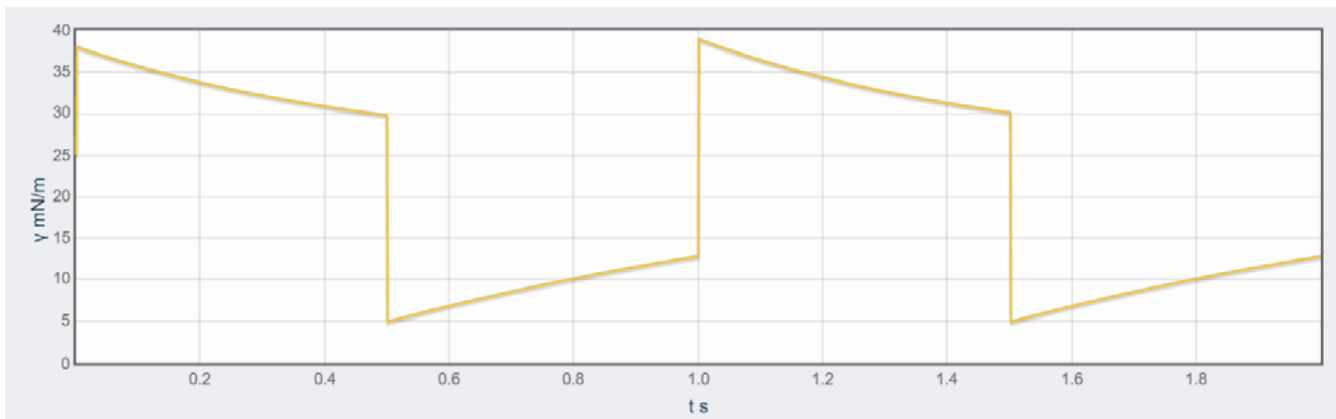
**Figure 1-7 The same CRM setup but with a square wave.**

What parameters should be used in any given situation? The default parameters are those from the CRM papers and fit the leisurely timescale of lung surfactants (though in emergency situations deliberately rapid assisted breathing is required) and relatively high elasticities. Things like liposomes are likely to have similar high elasticities and slow relaxation times. Because the k values are /s and the frequency is also /s, if you had a system where k were 1000x faster you would simply imagine that the frequency scale was milliseconds instead of seconds.

For those of us who aren't familiar with lung surfactants, why is it important for the surface tension to go so low on compression? The answer is complex, but one element of it is that the effort needed to open up the lungs after compression depends directly on the surface tension – so a low surface tension means that it is easy to start expanding the lungs and bring in fresh air. Another element is that it helps the lungs expand more evenly – those that are still compressed are easier to open than those (by chance) that are already more open.

## 1.11 Partition coefficient

One key property of surfactants is often overlooked - the partition coefficient, or how much a given surfactant prefers to be in the oil phase rather than the water phase. If at equilibrium the concentration of surfactant in the oil is $C_{oil}$ and in water is $C_{water}$ then the partition coefficient is given by $K_p = C_{oil}/C_{water}$.
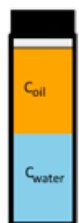


Why does the partition coefficient matter? If you believe in Bancroft's rule ("The phase in which an emulsifier is more soluble constitutes the continuous phase") then it is useful information, even though there are plenty of exceptions to Bancroft. More generally, if the partition coefficient is wildly in favour of one phase rather than the other then the chances are that it won't be such a good interfacial molecule because any surfactant molecule deep inside one phase is doing nothing useful at the interface.

The key problem with partition coefficients is that they say nothing about absolute solubility. If one surfactant is soluble to a level of 0.001% in both phases and another is soluble to 10% in both phases, they each have a partition coefficient of 1, but will possess totally different practical properties. This problem becomes more severe when you try to understand trends in partition coefficients as the surfactant tail is systematically changed from short to long. If this results in an increase of the partition coefficient it is "obvious" that the longer tail means that the surfactant is more soluble in the oil. However, this is often wrong - the surfactant with the longer chain may be less soluble in the oil than the one with the shorter chain. The increase in the partition coefficient is therefore not due to increased solubility but to a smaller decrease in solubility in the oil relative to the decrease in solubility in the water.

Another issue is that partition coefficients are meaningfully measured at concentrations below the CMC so that the effects of micelles are more-or-less eliminated. If the concentration of surfactant used in a given application is much higher than CMC then the sub-CMC partition coefficient might not be directly applicable. Despite these complications, it is still the case that partition coefficients should be the sort of numbers we routinely know for our surfactants as they tell us a lot about the inherent nature of the molecules. Sadly this is not the case. An excellent paper[13] from the U. Stuttgart's Colloid & Interface Science group provides the reasons why measurement should be routine, why the trends they observe (changing surfactant head group, surfactant tail length and oil) are happening (the oil partition example above is taken from the paper) and, importantly, why the partition coefficient is relatively easy to measure.

### 1.11.1 Measuring partition coefficients

Take a 50:50 oil water mix, add surfactant to the water at a level somewhat below its CMC, shake and allow to stand for a day or so. Commercial surfactants (as the paper shows) can take some time to equilibrate while chemically pure ones will equilibrate relatively quickly. Then measure the surfactant concentration in the water phase.

But even with sophisticated analytical equipment it is relatively hard to measure surfactant concentrations. Fortunately there is an elegant trick. Simply take a small sample of the aqueous phase and measure its surface tension. From the graph of surface tension versus concentration (the Langmuir-Szyszkowski isotherm) that is routinely used to identify CMC and to find important properties such as $\Gamma_m$, it is trivial to go from the measured surface tension of the equilibrated solution, $\gamma_{equil}$ to the actual concentration $C_{equil}$. (If you are measuring at concentrations above CMC you first have to dilute the samples by a known amount to bring them into the sub-CMC range.) After that, calculating the

13   G.Catanoiu, E.Carey, S.R.Patil, S.Engelskirchen, C.Stubenrauch, Partition coefficients of nonionic surfactants in water/n-alkane systems, J. Colloid & Int. Sci., 355, 150-156, 2011, 27, 14783–14796

partition coefficient is merely arithmetic. Knowing the original concentration $C_{orig}$ and $C_{equil}$ then:

$$K_p = \frac{C_{orig} - C_{equil}}{C_{equil}}$$

1-23

The app lets you try out the idea using the simplified isotherm from the CMC app.



**App 1-7** https://www.stevenabbott.co.uk/practical-surfactants/Partition.php
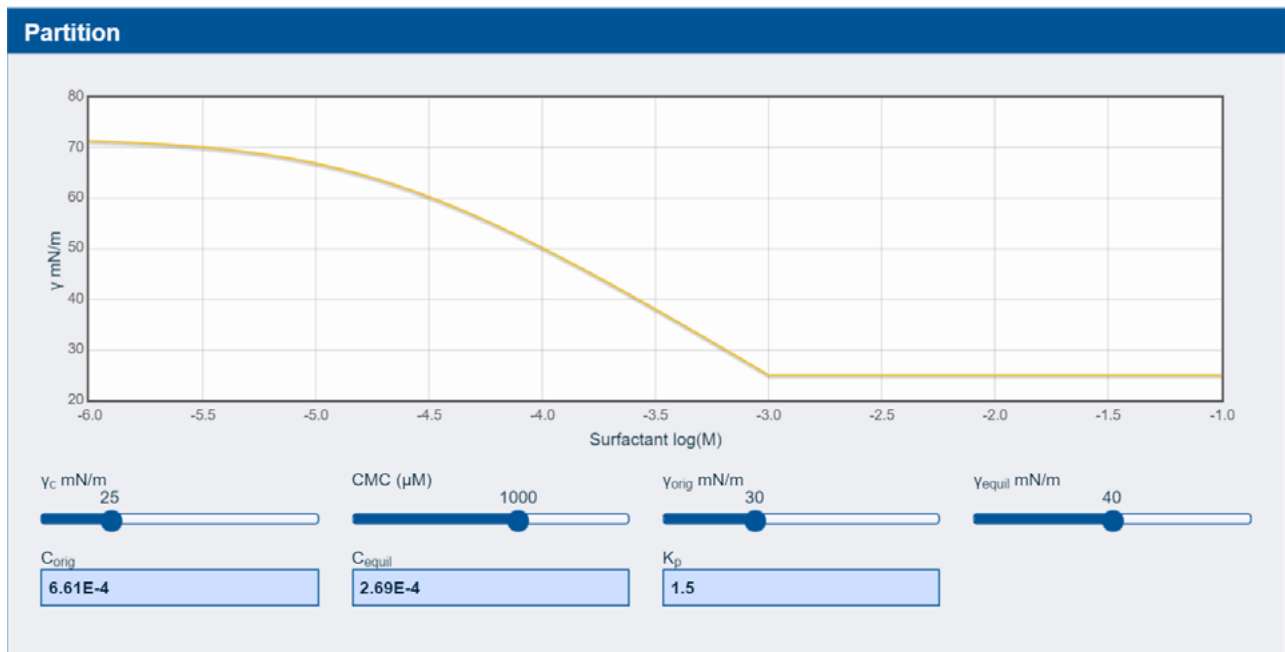
What does the paper find for the three types of surfactant studied: ethoxylates, phosphine oxides and APGs? The answer is that "it's complicated". There is no obvious correlation with other properties such as HLD (and certainly not HLB!) or CMC. Dependencies on surfactant chain lengths are small for the relatively hydrophilic (low $K_p$) APGs, large for the DMPOs (from ~1 at C10 and ~55 for C14). And there is a relatively small oil dependency for the specific APG and DMPO surfactants and relatively large dependency (60 down to 30) when the oil changed from hexane to dodecane for $C_{12}E_6$.

## 1.12   The basics summarised

In summary, our surfactant formulation lives would be much better with a shift of emphasis:

- The routine data for our surfactants should be supplied as $\Gamma_m$, $\gamma_c$, K (or a) rather than the less-significant CMC.

- Dynamic surface tension data should be provided in Rosen or Stebbe terms.

- Elasticity should be a more normal part of our vocabulary, even if it is difficult to measure.

- Ideas such as HLB, CPP and Bancroft should be abandoned in favour of curvature-based whole-system theories such as HLD, the topic of the next chapter.

## 2     HLD – towards rational emulsions

We can all agree that understanding the science of a hydrophilic-lipophilic balance is vital to many aspects of emulsion formulation. Because HLB is reserved for an idea that is fundamentally flawed (it ascribes the balance to the surfactant, not the system) we have to choose a different term. HLD, Hydrophilic-Lipophilic Difference[14], conveys a similar sentiment, without the connotations of the discredited system. If we say that in many ways a formulation is "optimal"[15] when things are balanced then this means that the definition of balance is when HLD=0. We can readily write down a formula for HLD that includes the key aspects of any formulation:

$$HLD = SurfNo + OilNo + f(T) + f(S) \qquad \text{2-1}$$

All this is saying is that if we have some number that characterises the surfactant, another number that characterises the oil and then have some functions f() of temperature, T and salinity S then we can calculate the HLD. To reach an optimal state, the surfactant number, oil number, f(T) and f(S) must add up to 0.

The modern way of expressing this is:

$$HLD = Cc - k.EACN - \alpha.\Delta T + f(S) \qquad \text{2-2}$$

As explained earlier, calling the surfactant number Cc is an historical accident. At one time it seemed a good idea to call it a characteristic curvature, via some analogy to CPP. A possible mnemonic is to think of it as a **C**haracteristi**c** value.

EACN is the Effective Alkane Carbon Number. For hexane this is 6 and for hexadecane this is 16 – after all, each of those is an alkane carbon number. Having 6 carbon atoms does not guarantee an EACN of 6. Benzene has an EACN of 0 and Cyclohexane has an EACN of 3. Squalane has 30 carbon atoms and happens to behave as if it has 24. How can benzene behave like an alkane with 0 carbons? These are numbers that fit beautifully on a line defined experimentally from pentane to hexadecane, extending in either direction. Although it makes no chemical sense to say that something is equivalent to an alkane with 0 carbons, in terms of balancing emulsions these numbers make perfect sense. The scale continues beyond 0; for example, trichloroethylene behaves like an alkane with -4 carbons. The EACN scaling factor, k, is generally taken to be 0.17.

---

14  The D can also stand for Deviation or Distance. All three terms are fine; I happen to marginally prefer Difference.

15  The word "optimal" is routinely used to describe the "balanced" state. It is optimal only in the sense that it provides a useful, unambiguous reference point. The Greenwich Meridian is not (for most of us) the optimal location on the planet, but navigation is much easier for knowing where we are in relation to it.

Note that the Cc and EACN work in opposite directions as their signs are different.

The α term goes with ΔT, the difference in temperature from the standard state of 25°C. For typical anionics α=0.01 – the term gives a relatively slow decrease (because of the –sign in front of α) in HLD with temperature. For typical ethoxylates α=-0.06, the term gives a relatively rapid increase in HLD with temperature (because two negatives make a positive). For typical sugar surfactants such as the APGs, α=0, i.e. the balance does not change with temperature.

The salinity term is 0.13S for nonionics and ln(S) for ionics, where S is expressed in g NaCl /100ml. It is another historical accident that g/100ml is readily confused with %wt/wt and is relatively unscientific compared to a function expressed in terms of molarity. For monovalent salts other than NaCl then S=(58/MWt)g/100ml, and for divalent and trivalent salts the value of S is changed via the ionic strength which requires individual MWts of the anion and cation as well as their charges. For those who need such calculations, an HLD-Expert app is available on the Practical Surfactant site.

So *if* we know the Cc value for a surfactant and the EACN value for the oil, it is trivial to work out whether any reasonable combination of temperature or salinity will bring the formulation to balance (HLD=0) or whether the formulation is likely to be in the O/W regime (HLD<0) or in the W/O regime (HLD>0).

Thanks to heroic efforts mostly from (in alphabetical order) Acosta, Aubry, Harwell, Sabattini and Salager there is a reasonable public domain list of Cc and EACN values (used in the apps described below). Just as importantly (and unlike HLB) there are relatively simple and objective methods for measuring the Cc value of any surfactant and the EACN value of any oil.

So much for the background. Now it is time to start playing with HLD to see how it can be used for formulations – with one word of caution. No one claims that HLD is a perfect system. It is an obvious simplification of a complex system. Cc and EACN values are subject to errors because the "same" surfactant or the "same" oil from two different suppliers can be very different. Indeed, we will later see that HLD is particularly good in spotting such differences and allowing suppliers and customers to adjust rationally to these differences.

## 2.1    Specific HLD examples

Two apps allow us to track down Cc and EACN values. In addition to the lists of the values for nominally pure materials the apps take into account that rarely do we have the luxury of a single surfactant or a pure oil.

**App 2-1** https://www.stevenabbott.co.uk/practical-surfactants/CC.php
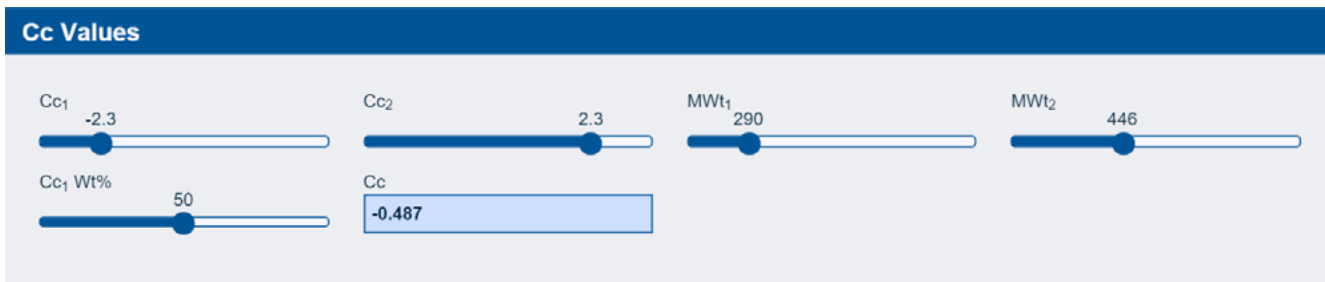
Here we have two surfactants. The first has a large, negative Cc value of -2.3. This, and its MWt indicate that it is SLS (see the preface for surfactant abbreviations). The second has a large, positive Cc value of 2.3, making it close to AOT. A 50:50 mix of them have a Cc of -0.492. Cc=0 is found with a 39:61% mix. The mixing rule is simple: it is the molar-weighted average. So a 50:50 wt:wt mix contains more moles of the SLS (MWt=289) with its low Cc than the AOT (MWt=445), so the mix has a negative Cc.

The comment in the previous chapter about most surfactants being useless can now be brought to life. A surfactant like Polysorbate 80 has a Cc of ~-3.7, massively (and uselessly) in the hydrophilic zone. Sorbitan monolaurate has a Cc of 3.5, massively (and uselessly) in the hydrophobic zone. Each has to be "blended" (or, rather, diluted) with another extreme surfactant in order to get into the practical range of formulations which will tend to hover around Cc=0, say from -1 to +1. Why Cc=0? Because the HLD theory was calibrated with relatively normal oils under relatively normal conditions so that 0 was comfortably in the middle.



**App 2-2** https://www.stevenabbott.co.uk/practical-surfactants/EACN.php
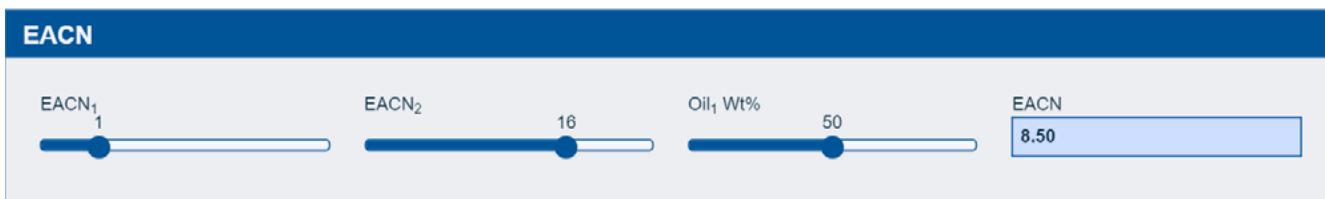
The EACN example is a mix of toluene (EACN=1) and hexadecane (EACN=16). The mixing rule is a simple wt:wt approximation – there is no compelling evidence to use volume or mole ratio instead. So the 50:50 toluene:hexadecane mix has an EACN=8.5. The reason for including this pair of oils is described below.

So we know our Cc and EACN values. We can now put them into the HLD equation using the next app:
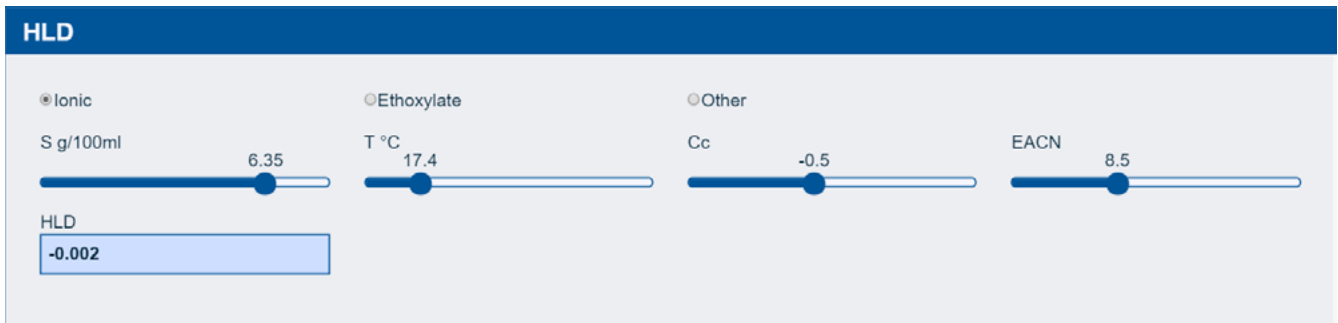
Given the Cc ~ -0.5 and the EACN of 8.5, it takes (for these anionic surfactants) a combination such as 17.4°C and 6.36g/100ml of NaCl to get a balanced formulation.
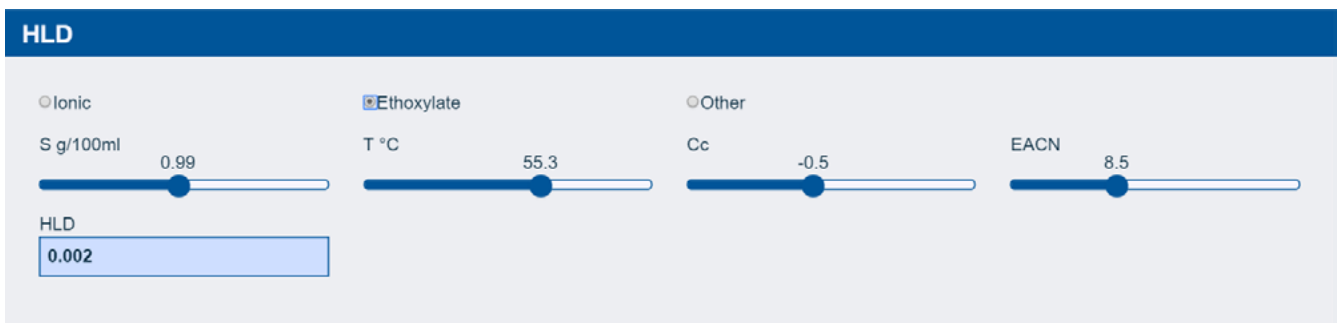


**Figure 2-1 The same Cc from an ethoxylate system is balanced at a very different salinity and temperature.**

Changing to an ethoxylate, at 1g/100ml NaCl the formulation is balanced at 55°C.

This is immensely liberating. Given just about any starting surfactant combination and oil combination the formulator can either adjust things to get an optimal formulation at some desired salinity or temperature, or, just as importantly, realise that the given surfactants and oils will *never* give a reasonable balance. A clear negative result is as important as a clear positive one.

Hopefully you are currently protesting: "All this has told me is how to get a balanced formulation – but what is a balanced formulation, and do I really want it?"

The answer to the second question is generally: "No". A balanced formulation for a classical emulsion would be a catastrophe because, as we will see, the definition of balanced is "lowest possible interfacial tension" – which means that colliding drops will rapidly coalesce and cream. In general for an O/W emulsion the best place to be is in a modest regime of HLD<0 and for a W/O emulsion in a modest regime of HLD>0. Why modest? See below. However, if you want a

crystal clear microemulsion formulation HLD=0 is the exact place to be. And if you want to make a classical emulsion using ethoxylates this is the exact place to start from, though you will quickly want to move to a lower HLD.

The last remark demands explanation. As we will see in the Emulsions chapter the theoretical amount of energy required to create an emulsion is remarkably small. Our practical experience is that it requires a lot of energy to break the oil into small drops. However, if the interfacial energy is very, very low then it requires very little dispersion energy to break up the oil drops. A reason for this will appear when we discuss critical capillary numbers. So one strategy for making an emulsion is to take the mix into a domain where the interfacial energy is low, easily disperse the drops within that domain then, as rapidly as possible, return the mix to a domain with relatively high interfacial energy so the drops cannot so easily coalesce.

With ethoxylates the trick is that at higher temperatures (56°C in the example above) the formula is balanced.
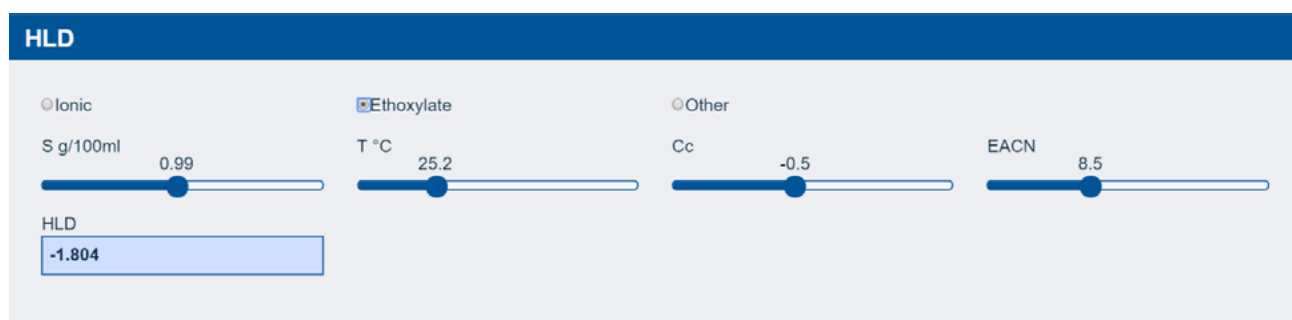


**Figure 2-2 Back at room temperature the HLD is negative – this is an O/W emulsion**

When the formulation returns to 25°C the HLD becomes negative, which means a "hydrophilic" balance, i.e. an O/W emulsion. Had we gone above 56°C we would have found a +ve HLD value, i.e. a W/O phase. The balanced temperature is called the Phase Inversion Temperature, PIT, and this ethoxylate emulsification trick is called the PIT technique.

Unfortunately the PIT technique only applies to ethoxylates. Heating ionics decreases the HLD and heating APGs causes no change. So to make efficient emulsions requires other ways of getting to HLD=0. We need to understand the NAC approach and fish diagrams before we can find some fascinating ways to do this effectively. But one obvious way is to use the Cc mixing rule. Start with a high Cc surfactant which gives HLD=0 then rapidly add the low Cc surfactant to get HLD to a negative O/W regime. One of the many problems of the dominance of HLB thinking, along with the notion that the PIT technique is something special, is that this rather obvious strategy for making efficient emulsions with ionics or APGs has hardly been explored. I have termed the generalised methodology PIF (Phase Inversion Formulations), of which PIT is just one

subset. I don't care if the PIF terminology catches on or not. I do care that the generalised methodology catches on.

## 2.2    A waste of good surfactant

Apart from the times when the surfactant is being used to carry some oil within an aqueous environment (laundry and other solubilization activities), any surfactant that is not at the interface is a waste.

Surfactants with low $C_c$ values are a delight to use in water, because they are so soluble or, rather, they happily create large micelles. Surfactants with high $C_c$ values are difficult to use with water, and are much more readily added by dissolving within the oil, where they happily sit as inverted micelles. Such happiness in aqueous or oil environments is precisely what we do *not* want. We want the surfactant to be unhappy in both environments and therefore forced to sit at the interface doing interesting things, such as creating emulsions.

Because "water" and "oil" mean very different things (depending on salinity and EACN respectively) there is no possibility of a perfect surfactant that can always be balanced at the interface. The failed ideas of CPP implied that the correct tail volume, head area and tail length would give the perfect balance. This ignores the fact that the balance is required in the presence of the oil which likes to go into the tail region and swell it. So the oil changes the packing parameter, PP and the idealised, oil-free CPP is totally irrelevant.

The reason for bringing in PP is that when HLD=0 the interface has minimum curvature, equivalent to PP=1. This makes intuitive sense. Strongly O/W emulsions have a large negative curvature (using the HLD sign) and strongly W/O emulsions have a large positive curvature. So when HLD=0 there is no curvature. This is rather hard to imagine. It does not mean that a test tube has a line down the middle between the oil and water – that is just a very bad emulsion. At the limit of HLD=0 the oil and water form a totally homogenous mix with no obvious interface. Instead there is a vast fractal network with a net curvature equal to zero. For all the oil to be in contact with all the water requires a lot of surfactant at the interface. Which brings us back to the point of this section – at HLD=0 the surfactant is equally unhappy in oil and water so tends to be, efficiently, at the interface.

Unless we require crystal clear microemulsions that require lots (10's of % for 50:50 O:W mixes) of surfactant, or specifically require ultra-low interfacial tensions (as discussed in the section on EOR, Extended Oil Recovery) the HLD=0 state is useless in itself. Its main use for those formulating real emulsions is as a reference point in complex surfactant space. We know three things when HLD=0:

- we know unambiguously where we are;

- we are using surfactants most effectively;

- we know which direction in which to move (by altering any of the parameters in the HLD equation) to achieve the goal of an efficient emulsion.

There is a lot of irrational dislike of HLD theory and of microemulsions. One common abusive remark is that "HLD is only good for microemulsions, and microemulsions are useless because they require a lot of surfactant". The first part of the statement is simply false. The second part is true not because microemulsions are weird and demand lots of surfactant at the interface but because microemulsions are made of very small emulsion particles and the surface area to volume ratio of small particles is greater than for large particles. Nanoemulsions, which are intermediate in size between microemulsions and conventional emulsions require quite a lot of surfactant because their surface area ratio is intermediate. The [https://www.stevenabbott.co.uk/practical-surfactants/EmuSA.php](https://www.stevenabbott.co.uk/practical-surfactants/EmuSA.php) app discussed later in the emulsions chapter performs the necessary calculations.

## 2.3    Extended Oil Recovery, EOR

Funding for academic research on surfactants is relatively difficult to find as the subject is no-longer glamorous, even though it remains of great importance. Funding for HLD work has, at times, been relatively plentiful because it can be of huge importance to the oil industry – though its importance depends on the price of a barrel of crude.

Approximately 1/3 of the oil in a typical well can be extracted by pumping. The other 2/3 will not come out unaided. It might seem a good idea to drill a hole on the opposite side of the well and push the oil out with gas or water. This technique rapidly hits a problem. Although the fluid can sometimes cause a "plug flow" of oil, or mix with the oil and produce a mixture that comes out of the ground, more often the fluid punches a hole through the oil and the only thing that comes out of the ground is the fluid that was pumped in.

As instinct would suggest, if the water is flowing faster and/or it is more viscous (e.g. from a polymer additive) then more oil is pushed out. This requires considerable energy once the "easy" oil has been pushed out. There is a graph that describes the efficiency of this removal. It plots fraction-remaining versus Capillary Number which is a much-used concept in fluid flows. The capillary number for a flow of velocity U, viscosity $\mu$ and interfacial tension $\gamma$ is given by $U\mu/\gamma$:
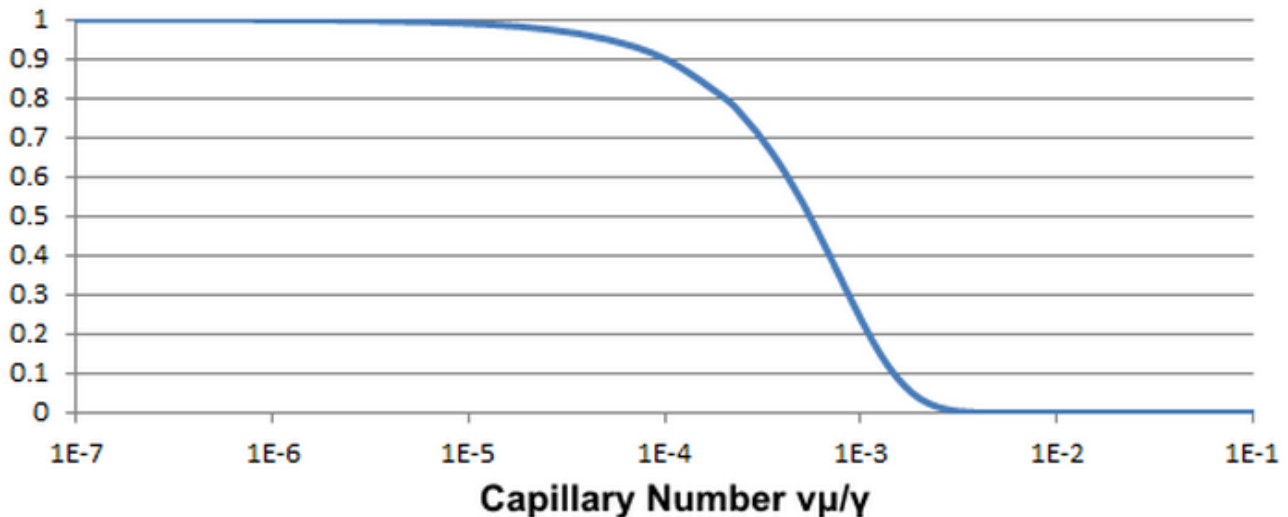
## Fraction Remaining versus Capillary Number



**Figure 2-3 For good EOR a Capillary Number greater than ~1E-4 is required. This can be done (at the penalty of high energy costs) with high velocity or viscosity or elegantly with a super-low interfacial energy.**

Only above a critical capillary number does the fraction remaining fall to a low level. Although this high capillary number can be reached with large v and/or μ, the alternative strategy is to reduce γ to a very low number in order to create a very high capillary number. At HLD=0 γ can be super-low, below 0.01 mN/m, so with relatively low viscosity and flow velocity (i.e. low energy input) it is possible to extract a lot of oil.

Suppose you used inspiration or HLB to come up with a choice of surfactant. You would then spend 3 months pumping a lot of surfactant solution down a hole and wait for the results. Some millions of dollars later you will find a rather dirty surfactant solution coming to the surface with little oil. Your popularity with the accountants might be diminished.

The alternative is to use HLD. You know the temperature (relatively high) and the salinity (definitely high) of the well and you have carefully measured the EACN of the oil. It is then relatively straightforward to create a blend of your favourite (low cost) surfactants tuned for maximum EOR efficiency. 3 months after you started to pump this surfactant blend down the hole you find, to the satisfaction of the accountants, a copious stream of highly oily water coming to the surface. There is one remaining challenge: to separate the oil and water. This requires the lowest possible interfacial energy which, conveniently, you have already created in order to extract the oil. The oil separates out nicely and everyone is very happy.

This is a gross simplification of a deeply serious research and engineering challenge of huge proportions. However, the essence is correct. Each well has different temperatures, salinities and oils so each EOR operation requires a fine-tuning of the surfactant via HLD. It is no coincidence that the leading academic

lab for research on HLD has been the Universidad de Los Andes in Mérida in Venezuala. The lab has for many years been run by Salager who first developed HLD theory and the lab is funded by the world-wide oil industry.

It is interesting to note that where surfactant theory is really important (billions of dollars at stake) the world-wide community uses HLD theory. Where the theory is not so important (e.g. in the cosmetics industry where many formulations are copies of other formulations) the world-wide community uses ideas long-since shown to be erroneous.

The final chapter returns to the topic of EOR and provides an app discussing the critical capillary curve (along with Bond number effects).

## 2.4    Measuring Cc and EACN

A key strength of HLD is that the Cc and EACN values can be measured relatively easily. For those who have a high-throughput system it is particularly easy and the EOR industry makes good use of such automated methods – they simply have no choice.

The discussion on the measurement techniques also reveals a lot more about the meaning of HLD theory.

Let us, at first, cheat. Take a nonionic surfactant known to have a Cc of 1.4. Let us choose a temperature of 25°C so there is no temperature effect, and choose pure water with a salinity of 0. Using the HLD calculator it is very easy to find that for this system HLD=0 if the EACN=8.5 because 0.17*8.5=1.4 and the HLD equation has a negative term for the EACN: 1.4 -1.4=0.

8.5 happens to be the average of 1 and 16. The EACN of toluene is 1 and that of hexadecane is 16. From the EACN mixing rule, a 50:50 blend has an EACN of 8.5. So we know that something special will happen in a test tube containing our surfactant, water and the toluene:hexadecane blend.

Now set up 9 test tubes, each with 5% of the surfactant and each with a 50:50 oil:water mix, but where the oil systematically varies from pure toluene on the left (EACN=1) to pure hexadecane on the right (EACN=16). Having given the tubes a good shake and allowed them to settle out, let's see what they look like. Although the app is an idealised version of this experiment, I and others have done exactly this experiment and the real-life tubes follow the general scheme:
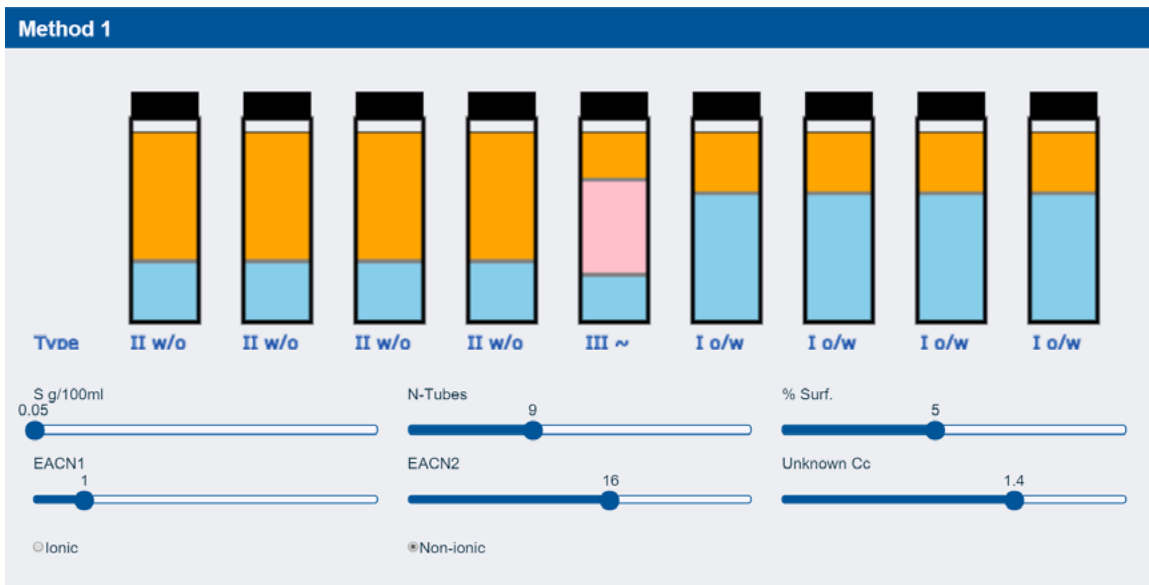
The oil phase is always on top and in the apps is always coloured orange. The water phase is on the bottom and is blue. The tubes on the left show two phases, as do the tubes on the right. The difference is that on the left the oil phase is larger and on the right the water phase is larger. This immediately tells us something significant. The only way the oil phase can be larger is if it contains lots of water. And the only way the water phase can be larger is if it contains lots of oil.

This means that the tubes on the left contain a W/O emulsion and those on the right contain an O/W emulsion. We have to be careful with our words. "Emulsion" tends to mean a white cream of oil in water or water in oil. Here we have clear phases, though undeniably either the oil contains lots of water or water contains lots of oil. These are really microemulsions and the O/W are called (Winsor) Type I while the W/O are called (Winsor) Type II.

What about the tube in the middle? Here there are three phases and clearly both the water and oil phases have lost out to whatever is in the middle. This is the (Winsor) Type III microemulsion. Because HLD~0 and the interfacial tension is ~0 there is no curvature so there is no obvious way to describe what this phase is. The best description is a fractal blend of surfaces slightly curled one way or the other, leading to an average curvature of 0. If a lot more surfactant was present then all the oil and water would be in the single phase which would be called Type IV – though such high levels of surfactant are inappropriate for measuring Cc values.

In reality, when we're not cheating by knowing the answer, the Type III might be in tube #2 and via the Cc slider it is easy to confirm that the Cc must be 0.6. Or the Type III might be in tube #7, which implies a Cc of 2.1. All too often an experiment reveals all tubes looking approximately the same, with no sign of a Type III. What can be done then? It depends. If all the tubes have a larger

volume of the oil phase we at least know that the Cc is higher than the value equivalent to the right-most tube, so the scan has to be repeated with an oil with a higher EACN, such as squalane. If all the tubes show a larger water phase we know that the Cc is lower than the left-most tube and, perhaps, adding some salt (try the salinity slider) will just get the Type III into that left-most tube.

There are alternative scan techniques, described in the app. One known Cc and one known EACN can be used, with the scan being a step-wise variant of the surfactant blend. Or it can be done with a known EACN and a systematic addition of salt.

Although these scans are conceptually easy, it involves some judgement and experience to find the best scan for any given surfactant. For those of us who are formulators it is a big question whether it is worth developing the in-house knowledge or whether it should be outsourced to someone with experience and a suitable robotic system. Those who supply surfactants should be ashamed if they do not have a high-throughput setup to provide Cc values for each of their surfactant products. At the time of writing most surfactant suppliers are, for various reasons, refusing to issue the Cc values of their products. Happily, the mood of the surfactant community is changing and once a few more suppliers start to provide Cc values, the rest will have no choice but to follow.

For *internal* purposes, all surfactant suppliers are strongly recommended to adopt Cc measurement techniques. It is no secret that most commercial batches of surfactants are broad mixtures, somewhat ill-defined. There is nothing wrong with this, because none of us want to pay for pure surfactants. If each manufacturer keeps a modest stock of a deliberately low-Cc and deliberately high-Cc variant of each of their products then each fresh production batch could have its Cc measured very rapidly (just a few tubes, optimised to determine the Cc within the manufacturing range) and then some low- or high-Cc material added to bring each batch to the nominal value. This would greatly reduce the number of surprises their customers find when the "same" product happens to vary in a direction that creates problems for their own formulation.

It is similarly straightforward to measure the EACN of an oil.
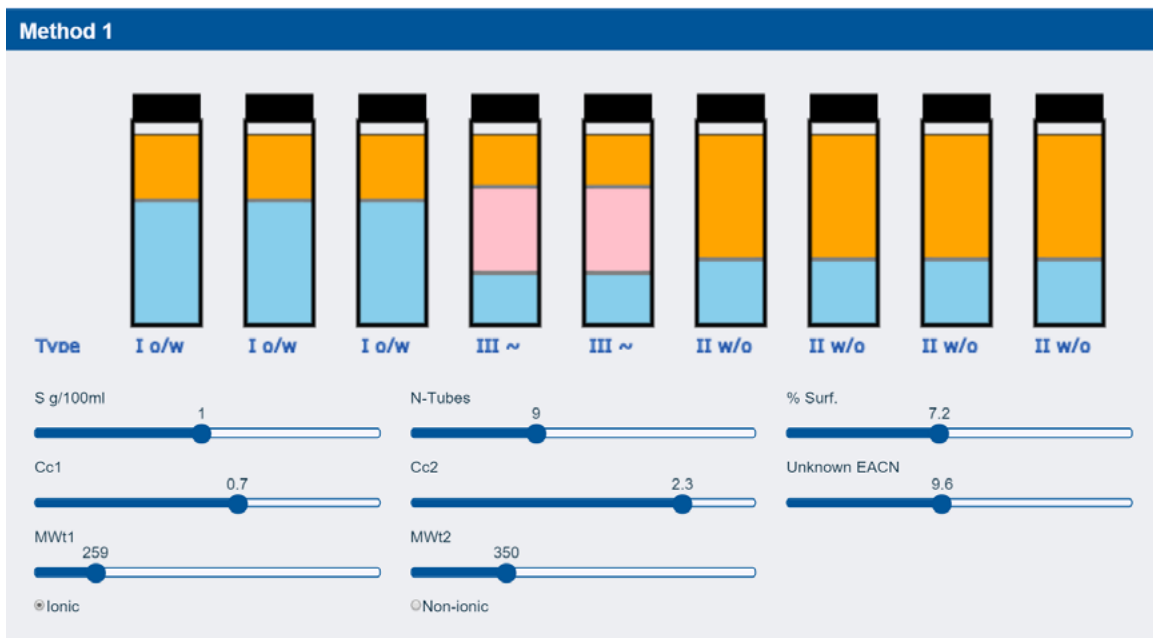
In this case the scan involves the ratio of two known Cc surfactants and the Type III straddles two tubes. The app shows that the scan can be done with a known EACN and known Cc or with a known Cc and a salt scan. The PIT-shift method shown further down that app page is really an admission of the limits of HLD and is discussed in a later chapter.

### 2.4.1 Scan tricks

It is a universal phenomenon that *your* specific scan, for which you had such high hopes, fails. You might fail because you made the wrong estimate of the Cc or EACN and so all tubes look the same, telling you almost nothing. Or your pair of surfactants might dislike each other intensely, as is the case with SLS and AOT, which is why SDHS-AOT scans are more common for ionics. So, what do you do? First, don't despair. We've all been there. And because you have the apps, salvaging something from the failure might still be possible (though if it's the SLS/AOT mess, just give up). Here are some of the standard tricks.

- It the volume of the "water" phase is slightly higher than the oil phase, this means that you have oil-in-water (and vice versa). Using the relevant app, change your guess for the Cc or EACN so that all your tubes show all O/W (or W/O if it's the other way round). Now see if a change in temperature (especially with ethoxylates) will shift things in your favour. Or (if ionics) can you add some salt to each tube to push them in the right direction? It's amazing how you can rescue some "failed" scans with a bit of temperature or salt. If a Type III appears, from your known adjustment you should be able to get a "good enough" value and, if necessary, create a new scan in the right area.
- Set up a scan with an easy oil and surfactant combination so you get a nice Type III in the middle tube. If you are trying to find a Cc, add a small amount

of your unknown to each tube and see which way the Type III moves. If you are trying to find an EACN, add 10% of the new oil (maybe you can start your scan with 40:60 O:W) and see which way the Type III shifts. In each case, with the app you can get a rough estimate of the value and can then plan a more precise scan. This sort of approach is often called a "composition scan" and is becoming increasingly popular for those who have a set of rather well-behaved systems which can be used as reliable bases for such scans. The Fast scan method, below, is usually a variant of this.

- If you have genuinely no idea where the Cc or EACN might be, then set up a very broad scan with just a few tubes to get some idea (even by seeing if the tubes are Type I or II at the extremes) before planning a precise scan.
- If your tubes are taking too long to separate (or maybe just do this as routine), put them in a centrifuge for a while. How fast and how long? That's a matter of judgement, but in a typical lab centrifuge you can't go super fast and most of us get bored waiting more than an hour - at which point you can at least check if things are getting better.
- Although it used to be a reasonable rule to "throw in a few % of sec-butanol" to help phases to separate without affecting the HLD, we now know that sec-butanol isn't always neutral. But it's better to lose some accuracy and gain insights if the sec-butanol takes separation time from weeks to hours.

### 2.4.2  Phase volumes

The tube with three phases gives some important information beyond the fact that HLD=0. Different surfactants have different solubilizing powers so, for a given % surfactant, the size of the middle phase can change from small to large. It is a good habit to take photographs of the tubes for later reference (robotic systems should do this automatically) and with good lighting and some simple image analysis the phase volumes can be determined automatically. The factor controlling these volumes, the $\xi$ parameter, is discussed later, along with an app to extract the key information from the volumes.

### 2.4.3  Fast scans

The scans using tubes and looking at the phases are the gold standard. But often they are too much work, especially if we want quick scans and/or only have small volumes of samples. There are two alternatives.

The first builds on the fact that many of us have noticed that the tube where HLD=0 can be spotted very quickly during the tests, though it can take hours or days to confirm this. It somehow behaves differently from all the others. Although many of us have noticed this phenomenon, it has never been formally supported as a method for determining the HLD=0 point.

Fortunately, a paper by the Acosta team[16] has formally shown (at least for ethoxylates!) that a method based on this common observation is much faster and just as accurate. The key to the fast method is the fact that when HLD=0, the interfacial tension (IFT) is minimum and therefore the stability of the emulsion (not the microemulsion!) created by mixing the oil/water/salt/surfactant system is minimum. Indeed for demulsification this is standard trick is to tweak the system to HLD=0 and the Emulsion Stability theory discussed later in the book (based on work from the Acosta team) shows the dramatic reduction when HLD=0.

So what is the fast technique? Take your N tubes, shake them all at the same time then look at the tubes for seconds or a few minutes to see which one (if any) is showing distinct signs of emulsion instability. Or, to quote the paper itself and adopt a more rigorous and formal definition of the technique:

*To determine S\* using the emulsion stability method, all vials contained in one salinity scan were simultaneously mixed and left to equilibrate at room temperature on a flat surface. The time to separate the excess phases (i.e., the time when the separation front reached its stable position) in each formulation was recorded, and the optimal salinity (S\*) was identified as the middle phase bicontinuous IE where the excess phases separated in the shortest time. It is important to note that all the vials contained in the scan must be subjected to the same mixing conditions since the coalescence of particles can be affected by the mechanical energy input.*

The paper methodically carries out all the standard techniques (including measuring IFT) to determine the HLD=0 points by those accepted methods then compares the results to those obtained by the fast method. The only difference seems to be that the fast method is, if anything, slightly more accurate. One can readily imagine circumstances where it will be slightly less accurate and, of course, there is a loss of the extra information (phase volumes) from the classical methods.

Now that this fast method has been validated, the implications are game changing. There are 1000s of surfactants and oils that need to be measured and the amount of work via the classic methods discourages most of us who have cost and time constraints. The fast method should allow a well-set-up lab to measure large numbers of values in a comparatively short time. Although the paper is "only" about measuring Cc of ethoxylates, the basic science tells us that it will work for other surfactant classes and for measurement of EACN.

---

16    Silvia Zarate-Munoz, Felipe Texeira de Vasconcelos, Khaing Myint-Myat, Jack Min-chom, Edgar Acosta, A Simplified Methodology to Measure the Characteristic Curvature (Cc) of Alkyl Ethoxylate Nonionic Surfactants, J Surfact Deterg, 19, 249-263, 2016

There is a trick within a trick. Some surfactant systems don't give clean separations with the oil and salinity used in the tests. In these cases it is best to use a fast scan system which already shows clean, quick results. Here is how to do it.

Set up a scan with a known, well-behaved, surfactant and with either a salinity scan or an EACN scan, adjusting parameters such that the fast method gives you separation in the central tube when you use 100% of your known surfactant. Now repeat with a replacement of, say, 50% of the known surfactant with the unknown. Does this send the HLD=0 point out of the range of the scan? No problem, just try it with 25%. Does it, instead, not change the HLD=0 point? No problem, its Cc is identical to the reference. More likely is a shift to a higher or lower tube from which you can quickly determine the Cc value of the blend by using the known salinity or EACN in that tube and using the HLD app. From the Cc of the blend it is a simple matter to find the Cc of the mix and then from the Cc calculator app and from the known Cc, the known MWts of the two surfactants and the known ratio, the Cc of the unknown can be found by adjusting the slider of the unknown till the calculated value equals your measured value. To be sure (and to accommodate any non-ideal mixing behaviour) you can try a different % of the unknown and use the new tube position to re-calculate the Cc.

The second fast method is to use a conductivity scan. If you have a small conductivity probe (or impedance probe) the experiments can be done on small volumes.

Here, in general, is how to do it. The assumption is made that the water contains at least a small amount of salt so that a continuous aqueous phase is easily identified by its significant conductivity.

We know that in Type I, oil in water, we have a conductivity of the bulk (saline) water reduced somewhat (we might not be able to measure the small difference) by the non-conducting oil drops. We know that in Type II, water in oil, we have no significant conductivity. And at Type III we have some weird phase that should show conductivity significantly different from Type I.

So all you do is set up your tubes with little stirrers to allow general mixing, put in your probe and see which phase you are in. This is especially useful for scouting scans when you have no idea where you are in HLD space. When you think you have found the HLD=0 point then you can make up a few conventional scan tubes around that space to look more closely at what is happening and, after settling, to get the phase volumes which tell you about how efficient the surfactant is.

### 2.4.4 Silicone oils

In terms of contact angles, silicone oils are rather more remote from water than, say, hexadecane. But this does not explain the difficulty in general of emulsifying silicone oils except through the use of silicone surfactants. The key is that the EACN is, to a large extent, a measure of the free energy of interaction of the surfactant chains and the oil. This depends on two factors. The entropic factor explains most of the effect of the oil chain length. Hexadecane has a higher EACN than hexane because there is less increase in entropy when one of the molecules is larger. The enthalpic factor depends on whether oil-oil interactions are preferred to oil-surfactant interactions. Obviously an ester is more compatible with itself than with a surfactant hydrocarbon chain so an ester is less likely to mix with the surfactant. For those who are familiar with Hansen Solubility Parameters, the reason that cyclohexane is different from hexane (EACN=3 instead of 6) is that the δD value (Dispersion) of cyclic molecules is higher than linear ones, and benzene has an EACN=0 because the aromatic δD is even higher. This leads us, at last, to the silicones. The siloxane chains show very little Dispersion energy and their δD value is much lower than any alkane. So they have little compatibility with the surfactant chain. Of course they have little compatibility with water so given a choice they will prefer the surfactant chain. But their interactions with the chains is not a happy one and when we come to the ξ parameter we will see that silicones cannot be effectively controlled by the surfactant chains so the emulsifying potential is very low. A silicone surfactant chain interacts nicely with silicone oils so the ξ parameter is large, which explains the general truth that silicone oils need silicone surfactants.

### 2.5    Winsor R ratio and Packing Parameter

HLD is a quasi-empirical formula (though firmly based on thermodynamics) that allows formulators to formulate. Because it has plenty of problems (which have never been masked by those who have developed the technique) it sometimes gets dismissed as either inferior to or "just an equivalent of" the Winsor R ratio or the Packing Parameter. It is certainly inferior to neither of them and if the word "just" is removed from the accusation then it can proudly say that it is at least intellectually equivalent to the R ratio and/or PP. Its advantage is that the key parameters, Cc and EACN are readily measured and so a practical formulator can use it. The R ratio and PP are both splendid ideas. It just so happens that none of us can really formulate using them.

The R ratio is the balance of the "interaction strengths per unit area" A that compare what happens to surfactant S, oil, O, tail portion t and head portion h given by:

$$R = \frac{A_{SO} - A_{OO} - A_{tt}}{A_{SW} - A_{WW} - A_{hh}} \qquad \text{2-3}$$

An initial guess at R would take the ratio $A_{SO}/A_{SW}$, i.e. the relative attractions of the surfactant for the oil and the water. That was Winsor's original R ratio. The extra terms represent modifications to each of those two terms. The effect of the surfactant-oil attraction is reduced by the size of the oil:oil and tail:tail terms while the surfactant-water attraction is reduced by the size of the water:water and head:head terms.

This makes good scientific sense. As R changes from <1 to >1 the system changes from O/W to W/O with R=1 representing the balanced, zero curvature state. Sadly no practical implementation of this formula has been possible. An important paper by Fraaije[17] that explores the relationships between different type of curvature and torque models is more critical; it states that "Many other quantifications of R have been attempted over the last few decades … but so far all have failed" The paper then points out that, for a reason unknown to Winsor at the time, R *cannot* work. This reason is rather important and is one reason that the Fraaije paper is mentioned several times in this book.

All our diagrams of surfactants have nice heads and nice tails. It then becomes natural to think that we can divide the surfactant into a head region and a tail region, as required by the R formula. The paper points out that this is wrong for two reasons, which means that the R formula is fatally flawed. First, even a simple surfactant presents a rather disordered interface so it is not clear where to draw the line between head and tail. Second, many surfactants have no clearly-defined head or tail region. So, for example, if you want to think about the balance of forces around the head/tail interface, you need to define where the interface is and how far each of your forces is from that interface. The Helfrich torque approach bypasses the issue because the torque is an average of all forces-times-distance and any arbitrary reference line is sufficient for the average to be a constant. The implementation of such ideas require the bending constant $k_c$ and the splay constant $\bar{k}_c$ discussed in the first chapter and the problem of realising the potential of these ideas is our lack of independent knowledge of these values. The point of raising the issue here is to highlight the fact that surfactant science is still feeling its way towards understanding the key issues even of simple oil/water/surfactant systems. The Helfrich torque approach might turn out to be useful in the context of another "curvature" approach, that of Packing Parameters.

So, sadly, Winsor R, although endlessly discussed in the surfactant literature, is of no practical use to us.

---

17   Johannes G. E. M. Fraaije et al, Method of Moments for Computational Microemulsion Analysis and *Prediction in Tertiary Oil Recovery*, Langmuir 2013, 29, 2136−2151

The PP approach takes the same formula as the discredited CPP and brings it to life:

$$PP = \frac{V}{A.l} \qquad \text{2-4}$$

The insight from Tchakalova[18] working at Firmenich is that using some basic assumptions about how oils (plural) and water interact with the surfactant tail and head regions, the overall curvature can be modelled systematically. By actually measuring (e.g. using NMR) where various species are within a microemulsion system the formula can be brought to life. This approach will be discussed in detail when we need to get deeper into the problems and issues with HLD. The fact that the different components are interacting not just with "head" and "tail" but with different parts of the system (e.g. in the region where head meets tail) emphasises the point about torque moments that do *not* require an arbitrary dividing line between head and tail, even though, for simplicity, the model assumes the usual sharp distinction between them. The extra importance of the Tchakalova approach is that adding an extra oil, such as a fragrance molecule, has a double effect on PP. First, it changes the curvature by, for example, going into the tail. Second, by going into the tail it drags more oil into the tail, changing the curvature further.

So although PP is a very powerful idea, it is in practice so difficult to implement that it is generally wiser (especially for practical formulators) to return to HLD. Indeed, Tchakalova makes it clear that the work required for her experiments was so enormous that it is unlikely to become a routine tool and her follow-up papers represent ingenious variations, discussed later, around EACN.
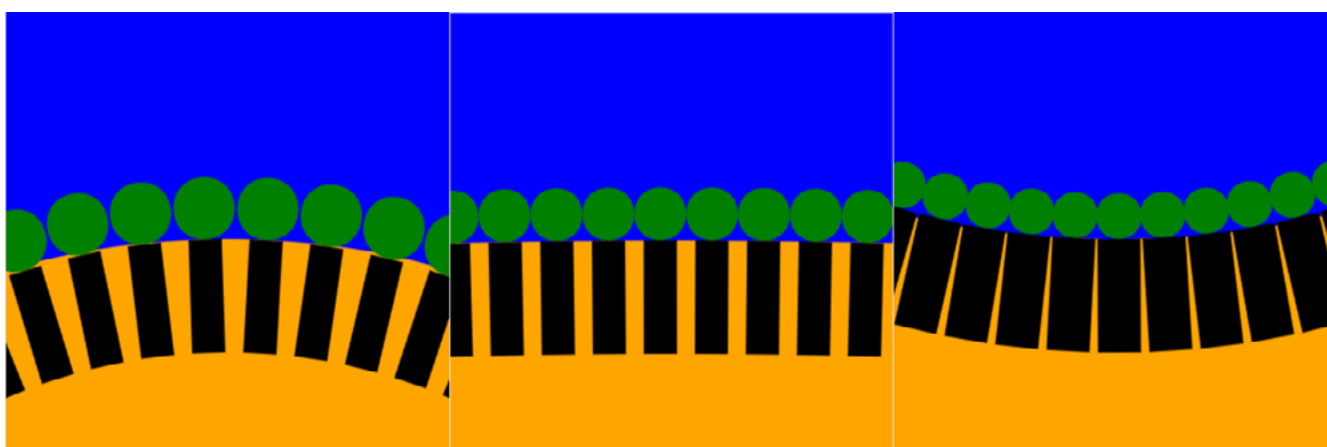


**Figure 2-4 The PP changes from 0.85 thru ~1 to 1.15 via systematic changes to the tail volume.**

18   Vera Tchakalova, et.al. Solubilization and interfacial curvature in microemulsions I. Interfacial expansion and co-extraction of oil, Colloids and Surfaces A: Physicochem. Eng. Aspects 331 (2008) 31–39

## 2.6      Who cares about microemulsions?

First, what is a microemulsion? It is unfortunate that the "micro" term was applied to them when they were first discovered. In fact they are nanoemulsions – nano-sized droplets of oil in water or water in oil. We cannot call them "nanoemulsions" because that term is now understood to mean conventional emulsions with nano-sized drops. Only when there is an especially efficient system do the droplets reach a size large enough to give some scattering effects, a sort of opalescence along with an attractive blue or yellow tint. When measuring Cc or EACN values it is often a useful idea to shine a laser through a phase in a test tube to confirm the significant scatter that is characteristic of a big microemulsion.

Standard emulsions (and that includes nanoemulsions) are thermodynamically unstable – in the long run all such emulsions will separate into oil and water phases. Microemulsions are thermodynamic. They can be created with a very modest amount of mixing (a simple inversion of a test tube often suffices) and will never separate.

As mentioned previously, for some reason, microemulsions tend to attract derision. "They need large amounts of surfactant" and "They are only stable at a constant temperature" are insults regularly thrown at them. As discussed, the first is trivially true because there is so much surface area when droplet sizes are so small, though by understanding NAC theory and tuning the system the amount of surfactant can be greatly reduced. The second is strongly true for ethoxylates (with their $\alpha=-0.06$ thermal coefficient), weakly true for ionics ($\alpha=0.01$) and untrue for many sugar surfactants such as APGs.

Some industries positively require microemulsions. EOR cannot function without the super-low interfacial energies. An increasing number of cosmetic formulations (which have to contain lots of surfactant anyway) use the opalescence of a microemulsion as a symbol (entirely irrational) of quality. The world of "solvent" cleaning has been transformed by the use of safe, water-insoluble oils such as dibasic esters in a single-phase microemulsion format where the oil, surfactant and water each contribute to the overall cleaning performance.

Emulsion formulators in industries such as cosmetics may have never encountered the idea of microemulsions yet totally rely on the PIT method to manufacture many tons of product. The PIT method is simply a subset of the general PIF (Phase Inversion Formulation) technique which takes an emulsion to the Type III state with low interfacial energy in order to create small emulsion particles with relatively low mixing energies, before quenching the system into a conventional emulsion format within the Type I (O/W) or Type II (W/O) regions.

Those who need to break emulsions (e.g. in separation of crude oil from the initial oil & water mix that comes out of the ground, or those creating bio-diesel from palm oil) have found by luck or judgement that coming to the Type III state produces rapid coalescence because the interfacial energy is so low.

Those who just make real emulsions can still learn much from HLD. After all, HLB was invented not for creating microemulsions but for creating real ones. Anything HLB can do, HLD can do much better. So if you want an O/W emulsion, make sure that the Cc of the surfactant blend and the EACN of the oil place you in the "Type I" region. From now on, the scare quotes will not be included, though it is important to explain why they have been added in this paragraph. Type I is a terminology that rightly belongs to microemulsions. So strictly speaking you cannot have a Type I macroemulsion or nanoemulsion. However, because the natural direction of curvature in an O/W emulsion is the same as in an O/W microemulsion (and similarly for Type II and W/O) it seems reasonable to use the same terminology for both. After all, Winsor R ratio and PP are used in the real emulsion communities, unaware that these, strictly speaking, are thermodynamic ideas that do not apply to the kinetic stability of typical emulsions. The fact that the PIT technique (very much a microemulsion approach) is used every day to make standard emulsions shows how useful it is to use a technically imprecise language across all emulsion phenomena.

The intuitive link between microemulsion thinking and macroemulsions has had its ups and downs. At one time a naïve link was made (by the great Langmuir) between microemulsion curvature (though it wasn't called that at the time) and the curvature of macroemulsions – the so-called Wedge theory. A devastating critique (by the great Hildebrand) showed that this link was totally false. But acting as if the link was correct seemed to work well. A more recent take on Wedge theory, discussed later, explains why there really is a link between the tight curvature of a microemulsion and the gentle curvature of a macroemulsion drop. So we can be confident via multiple chains of reasoning that HLD-inspired macroemulsion formulation is a valid approach.

## 2.7    HLD is not enough

After having said so many good things about HLD, and having shown how poor its simpler competitors are, it might seem strange to announce that it is not enough.

The truth is that all those Cc and EACN measurements have been done with nice 50:50 O:W mixes and the idea that anything different might happen at other O:W ratios is inconceivable within the HLD equation – there are only terms for the surfactant (blend), the oil (blend), salinity and temperature.

HLD, therefore, is an excellent starting point for navigating through the complexities of surfactant space. It is *necessary* but very far from *sufficient*. To

increase its power requires some deeper insights into the meaning of curvature. For that we need to add the idea of Net Average Curvature, NAC.

# 3 NAC – Net Average Curvature

The HLD theory took a significant leap forward with the publication by the Sabatini group in Oklahoma of the first paper[19] describing the Net Average Curvature model, NAC. The model has then been developed further particularly by Acosta[20]. It allows many more aspects of the interactions between water, oil and surfactant to be modelled using just a few extra ideas along with two parameters we have already encountered, head area A and tail length l, plus a new one, ξ (xi).

What is curvature? It is simply 1/Radius. Usually we discuss things in terms of radius, but on other occasions it is more intuitive to talk in terms of curvature. So for an oil drop of radius $R_O$ the curvature is $1/R_O$ and for a water drop of radius $R_W$ the curvature is $1/R_W$.

The starting point for NAC is the net curvature, H, of the interface, i.e. the difference between the curvatures $1/R_O$ and $1/R_W$. H is equal to HLD/L where L is a length that is proportional to the length of the surfactant tail. There are good thermodynamic reasons why this should be so:

$$H = \frac{1}{R_o} - \frac{1}{R_w} = \frac{-HLD}{L} \qquad \text{3-1}$$

Indeed, it has long been known that in the Type I domain where HLD<0 $R_O$ (measured, for example, via light scattering) correlated well with HLD. The factor of L which is typically 1.2-1.4 longer than the extended chain length $L_c$ is no great surprise – it is intuitive that a longer chain could enable more oil to be trapped, and the longer L the smaller the curvature so the larger the radius. The chain length $L_c$ in Å is readily estimated by the following equation based on the number of carbon atoms in the chain, $n_c$:
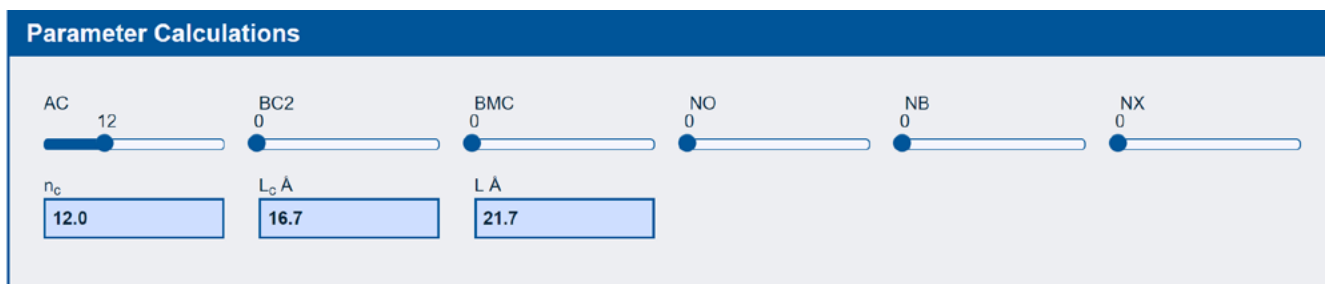
$$L_c = 1.5 + 1.265 n_c \qquad \text{3-2}$$

Many surfactants do not have chains made just of linear alkanes, so $n_c$ is calculated from the number of Alkane Carbons in the chain AC, the number of Branched Carbons at the 2- and middle positions, BC2 & BCM, the number of EO or PO groups, NO, and the number of Benzene (NB) and o-xylene (NX) groups. There is some disagreement about the precise formula, but an indicative version is:

---

19   Edgar Acosta et al, Net-Average Curvature Model for Solubilization and Supersolubilization in Surfactant Microemulsions, Langmuir 2003, 19, 186-195. The terminology is based on the old SAD form of HLD

20   Edgar J. Acosta, The HLD–NAC equation of state for microemulsions formulated with nonionic alcohol ethoxylate and alkylphenol ethoxylate surfactants, Colloids and Surfaces A: Physicochem. Eng. Aspects 320 (2008) 193–204

**Parameter Calculations**

| AC | BC2 | BMC | NO | NB | NX |
|----|-----|-----|----|----|----|
| 12 | 0 | 0 | 0 | 0 | 0 |

| $n_c$ | $L_c$ Å | L Å |
|-------|---------|-----|
| 12.0 | 16.7 | 21.7 |

**App 3-1** https://www.stevenabbott.co.uk/practical-surfactants/NACMore.php

As an example, a linear nonylphenyl tail would give $n_c$ = 9+3.5 = 12.5 so $L_c$ = 17.3 (so L ~ 23Å) and a mid-ethyl-branched nonylphenyl would be 7+2*0.71+3.5 = 11.9, giving L ~ 15.8Å.

One problem with the net curvature equation is that the water is a continuous phase which doesn't obviously have a radius. So what *is* the radius of the continuous phase? The answer is to say that for a given volume of water, $V_w$ and a given Total Surface Area of surfactant, TSA, the *equivalent* radius can be calculated via:

$$R_w = \frac{3V_w}{TSA} \qquad \text{3-4}$$

In turn the TSA can be calculated via:

$$TSA = C_s.V_w.N.A \qquad \text{3-5}$$

$C_S$ is the surfactant concentration, $V_w$ is the volume of water (so in fact $V_w$ cancels out), N is Avogadro's number and A is the area per surfactant molecule. This finally gives us:

$$R_w = \frac{3}{C_s.N.A} \qquad \text{3-6}$$

At low TSA, $R_w$ is so large that $1/R_w$ is essentially zero, which means that the worries about the $1/R_w$ term are misplaced. The reason for including it is that it makes the system consistent across the whole range of O/W and W/O – the same equation applies at the other extreme of W/O emulsions where the worry is the meaning of $R_o$.

Just as we have the question of what L is, the estimation of A is also an issue. For simple systems A can be estimated satisfactorily. When "hydrophilic linkers" (e.g. alcohols) are added the A tends to increase (by definition a hydrophilic

linker goes to the head area), usually to the detriment of overall surfactant efficiency. The whole linker issue is discussed in the Problems chapter.

Armed with the known HLD, an estimate of L and the calculated $R_W$ the theory gives us $1/R_O$ when in the Type I regime (and via a continuous phase $R_O$, $1/R_W$ in the Type II regime). There is an obvious problem with this NC (Net Curvature) theory – when HLD=0, the Type III regime, H=0 so 1/R becomes infinite, not a meaningful value. Clearly something has to ensure that the curvature never reaches zero so that the radii never approach infinity.

To solve this problem it is necessary to add another fact about surfactants. They tend to have a characteristic length ξ which means that the *average* curvature cannot be lower than 1/ξ. So:

$$\frac{1}{2}\left(\frac{1}{R_o}+\frac{1}{R_W}\right)\geq\frac{1}{\xi} \qquad \text{3-7}$$

We thus have a Net Average Curvature model, NAC, which combines the Net curvature (which links to HLD) and Average curvature which links to ξ, which itself can be measured directly via neutron or X-ray diffraction. It can also be measured indirectly from measurements of phase volumes (as discussed earlier in the context of the scans), where a large ξ means a greater relative volume of the middle phase. The average curvature restriction gives meaningful (and correct!) predictions of the amounts of co-dissolved oil and water in the Type III domain.

These equations are so simple and obvious that it is hard to imagine that they can be put to much use. The good news is that they are very useful – and that the apps do all the calculations for you.

There is one further important point. The descriptions of oil drops and water drops are an aid to understanding the theory. They are not to be taken too literally once HLD gets close to 0. The whole system in this region is a complex geometry. NAC cuts through the complexities (which otherwise need discussions of things like Gaussian g numbers) with a simple equation that does a remarkably good job of predicting complex behaviour from the simplest of ideas. NAC *can* be described in terms such as Helfrich torque theory and it is possible that in the long term such approaches will prove to be a superior way of predicting emulsion behaviour. Till then, NAC is more than good enough for many of us.

## 3.1    What can NAC *do*?

A good way to answer that question is to show that adding NAC to HLD makes it possible to calculate a fish diagram. The app shows us what a fish diagram

is, and the power of the approach is clear because the data are calculated purely via HLD-NAC. I had already been convinced by the power of HLD; it was reading an Acosta paper on predicting fish diagrams[21] that made me convinced of the power of NAC. To understand what it going on, first the diagram is explained then the way of calculating it from HLD-NAC is described.



**App 3-2** https://www.stevenabbott.co.uk/practical-surfactants/Fish-and-More.php
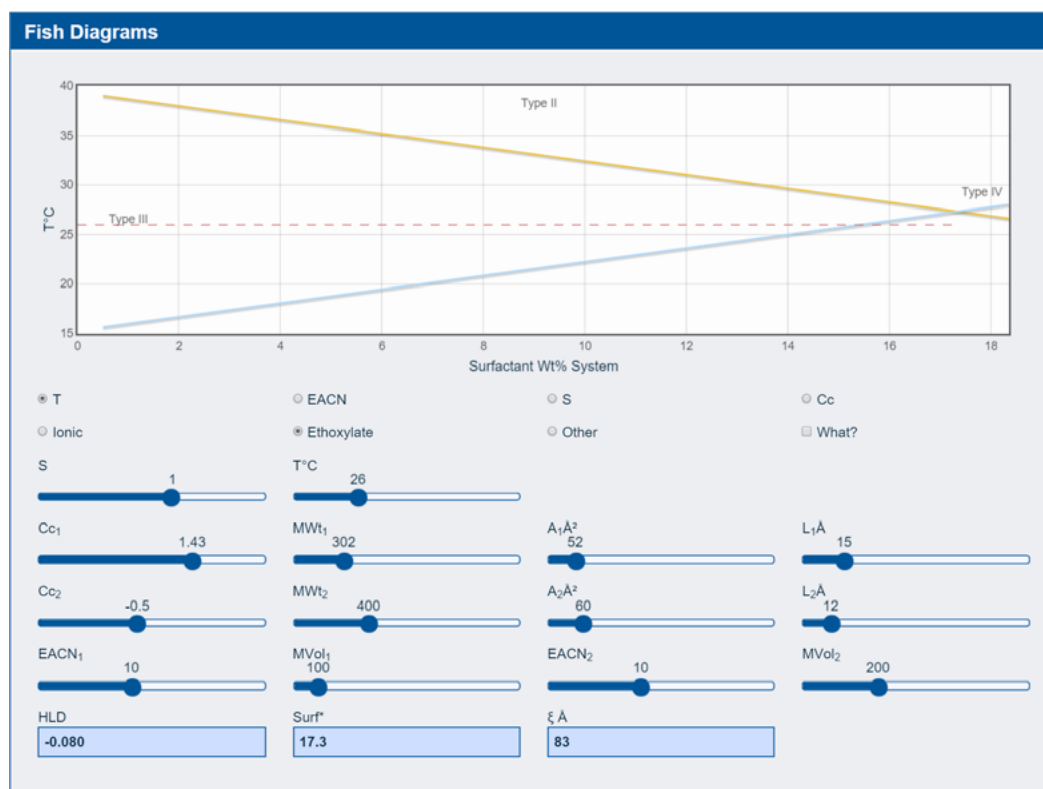
This app happens to be rather powerful, which means it is rather complex. For our current purposes ignore the $Cc_2$ line, the $EACN_2$ and $MVol_2$ values and the temperature T. So the setup is a surfactant with a Cc of 1.43 and MWt of 300 with a head area of 52Å² and an effective tail length of 15Å. The salinity is 1g/100ml.

If this were an experimental fish plot, there would be a set of tubes of 50:50 O:W with 1%, 2% …20% surfactant. The tubes would be cooled to, say, 10°C and the temperature gently raised, observing what is happening between the phases. Imagine the tubes shown in the Cc and EACN measuring apps. In this example, as the 2% surfactant tube is warmed, at 15°C the neat Type I system, with a larger water phase as this is O/W, becomes unstable and at 37°C it finally flips into a neat Type II system with a larger oil phase as this is now W/O. The experiment around 17% surfactant shows a distinctive behaviour – at 26°C the whole tube becomes a single phase.

---

21   Edgar J. Acosta, The HLD–NAC equation of state for microemulsions formulated with nonionic alcohol ethoxylate and alkylphenol ethoxylate surfactants, Colloids and Surfaces A: Physicochem. Eng. Aspects 320 (2008) 193–204

Missing from this plot is what happens at very low % surfactant. In real fish diagrams the lines curve in towards a point that marks the limit of solubility of the surfactant, i.e. the CMC. When these plots were explored by Cates et al[22], they decided that they looked like a fish with the head to the left (curving to the CMC) and a tail to the right (starting where the blue and orange lines cross). They have been called fish diagrams ever since, even if many of them look like rather distorted fish.

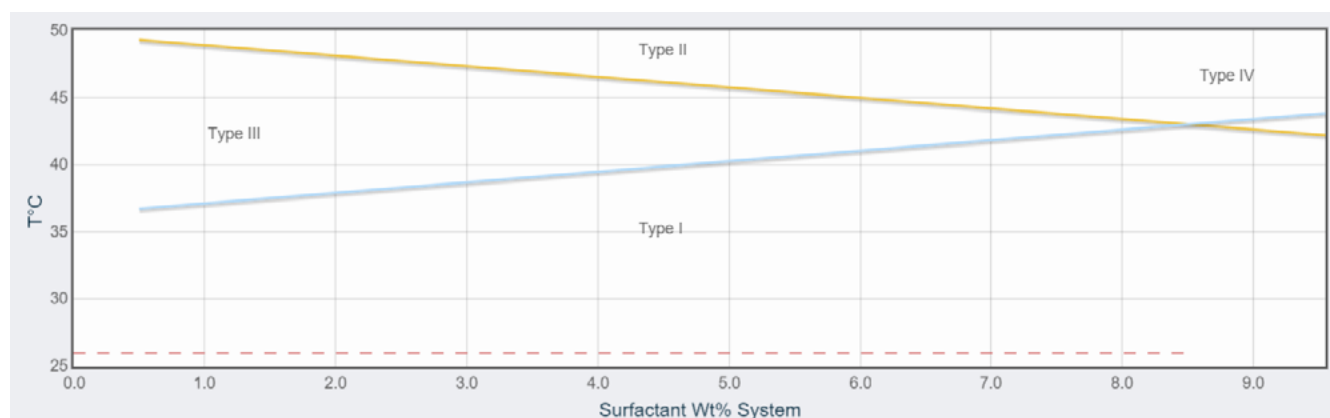If the Cc is changed to 0.5 and the tail is changed to 17Å (please do this live in the app) then the diagram changes:



**Figure 3-1 The same fish diagram but with Cc=0.5 and L1=17.**

It is no surprise that the temperature at which things change has shifted to a higher value; the tail crossing, for example, is now at 42°C. Because the Cc has changed and oil and salinity have remained unchanged, the only way the system can balance and become Type III is (because this is an ethoxylate) via an increase in temperature. That change could be predicted from HLD without NAC. The other change is just as significant. The crossing point is now at 8.5% surfactant. This means that increasing the tail length (other things being equal) makes the surfactant *much* more efficient. It takes NAC to be able to calculate such a change as HLD knows nothing about the surfactant other than its Cc value.

How can all this be calculated from the trivially simple NAC equations?

Assuming (and this is a *big* assumption which is discussed later) that we know the value of ξ, the calculations are remarkably straightforward. Starting at low temperature in the Type I zone we can calculate the radius of an oil drop because we know the effective radius of the water drop from the % surfactant (suitably translated into molar concentration, which is why the MWt is a required input) and from the knowledge that $1/R_O - 1/R_W = HLD/L$. We know HLD because

22  ME Cates et al, Theory of Microemulsions: Comparison with Experimental Behavior, Langmuir 1988,4, 802-806

we have chosen a temperature and know the Cc, salinity and EACN. We also know that if we were close to the transition zone then $0.5(1/R_O+1/R_W) = 1/\xi$. Because we have started at a low temperature, there will be no match between the two $R_O$ values. We increase the temperature and try again. Eventually the temperature gives an HLD and $R_O$ (i.e. from Net Curvature) which match the value from the $1/\xi$ equation (i.e. the Average Curvature) so we know that this is a transition temperature and can mark that point accordingly in the graph. The process can be repeated from high temperature downward and also at increasing surfactant concentration so we get both the curves (transition from Type I to Type III and transition from Type II to Type III) and the crossing point. So by using the Net Average Curvature in a systematic manner, the whole fish diagram can be mapped out.

The question of how the $\xi$ value is calculated from the other inputs is answered later. Accept, for the moment, that it is a known value.

There is much more to say about fish and fishtail diagrams. But first, what else can NAC do?

It is possible to calculate the drop size and viscosity of a microemulsion. This at first seems an astounding claim. How can viscosity be calculated from such a simple theory?



**App 3-3** https://www.stevenabbott.co.uk/practical-surfactants/Drops.php
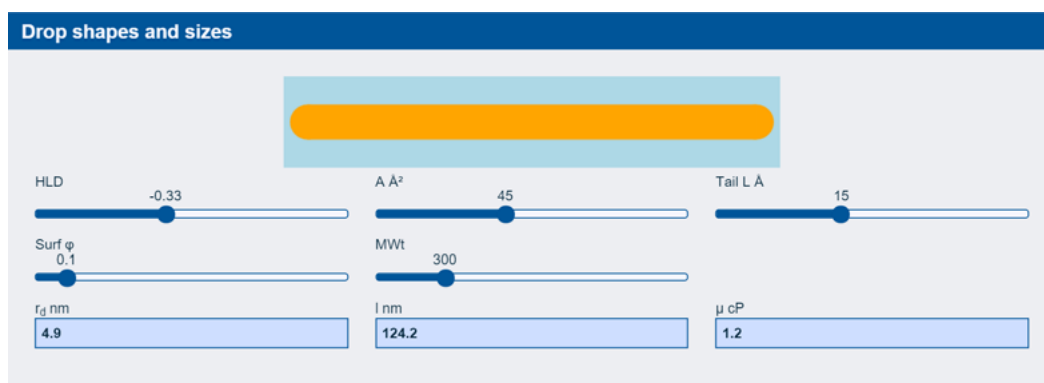
The calculation[23], again from the Acosta group, is rather straightforward though is not described in detail here. The calculated radius of the drop, say $R_O$ for the oil, is a nominal radius. A real drop can only accommodate so much curvature within a sphere. Above a certain volume the drop must generate more curvature by expanding into a spherical cap cylinder. The app calculates the radius of the cap, $r_d$ and the length of the cylinder l. By invoking the Einstein viscosity equation which takes into account the effects of overlapping cylinders it is possible to calculate the viscosity.

23   Sumit K. Kiran and Edgar J. Acosta, Predicting the Morphology and Viscosity of Microemulsions Using the HLD-NAC Model, Ind. Eng. Chem. Res. 2010, 49, 3424–3432

Because the diameter of the oil droplets in water or water droplets in oil can be calculated, it is also straightforward to calculate the overall solubility of oil in water or water in oil using the customary units ml/g. i.e. ml of oil (or water) per gram of surfactant.



**App 3-4** https://www.stevenabbott.co.uk/practical-surfactants/Solubility.php
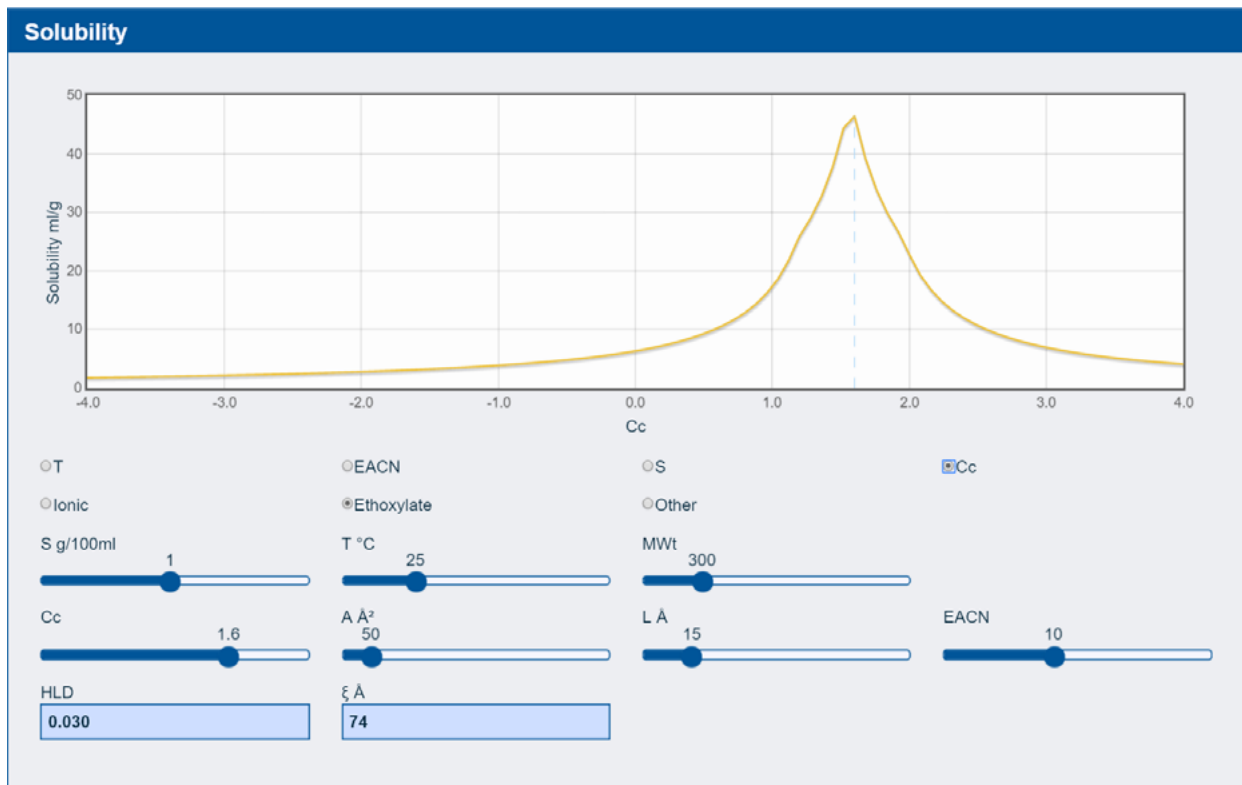
The calculation follows a similar pattern. From the HLD and ξ, $R_O$ or $R_W$ can be derived as above and the solubility calculated using the following formula, where R is either $R_O$ or $R_W$ depending which side of HLD the formulation happens to be:

$$Sol = \frac{\frac{4}{3}\pi R^3}{4\pi R^2 . \frac{MWt}{N.A}}$$   3-8

The term along the top is simply the volume of each drop. The term on the bottom is the surface area of the drop multiplied by the mass of the surfactant necessary to cover that surface area, i.e. MWt/(N.A) where, as before, N is Avogadro's number and A is the area of each surfactant molecule.

A very similar app allows the IFT (interfacial tension) to be calculated:
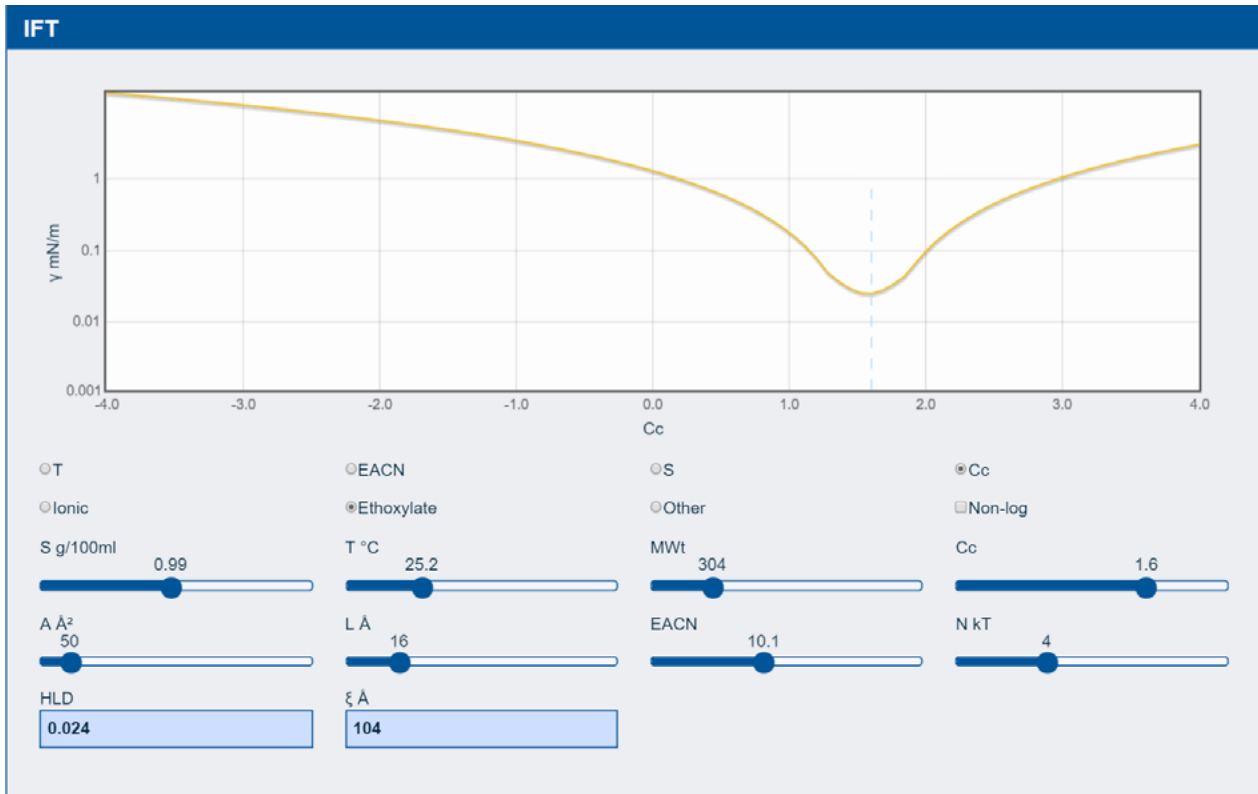
The default mode here is to plot IFT in log format. This highlights the differences between low (0.1), very low (0.01) and super low (0.001). In non-log form it is just a gentle parabola. The calculation is rather different. As before, $R_O$ and $R_W$ are calculated and then the IFT is calculated as:

$$\gamma = NkT\left(\frac{1}{4\pi R_O^{\ 2}} + \frac{1}{4\pi R_W^{\ 2}}\right) \qquad 3\text{-}9$$

where NkT is the rigidity of the interface in Boltzmann units. Here N just means "number of" and is nothing to do with Avogadro. For an ionic surfactant this is typically 1kT (less rigid, less robust) and for ethoxylates this is typically 4kT (more rigid, more robust). There is nothing surprising about this formula; all it is saying is that curvature imposes a force (tension) which has to be countered by rigidity. The less curvature, for a given rigidity, the less tension is required, or to put it the other way round, large radii mean low IFT.

It is convenient to describe here how to measure such low IFTs using a spinning drop tensiometer – they are too low for conventional measurement techniques. A drop of the oil is injected into in a cylindrical tube containing the aqueous phase. The tube is rotated around its long axis and the spherical drop starts to elongate. At a suitable rotation speed the drop is long enough to measure accurately:

**App 3-6**  https://www.stevenabbott.co.uk/practical-surfactants/SDT.php
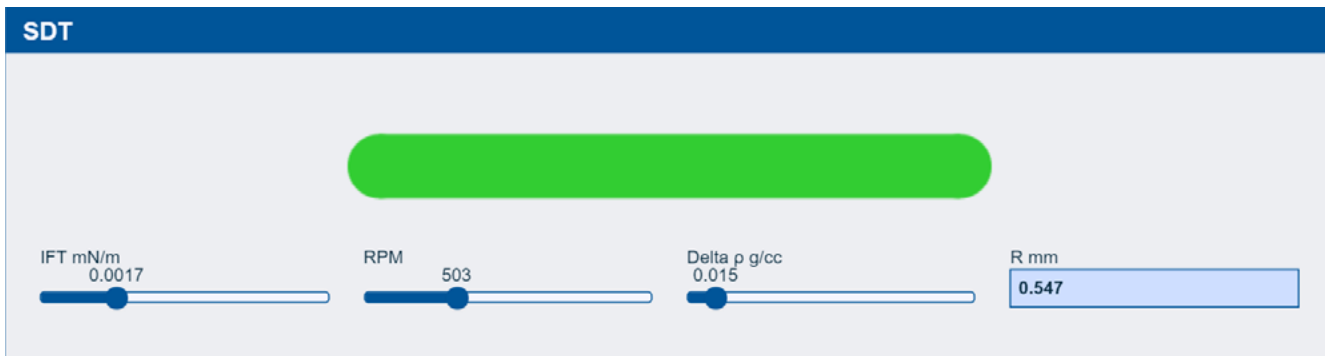
If the difference in density between the oil and water is Δρ, if the rotation speed is ω and the radius of the squeezed cylinder is R then the interfacial tension, γ is given by the Vonnegut formula:

$$\gamma = \frac{\Delta\rho\omega^2 R^3}{4} \qquad \text{3-10}$$

The reason the spinning squeezes the drop is that the heavier aqueous phase is being pushed to the outer part of the spinning tube, which of necessity pushes the oil phase towards the centre of the tube – and the oil has no choice but to elongate.

A spinning drop tensiometer should be on every formulator's workbench because identifying low interfacial energies is equivalent to identifying HLD=0, which places one's formulation nicely within surfactant space. Unfortunately their simple principle is not matched by simplicity in practice so these devices tend to be for academics rather than formulators.

## 3.2    More fish power from NAC

It is impressive that NAC is able to generate a standard thermal fish diagram. Now we can explore its power more fully, starting to use all the extra input options within the app.

The core problem for formulators is that surfactant space is very complicated. Surfactants and surfactant blends can change, oils and oil blends can change, salinity can change and temperature can change, all within the context of the concentration of surfactant changing. Visualising such a complex space is impossible, so to see what is going on we tend to look at convenient slices through such a space – keeping most things constant and varying a single parameter. The most usual such slice is the thermal fish – salinity, surfactant and oil are constant, just the temperature is allowed to change along with the surfactant concentration.

By analogy, a Salinity fish keeps everything constant except for salinity, a Cc fish keeps everything constant except Cc and an EACN fish keeps everything constant except for EACN.

HLD-NAC knows nothing of the specific scan – all it does is systematically vary the required parameters to find where the Type I/III and Type II/III transitions take place.

The normal rule for the surfactant apps in this book is: "Play around to see what happens". This is true for the fish app, along with the further instruction: "Don't panic". To see why this second instruction is required, select the thermal, ethoxylate fish from the examples above and click the Other option. The plot goes blank. There is no fish to be seen. To get a strange-looking plot is one thing – sliding some sliders might show some trends that can get the plot looking more normal. With a blank plot there is no hope.

What happened when Other was selected? Remember what the thermal fish is doing. It is (effectively) calculating what happens to the HLD over the designated temperature range. For ethoxylates quite a lot happens because the thermal term, α, in the HLD equation is relatively large at -0.06. The Other class is typically something like an APG surfactant where α is effectively 0. So there are *no* changes to the HLD over the temperature range and therefore no possibility of a fish diagram.

As a complete change, set up a salinity scan (S) with the settings shown:



App 3-7  https://www.stevenabbott.co.uk/practical-surfactants/Fish-and-More.php
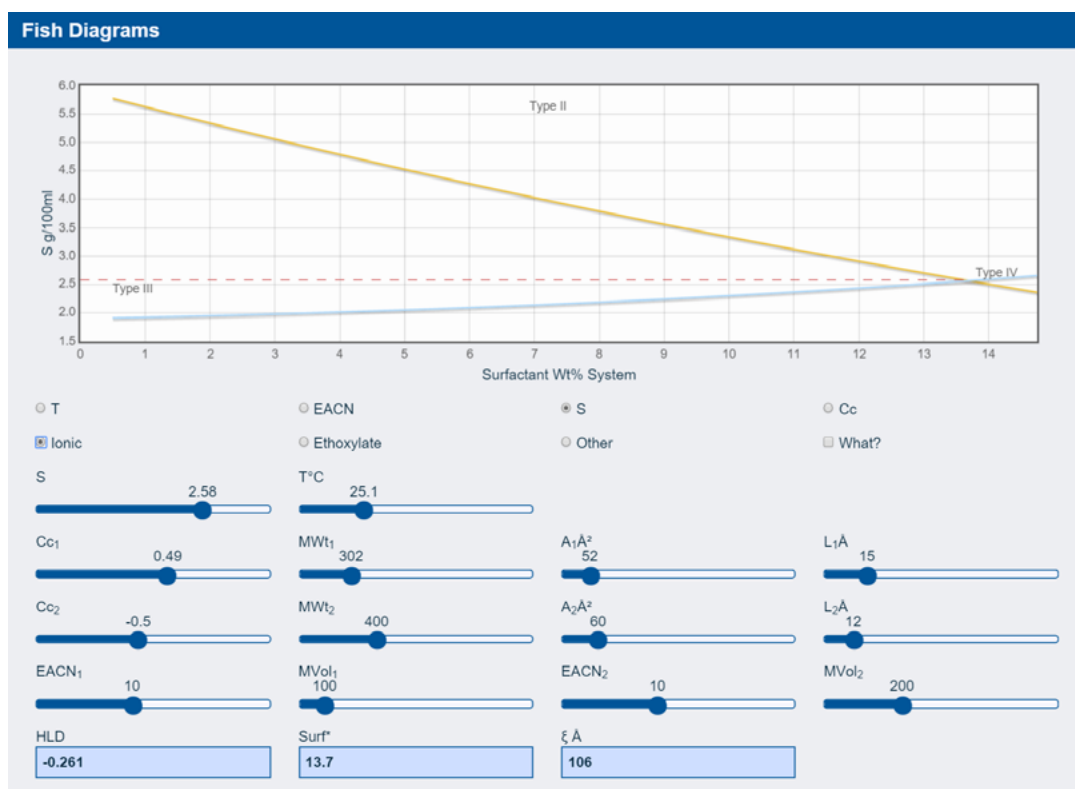
Because this is an ionic surfactant two effects are apparent. First, the whole fish is covered over a rather small range of salinities, 1.5 to 5.5 g/100ml. This is because of the log(S) dependence for ionics in the HLD equation. Second, the fish is highly skewed. This is because the surfactant itself acts as a salt, so the vertical (salinity) scan effects are affected by the horizontal (surfactant concentration) effects. Switching to Other mode causes no panic.
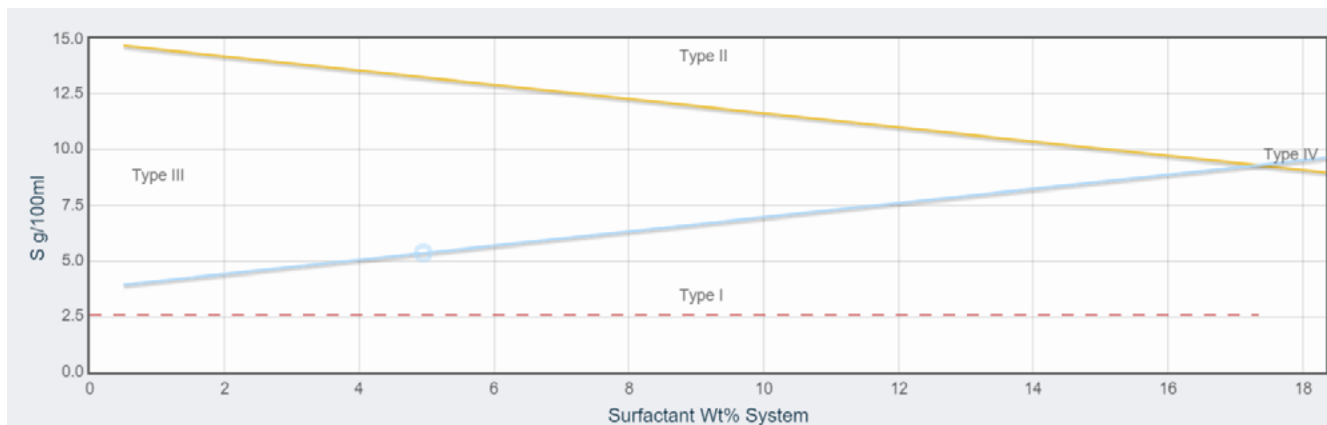


**Figure 3-2 The same fish diagram but with an Other (APG) surfactant.**

Instead a nice symmetrical fish emerges (there is no extra salinity from the APG), though it requires a large salinity range (3-15 g/100ml) because HLD depends linearly on S.

Switching to an EACN scan instantly produces a blank plot. This time it is the fault of the slider settings. The scan goes between $EACN_1$ and $EACN_2$ and these happened to be set to the same value of 10. By changing them to the familiar 1 (Toluene) and 16 (Hexadecane) a reassuring plot emerges:

**Figure 3-3 An EACN fish between Toluene and Hexadecane.**

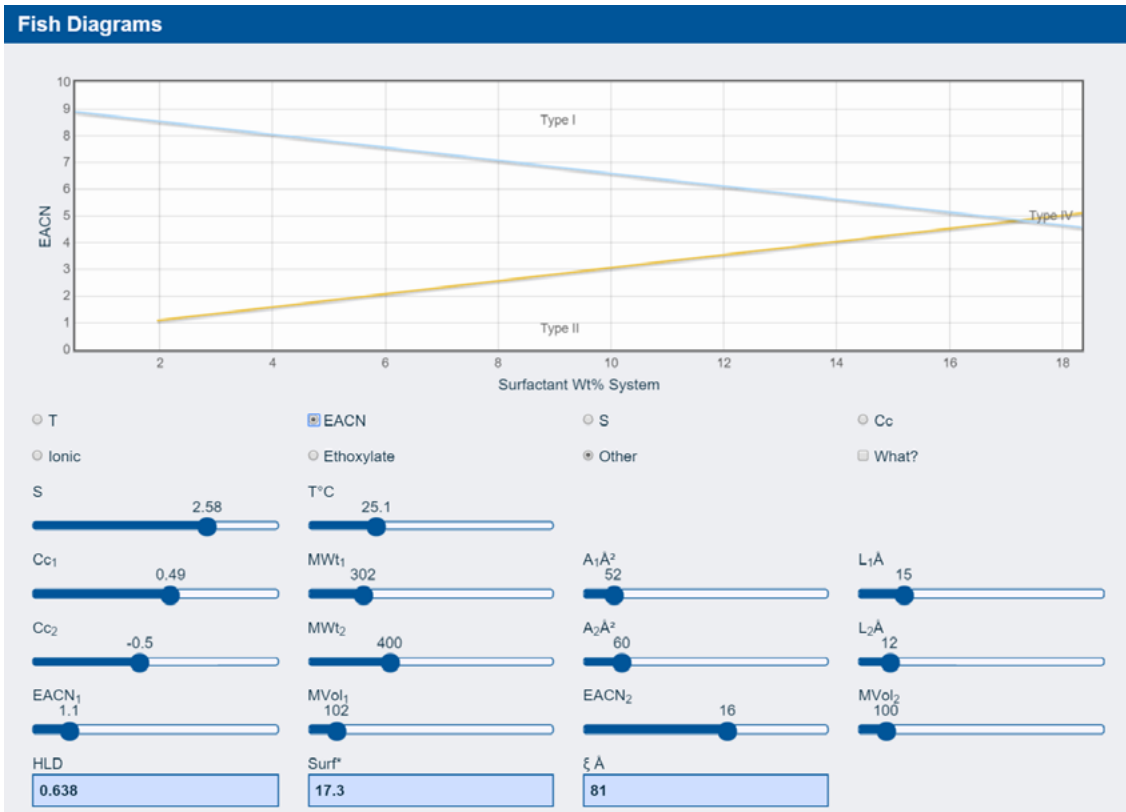Unfortunately this is *not* an accurate calculation. The molar volume (MVol) of each oil has been set to 100 giving a nice symmetrical plot. If you change to the real MVol of hexadecane (~300) the plot becomes very skewed, for reasons to be explained with the big discussion on ξ.

Although the NAC approach is an oversimplification and although ξ poses some significant problems, the HLD-NAC approach provides an impressive ability to explore the huge complexity of surfactant space with a few simple choices, slider values and determination to not panic.

There is still one big issue not yet addressed. All those fish diagrams are for 50:50 oil:water blends. These are great for scientific experiments and largely irrelevant for many real-world formulations. That is why we must go to the fishtail plot.

## 3.3     More power from NAC – the fishtail plot

The fish plot has a single "tail" point – the crossover where a 50:50 O:W mix becomes a single Type III/IV emulsion. Suppose it takes 20% surfactant to reach that point. It is rather obvious that a 20:80 O:W blend would not need 20% surfactant to be able to solubilize the oil. So if we *could* generate a fish plot at 20:80 it should show a crossing at a much lower % surfactant. Here it is:

**App 3-8** https://www.stevenabbott.co.uk/practical-surfactants/FishTail.php
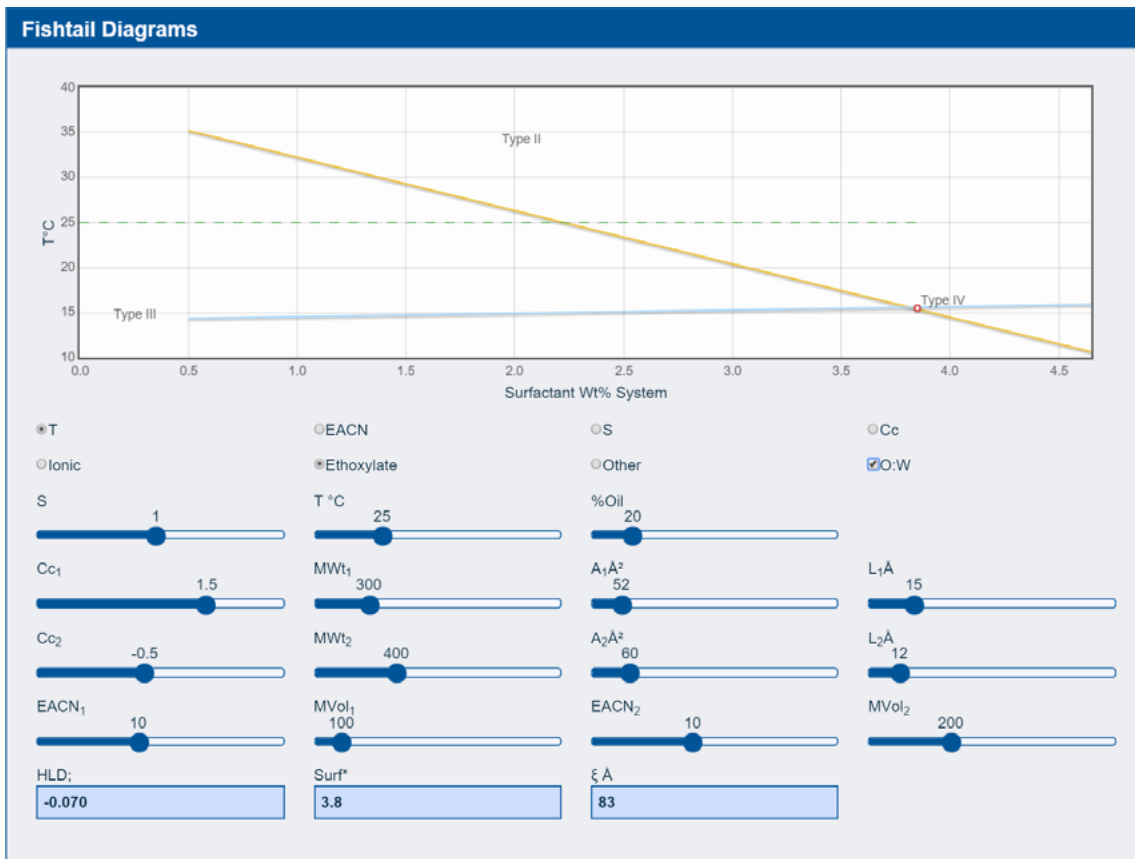
This is the same setup as the original, symmetrical fish which had a tail crossover at 17.2% surfactant. Here we see an asymmetrical fish and the crossover is at 3.8%. Note that the O:W option is selected and the %Oil slider is at 20%. If we de-select the O:W option then the familiar 50:50 fish lines appear, plus the fishtail plot with data points at regular intervals showing where the tail crossing is at the various O:W ratios.



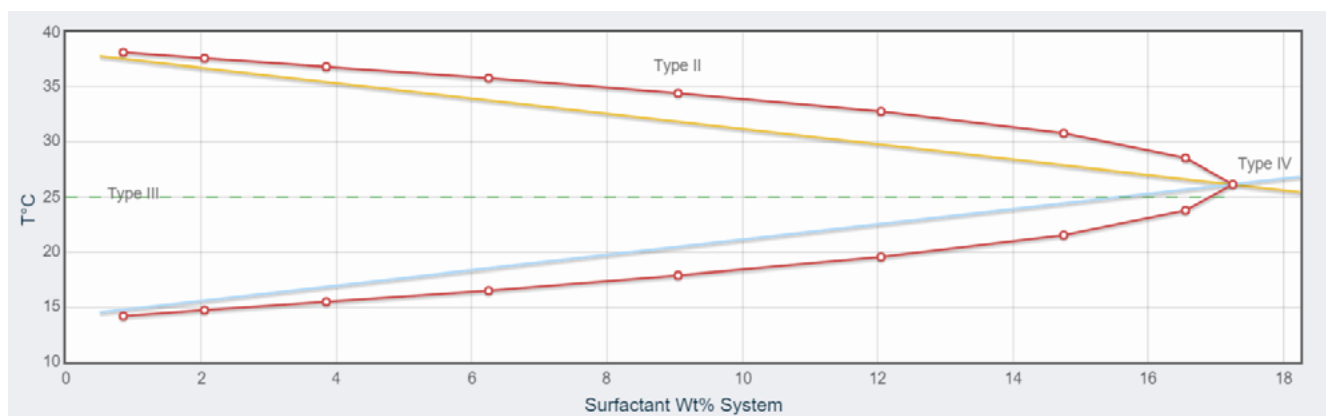**Figure 3-4 The fishtail plot showing the crossing points at regular O:W intervals. The mouse is hovering over the 20:80 crossing that was found in the app with the O:W option selected.**

Whether you look at the plot with individual %Oil settings or at the whole plot with the fishtail points, this is a powerful way to explore surfactant formulation space. The calculation follows the same plan as the conventional fish plot,

adjusting the calculations of the oil and water radii using the given oil:surfactant (or water:surfactant) ratios.

Because this is an ethoxylate plot, it is a good way to find the PIT for a specific oil:water ratio. Now we have more power at our disposal we can now discuss PIF.

## 3.4    PIF – Phase Inversion Formulation

There is no doubt that the PIT technique is a useful way to create real emulsions using the low interfacial tension created at the phase inversion temperature. Yet PIT is only of use for ethoxylates with their large negative thermal coefficient in the HLD equation. The fishtail plot allows us to be far more adventurous. How about a PICc, phase inversion Cc technique?
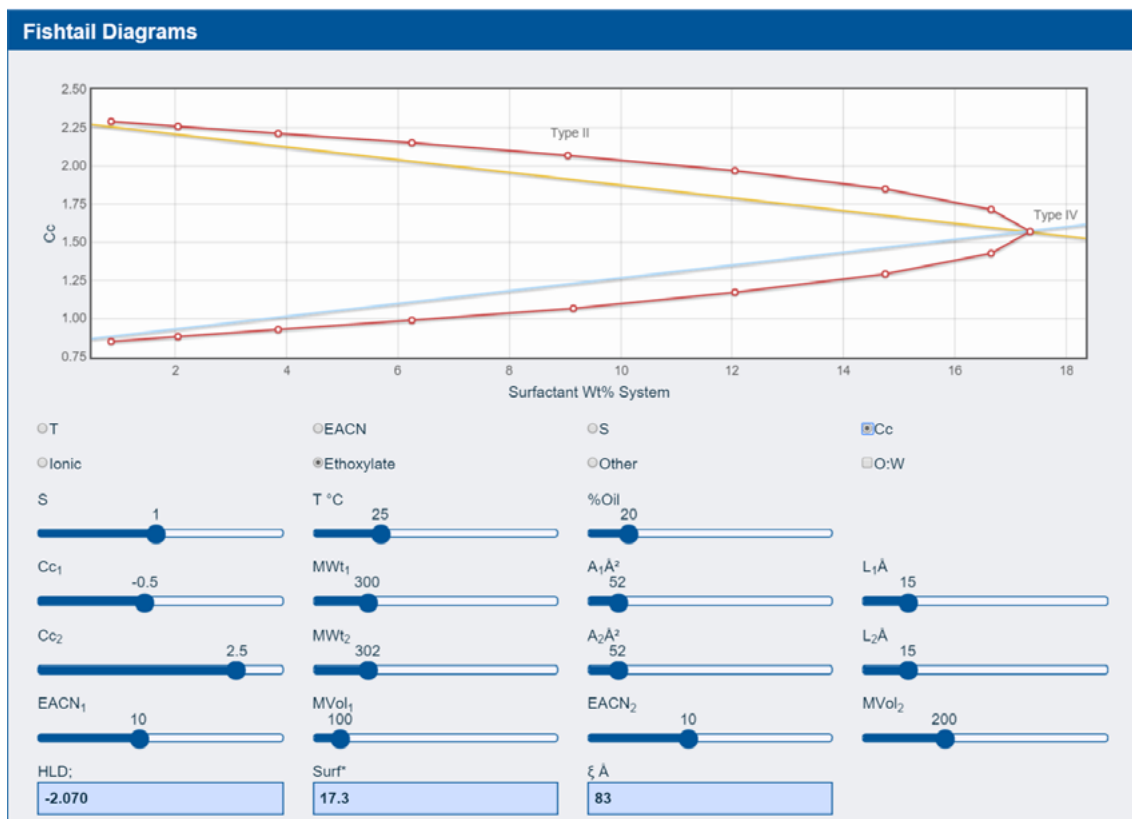


Figure 3-5 Everything is set up for a PICc.

For instructional simplicity the setup is greatly simplified. Two surfactants, with Cc of -0.5 and 2.5 happen to have the same MWt, A and L values. This is obviously unrealistic; with more realistic differences the principles are unchanged, the plot just looks uglier.

The mouse readout tells us that if we want a 25% O:W ratio then at Cc=1 and 6.3% surfactant we will be nicely in the tail crossover region so the whole formulation will be at low interfacial tension. Note that this is at 25°C – no wasteful heating of the emulsion mix is required. Now by throwing in some more

of the low Cc surfactant, the system will be moved into the Type I regime, just as cooling from the PIT moves the formulation into Type I. Because throwing in some low Cc surfactant is very easy, the whole process is quicker, cheaper, greener than the classical PIT technique. And although we have done this with an ethoxylate system, the same thing can be done with an APG or ionic system.

There are some problems with this simple scenario and in any case the assumptions are relatively naïve. No suggestion is being made that you can sit at your desk, throw in some numbers and instantly transform your ability to create emulsions via a greener approach.

Instead, what this example highlights is that using HLD-NAC it is possible to start to imagine totally new ways of doing phase inversion formulations. It is a case of thinking in PIF terms. If the best form of PIF is a PIT, that's great, it is a proven technique. More often it will be a PICc or PIEACN. It might even be a salinity version, though the acronym is less helpful.

It can even be a PIO – using the oil ratio. It is well known that phase inversions can occur when more oil is added. Without HLD-NAC it has been hard for the formulator to know where to begin – most additions of oil will certainly *not* give the desired effects – the whole system, surfactant type, temperature, salinity, Cc, EACN has to be right for the technique to work. So-called catatrophic inversions at 75:25 and 25:75 O:W ratios are discussed more fully later on.

Using the fishtail plot, with sufficient precautions about its assumptions and limitations, allows the formulation team to come up with ideas that should at least be in the right formulation zone, even if it cannot possibly be highly accurate.

## 3.5    Going on for ever

In the next chapter about real emulsions we will see that it is not too hard to have a full surfactant coverage of the emulsion drops (though this is not always necessary), making a real emulsion look similar to a microemulsion. So the question naturally arises as to why microemulsions don't go on for ever; why do we not have 1µm microemulsions? It would make life much easier because emulsions are thermodynamically unstable and hard to make, while microemulsions are thermodynamically stable and, in principle, trivially easy to make.

There are, as always, multiple ways to think about this. A particularly clear view is expressed by McClements[24] who in the same paper also addresses some of

24  David Julian McClements, Nanoemulsions versus microemulsions: terminology, differences, and similarities, Soft Matter, 2012, 8, 1719

the problems about terminology (microemulsions are really nanoemulsions, etc.) discussed earlier.

As we know, microemulsions are thermodynamically stable, which means that they have negative free energy which in turn is made up from two terms – enthalpic and entropic. We know that the enthalpic term must always be positive (bad) because it takes energy to create an interface. We also know that the interfacial area (for a given volume) increases linearly with 1/r – halving the radius doubles the area and doubles the enthalpic penalty. On the other hand, the more particles (which, for a given volume, means smaller r) the higher the entropy, which is a negative term in free energy calculations. A microemulsion appears spontaneously at some suitably small radius where the entropic term beats the enthalpic term. So far so good. But from this it sounds as though any small emulsion drops could be thermodynamically stable, yet we know that nanoemulsions of radii in the microemulsion domain are thermodynamically unstable. We also know that microemulsions have to be specifically tuned to the oil/water/salt/temperature environment whereas nanoemulsions require far less tuning.

The final piece of the puzzle is found from the clue about tuning. Microemulsions are at their most stable when the interfacial energy is super-low, i.e. when the curvature is close to zero. With such low interfacial energies, the enthalpic term is brought low enough that the entropic term can win and give a negative free energy.

So microemulsions are always at a sweet spot of enthalpy reduced to a minimum via curvature and entropy raised to a maximum via small drops.

Clearly, if you try to tune the system to give a bigger r at the minimum interfacial tension the entropic term is going to be reduced and although the free energy will still be negative (if the IFT is made super-low), once it is less than kT the emulsion will be unstable in practice because the thermal fluctuations are, by definition, bigger than the enthalpic term. McClements even links the instability to Kabalnov's revitalised "Wedge" theory of emulsion instability discussed later. So the factor that stops us from producing indefinitely large microemulsions is the random thermal energy in our formulations. Because kT is expressed in °K, modest reductions in temperature, say from 300 to 280 make little difference to the effect, so you can't make significantly larger microemulsions by cooling down the formulation.

It turns out that this practical limit to microemulsion efficiency is captured in the ξ parameter, so it is now time to come clean about why ξ is not as straightforward as we might like.

## 3.6      The problem of ξ

ξ is an idea first developed by Nobel Prize Winner genius Pierre de Gennes. His genius consisted in coming up with simple ideas with profound consequences. The ideas were often found to be not as neat and simple as his papers implied. He was happy for this to be so. All he wanted was to set a general direction (which others failed to do) so that others could fill in the details (which they were often delighted to do, a task he found far too boring).

In the FishAndMore app there is a little What? Button. Click it and a graphic appears explaining how A and L affect the head and tail of the surfactant (obviously) and how ξ, which depends on A and L affects the size of oil drops. Here is a view of the ξ part, showing how large lengths from ξ lead to polygons with large circumferences:
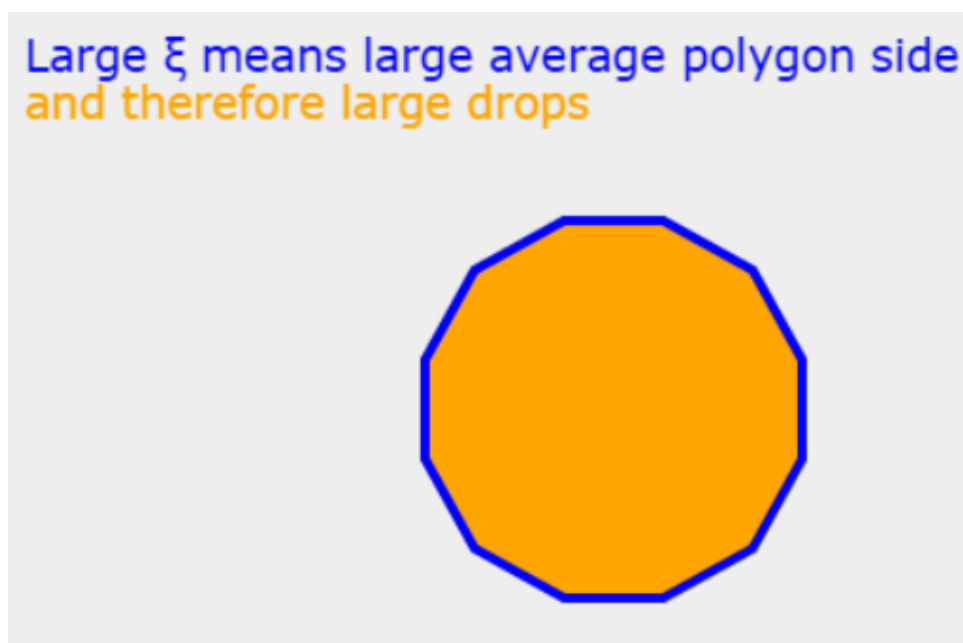


**Figure 3-6 ξ can be thought of as a coherence length for chunks of surfactant. The longer that length the longer the side of the polygon made from those lengths, so the larger the drop of oil, i.e. the higher the efficiency of the surfactant.**

That is de Gennes' big idea. If ξ were a constant for each surfactant, and if we had a list of such values, life would be very easy and many of the hesitations about the power of HLD-NAC would disappear.

The fact that ξ is not a constant is easily demonstrated. It is a universal truth that it is easier to fully emulsify a small oil like hexane than a large oil like hexadecane. Of course many people who try to emulsify hexadecane with a surfactant optimized for hexane will fail because they forgot to adjust the HLD via salinity, temperature or Cc. Even those who correctly optimize the HLD still find that larger oils are emulsified less efficiently, i.e. ξ is lower for larger oil

molecules. To explain this, and much else about ξ, Acosta created a ξ predictor, which[25] I was able to put into app format:



App 3-9  https://www.stevenabbott.co.uk/practical-surfactants/Xi.php

The app shows ξ in a rather different manner; rather than a polygon there is a bigger or smaller oil zone depending on whether ξ is larger or smaller. This is because the system is shown at HLD~0 and on average there is no curvature. Instead the oil is more or less associated with the tail region. As L increases and A decreases (i.e. tighter packing) more oil can be "captured" by the tail region so the solubility of the oil increases. The effects of L and A are intuitive. The problem is the third factor: the molar volume (MVol) of the oil. MVol is MWt/Density, i.e. the volume occupied by 1 mole of oil. As most of us are not familiar with MVol the input to the app is via the oil density and the MWt. Increasing MWt for a given density decreases ξ, i.e. larger oils are "controlled" less by the tail of the surfactant. This makes some intuitive sense. And it explains why it is so much harder to create a microemulsion of a large oil.

There is, at present, no precise theory for this. Instead, Acosta has created a correlation built on the idea of an Overlap Factor (OF), and has, in turn, created a simple way to estimate OF. No doubt with more data from the surfactant community these ideas can be refined. For now they allow the apps to give a reasonable estimation of ξ – it is the method that is used to provide the values shown in, say, the fish diagram apps.

OF looks somewhat like a packing parameter:

$$OF = \frac{V_{SurfTail}}{A_{SurfHead} V_{Oil}^{1/3}} \qquad \text{3-11}$$

A careful analysis of the available data allows two correlations to be made, one for ethoxylates and one for ionics:

---

$$\xi_{ethoxylates} = 0.4e^{6.6OF} \qquad \text{3-12}$$

$$\xi_{ionics} = 40e^{1.2OF} \qquad \text{3-13}$$

Finally, to make it easier to calculate from standard inputs, OF is re-defined:

$$OF = \frac{13\dfrac{L}{A}}{\dfrac{MWt_{oil}^{1/3}}{\rho_{oil}}} \qquad \text{3-14}$$

The key issue for formulators is the exponential dependence on OF. Relatively small changes in L, A or MVol can lead to large changes in ξ and therefore large changes in the ability for the system to solubilize oils. In the day-to-day world of formulations one surfactant looks much like another. But relatively modest differences in surfactants can make relatively large differences to the resulting formulation. This is not obvious to many formulators because changes in a surfactant affect *both* HLD and ξ so it has been hard to disentangle the various effects. A surfactant supplier might claim that their surfactant molecule is especially efficient – and their claim might be right if it is matched with the right oil at the right HLD. Change the oil or the HLD and the claim is worthless.

### 3.6.1 Phase volumes

There was a large gap between my claims, above, that HLD-NAC is straightforward, and my ability to implement it intelligently. Thanks to the inspiration and inputs from Alejandro Gutierrez at VLCI, and further advice from Prof Acosta, I finally managed to implement a phase volumes app that not only was easy enough that even I could understand it, but which came with step-by-step instructions so that anyone can do it. First, let us look at the app:
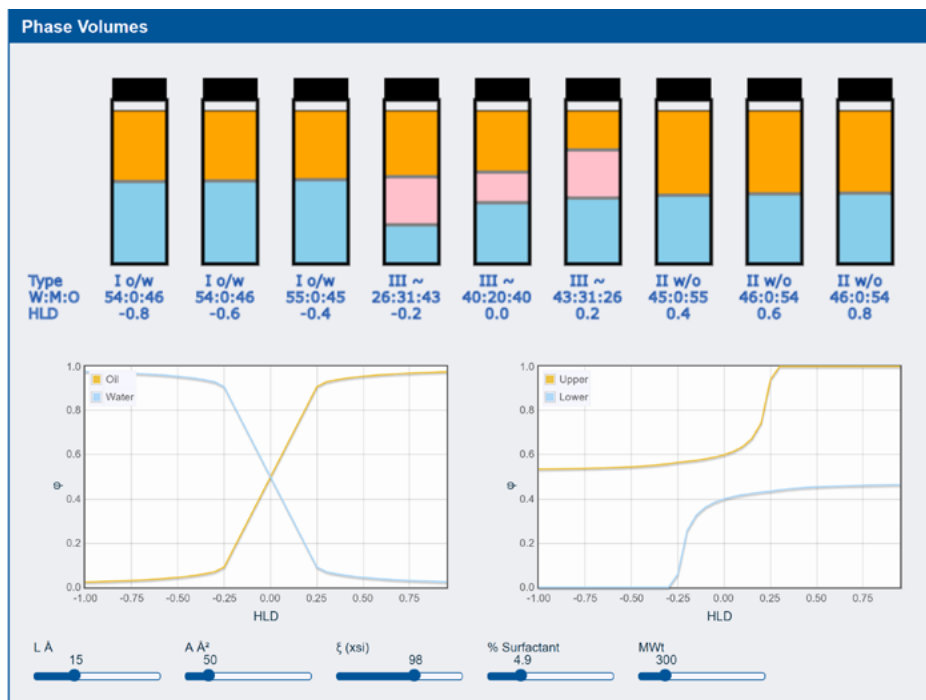
The setup is similar to the tubes used to measure Cc and EACN values, but here HLD = 0 is always in the middle. What interests us is what the various phase volumes are as we go from -0.8 < HLD < 0.8, a convenient range to see the big picture and the local detail. The tubes show what we actually see and the volumes of the Water, Middle and Oil phases are indicated below in the W:M:O row.

Then we have two graphs that appear in the academic literature and which, I confess, I have always found difficult to interpret. On the left we have the phase volume fraction, φ, of oil in the water or water in the oil. Normally the plots end around HLD = 0 but I have extended them across the full range. They make sense after a while. The second plot shows where the lower and upper interface boundaries are within the tube. On the left you only have O:W boundary, as you go to the middle this rises (hence it is called the Upper boundary) as the Lower boundary, between water and middle phase appears till, finally, the upper boundary disappears because you now only have the W:O boundary.

In real experiments you measure the tubes and work out the key parameters which are L, A and ξ. Here we do it the other way round so you can see what happens as you change the parameters. I have to admit (again!) that the results surprised me. They are logical but not at all intuitive. And that is why the app is so important! We all need to grasp the effects of having longer or shorter chains, bigger or smaller head areas and, of course, larger or smaller ξ values.

So how do we do the calculations. Here is what I wrote in the app. Hopefully it will start to make sense as you slowly go through it. To help you (it certainly helped me, because I created it before I could write the app), there is an Excel

sheet you can download from the page which does all the steps for you. Those who read Javascript can get the code from my page. Here goes:

- $Hn = -HLD/L$. This is the Net Curvature. In the calculations below we use the absolute value, ignoring the sign.
- $1/Ro = 1/Hn+1/Ref$ if HLD < 0 or else $Ro = Ref$. This is the basic oil radius
- $1/Rw = 1/Hn+1/Ref$ if HLD > 0 or else $Rw = Ref$. This is the basic water radius
- $Ha = 0.5(1/Ro+1/Rw)$. This is the Average Curvature.
- $1/Ha$ is a sort of average radius
- If $1/Ha > \xi$ then we're Type III. Otherwise our type is I if HLD < 0 or II if HLD > 0. You can see that $\xi$ is crucial for knowing about the Type III phase.
- If we are in Type III then we need to calculate the microemulsion (me) radii Rwme and Rome. If not then they are simple Rw and Ro. $\xi$ is again needed in the Type III.
- $1/Rwme = HLD/(2L)+1/\xi$
- $1/Rome = 1/Rwme-HLD/L$
- We need to convert radii into volumes via the total head area, A. The factor 3 appears because volume is $4/3\pi r^2$ and area is $4\pi r^3$ so when you divide them you end up with 3 remaining.
- $Vome = Rome*A/3$
- $Vwme = Rwme*A/3$
- $Vtot = Vome + Vwme$
- It's now easy! We first start with the phase volumes. These can be plotted directly
- $\varphi o = Vome/Vtot$ if HLD < 0 otherwise a pseudo-volume, $1-\varphi w$
- $\varphi w = Vwme/Vtot$ if HLD > 0 otherwise a pseudo-volume, $1-\varphi o$
- Then we need the true volume fractions of oil, middle and water phase, Vo, Vm, Vw. These get summed to Vt and then turned into fractions Vof, Vmf, Vwf, but as that's an obvious step it's not spelled out here.
- $Vo = VolOil - Vome$ if HLD < 0 orelse $Vwme + VolOil + VolSurf$.
- $Vm = Vome + Vwme + VolSurf$ when we're Type III or 0 otherwise
- $Vw = VolWater - Vwme$ if HLD > 0 orelse $Vome + VolWater + VolSurf$
- The phase boundary lines can now be calculated as fractions of a tube. These are not at all obvious, at least to me.
- $Bound1 = Vmf + Vwf$ in Type III and Vwf if Vof < Vwf, otherwise 1
- $Bound2 = Vwf$ in Type III and Vwf if Vwf < Vof, otherwise 0

In the earlier version of this book, adding this section was both impossible (I couldn't do it) and a waste of time (no one would have used it). A few years later there is now a real hunger to use NAC more intelligently and I already know that the app and the step-wise instructions have been used by some key surfactant formulators.

### 3.6.2 Other ways to look at ξ

Whatever ξ is, it is capturing something important about curvature, about persistence length, about self-association. The idea that it is important has been in the surfactant community for decades and ξ has from time to time been measured using X-ray or neutron diffraction. It keeps resurfacing in somewhat different guises. A recent example is the paper mentioned earlier from Fraaije[26]. He explores curvature, packing parameter and HLD using Helfrich's (1990) surface torque density approach. If I properly understood torque, moments, splay constants and bending constants I might be able to explain Fraaije's approach. In fact, all I can say about it is that there is hope that ideas (curvature, PP, HLD) that have been floating somewhat disconnectedly for decades might be coming together to provide improved ways for tackling the problem of predicting surfactant behaviour.

### 3.6.3 Creating a large ξ for efficiency

For a given oil, salinity and temperature we know we have to choose a surfactant (blend) with the correct Cc value. Yet two surfactants with the same Cc might show very difference solubilisation capacity – in other words they have different ξ values. Acosta's method for estimating ξ shows what we have seen in other contexts, that a large tail and small head lead to greater efficiency. So why don't we make surfactants with enormously long tails? The answer is that in general, longer tails, even when branched, tend to pack together too well and the surfactant becomes a horrid viscous solid that is difficult to use. The phase diagram chapter shows that hexagonal and cubic phases are particularly hard to deal with as they are highly viscous. And if the MPt of the surfactant is too high or the solubility too low, the main "phase" is simply solid surfactant, or, to put it in common parlance the solubility becomes lower than the CMC so the surfactant is below its Krafft temperature and is incapable of forming micelles.

In any case, for a given ethoxylate Cc, a longer tail implies a larger head so the ξ effect of the tail might be somewhat diluted by the larger head area.

The EOR industry, starting with work by Salager, has found a way around this dilemma[27]. They keep conventional heads and tails and stick a "neutral" element between them to create an "extended surfactant". We know that alkanes are tail-like and the ethylene oxides, EO, are head-like. The trick is to use propylene oxides, PO, that are "neutral", neither head nor tail. What these provide is extra solubilisation capacity without the need for larger heads or the problems of difficult phases with long tails. They have been likened to "built-in linkers" for

---

26   Johannes G. E. M. Fraaije et al, Method of Moments for Computational Microemulsion Analysis and Prediction in Tertiary Oil Recovery, Langmuir 2013, 29, 2136−2151

27   Miñana-Pérez M, et al, Solubilization of polar oils with extended surfactants, Colloid Surf A 100, 1995, 217-224

those who are familiar with the idea of linkers. For those who are not, linkers are discussed in the Problems chapter as they provide much confusion along with, sometimes, increases in solubilisation capacity.

The graphic shows the general idea.

Alkane tail      Head          Standard Surfactant

Alkane tail      Neutral    Head       Extended Surfactant
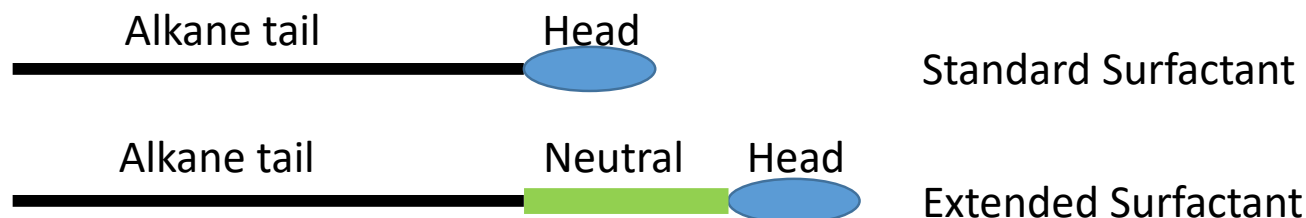
e.g. Tail=C12-C13, Head =$(EO)_6$ or $SO_4Na$, Neutral=$(PO)_4$

**Figure 3-7 An extended surfactant is made by inserting a neutral portion (e.g. polypropylene oxide) between the tail (a typical alkane) and the head which could be ethoxylate, sulfate, etc.**

A paper describing the properties of a range of extended surfactants[28] comes from a fruitful collaboration between the Acosta academic group and the Sasol EOR team under Dr Charles Hammond. The higher solubilization obtained by their molecules arises from, as expected, a higher ξ and this is related, as expected, to interfacial rigidity – the extended surfactants are more rigid. This gives rise, unfortunately, to a major downside of this approach. Rigid surfactants are wonderful for everything except kinetics. They are slow to equilibrate and can readily get stuck in unwanted configurations. The trick, therefore, is to blend them with co-surfactants that reduce rigidity enough to improve the kinetics without compromising too much on ξ. This is a tough challenge and helps explain why extended surfactants are not more widely known and used outside the world of EOR. The PO unit is also reputed to be less biodegradable, hindering their use in consumer products where their increased efficiency would be much appreciated.

### 3.6.4  CµC

We can now come back to the useful concept of the critical microemulsion concentration. As already mentioned this is generally more significant than CMC. It is the concentration of surfactant at which a microemulsion can be formed. Because microemulsions are formed when HLD~0, CµC is meaningful only within the effective microemulsion range, so any measurement of CµC has to be referenced to salinity, EACN and temperature. This helps explain why CµC is not much known – it is not some simple value that can be applied to a surfactant. As pointed out previously, neither is CMC the simple number it is

---

28   Edgar J Acosta, Sumit K Kiran, Charles Hammond, The HLD-NAC Model for Extended Surfactant Microemulsions, Journal of Surfactants and Detergents, 15, 2012, 495-504,

always assumed to be. In any real application where oil is involved, the CMC will increase depending on the partition coefficient of the surfactant between oil and water. The more the surfactant partitions into the oil the larger the concentration required to generate micelles in the bulk water phase. An example of this[29] in the context of CμC is that for a specific surfactant (sodium dihexylsulfosuccinate) and salinity (5%) the CMC in the presence of tetrachloroethylene is ~0.02% and for trichloroethylene it is ~0.1%. The respective CμC values are ~0.06% and ~0.3%. The quoted "standard" CMC is 1.5%. Because I argued that the oil should *increase* CMC why are these CMCs so much *lower* than the standard value? Because the standard value is in pure water and these values are in the presence of salt which makes the anionic surfactant more hydrophobic so lowers the CMC, presumably to <0.02%.

### 3.6.5  Silicone oils

As discussed in the HLD chapter, silicones are difficult to emulsify with conventional surfactants not because of surface tension or contact angle reasons but because at the thermodynamic level they interact poorly with surfactant alkane tails. Silicones are much happier interacting with silicones than with alkanes. So even though the choice between water and alkane is straightforward, the silicones do not form any long-range interactions with hydrocarbon surfactant tails, so ξ is small and the emulsifying power is very small.

### 3.7　　The point of this chapter

The previous paragraph brings us to the point of this chapter. To repeat: no one claims that HLD, NAC or ξ calculations are perfect. A later chapter will explicitly discuss the problems and challenge the surfactant community to come up with improvements.

On the other hand, no other approach has given the formulator the ability to comprehend and map out so much of surfactant space with so few parameters, and thereby bring a lot of order from a lot of chaos.

I need to inject a couple of personal views. First, a large amount of the published surfactant literature is a waste of time, for a very simple reason. A group reports that surfactant X does something interesting with oil Y. If you happen to be interested specifically in X and Y then the paper is really useful. Generally your interest is something close to X or close to Y. If these papers had clearly mentioned temperature, salinity, Cc and EACN then it would be very easy to extrapolate from the paper to a neighbouring case. Most papers miss out such key information so it is impossible to apply their work to any other case. This

---

29  Edgar J. Acosta, Jeffrey H. Harwell, and David A. Sabatini, Self-assembly in linker-modified microemulsions, Journal of Colloid and Interface Science 274 (2004) 652–664

is a tragedy. If HLD had displaced HLB in the late 1970's, and if everyone as a matter of routine had recorded the Cc of the specific surfactant and the EACN of the specific oil then 30+ years of research would be a rich treasure-trove for "big data" mining. Instead, the literature is largely a wasteland of isolated papers of no use to the wider community.

Second, Professor X gets up at a conference and talks about important research in some area that requires an optimized surfactant system. Invariably the professor announces that some obscure mix of two surfactants has wondrous performance. The implication is that the professor has generated this blend through pure genius. As I sit there doing a rough calculation I work out that a competent researcher armed with HLD-NAC would have reached the same conclusion in about 5 minutes and could have found an even better blend after 10 minutes. The trouble is that it wasn't Professor X who produced the magic formula. It was a PhD student who had tried 100's of formulations over months or years before homing in on this particular blend. This is not hypothetical. I have sat in such conferences and spoken to the poor PhD students.

The two previous paragraphs are my lament for the massive waste of effort from the surfactant community in the past 30 years. Not only would HLD (+NAC) have reduced the waste of all those unhelpful papers and all those PhDs with bad formulations. By now we would have so much great data that we would have theories that went beyond the limitations of HLD-NAC.

The surfactant suppliers, in particular, are guilty of double standards. By now they all know that the field of EOR is a huge opportunity for specialist (which hopefully means profitable) surfactants. In this area they are *all* using some form of HLD-NAC. It is not possible to gamble on $million oil-field tests using intuition or HLB. They *all* have robotic systems carrying out forms of Cc and EACN measurements. Yet in most other areas of surfactancy they pretend that Cc and EACN do not exist. They continue to provide tables of HLB and CMC that are of very little use to their customers. They may not even be controlling the quality of their surfactant blends via Cc tests.

The point of this chapter is that HLD-NAC is a pretty good tool for navigating through surfactant space. With the apps, or your own implementation of the theory, a lot can be achieved for very little effort. It is a great starting point for any formulator and it suggests new approaches such as PIF that can be applied to formulation tricks (phase inversion) that have tended to require luck or genius to develop. The fishtail app is especially powerful because it helps to see how to change key parameters away from the academic 50:50 O:W setup to the real-world formulation with, say, 20% oil. An experienced formulator will be able to use HLD-NAC to make sense of many previously unexplained observations and perhaps get round some previous formulation bottlenecks. Above all, systematic use of HLD-NAC by suppliers (publishing Cc and EACN values) and by those

who write papers or speak at conferences will allow a critical mass of data to be assembled to allow the theory to tackle the tough problems that remain.

Before addressing those problems we need to explore the area of conventional emulsions where the tools of HLD-NAC do not apply so directly, yet still exert a powerful guiding hand.

# 4 Emulsions

A standard criticism of HLD-NAC is that it is merely a microemulsion theory that is irrelevant for real emulsions. Initially this diagnosis looks accurate. Microemulsions are thermodynamically stable and emulsions are only kinetically stable. So the techniques of HLD-NAC cannot be used directly to understand everything about emulsions. Plenty of other factors such as charge and steric stabilisation via DLVO theory, and the kinetics of coalescence are required.

Yet the formulation guidance provided by the surfactant community tends to be based on ideas such as HLB, Bancroft's Rules, CPP, Winsor R-Theory all of which are curvature-based theories which are better-described via HLD-NAC. One of the challenges, therefore, is to know how much to use HLD-NAC, how much to rely on colloid chemistry and how much to look for other design tools. A starting point for the discussion is some numbers.

## 4.1 The numbers behind an emulsion

When a volume of oil is converted to a typical emulsion there is a huge number of drops created and a huge increase in surface area, which requires energy. The app does the calculations for you:



**App 4-1** https://www.stevenabbott.co.uk/practical-surfactants/EmuSA.php
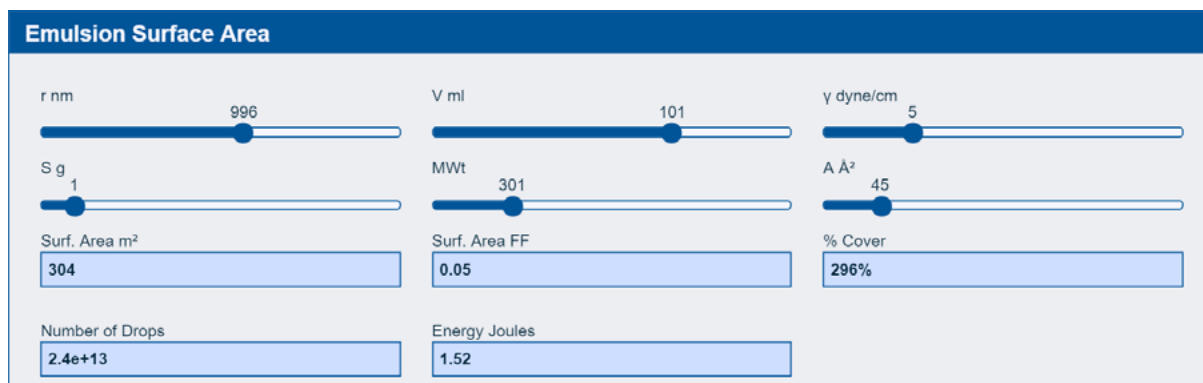
The default settings show that 100ml of oil is converted to 2.4e13 drops of 1µm radius, creating a surface area of 300m². It is popular to translate this into a more familiar unit of area. In this case it is about 5% of a football field (FF taken as the average of the American and soccer versions). The surfactant reduces the interfacial tension from 40-50 mN/m to 1-5 mN/m. Because we have a vision of lots of energy going in to an emulsion mix and because we know that the surface area is high it is easy to think that most of that energy has gone in to creating that surface area. But 5mN/m is 5mJ/m² so 300m² requires only 1.5J to create it, which is negligible. Typical emulsions require kJ, not J! So most of the energy used in making an emulsion is ending up as heat, which often requires further energy to remove it. The reasons for this inefficiency, and what to do about it, are discussed later.

The numbers also tell us how much of the added surfactant is available to be at the interface – the % Cover. With 1 g of surfactant (with an effective area of 45 Å²/molecule) for 100ml of oil, and drops of 1µm radius, there is a 3x excess (300%). Sliding the radius input shows that for 10µm radius there is a 30x excess and for 100nm there is 30% coverage – i.e. the effect is linear in radius and reflects the intuition that finer emulsions need more surfactant to gain maximum stability or, to put it another way, for a given amount of surfactant the eventual radius that can be achieved after efficient emulsification is that which gives 100% coverage. My experience is that the idea of a limiting size controlled by the amount of surfactant is either "well known" or a complete revelation. If you are already familiar with it, the app makes the calculation easy. If you have not encountered it before, you now have good reason to make sure that head area and MWt are a natural part of your information pack on potential surfactants.

The criticism that microemulsions need lots of surfactant merely reflects the fact that their radii are in the 20-50nm range so the same amount of surfactant as in the example above would only cover 6-15% of the drops, not enough to create a microemulsion.

## 4.2    Size distributions

Often it is good enough to say "The emulsion drops have a size of 200nm". It may not matter whether they are actually 100nm or 300nm, nor that they cover a range of sizes. We all know that a 200nm emulsion is rather different from a 2µm emulsion. At other times the size and the size distribution are of great importance. The trouble is that defining "size" and specifying the distribution are not trivial tasks. Even well-intentioned scientists can get confused if they aren't aware of the type of size distribution being discussed. And we can easily show that unscrupulous academics and suppliers can honestly say "90% of the emulsion drops are below 200nm" when in fact, from the user's point of view 90% of the emulsion drops are above 200nm. Note that we haven't said whether radius or diameter is being used – such ambiguity is all too common and means that we are often out by a factor of 2 even before asking deeper questions.

This all becomes clear from the app, where inputs (and the graph) are in terms of radii and the outputs (because that seems to be the confusing norm) are in terms of diameters.

**App 4-2** https://www.stevenabbott.co.uk/practical-surfactants/Distribution.php

Focus first on the "raw" curve. You synthesise a size distribution of your own choice by specifying two radii, $r_1$ and $r_2$ which are the peaks of two distributions, then you specify their peak heights (from 0 to 100) and their widths via the two w parameters. The w parameters are not intuitive because the curves are "gamma distribution functions" that can generate a characteristic tail. Simply play with r, h and w values till you get a raw distribution that interests you.

Now look at the Cum. N curve. This is the Cumulative Number distribution which just adds up the number of drops of a given radius. You can readily see that ~90% of the drops are below 200nm, and there is a 10% rise due to the small peak near 200nm.

Looking at the Cumulative Mass curve the story is very different. All those small drops only add up to ~10% of the total mass (or volume). The "small" peak at 200nm actually contains 90% of the mass. This is because mass goes as $r^3$ so a single drop at 200nm contains the same mass as 1000 drops at 20nm or 1 million drops at 2nm.

If you were an unscrupulous academic or seller of emulsions you would show everyone your cumulative number distribution curve to show how small your particles are. But as a user you would be rather upset when you discovered that most of your emulsion was in the form of rather large drops.

Similarly, a supplier can quote that the Diameter of their particles is 93nm, but if you measured it yourself you might find that it is 321nm. This is because of the bewildering variety of size definitions. The first diameter is D[1,0] and the other is D[3,2]. They might equally quote that D50=200nm, with the possible confusion

that D̄32 is another way of talking about D[3,2]. D[1,0] is the Number mean, D[3,2] the volume/surface (or Sauter) mean and D[4,3] is the mean diameter over volume (DeBroukere). They are defined as:

$$D[x,y] = \frac{\sum D_i^x . n_i}{\sum D_i^y . n_i}$$  4-1

and D50 is the Median Volume, i.e. the diameter where half the particles have less volume and half have more volume.

Each of these numbers has its uses and there isn't a right or wrong way of specifying drop size. The key is to play enough with the app to get the general idea and to be able to ask the right questions when someone says "The emulsion has drops of size 200nm". At the very least, ask if the size is a radius or a diameter     .

## 4.3     Viscosity

The viscosity of an emulsion makes a difference to the energetics of creating it and also to its long-term stability. There are many formulae for how the viscosity η of an emulsion depends on the volume fraction of the drops, φ and the viscosity of the bulk phase, $\eta_0$. The most basic is the Einstein formula:

$$\eta = \eta_0 \left( 1 + 2.5\varphi \right)$$  4-2

This works only for small values of φ. Because it is designed for hard spheres it over-estimates the viscosity for relatively low viscosity drops. The Taylor formula is equivalent to Einstein at high oil viscosities (i.e. near-solid drops) and incorporates a factor k=$\eta_{drop}$/$\eta_0$, i.e. the ratio of the viscosity of the fluid in the drop to that of the bulk phase:

$$\eta = \eta_0 \left( 1 + \frac{5k + 2}{2(k + 1)} \varphi \right)$$  4-3

Another popular formula is Doughery-Krieger:

$$\eta = \eta_0 \left( 1 - \frac{\varphi}{\varphi_c} \right)^{-2.5\varphi_c}$$  4-4

Here $\varphi_c$ is the limit of close-packing of spheres, typically 0.74 which is the value for ordered spheres though 0.64 is also used which is the value for random packed spheres. The factor of 2.5 is the standard assumption for an "intrinsic viscosity". While Einstein and Taylor greatly under-estimate viscosities at high φ,

Doughery-Krieger starts to over-estimate them because it is a formula for solid spheres, not emulsion droplets.

For our needs, the Yaron, Gal-Or formula is provably the best. In addition to the volume fraction φ (along with $\lambda = \varphi^{0.333}$) it uses the viscosity ratio k to make a (modest) difference to the overall viscosity. The formula is a monster:

$$\eta = \eta_0 \left[ 1 + 5.5\varphi \frac{\left[ 4\lambda^7 + 10 - \frac{84}{11}\lambda^2 + \frac{4}{k}\left(1-\lambda^7\right) \right]}{10\left(1-\lambda^{10}\right) - 25\lambda^3\left(1-\lambda^4\right) + \frac{10}{k}\left(1-\lambda^3\right)\left(1-\lambda^7\right)} \right] \qquad \text{4-5}$$

The app becomes less reliable when φ>0.6. Up to that point the radius of the emulsion drops makes very little difference. Above that things get too complicated for the app to be reliable. Doughery-Krieger is too strongly dependent on a rather subjective valuation of what the close-packing limit should be. More importantly, smaller drops showing a bigger increase in viscosity[30] and non-Newtonian behaviour (not significant below 0.6) becomes significant. Intriguingly, at high shear rates the Yaron, Gal-Or formula is remarkably accurate because the non-Newtonian effects over-ride the strong particle-size dependent increase – that is why the app has been allowed to calculate at high volume fractions.



**App 4-3** https://www.stevenabbott.co.uk/practical-surfactants/EmuVisc.php
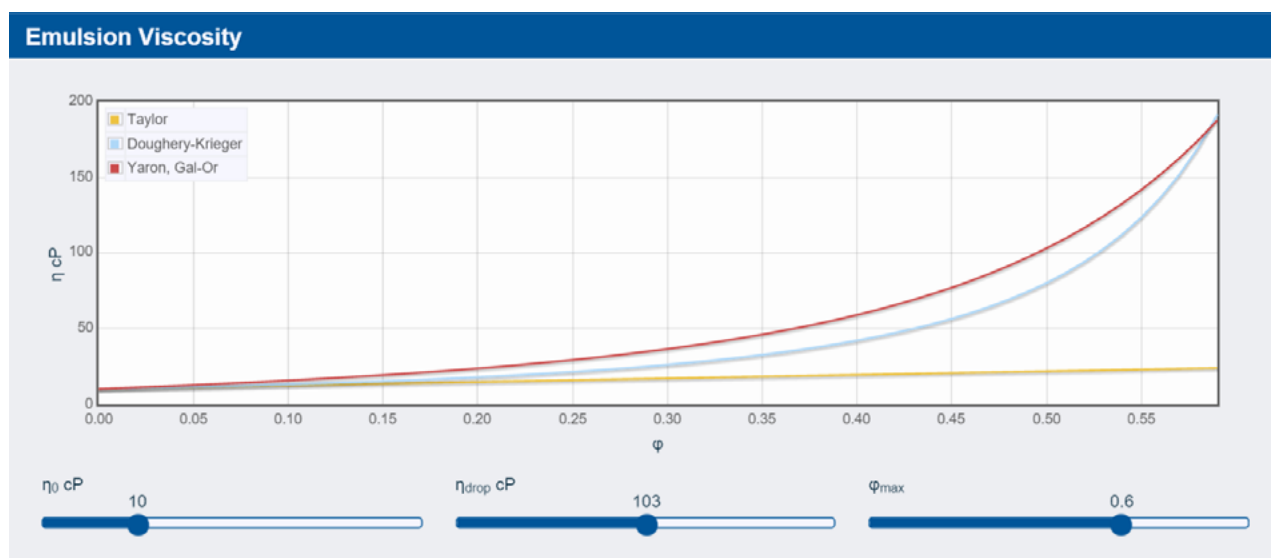
The app shows two things. First, for those wanting a high viscosity emulsion the best strategy is to start with a high viscosity – either thickening the water with glycerol or some polymer or by having a water in (viscous) oil emulsion. Second, the effects of φ aren't all that significant. By the time φ reaches 0.6 there are probably more important issues such as phase inversion, and, as hinted above,

30  Svetlana R. Derkach, Rheology of emulsions, Advances in Colloid and Interface Science 151 (2009) 1–23

over 0.6 things become complicated as particle size becomes important and particle size distributions play a big role in defining the close packing limit for Doughery-Krieger.

The Yaron, Gal-Or formula will turn out to be important in a discussion on why high φ emulsions can, in the right circumstances, naturally produce smaller drops.

## 4.4    Avoiding destruction

The astonishingly low amount of energy required (in theory) to make an emulsion points to the real problem of making them. The drops, once formed, readily return to their original state. Any emulsion-forming process will tend to re-make drops many-times over. The "once formed" bit is also not as easy as it sounds – the chances that the energy being delivered will actually cause a small drop to form are usually low and vary enormously throughout the emulsification system. The transfer of energy into drops is a huge subject within fluid dynamics and mixing and is discussed later. For the moment note the oft-quoted point from Braginsky[31] that a serious mismatch (in either direction) from a viscosity ratio of 1 between dispersed and continuous phase leads to a significant decrease in efficiency and therefore a larger drop size, with a ratio >3.7 making the process impossible. We will explore a generic "drop size" app later on. For now, the assumption is that a drop of some radius has formed and the question is: "What happens next?" to which the answer usually is "It goes back into the bulk oil or water". The following sections assume that the instant destruction has *not* taken place and instead describe the various ways in which emulsions self-destruct over somewhat longer timescales (seconds to hours) during and after manufacture. The most difficult of these, coalescence, is discussed after the easier ones have been dealt with.

### 4.4.1  Creaming

Gravity causes lighter oil drops to rise. If a lot of drops reach the surface and coalesce then there is a layer of "cream" on the top, so this failure process is called creaming.

The same physics causes heavier particles to fall. This is called sedimentation. Confusingly, many papers on creaming call the process sedimentation – to a physicist, sedimentation with a negative velocity (i.e. upwards) is still sedimentation. I will stick to more intuitive nomenclature because O/W emulsions are more common. For those who routinely deal with W/O emulsions then "sedimentation" is indeed the normal process.

---

31  Braginsky, L. M.; Belevitskaya, M. A., In Liquid-Liquid Systems; Kulov, N. N., Ed. Nova Science: Commack, 1994

The velocity, v, with which a drop moves is a balance of forces. The upward motion depends on the volume of the drop (those who remember Archimedes remember that "a body immersed in a liquid experiences an upthrust equal to the weight of fluid displaced"), so is proportional to $r^3$ and the resistance to flow is proportional to the diameter, i.e. $r^1$ so on balance creaming depends on $r^2$, therefore doubling the size of a drop quadruples the rate at which it creams. The density difference between water and oil ($\rho_w - \rho_o$) and the viscosity of the fluid, $\eta$ are the other inputs required for the Stokes' equation which tells us that:

$$v = \frac{2gr^2\left(\rho_w - \rho_o\right)}{9\eta} \qquad 4\text{-}6$$

**Creaming**

| Radius nm | η cP | ρo g/cc | ρw g/cc |
| --- | --- | --- | --- |
| 4972 | 5 | 0.8 | 1 |

| Height cm | Velocity mm/hr | Hours | |
| --- | --- | --- | --- |
| 5 | 7.8 | 6.4 | |

**App 4-4** https://www.stevenabbott.co.uk/practical-surfactants/Creaming.php

The specific example shows that drops of 5μm radius of an oil with a density of 0.8 in a 5cP dispersion will rise at 7.8mm/hr, so for a small 5cm sample in a pot, the whole thing will have creamed in 50/7.8=6.4 hours. The numbers are only illustrative, many other things will have happened in those hours. The point of the app is to give a feel for the effects. Everything is linear except for r, so tweaks to density or viscosity have a smaller effect than tweaks to the radius. However, many long-term stability issues are solved by having a high viscosity, and it has often been noted that concentrated emulsions (with, as we have seen, viscosities 10x higher) can be more stable than dilute ones.

Indeed, given enough time, any emulsion stabilized solely by viscosity will slowly destabilize into a two-phase system due to creaming. Using the yield stress of a fluid instead of its viscosity could be a wiser way to stabilize against this phenomenon. The yield stress is the minimum stress required to initiate flow in a fluid (or for an object to move through it). It can be measured using a classical cone-plate rheometer and can provide you with the maximum size of droplets that can remain fully suspended within it. In fact, if the yield stress is higher than the creaming stress induced by gravity ($\sigma = \Delta\rho.g.r$), your droplets will remain indefinitely suspended in your liquid, regardless of the product's lifetime. Physically, this yield stress can arise from the presence of polymers that create a lattice in your external phase or be directly induced by the high concentration of droplets in your emulsion. For a classical polydisperse emulsion, large yield stresses appear when the internal phase exceeds 70%. This is captured in an app: https://www.stevenabbott.co.uk/practical-rheology/Emulsion-Yield.php.

However, as pointed out by Dr Duncan Gilbert who kindly alerted me to this issue and provided the text above, the yield stress to stop a typical emulsion from creaming is in the range of 0.1 Pa. Although this can be measured, I know of no way to predict or control this small amount of yield stress. As is so often the case, being able to work on a good rheometer and knowing what you are looking for, can open up formulation possibilities that might otherwise be missing.

A more complex/realistic creaming model is discussed in the flocculation section.
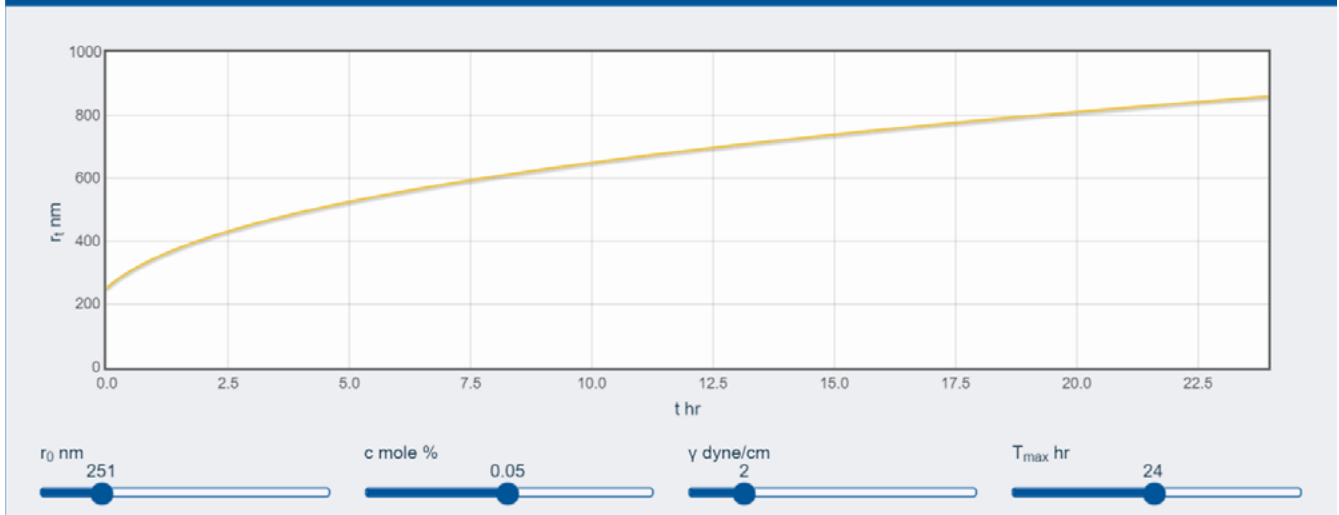
### 4.4.2 Ostwald Ripening

For a given interfacial tension, γ, the pressure inside an emulsion droplet of radius r is proportional to γ/r. If two drops of different radii were somehow connected via a thin pipe (the same effect is shown with bubbles where a pipe connects a smaller and larger bubble) then, contrary to intuitions, the two bubbles do not change to some intermediate size. Instead the smaller one gets smaller and the larger gets larger – driven purely by the pressure difference which gets larger (the effect accelerates) as the process continues. In real emulsions the drops are connected indirectly via any oil molecules that happen to be soluble in the water between the drops.

The process by which the smaller drops move over time to larger ones is called Ostwald Ripening and the process is somewhat difficult to model. A good-enough approximation is the LSW (Lifshitz-Slyozov-Wagner) model which states that for an emulsion with average radius $r_0$ at t=0, interfacial tension γ, with an oil diffusion coefficient D, concentration (or solubility) c and molar volume V then the radius at time t, $r_t$ is given by:

$$\langle r_t^3 \rangle - \langle r_0^3 \rangle = \frac{8\gamma DcVt}{9RT}.$$   4-7

The <brackets> highlight that these are average radii. An optional extra term f(φ) can be added to reflect the fact that higher volume fractions, φ, tend to show faster ripening; so f(φ) can vary from 1 when φ~0 to 2.5 when φ~0.3. The equation (minus the φ effect) is brought to life in the app:

**Ostwald Ripening**

$r_0$ nm — 251
c mole % — 0.05
$\gamma$ dyne/cm — 2
$T_{max}$ hr — 24

**App 4-5** https://www.stevenabbott.co.uk/practical-surfactants/Ostwald.php

Although the process is relatively slow in the long term, the initial increase in average particle size is quite rapid, so Ostwald ripening can render a desirable fine emulsion into a relatively coarse one, even if that coarse one only slowly becomes coarser. Moving the mouse over the graph gives a precise readout at each point (and you can zoom in by changing $T_{max}$). In this example, the radius has doubled to 500nm in just over 4 hours, even if it then takes 40 hours to quadruple to 1µm.

The situation is even more severe when nanoemulsions are desired. Using the same conditions as above but starting with a 50nm radius, the drops are over 200nm radius after a mere 20min. As is often pointed out in reviews of nanoemulsions, the challenge of making them (which can be tricky) is often less than the challenge of stopping them from becoming macroemulsions via Ostwald ripening.

The formula tells us that a low interfacial tension, which is obtained near HLD=0, reduces Ostwald Ripening, though unfortunately it greatly increases coalescence.

The concentration term, c needs some explanation. At first this tells us that oilier oils ripen more slowly because they are less soluble in water and cannot so easily shuttle between small and large drops. High EACN values are, in this respect, an advantage, so for a given surfactant very far from HLD=0, decane emulsions ripen much faster than hexadecane ones. If our mission is to formulate with a given oil then we have no control over c in pure solubility terms.

Given that the solubility of the oil in any micelles is likely to be rather high it is has often been suggested that the more surfactant that exists as micelles, rather than at the emulsion interface, the higher the effective c is going to be. The oil can go from a smaller drop into a micelle and out of the micelle into a

larger drop. Given that micelles are in rapid statistical equilibrium this can be a relatively rapid process, though admittedly the diffusion coefficient, $D$, will be less than that of the pure oil. The literature on such effects is full of confusion and contradictions because the whole system is complex. Clearly there are no simple rules. It seems reasonable to assume that higher micellar concentrations (i.e. "wasted" surfactant) might increase Ostwald ripening. A much-referenced paper by Kabalnov[32] expresses confidence that simple ionics like SDS show no significant effect, even at very high surfactant concentrations and counter-examples do not seem to be plentiful. For nonionics the effects seem to be surprisingly small. Theory suggests that micelles should give 100x increases in rates of ripening and experiments suggest, at most, factors of 3-5.

For those who have to use relatively water-soluble oils, such as those doing emulsion polymerisation of vinyl acetate or styrene, Ostwald ripening is a severe problem, easily solved with a very simple trick. Adding 1-2% hexadecane (or even better, squalane) to the oil brings the ripening to a complete halt. There are competing descriptions of how this effect works, one based on concentration gradients and another on osmotic pressure effects, though the two are thermodynamically equivalent. What happens is that the hexadecane hardly wants to leave an oil drop. As the more soluble oil leaves, the hexadecane concentration rises rapidly, creating a chemical potential imbalance. This makes it impossible for the drop to shrink any further and ripening halts. At first it seems surprising that such a small amount of hexadecane can have such a large effect. There is no doubt, though, that the trick works very well.

For W/O emulsions Ostwald ripening can be severe because the small water molecule can rapidly diffuse and its solubility in oils other than alkanes is significant. The equivalent to the hexadecane trick is to add some salt to the water. As the salt is totally insoluble in the oil it brings ripening to a halt via the same type of explanation based on concentration and/or osmotic pressure.

A paper[33] that combines a number of themes discussed so far makes some beautiful, small, stable nanoemulsions. They use triglycerides (peanut oil) of high viscosity, so to get below the 3.7 viscosity ratio domain (discussed below) they have to thicken the water with polymer. And because the triglycerides are highly insoluble in water there is effectively zero Ostwald ripening.

There is one other strategy for reducing Ostwald ripening. If smaller drops automatically had a lower surface tension then there would be no driving force. At first this seems impossible. But recall the discussions around the DST behaviour of lung surfactants. If a drop shrinks and compresses the surfactant

32   Alexey S. Kabalnov, Can Micelles Mediate a Mass Transfer between Oil Droplets?, Langmuir 1994, IO, 680-684

33   Tim J. Wooster, Matt Golding, and Peerasak Sanguansri, Impact of Oil Type on Nanoemulsion Formation and Ostwald Ripening Stability, Langmuir 2008, 24, 12758-12765

the surface tension will go down. Normally the compression would eject the excess surfactant to restore the surface tension to normal. But if the "relaxation" parameter is near zero, i.e. if the surfactant does not want to exit into the aqueous phase, then the surface tension will remain low and the Ostwald driving force is removed. There are plenty of examples where this strategy of using insoluble surfactants works, if not perfectly at least well enough.

A more sophisticated Ostwald model which also follows the change of the size *distribution* (rather than just the average size) can be found in the Foams chapter.

### 4.4.3 Flocculation

When two droplets join together yet remain as individual drops, that is flocculation. If they join together and the interface between them breaks to form a single drop, that is coalescence which is discussed in the next section. Whether coalescence goes via flocculation or is a single-stage process depends on many things. In this section, it is assumed that the drops have a really tough shell (which is often the case in food science where the surfactants are large proteins) so there is no coalescence.

Flocculation requires particles to bump into each other. It is common to distinguish between two processes which operate on different sizes of drops:

- Perikinetic: drops bumping into each other because of Brownian motion, i.e. energies in the kT regime.

- Orthokinetic: drops bumping because into each other due to other motions such as stirring or via buoyancy which in turn is caused by flocculation from perikinetic effects...

Some numbers can be put to this. The rate of orthokinetic flocculation depends on r³ and on a velocity gradient G, i.e. the velocity difference between two particles divided by the distance between them. Perikinetic flocculation depends on kT, Boltzmann's constant times absolute temperature. The *relative* rate for the two processes is therefore given by:

$$\frac{Ortho}{Peri} = \frac{4\eta Gr^3}{kT} \qquad \text{4-8}$$

In a typical emulsion, η might be 10cP and a 1µm drop of an oil of density 0.8 will have a velocity relative to a 0.5µm drop of ~50µm/s. So the velocity gradient over a 100µm distance is 0.5/s. This means that the Ortho/Peri ratio is ~5, so already drops with r=1µm will start to be controlled by creaming motions. It is common to see statements that the theories of flocculation and coalescence break down for emulsions larger than, say, 1µm and the crude ortho/peri ratio

calculation shows that this is a plausible notion. Despite this, it is also common to see the theories applied to larger droplets.

Flocculation can produce some desirable effects; for example the emulsion might gel and gain long-term stability against creaming. Here, though, the assumption is that flocculation is a bad thing because the effective radius of two drops that have joined together is much larger so they cream faster. As these conjoined drops rise they can bump into others (single or joined) and flocculate further to give accelerated creaming. At the same time, the presence of flocculates reduces the free space for motion and decreases the vertical velocity so creaming decelerates. The whole process is complex. The approach adopted by Gurkov's group in Sofia[34] is sufficiently sophisticated to capture much of the complexity while being sufficiently simple for the principles to remain clear and to be implemented in an app. The theory can be matched to experimental data with a minimum of fitting parameters which in any event provide a lot of insight. Remember that there are likely to be plenty of exceptions to any relatively comprehensible theory. The approach explicitly rules out coalescence and Ostwald ripening, i.e. it works best for tough emulsion shells (the paper used protein surfactants) and relatively insoluble oils.

There is another restriction for the app. The assumption is that the starting emulsion is relatively mono-disperse. This tends to give a "clean" creaming where the boundary between a creamy phase and a clear phase is sharp. Highly polydisperse emulsions have super-fast creaming of the large drops (as we saw in the simple creaming section, a 20μm drop rises 100x faster than a 2μm drop) and the smaller ones take longer to catch up to the interface. Or, to put it another way, in polydisperse emulsions there are two mechanisms, orthokinetic for large drops and perikinetic for small drops rather than a single mechanism assumed here.

The starting point is Stokes law for the upward motion of the individual particles, $U_0$. For particles of radius r and density difference Δρ in a medium of viscosity η it is given by:

$$U_0 = \frac{2gr^2\Delta\rho}{9\eta} \qquad \text{4-9}$$

The next point is that instead of the simple viscosity assumed in the ideal creaming case there is a crowding viscous drag that is conventionally assumed to have a functional form of $(1-\varphi)^{5.4}$ where φ is the local oil volume fraction.

Then there is the effect of the crowding during creaming on the vertical velocity. The idea is that the crowding creates a "pressure" that slows the process. This

34   Tatiana D. Dimitrova, Theodor D. Gurkov et. al., Kinetics of Cream Formation by the Mechanism of Consolidation in Flocculating Emulsions, Journal of Colloid and Interface Science 230, 254–267 (2000)

pressure, p, depends on φ and is assumed here to be proportional to $φ^4$, starting with an intrinsic, dimensionless pressure $p_0$ which is typically $10^{-3}$.



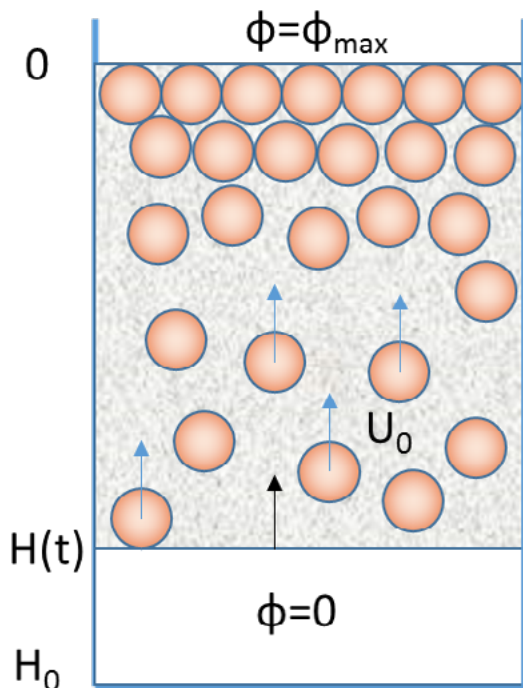**Figure 4-1 The key elements in the creaming by flocculation in an emulsion of initial height Ho and particle velocity U0. The boundary at time t is at H(t), with φ=0 below it. At the top the oil creams to a density φmax.**

With the basics in place the calculation starts with height $H_o$ and volume fraction $φ_o$ throughout the tube other than the boundary condition concentrations of 0 at the bottom and the close-packed equilibrium value of $φ_{max}$ at the top. At each timestep the particles rise with velocity $U_0$ so it is possible to go through each layer tracking the increase of concentration from particles below and decrease from particles rising – a classic differential equation problem. The complication is that as concentrations rise the φ effects on pressure and friction affect everything and the integration becomes very complicated. Fortunately the team found a suitable methodology and even more fortunately for me they kindly sent me their C code from which I was able to create the Javascript code for the app – which acknowledges their assistance.

An interesting feature of the *calculation* (rather than the outputs) is that it is independent of most of the input values. What this means is that for a given $φ_o$ and $φ_{max}$ there is a single form of the solution which works in dimensionless space. The input values are merely used to translate the generic solution into real-world outputs. The key to the translation is the ratio of height to particle velocity, $H_0/U_0$, which makes intuitive sense.

**App 4-6** https://www.stevenabbott.co.uk/practical-surfactants/Flocculation.php
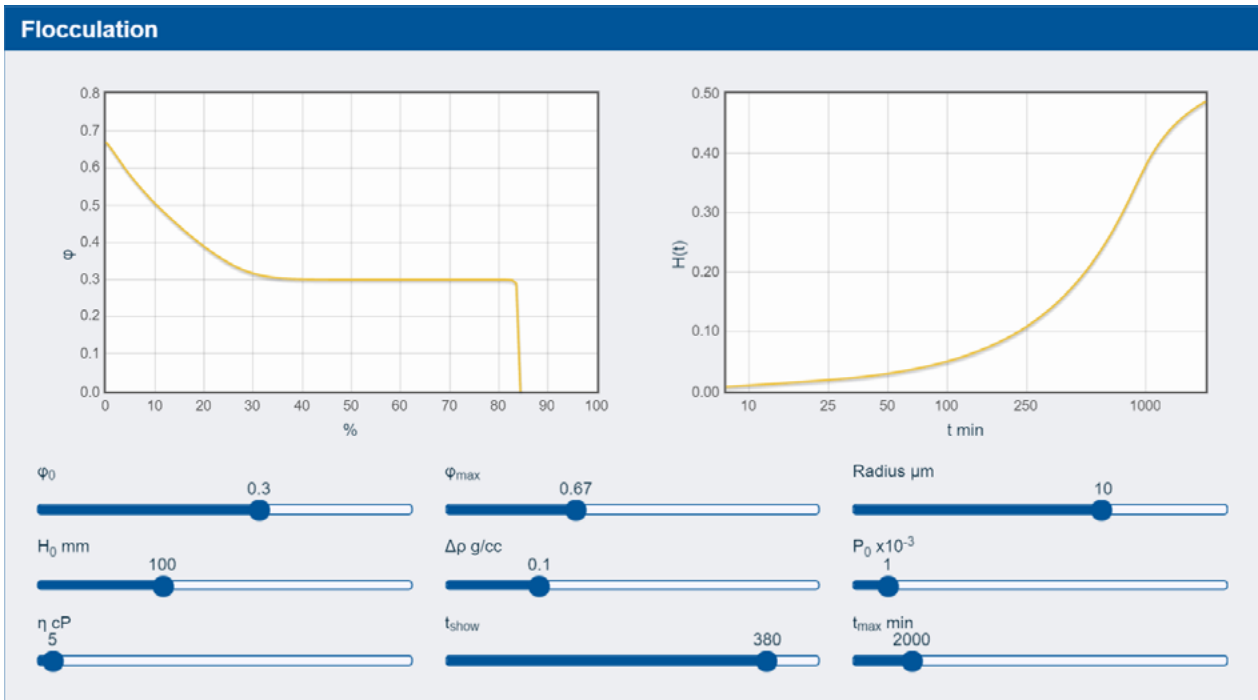
The app shows two views of the data. The graph on the right plots the boundary height H(t) versus time. In this example very little happens for the first hour, then gradually the interface between emulsion and clear water starts to rise, approaching the final value after 2000min. The graph at the left shows the distribution of the particles throughout the column at the chosen time ($t_{show}$) of 380min. There is clear water from the bottom up to 80% of the column depth, then there is a plateau at the original 0.3 volume fraction reaching 40% and after that the concentration rises rapidly to its maximum value.

One function of the app is to explore the effect of the two variables directly under your control, r and η. Not surprisingly, halving r provides a more dramatic reduction in flocculation than doubling η because $U_0$ depends on r².

### 4.4.4  Coalescence

Having looked at pure flocculation, we now look at pure coalescence – ignoring the issue of whether the coalescence proceeds via flocculation. The key difference between the two phenomena is that the surfactant monolayers easily fuse together in coalescence to give a single drop; whether that fusion is instantaneous or goes via a metastable flocculation is of no direct relevance to the calculations.

The first part of any discussion on coalescence is easy. The more collisions there are between emulsion particles, the more coalescence there will be. The number of collisions in turn depends on the number of particles per unit volume and their average velocity which derives from thermal energy, i.e. Brownian motion, which is calculated, as before, in terms of kT, Boltzmann's constant times absolute temperature. This means, as defined above, that we

are considering perikinetic effects. Instead of velocity it is more usual to use an equivalent measure, the Diffusion Coefficient, D which depends on radius (resistance to motion in a fluid is proportional to radius) and viscosity. It is given by:

$$D = \frac{kT}{6\pi\eta r} \quad \text{4-10}$$

Thanks to von Smoluchowski we know that the rate at which particles collide depends on R, the collision distance which we can take as 2r and the number of particles per unit volume, n:

$$CollisionRate = 4\pi DRn^2 \quad \text{4-11}$$

We also know that the surfactant provides a barrier that reduces the chance of coalescence. The assumption is that this barrier is of an energy E so the coalescence rate is given by:

$$CoalescenceRate = 4\pi DRn^2 e^{-\frac{E}{kT}} \quad \text{4-12}$$

Because we know D and because R=2r we can get:

$$CoalescenceRate = \frac{4kT}{3\eta} n^2 e^{-\frac{E}{kT}} \quad \text{4-13}$$

Note that the coalescence rate does not appear to depend on the radius – though for a given fixed volume of oil the initial number of particles, n, is higher when they are smaller. Integrating this gives the final equation in terms of the particle density at time t $n_t$ given an initial density $n_0$ at t=0:

$$\frac{1}{n_t} = \frac{1}{n_0} + \frac{4kT}{3\eta} t e^{-\frac{E}{kT}} \quad \text{4-14}$$

There are many problems with this formula. Generally we have no idea what E is and as the drops coalesce the interfacial surface area goes down so E may increase if the original drops were starved of surfactant. E also depends on whether the particles are deformable or, effectively, hard spheres. The rule of thumb is that the model is easy when particles are below a few μm and (because of curvature) are relatively hard, and difficult for large drops where deformability adds many complications. In any event, for larger drops orthokinetics probably dominate so this perikinetic theory is irrelevant. As mentioned above, there are extra complications about whether the drops simply coalesce or whether they first flocculate - i.e. clump together while remaining as individual drops. Furthermore, in general we have no interest in n and every interest in particle size, and going from n to particle size is not trivial.

Despite all the problems, it is still better to have a crude model rather than none at all. The inputs to the app are oil concentration and drop particle size (from which $n_0$ can be calculated), the viscosity $\eta$ and then E in terms of kT, remembering that a barrier of just 7kT makes the exponential term equal to <0.001. The conversion of n to drop size uses the simplest of algorithms that distributes the original volume over $n_t$ drops, each of the same radius. You can choose to look in n-mode (number of drops versus time) or r-mode (radius of drops versus time).



App 4-7  https://www.stevenabbott.co.uk/practical-surfactants/Coalescence.php
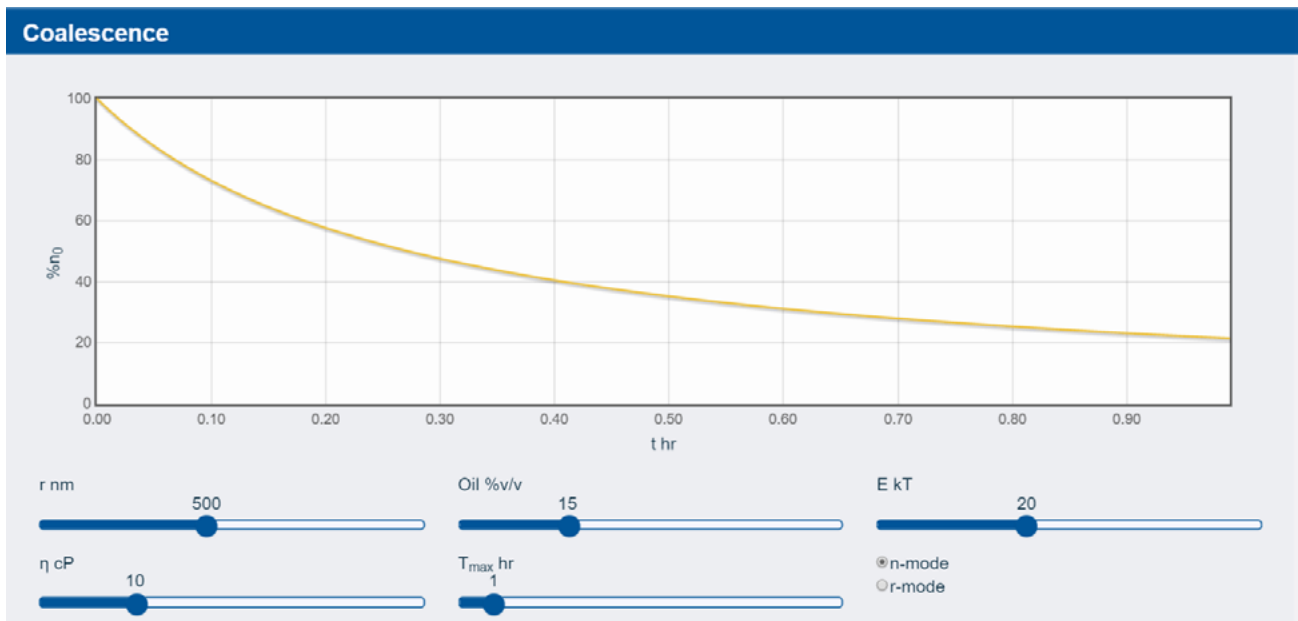
In this example, the number of particles has reduced to 20% of the original after 1hr. In radius mode, the original 500nm radius has increased (when viewed in r-mode) to 830nm in the same time.
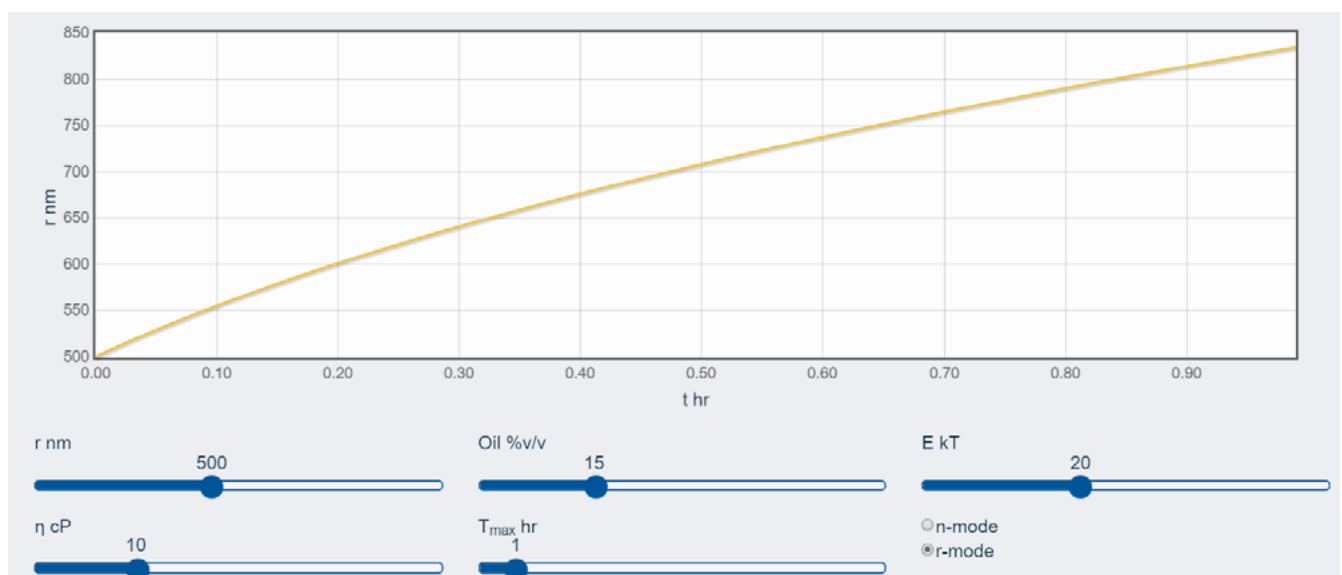


Figure 4-2 The same conditions, plotted in terms of the simplistic drop radius.

The examples use a barrier of 20kT. Because of the exponential nature of the E term, a small reduction to just 17kT means that the number of particles reduces to 20% after 3min rather than 1hr.

So forgetting about all the other issues behind this simplified model, the key message is that coalescence is strongly dependent on the barrier between particles. So what are the barriers to coalescence? The value of 20kT used in this example was not chosen at random; as we will see, 20kT is generally considered to be the dividing line between unstable and stable colloidal particles.

Note the unfortunate contradiction between flocculation and coalescence. Flocculation is driven by particle *velocity* which decreases rapidly (radius²) with decreasing particle radius. Coalescence is driven by particle *numbers* which, for a given volume fraction, increase rapidly (radius³) with decreasing particle radius. Given that the effects are likely to be coupled and that other factors (such as anti-flocculation "pressure" and coalescence kT barrier) may have controlling influences, whether a small radius is "good" or "bad" is a question with no simple answer.

### 4.4.5  DLVO

This is not the book for describing DLVO (Derjaguin and Landau, Verwey and Overbeek) theory. An app on my website [https://www.stevenabbott.co.uk/PracticalSolubility/DLVO.php](https://www.stevenabbott.co.uk/PracticalSolubility/DLVO.php) makes some attempt at describing the full theory. In truth the theory isn't all that useful to most formulators because many of the input parameters are unknowable and, in addition, there are plenty of effects other than DLVO to complicate the story. Still, it is better than nothing and the simplified version here gives some feeling for what levers we can pull to stabilise an emulsion.

The aim of DLVO is to calculate the net force $V_T$ at an inter-particle distance h, from a combination of one attractive force, $V_H$ and two repulsive forces, $V_D$ and $V_S$.

The key problem is that all droplets are attracted to each other via the van der Waals force. The size of the force, $V_H$, is described by the Hamaker constant, $A_{12}$ which is somewhere in the region of $10^{-20}$J (10 zeptoJoules) for typical drops. As the equation shows, the bigger the radius, r, the larger the force.

$$V_H = -\frac{A_{12}r}{12hkT} \quad \text{4-15}$$

For droplets with a charge around them there is the Debye repulsive force. This is a complicated behaviour involving multiple factors:

$$V_D = \frac{2\pi e_0 \varepsilon r \varphi^2 \ln\left(1 + e^{-\frac{h}{k^{-1}}}\right)}{k_B T} \qquad 4\text{-}16$$

Most of the complexity has been removed for this app and the key terms are r (so larger particles repel each other more) and φ which can be assumed to be equal to the Zeta potential (see https://www.stevenabbott.co.uk/PracticalSolubility/Zeta.php for a fuller explanation) which is a measurable parameter describing (simplifying greatly) the net charge at the surface. The curious $k^{-1}$ term is the Debye length and it depends on the concentration of ions in the solution. Adding salt reduces the overall repulsion because the charges at the surface get swamped by those in solution.

Finally, droplets with chains sticking out from the surface (the standard example is EO chains in ethoxylate surfactants) of length δ exhibit a steric repulsion, $V_s$:

$$V_S = 30 N_A 4\pi r \Gamma^2 \left(0.5 - X\right)\left(1 - \frac{h}{2\delta}\right)^2 / \left(\rho^2 MVol\right) \qquad 4\text{-}17$$

which, fortunately for us, simplifies to 0 when the distance is larger than 2δ and "very large" when the chains have a chance to touch. In the simplified app all the other parameters are set to reasonable values, though it is worth noting that the repulsion depends on Γ which is represents the density of chains sticking out – so a low density means a low repulsion.

The app does all the hard work for you:



**App 4-8** https://www.stevenabbott.co.uk/practical-surfactants/DLVO.php
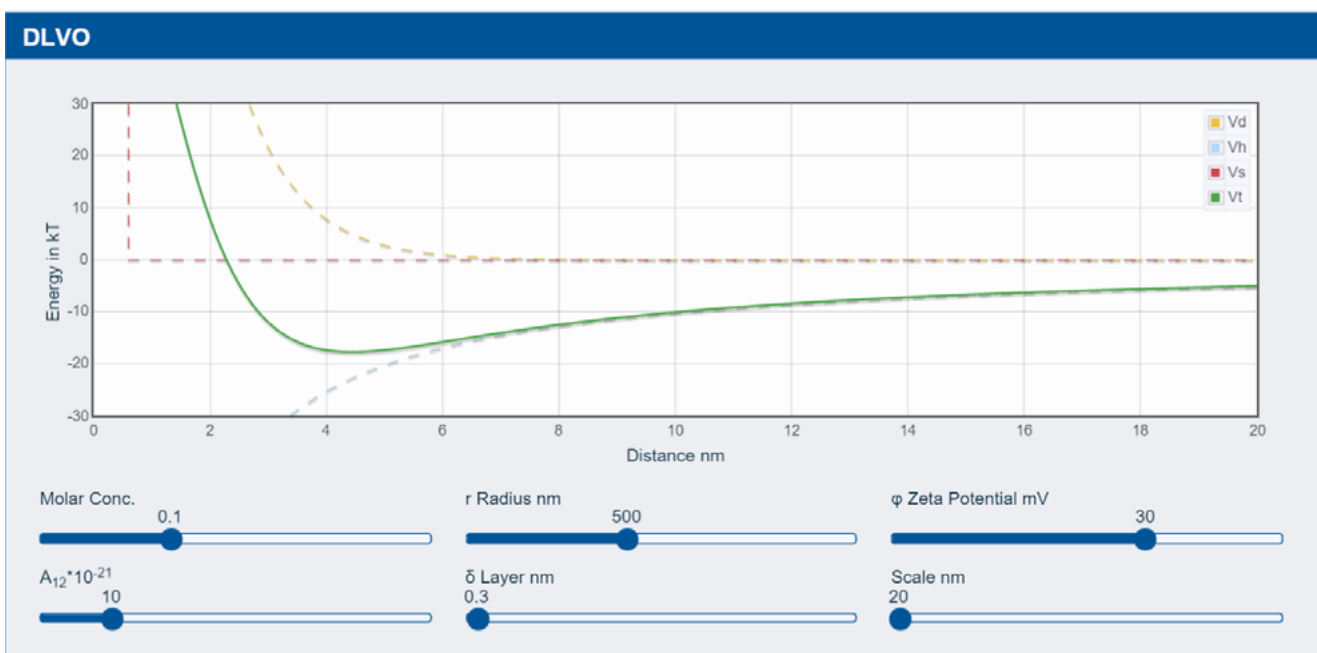
The dotted lines are the individual terms and the Y-axis is plotted in terms of kT. $V_H$ is the blue dotted line going negative (attractive), the red dotted line that suddenly turns upwards at 2δ is $V_S$ and the yellow dotted line is $V_D$. The total is in green and shows that the barrier is above 20kT by 1.6nm. In the world of colloids and nanoparticles, 20kT is generally considered to be the dividing line between stable and unstable formulations. Is this good enough to provide adequate protection? In theory, yes – the drops have to get to 0nm before they coalesce. In practice, who knows? A real emulsion drop is not a perfect sphere, it might well be swirling and pulsating from the mixing energy, dynamic surface tension effects might mean that there is insufficient surfactant to deliver a 30mV zeta potential, there might (randomly) be incomplete surfactant coverage at the point of collision, etc. etc.

The point of the app is not to create worries about whether zeta should be 30 or 31mV. The point is to show the general ideas that are used in the emulsions and colloids domain, and highlight the fact that having a large δ and/or a large zeta potential is necessary to give the sorts of barriers demanded by coalescence theory.

Because modern technology makes it easy to measure zeta potentials it is especially important to never measure the potential of a *single* formulation. Small impurities, small changes in pH, small formulation changes can flip zeta from a healthy 35mV to an unhelpful 10mV. By measuring zeta on some formulations deliberately tweaked to represent the sorts of changes that might be possible in real life it is easy to find whether you are in a safe zone where zeta is only modestly affected by formulation changes or whether you happen to be near a knife-edge where disaster could strike unexpectedly.

The steric term for O/W emulsions is ~0.1nm/EO unit, so it takes 10 EO units to produce a 1nm chain. APGs are ~1nm/ring and sorbitan is somewhat smaller. For W/O emulsions, the steric barriers are of a similar size. A C10 alkyl chain is ~1.5nm. However, these numbers should not be taken too seriously. If the emulsion drop is partially covered or if the packing of the portion sticking out (EO, sugars etc. for O/W, alkyl chain for W/O) is rather poor then the surfactant is only a weak barrier. This brings us once again to the notion that we should be attending at least as much to elasticity ideas (which are relatively little discussed) as to ideas of the more popular DLVO theory.

### 4.4.6  Dynamic barriers: Elasticity, Gibbs-Marangoni and Wedges

A fully-formed surfactant shell around an emulsion drop may well provide a barrier against coalescence. The problem during the production of an emulsion is that there is a very high probability of ill-formed droplets bumping into each other. This happens when a larger drop has just been broken into two by whatever dispersion process is being used. The drops are close to each other,

and they don't have a complete shell, so why don't they simply fuse if a slight movement causes them to bump back into each other?

The answer is that they frequently do – emulsion creating is massively inefficient in terms of energy applied. Still, many of them don't. And there are a range of possible answers for why they don't.

The first is based on an elasticity argument. As we have discussed previously, the pull at an interface is resisted by an elastic force from the surfactant system. If, however, there is a vast supply of surfactant (i.e. within the phase containing the surfactant) then as the interface is stretched there is no resistance because a fresh surfactant molecule quickly fills any gap, restoring the tension to its original value. It is like stretching a spring where fresh pieces of metal are continuously being added to compensate for the extra length – no force is required. So there is no resistance to drop recombination in the phase containing the surfactant, and only drops in the other phase survive. This is a typical explanation of Bancroft's rule (or, as discussed before, rule of thumb) that the phase in which an emulsifier is more soluble constitutes the continuous phase.

Another favourite answer is based on Gibbs-Marangoni effects.

Let us suppose that the two drops are starting the process of coming together. Given the hypothesis that the surfaces have an incomplete coverage of surfactant, as they move together they will attract any surfactant molecules in their path. Very quickly the zone between the drops becomes depleted of surfactant so that parts of the surface that are meant to touch have the least surfactant concentration. This means that there is a surface tension gradient (Gibbs) which drives surfactant molecules in the direction of the gap and these molecules sweep fluid (Marangoni) along with them, producing a force which actively separates the drops.

Given that surface energy forces are rather weak and that forces pushing droplets together can be rather strong, this doesn't sound too convincing. Another way to look at it is to imagine that the bringing the drops together sweeps surfactant from the surface of the drop. This creates a surface tension gradient of, say, 10 mM/m. Without some dramatic magnification this is of no significance. The magnification comes because the drop is curved with a radius r so a stress is created equal to 2.10/r, with the factor of 2 coming from the two drops. If r is 0.5µm this is 40kPa which is 0.4atm, a significant stress.

Gibbs-Marangoni provides an explanation of the Bancroft rule that is complementary to the Danov explanation discussed earlier. The "sweeping away" of the surfactant can only take place in the phase where the surfactant is most soluble, so the drops themselves are formed in the phase where the surfactant is less soluble.

One problem with "Bancroft explanations" is that there are so many of them. It is not clear to most of us whether the above explanations are basically the same or apply only under certain conditions.

Another problem is that they don't help us identify which specific surfactants and concentrations would be most appropriate for our specific oils and emulsion requirements. There are contradictions built in. We want high elasticity, which means a lot of surfactant, because at low surfactant concentrations there just cannot be a plausible amount of coverage to generate a strong-enough surface tension gradient. Yet too much surfactant means that the interface is totally saturated and can easily refill itself if the interface is stretched. We want a "fast" surfactant that goes rapidly to the surface to generate the elasticity, yet we want a "slow" surfactant that does not replenish the interface too rapidly upon stretching. We want an efficient (low CMC, i.e. high K) surfactant to get the most surface coverage with the least surfactant, yet large (polymeric) surfactants have desirable high elasticities as discussed earlier.

So although there is obviously a lot that is right with these sorts of explanations, they fail the definition of being great theories because they allow us to predict far too little.

Another problem with Bancroft explanations is that they say nothing about the intrinsic curvature at the oil/surfactant/water interface, the sort of curvature discussed within HLD-NAC. Indeed it is often stated with confidence that such curvatures *must* be irrelevant to the formation of conventional emulsions because those curvatures have radii in the range of nm while emulsions have curvatures in the range of µm. Yet again, the idea is that HLD-NAC is "only" to do with microemulsions and is irrelevant to real emulsions.

An early attempt to explain Bancroft via such curvature arguments was made by Harkins and Langmuir and was called the "oriented wedge" theory where the surfactants in the emulsion drop would want to orient themselves as if they were packed wedges, with the bigger part of the wedge on the outside. Although the theory came from the famous Langmuir, the famous Hildebrand, among others, was able to show that the theory must be nonsense because for a typical oil drop the interfaces (when seen from a molecular perspective) are effectively planar. So "microemulsion" thinking was proven to be irrelevant to real emulsions. And yet it cannot be denied that almost everything interesting that takes place in terms of understanding emulsions seems connected to curvature, with PIT emulsion techniques explicitly controlling curvature to achieve their aims.

Kabalnov[35] found an ingenious way to bridge the gap between the undeniable effects of molecular curvature and the undeniable fact that emulsion radii were too large for this effect to operate. The wedge does not operate at the level of the drop as a whole. Instead it operates at the one instant that matters: when a hole must be created in the bilayer between the two drops, so that the drops can fuse.
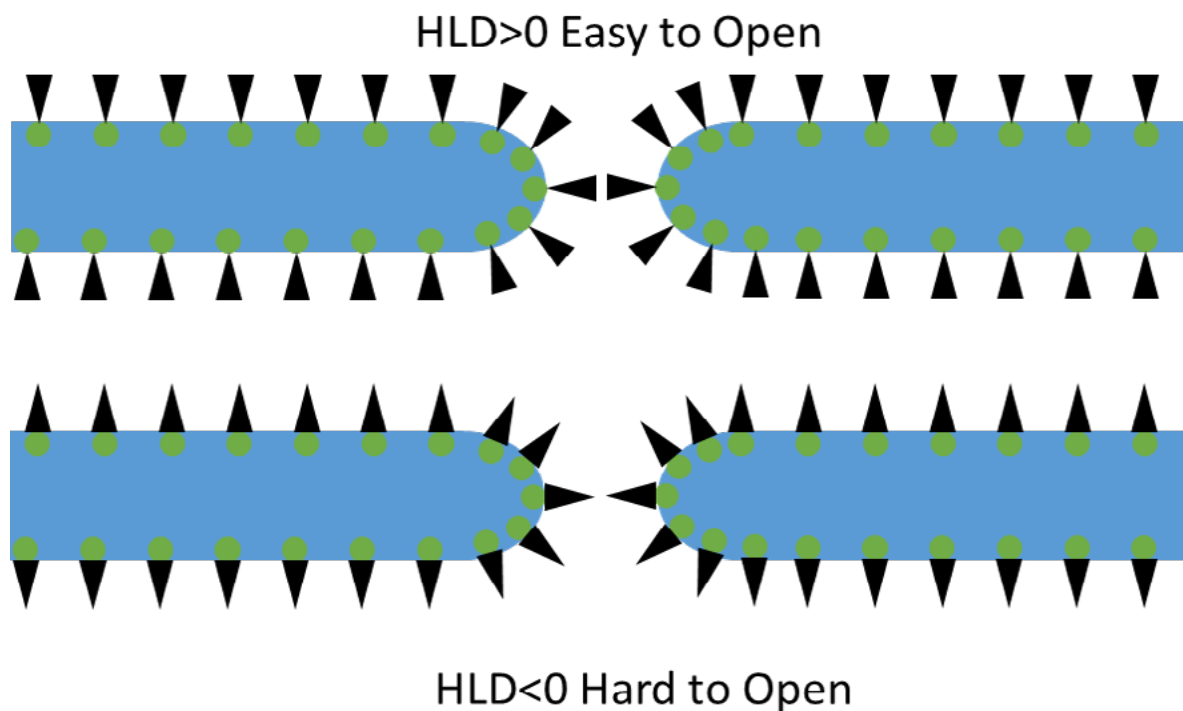


**Figure 4-3 The "wedge" explanation depending on natural curvature**

Suppose that water is the continuous phase. For two oil drops to fuse, a small hole must appear in the water, with surfactant flowing around the edge of the water. Now the radius of curvature is very small so molecular curvature effects *are* significant. If the curvature is the natural one (HLD<0) then the hole can be formed easily. If the curvature is the wrong way (HLD>0) then there is a large barrier to hole formation.

Kabalnov argues that although the various standard explanations are plausible, they lack the crucial link with curvature that *cannot* be ignored. Because, he argues, although there are plenty of exceptions to Bancroft, there are zero exceptions to PIT and, therefore, no exceptions to curvature.

Note that the graphics here, and those of Kabalnov use classic CPP shapes (large tails for HLD>0 and thin tails for HLD<0) for simplicity. But as we know the surfactant shape itself is not the key factor – it is the whole balance of oil, water, salinity and temperature.

---

35  Alexey Kabalnov and Håkan Wennerström, Macroemulsion Stability: The Oriented Wedge Theory *Revisited*, Langmuir 1996, 12, 276-292

A paper by Ruckenstein[36] acknowledges that the Kabalnov explanation might have some validity (calling it a "black film" explanation as it works for very thin films which, when they are in foam bubbles, are black), but only at high surfactant concentrations where the classic explanations break down because there is too small a surfactant concentration gradient. He also acknowledges that Bancroft breaks down at low surfactant concentrations when bulk mixing kinetic effects dominate because there is not enough surfactant to create any of the above effects.

A lot of this section makes no sense without a grasp of dynamic surface tension and elasticity. However, there is a word of caution about some aspects of DST. It has been argued[37] that the diffusion limiting aspect of DST might be irrelevant in emulsion formation because the mixing and turbulence give (convective) speeds faster than diffusion. This in turn takes us back to Danov's explanation which is based, in part, on interparticle velocities. No wonder there is no simple recipe for making an emulsion.

## 4.5      How to make an emulsion

As mentioned earlier, this is not a book about the science of creating droplets via impellers, high pressure nozzles, membranes, microfluidics, ultrasonics etc. Our question is: "Assuming we have made lots of drops of our desired size, how do we ensure they stay that size?"

After all our efforts at looking at the mechanism by which the drops grow larger, we are not much wiser in terms of practical choices.

Higher viscosities help (though they might make it harder to create the drops in the first place).

For ionics, large charges at the interface help in terms of zeta potential. There are multiple issues here. A nice charged groups such as a sulphate is also not too unhappy forming micelles in the aqueous phase so the partition coefficient with respect to the oil can be relatively low and the surfactant can be relatively inefficient. The charged groups repel across droplets but they also repel around a droplet, so the packing of surfactant is not all that tight and the interface is relatively non-rigid, making it more susceptible to coalescence (though no "rigidity" term has been explicitly added to the coalescence app, it is just one component of E). Any salts added to the formulation for other reasons will reduce the repulsion between drops. Any carboxylate surfactant will be affected by pH and/or the encounter with ions such as calcium which can help to form precipitates (scum).

36   Eli Ruckenstein, Microemulsions, Macroemulsions, and the Bancroft Rule, Langmuir 1996, 12, 6351-6353

37   Pieter Walstra, Principles of Emulsion Formation, Chemical Engineering Science, 48, 1993, 333-349

For nonionics the challenge is that it needs a lot of EO groups before steric repulsion is strong – the EO groups tend to coil rather than extend. The larger EO chains tend to require large alkyl chains to balance them, so the molecules become large and their dynamic surface tension behaviour becomes unsatisfactory as their diffusion coefficients are relatively low.

For food emulsions the good news is that the large "surfactants" (often large proteins) make very stable emulsions as they provide a robust shell around the drop. The bad news is that their large size makes them very slow to diffuse to the interface and also the elasticity is very low, so it requires a lot of work (i.e. making and remaking drops) to create relatively large emulsion drops and a very large amount of work (high pressure homogenizers with repeated cycles) to create smaller drops.

Although the preceding paragraphs are largely true, they are also largely unhelpful. They amount to saying "Too much of a good thing is a bad thing" which is merely a tautology. Can we do better? Yes. And the first thing to do is to avoid situations where it is impossible to create an emulsion, no matter how hard you try.

### 4.5.1  How *not* to make an emulsion

In the earlier discussion on EOR, mention was made of a critical capillary number. The capillary number is a fundamental (dimensionless) number used throughout fluid dynamics. For a flow of velocity U in a fluid of viscosity η and surface tension γ, the capillary number, Ca, is given by:

$$Ca = \frac{U\eta}{\gamma} \qquad \text{4-18}$$

Large values, by definition, imply that inertial forces (Uη) are more important than surface forces (γ). For removing oil from a well and for creating emulsions we clearly want large capillary numbers so we are not held back by surface tension. In each case there is a critical capillary number below which nothing much happens.

For standard dispersers there is a clear relationship between the critical capillary number and $\eta_D/\eta_C$ the ratio of viscosity of the oil, $\eta_D$ (or, more generally, the dispersed phase) to the viscosity of the water, $\eta_C$ (the continuous phase).
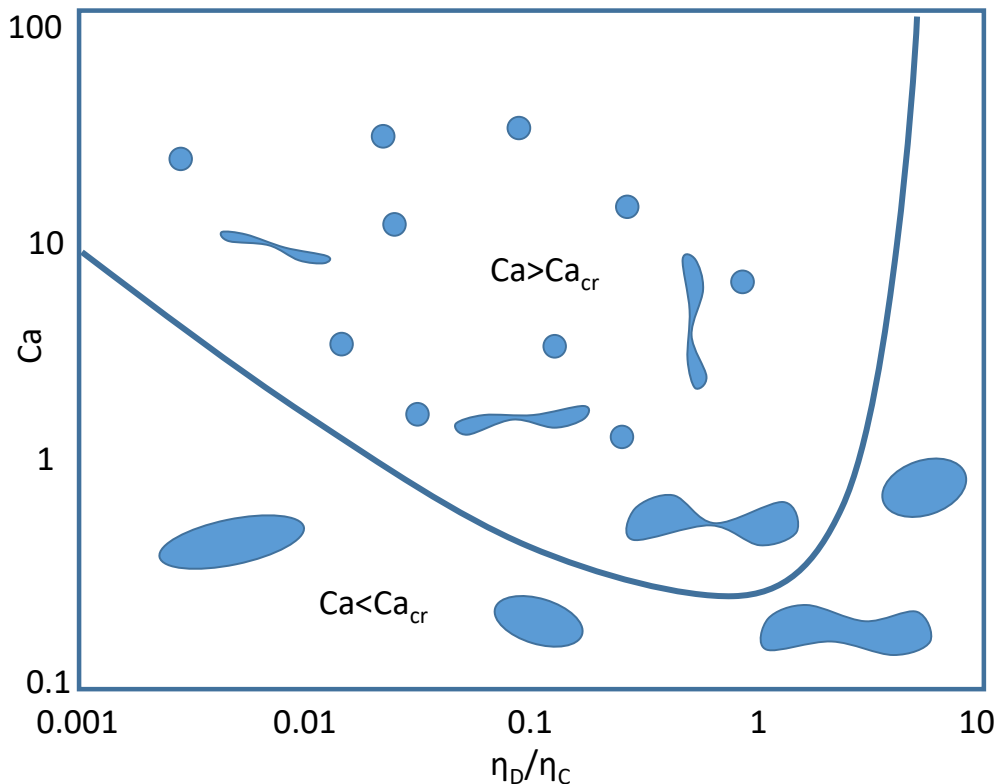
**Figure 4-4 The dependency of the critical capillary number on the viscosity ratio of the dispersed and continuous phases**

If the oil is much less viscous than the water then the process is relatively inefficient as it needs a large capillary number (3-10) to generate emulsion drops. When the oil and water are equally viscous it is particularly easy to create drops as the critical number falls to ~0.4. But there is danger of working in this zone. If the oil viscosity is 3.7x greater then it becomes impossible to create drops – the critical capillary number heads off to infinity. What is happening is that these viscous liquids absorb the energy by spinning and tumbling rather than splitting into smaller drops, and the more you hit them (higher capillary number) the more they just spin and tumble.

So we have some clear ideas on how to *not* make an emulsion.

We can take these negative ideas and extract three positive principle from this graph:

- If you have to make emulsions of viscous oils via dispersers then make the water is viscous too (via some convenient thickener) so that you are below the 3.7 ratio. This is the trick mentioned earlier for making peanut oil nanoemulsions.

- The limit applies only to shear processes. If you have to make emulsions with a ratio >3.7 then use something like a membrane emulsifier that works via extensional flow for which there are no such issues;

- A super-low interfacial tension (i.e. super-low γ) means a super-high capillary number and, therefore, easy dispersion even at lower disperser speeds. This is a key reason for using PIF techniques.

## 4.5.2 Predicting drop size

Making an emulsion involves putting energy density ε into a mix with a volume fraction φ of the dispersed phase, a viscosity of the continuous phase $\eta_C$ and of the dispersed phase $\eta_D$, with an interfacial energy γ and a density of the continuous phase ρ. From those inputs it is possible to estimate drop size. The energy is conventionally provided by a rotor of diameter D and rotation rate N. The standard approach to predicting drop size uses Hinze-Kolmogorov theory in various forms, and there are numerous parameterised variants which, regrettably, apply only to whatever limited range of systems were used for the parameterisation. What we want is a more universal predictive tool.

Fortunately, Tcholakova and colleagues at U Sofia[38] did a full-scale evaluation over a large range of viscosities, interfacial energies and volume fractions to create a set of equations with the maximum applicability with the minimum of parameters, each of which is readily determined experimentally with a few emulsions. The summary that now follows hardly does justice to the huge amount of work involved in arriving at the key formulae that are implemented in the app.

The basic Hinze-Kolmogorov approach involves formulae with ε, $\eta_C$ and ρ. One such is the derivation of the Kolmogorov scale, $\lambda_0$ which defines the smallest eddies that can be produced in the homogenizer. The equations which follow have a bewildering array of exponents which look rather odd; that is just the way it is:

$$\lambda_0 = \varepsilon^{-0.25} \eta_C^{0.75} \rho^{-0.75} \quad \text{4-19}$$

The central problem is that none of us knows the value of ε, the rate of energy dissipation per unit mass of the fluid, in J/kg.s. For a given type of homogenizer with rotation speed N and rotor diameter D, ε can be estimated via:

$$\varepsilon = b_1 N^3 D^2 \quad \text{4-20}$$

The paper provides values of 6 for $b_1$ for a lab-scale Ultra Turrax and 40 for a Magic LAB.

The dispersed phase can be in the *inertial* regime where the drops are larger than $\lambda_0$ and deform due to fluctuations in hydrodynamic pressure. Or the dispersed phase can be in the *viscous* regime where the drops are smaller than

38  Slavka Tcholakova et. al, Efficient Emulsification of Viscous Oils at High Drop Volume Fraction, Langmuir 2011, 27, 14783–14796

$\lambda_0$ and are ripped apart by viscous stresses. It is important to know which regime you happen to be in.

The drop diameter, d, in the *inertial* regime is governed by an awkward transcendental equation (the unknown is on both sides of the equation) for which an app is very useful. The equation is the Davies modification to Hinze-Kolmogorov that takes into account the viscosity of the dispersed phase. Note that in this section we are following convention and discussing diameters, not radii. The reason is that the measured values are one of the many variants of "average" diameter that particle sizers provide. Purists can read the paper to distinguish between $d_{32}$ and $d_{V95}$ and the rest of us can use a single d.

$$d = A_1 \left( 1 + A_2 \frac{\eta_D \varepsilon^{0.333} d^{0.333}}{\lambda} \right)^{0.6} \lambda^{0.6} \rho^{-0.6} \varepsilon^{-0.4} \qquad \text{4-21}$$

A good starting value for $A_1$ is 0.86 and for $A_2$ is 0.37. The volume fraction φ has no direct influence on the drop size. Instead, as we will shortly see, it is one of the factors in determining which regime the system is in.

There are various plausible ways to estimate d in the *viscous* regime. Unfortunately these did not give a good match to the experimental data. Fortunately a fitting function was found that does an excellent job:

$$d = A_3 SR^{-0.5} \left( \frac{\eta_D}{\eta_{EM}} \right)^{0.167} \qquad \text{4-22}$$

Here $A_3$ is a constant, SR is the shear rate (the more conventional γ symbol can get confused with interfacial energy) and $\eta_{EM}$ is the viscosity of the whole emulsion calculated using the Yaron, Gal-Or equation discussed in the viscosity section. The shear rate can be estimated from N, D and the rotor-stator gap h as:

$$SR = \pi \frac{ND}{h} \qquad \text{4-23}$$

So almost everything is in place. It just needs a calculation of $\varphi_{TR}$ the transition phase volume below which the inertial equation is used and above which the viscous equation is used. The equation looks somewhat similar to the inertial equation. That is because it is derived by combining the $\lambda_0$ equation and the inertial d equation. It uses IYG which is the Yaron, Gal-Or expression that calculates the viscosity at the given phase volume. This is another transcendental equation:

$$IYG\varphi_{TR} = 1 - A_1\left(1 + A_2 \frac{\eta_D \varepsilon^{0.25} \eta_C^{0.25} \rho^{-0.25}}{\lambda}\left(1 + IYG\varphi_{TR}\right)^{0.25}\right)^{0.8} \frac{\rho^{0.2}\lambda^{0.6}\varepsilon^{-0.2}}{\eta_C}$$

4-24

The app solves for $\varphi_{TR}$ and then for the given $\varphi$ it provides the appropriate calculation of d, with a note about whether it is using the Inertial, Transition or Viscous regime.



**App 4-9** https://www.stevenabbott.co.uk/practical-surfactants/EDSM.php

Although the numbers are important, it is equally important to have some understanding of the general trends. The paper provides that understanding:

- At low values of $\varphi$ the drop size depends on rotation speed, viscosities of both phases, interfacial tension and only slightly on $\varphi$. This is the "normal" Hinze-Kolmogorov-Davies domain.

- At high values of $\varphi$ the drop size has (over a surprisingly large range) no significant dependence on surfactant type, oil type or interfacial tension and depends only on the homogenizer shear rate and the relative viscosities of the oil and the emulsion (i.e. the viscosity taking into account $\varphi$).

- "Low" and "High" depends especially on the viscosity of the dispersed phase. When this is low, the system *never* makes it to the viscous domain so the small particle size remains constant across $\varphi$.

### 4.5.3 Using HLD-NAC

The Tcholakova app shows that even with lots of surfactant and very powerful homogenizers it is very hard to create small drops. One good way to do *much* better is to invoke PIF – the generalisation of PIT which makes it possible to do phase inversion formulations using not just the T of PIT but the S of salinity, the Cc of surfactants and EACNs of oils.

An alternative way to look at it is to take two facts on which everyone can agree and try to build an emulsion stability algorithm by refining the ideas behind them.

Fact 1. An emulsion is at its least stable when HLD=0. The interfacial energies are very low so drops can coalesce very rapidly. This is an unfortunate fact because it happens to be the ideal place to get easy break-up of drops for exactly the same reason.

Fact 2. It is hard to make a good emulsion when the HLD is very far from the balance point. This is a slightly more scientific way of capturing the relatively crude intuitions from HLB that you can't make a good W/O emulsion via very low HLB=1 or a good O/W emulsion via a very high HLB=17. And, of course, the HLB has to be in the right side of the balance point in order to create the desired type, O/W or W/O.

The simple way of explaining fact 2 is that (a) the right curvature is required to get O/W or W/O and (b) an extreme surfactant is probably so soluble in the bulk phase that very little of it will get to the interface.

Because HLD is a property of the system (T, S, Cc, EACN) it is possible to express an optimum distance for easily creating a stable emulsion. A good starting point for any formulator would be HLD=-1 for O/W and HLD=1 for W/O. This is sufficiently far from HLD=0 to avoid the problems of instability from low interfacial tension, and not so far that the surfactant is obviously happier in the bulk phase than at the interface.

An attempt to do better than such a rule of thumb is the CCS (Collision-Coalescence-Separation) model from the Acosta group. The HLD approximation says nothing about the other aspects of the surfactant and oil (in particular, A, L and $\xi$). The CCS model uses NAC to try to bring those factors into account. As with the Tcholakova model the precise predictions are less important than the insights gained by playing with the app. A full exploration of the ideas are contained in the thesis[39] by Sumit Kiran who carried out the bulk of the work when in Acosta's lab.

The starting point is one already discussed in the coalescence section – the classic Davies, Rideal model of small drops coalescing via Brownian motion. The key question is one of the activation energy for coalescence (E in the coalescence app) and the model uses two factors. The first involves the interfacial rigidity of the surfactant ($E_r$). The higher it is, the harder it is to remove the surfactant from the interface to allow coalescence. The second is the energy required to form a "neck" between drops and reach a critical neck length $t_{crit}$.

---

39   Sumit Kumar Kiran, Application of the HLD and NAC models to the formation and stability of emulsions, PhD Thesis, U. Toronto, 2013
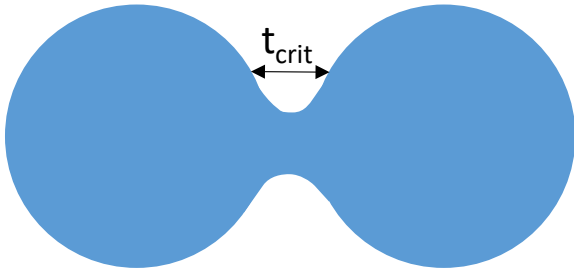
**Figure 4-5 The "necking" event in drops joining.**

$t_{crit}$ depends on the interfacial energy $\gamma_{OW}$ (lower energy means easier neck formation) and on a characteristic length below which nothing can happen. In its simplest form, $t_{crit}=2(r_{drop}+L)$ where $r_{drop}$ is the radius of a microemulsion drop (not the emulsion drop size) and L is the surfactant tail length. The activation energy for this process is given by $E_{neck}=0.73.t_{crit}^2.\gamma_{OW}$. Finally, DLVO terms derived from charge and steric stabilisation terms, can be added. These are calculable in principle via the DLVO app, though in this app they are covered via a DLVO distance which gets added to the $t_{crit}$ term. This leads to the overall activation energy that goes into the Davies, Rideal equation:

$$E = E_r + 2.9\left(r_{drop} + L + DLVO\right)^2 \gamma_{OW} \qquad 4\text{-}25$$

Readers might be puzzled about why $r_{drop}$ is the microemulsion drop diameter, not the emulsion drop. The answer is related to the Kabalnov ideas discussed earlier where holes can only be formed if the surfactants can curve into their natural shape – and their natural shape gives the microemulsion drop radius. As Kiran points out, the activation energy term could be calculated via Kabalnov's hole nucleation theory. However, the approach adopted here uses the elegant calculability of NAC theory to provide the required estimates.

The interfacial energy, of course, is worked out via NAC considerations. For the app the HLD is the variable along the X-axis so at each value of HLD the interfacial tension can be calculated based on surfactant tail length and head area, just as in the IFT app discussed in the NAC chapter. From this the activation energy for coalescence can be calculated and the time for the emulsion to coalesce follows from that.

The contradiction at the heart of the app is that from the Hinze-Kolmogorov approximation, low interfacial tension produces desirable small drops that are more stable to creaming. This is implicit within the app, but an explicit calculation is also provided:

**App 4-10** https://www.stevenabbott.co.uk/practical-surfactants/EStability.php
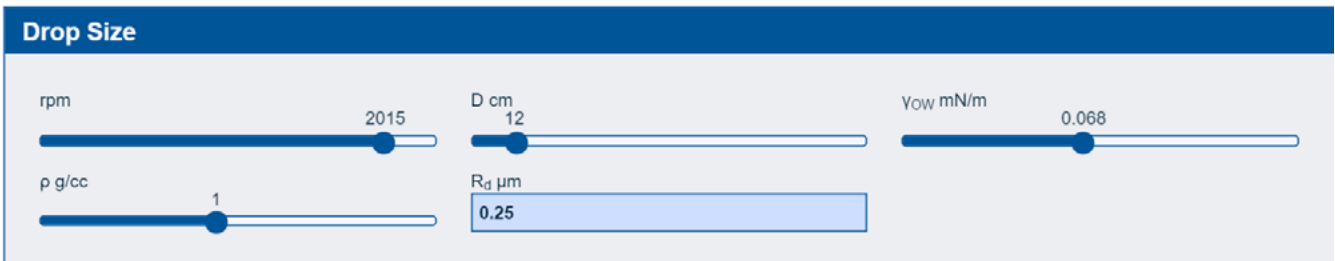
Yet at the same time, low interfacial tensions make it much easier for the drops to coalesce, and there are lots more of them to bump into each other to coalesce and cream. Everything about making an emulsion is a compromise. The app allows you to explore the tradeoffs.



**App 4-11** https://www.stevenabbott.co.uk/practical-surfactants/EStability.php
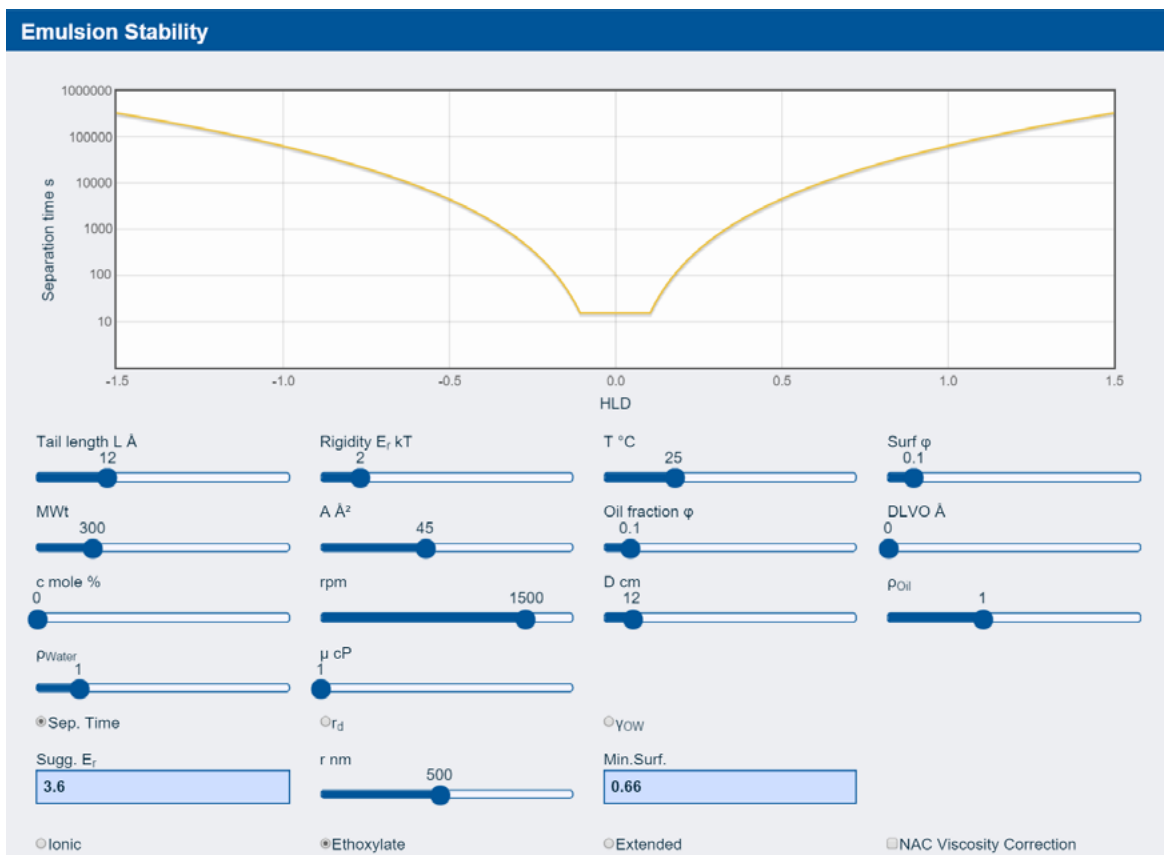
It will be immediately obvious from the screen shot that some extra calculations are also being included. The first thing that can be done is to set the density of the oil to a more realistic value. Now things look very different.

**Figure 4-6 The same setup but with oil density set to 0.85. Creaming sets in.**

The app takes into account other processes. So creaming is one possibility. At large HLD values the basic coalescence mechanism is very slow, making it look as though you need the worst possible surfactant. Yet a surfactant with a high interfacial tension means large drops and, as we know, the creaming velocity goes as r², so the emulsion will have creamed before it coalesces.

You can also increase the solubility of the oil in the water and start to see Ostwald ripening taking over as indeed it often does when trying to make nanoemulsions.

Some other calculations are provided such as the minimum required concentration of surfactant to cover the surface of the drops and an option to change the viscosity near HLD=0 using NAC calculations discussed in the NAC chapter.

This brief summary scarcely does justice to the power of this approach. The reader is urged to get to know the app and start to explore what it is saying within the context of their own formulations.

## 4.6 Emulsions by Catastrophic Inversion

So far the only type of inversion that has been discussed is the one around HLD=0 either via temperature (PIT) or EACN, Cc etc. (PIF). There is another type of inversion (often called 'catastrophic inversion' to distinguish it from the

HLD=0 type) that is hugely important for those who formulate emulsions. It can be important in a negative way because it can interfere with the manufacture of an expected O/W or W/O emulsion, often creating horrid messes (though sometimes these are desired) which might be o/W/O or w/O/W. You need to understand the inversion phenomenon to avoid getting these messes during scale-up from lab to production. Although inversion is often a problem, we will see that by harnessing it properly it can be used to create great emulsions that would otherwise have been too hard to produce.

We are all familiar with emulsion inversion from a simple example. Cream is an O/W emulsion. Left on its own (and kept free from microbes) it will stay as a stable emulsion for a long time. Yet impart some energy to it by beating or churning and it transforms to its thermodynamically stable W/O form: butter. The transition is easy to see because the relatively viscous oil drops inside the water contribute little to the viscosity, so cream is reasonably watery. Following the inversion, the low viscosity water is inside the high viscosity oil, giving us semi-solid butter. Such viscosity changes at phase inversion are relatively common. If you are aware of inversion phenomena they are a useful diagnostic. If you have never heard of catastrophic inversion the change in viscosity might be very puzzling.

There is another example familiar to most chefs. Take some oil, vinegar and egg and beat them with a spoon. The result is a horrible unstable W/O emulsion. Take the vinegar and egg and add the oil slowly so that it is always the minority phase. Now you will get a pleasing inverted O/W emulsion otherwise known as mayonnaise. If you are impatient, take the oil, vinegar and egg together and put into a high-speed blender – the result is the same mayonnaise.

And this is the problem with inversion. It involves kinetics as well as thermodynamics. Beat something slowly and you get one result, beat it hard and you get the other, add the ingredients one way and you get one type of emulsion, add them another way and you get the other type.

So, if you know what you are doing you can either avoid unwanted inversions or use inversions positively to your advantage.

"If you know what you are doing" is key. Reading the literature on inversion can be rather depressing because people are changing all sorts of parameters and getting all sorts of surprises that they can't explain. The approach of Salager[40] to inversion phenomena makes the whole system rather straightforward to understand. And, of course, the key is to use HLD to describe the phenomena. It is striking how easy it is with HLD and how incoherent it is without it.

---

40   The most beautiful exposition is in French: Jean-Louis Salager, Raquel Anton, Jean-Marie Aubry, Formulation des émulsions par la méthode du HLD, Techniques de l'Ingénieur, J2158.1-J2158.17, 2006
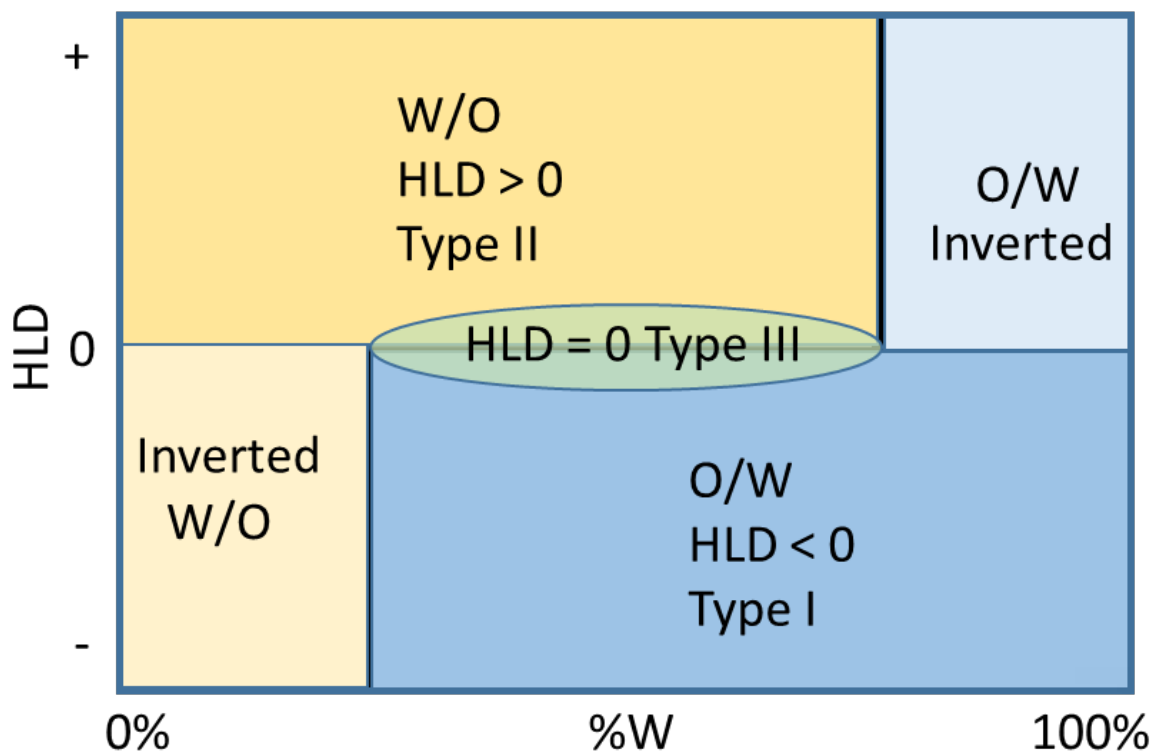
Here is the key Salager diagram:



**Figure 4-7 The essence of emulsion inversion is captured in this diagram, developed by Salager. At HLD=0 inversion is 'normal'. The other two inversions are 'catastrophic'.**

There is plenty going on in the diagram so we need to start with some familiar themes. The x-axis is the % water. The y-axis is HLD, with 0 in the middle. We already know a lot of this diagram. Below the HLD=0 line we have O/W emulsions and above it we have W/O. And where HLD=0, in the Type III or inversion domain, we have the ability to create emulsions using little energy because the interfacial tension is low, i.e. PIT or PIF emulsions.

This choice of y-axis is of crucial importance. Any change to any component or temperature in an experimental program moves you through surfactant space. If you try to compare two formulations without scaling the formulations to HLD you are lost. That is why it is depressing to read much of the inversion literature. Formulators have no idea where they are so are constantly surprised by what they find. This means that much of the literature provides little guidance to those whose systems differ modestly from those in the papers. As mentioned previously, if all papers gave enough information to deduce where they are in HLD space then we would be able to mine the literature for useful insights. Instead we have lots of disjointed papers from which little can be learned.

So, we have the HLD=0 line. What is new are the vertical lines at ~25% and ~75% water. They give a simple story: that even when HLD<0, if there is less than 25% water the resulting emulsion is W/O; and that even when HLD>0, if there is more than 75% water the resulting emulsion is O/W. So below 25% and

above 75% you get the inverse of what you would expect from HLD theory in the cases, respectively, of HLD<0 and HLD>0.

Why should the emulsions invert? There is no generally-accepted full answer and the most usual partial explanation goes along the following lines, using the <25% case and the resulting W/O emulsion, though the argument is the same for >75% and the resulting O/W emulsion.

Imagine the emulsion drops as spheres of oil in the water, so we have HLD<0. As the %oil increases (going from right to left in the diagram) there is no problem in packing spheres at 30, 40, 50, 60, 70% oil. But around 75% is the (ordered) sphere close-packing limit and it is no longer possible to have a water continuous phase around the over-packed oil drops. It is then easier to flip so that the minority phase, the water, forms 25% spheres in a continuous phase of oil.

The astute reader will notice a number of problems with that argument, but the argument contains a strong element of truth, is good-enough for our purposes and I have failed to find a better one.

That diagram contains the core of inversion science. But because inversion depends on *kinetics* as well as *thermodynamics*, reality is somewhat more complicated.
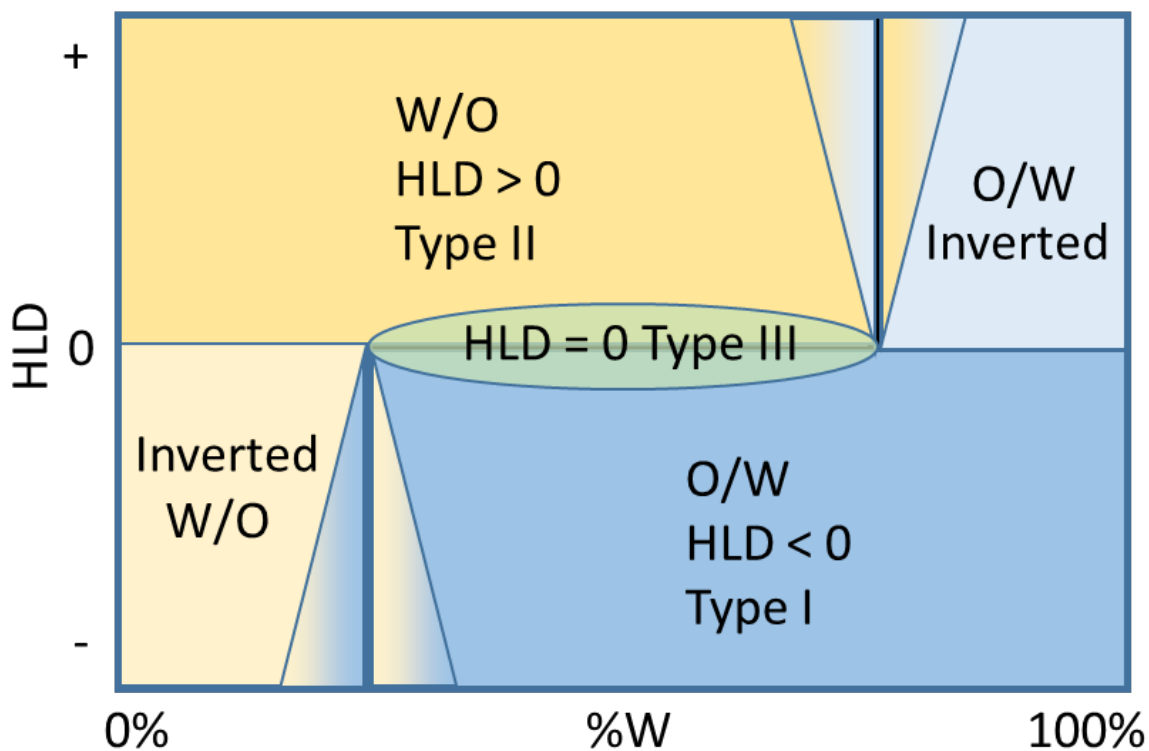


**Figure 4-8 Almost the same diagram, with some extra triangles**

The difference now is the shaded triangles. What they are conveying is the fact that if, for example, at large negative HLD you add more oil to an O/W emulsion, the inversion might not happen at 25% water but, say, at 15%. Going the other way, if you start with a W/O emulsion containing 15% water and add more water, it might not invert till 35% water. A similar effect happens at high positive values of HLD where the transition from W/O or O/W can be delayed, depending on the starting point.

There are more complications. Reverting to a diagram without the triangles we see that viscosity effects can be significant.
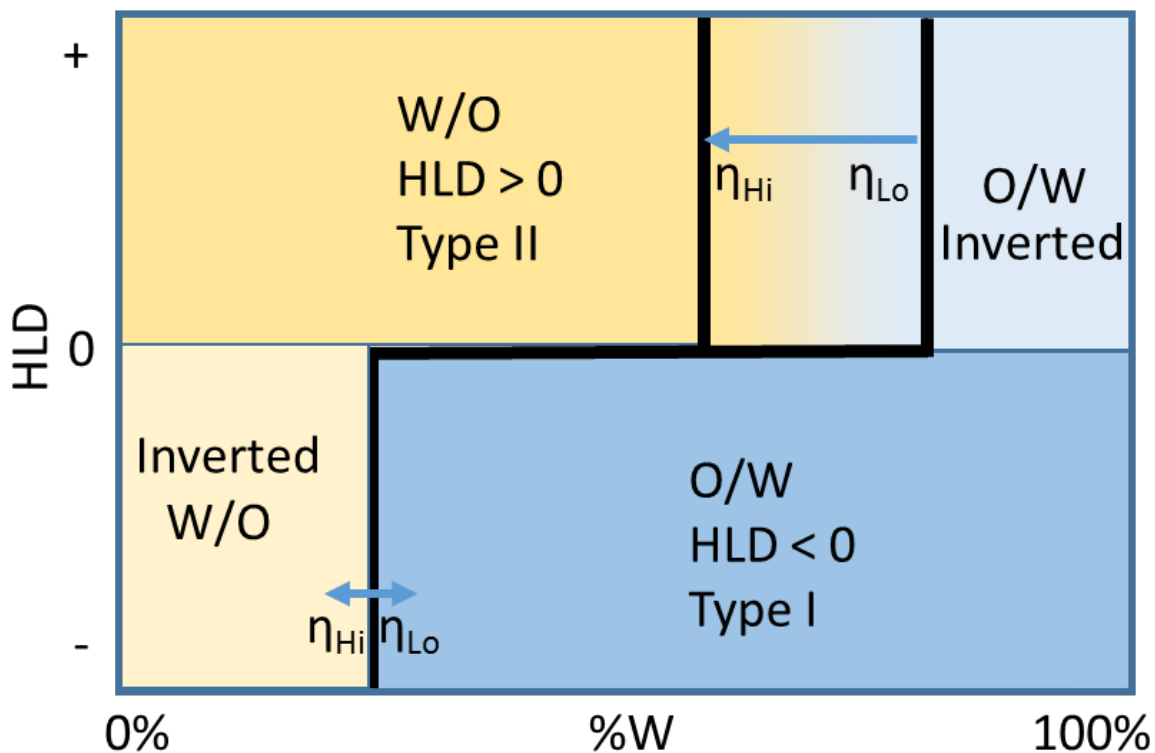


**Figure 4-9 Oil viscosity makes little difference to the transition when HLD<0 and a significant difference to the HLD>0 transition.**

What the diagram shows is that when HLD>0, the transition to an O/W emulsion can take place even at 55% water if the oil viscosity is high, whereas the W/O transition at HLD<0 is unaffected by viscosity. Presumably this is related to the "viscosity ratio" effects described earlier in the section on critical capillary numbers.

The three diagrams do a great job at describing the general rules. Of course real life is not as simple as this. But it is not as complicated as many people think. Because formulators change temperatures, surfactants and oils when looking at emulsion inversion they are simultaneously moving along the water axis and up and down the HLD axis. But if they don't know about HLD they have no idea where they are within the Salager diagrams so are constantly surprised by what they find.

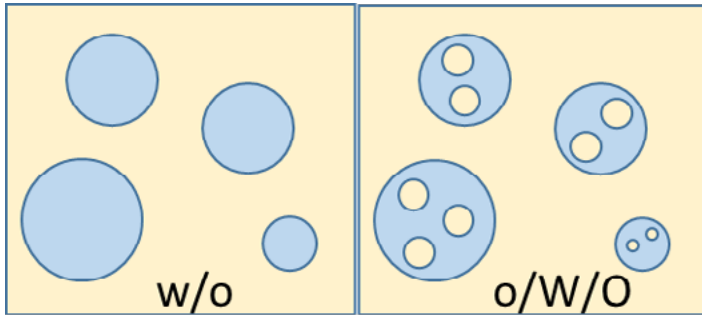There is one more piece of bad news before we get to the good news.



**Figure 4-10 The W/O phase might be a more complex O/W/O phase.**

The inversion to W/O might give a simple W/O emulsion. Frequently it gives o/W/O where the small o is the small drops of oil inside a W/O emulsion. The reason is that the system is trying as hard as it can to remain O/W in the face of the problem of being unable to pack all the oil between water drops. So some of the oil gets hidden inside the water drops before the system gives up and creates the W/O emulsion, with the o inside the W.

Sometimes these more complex emulsions are desired. Most of the time they are a nuisance because generally we want either an O/W or W/O emulsion, not the complex and unstable messes that are often found with o/W/O or w/O/W.

There seem to be no reliable rules for avoiding these complex phases. You have to be on the lookout for them (they are easy to spot under a microscope) and hope you can find some variant of mixing speed or addition rate to avoid them.

With one final diagram we can explore the good news about emulsion inversion.
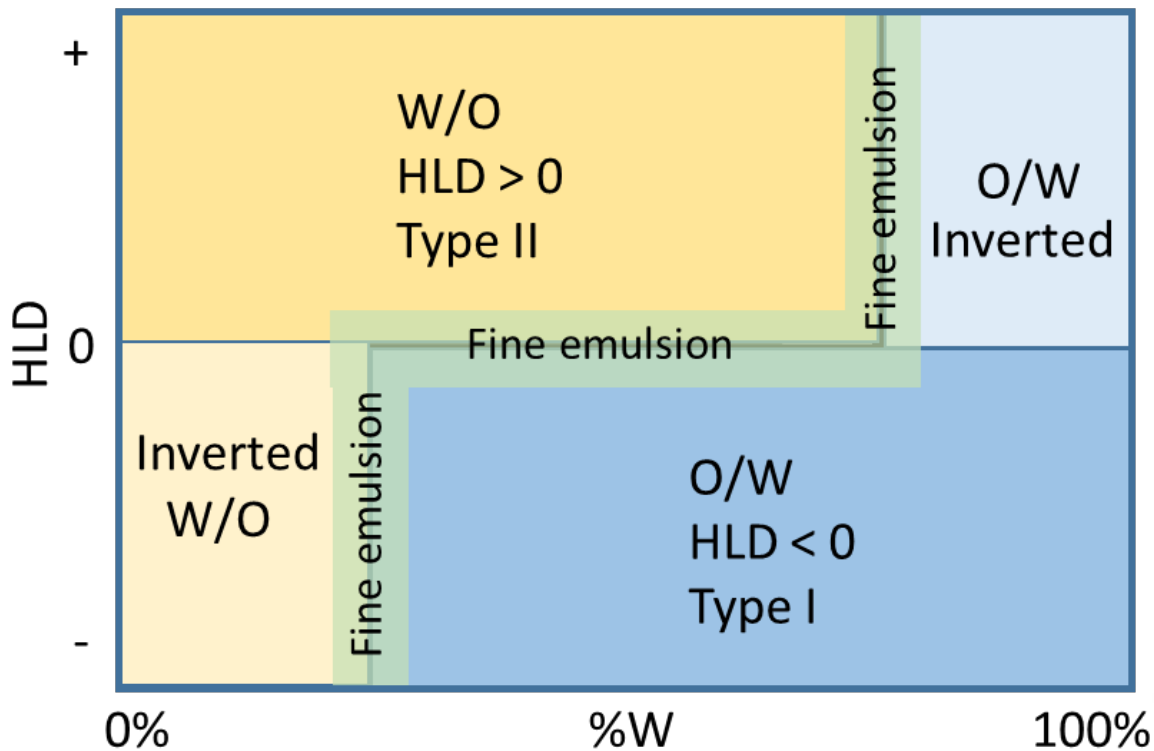
**Figure 4-11 Instead of having just one type of low energy emulsification system at HLD=0 we have two at the other inversion points.**

Whenever there is inversion it means that the effective interfacial tension is minimal which means, in turn, that it is easy to make an emulsion with small droplets without expending large amounts of power. The HLD=0 inversion point has been much-discussed. The two inversion points at 25% and 75% water (with all the shifts from these points described above) offer alternatives for those needing emulsions under circumstances where going via HLD=0 is not an option.

This inversion option is especially important for nanoemulsions. Although the majority of these tend to be restricted to ethoxylates so the PIT version of HLD=0 can be used, astute formulators can use phase inversion to obtain excellent nanoemulsions far from HLD=0. Not surprisingly the trick is used infrequently because most formulators do not have a clear idea of where they are in inversion space. Armed with the HLD approach it will be much easier for formulators to manipulate oils, surfactants and temperatures to make good use of the catastrophic inversion formulation region.

In addition to the fact that this methodology opens up formulations that do not use ethoxylates, it also makes it straightforward to solve the other issue of these inversion points – the fact that the emulsion rapidly fails unless the formulation is moved to a more stable region. With PIT the emulsion must be rapidly cooled. To make a fine O/W emulsion via the inversion technique, a crude W/O emulsion can be created at, say, 20% water, then water slowly added to get to the inversion point, with a final addition of water to take the formulation safely into

the classic O/W zone, where the formulator only faces the challenge of finding one of the techniques discussed earlier to resist Ostwald ripening.

Some readers may already be at ease with inversion emulsions and my hope is that the Salager graph helps them to formulate more wisely. But many of us are not at ease with inversion. There are two reasons to become more at ease with the theory and the practice.

- When formulations are transferred from lab to production it is often not practical to follow the exact lab sequences of water and oil. In the lab it might be easy to mix a small volume at the bottom of a flask yet in production the disperser head might not be able to reach down to that part of the large vessel. Although the final dispersion step in production will be on the same overall formulation, it is possible that there will be intermediate steps with different O:W ratios and if these are near the catastrophic inversion point then "bad" emulsions might form which fail to re-form to the desired type, or do so only with a lot more hard work. By being aware of the possibilities of inversion, the formulator can work with the production team to avoid the problems

- It is not very hard to simply play in the lab with inversion formulations. By doing so not only do you gain the confidence to avoid potential production problems, you also open up fresh ways to make interesting products via less conventional ways. No one (not even Salager) claims that inversions are easy. The triangles and the viscosity shifts in the diagrams show that there are lots of subtleties that are barely understood. Even the well-known inversion from cream to butter is subject to subtle changes in the milk, as anyone from the dairy industry will attest. But it *can* be made reliable for your specific area once you know where you are in HLD and inversion space.

To play creatively it takes little more than a microscope and a feel for the viscosity to know if you have inverted or not. An oil-specific (or water-specific) dye makes it very easy to identify phases under the microscope and if there is a significant difference in viscosity of the water phase and the oil phase then, just as with cream and butter, it is easy to identify if an inversion has taken place. As mentioned earlier, those with a conductivity probe can also quickly check if they are O/W (high conductivity) or W/O (low conductivity).

## 4.7    Pickering emulsions

Classical emulsions are made using small-molecule surfactants that are relatively mobile. This is good when you want them to come quickly to the interface in order to create an emulsion. This is bad when you want the emulsion to be stable over the long term.

A lot of food science emulsions are made with relatively "bad" surfactants that are slow to reach any interface and often pack so poorly that they aren't too good at producing low interfacial energies. On the other hand, once they've formed an emulsion they tend to be stable over longer timescales because the large proteins or phospholipids are too slow to move out of the way when there is a chance to coalesce – and they provide big steric barriers, and/or they are too insoluble in the water phase to allow any Ostwald ripening.

The logical continuation of this trend is to have solid particles that can assemble to create a rather stable shell around the internal drop. Purists dislike the term (W Ramsden found them four years before S Pickering), but they are commonly called Pickering emulsions.
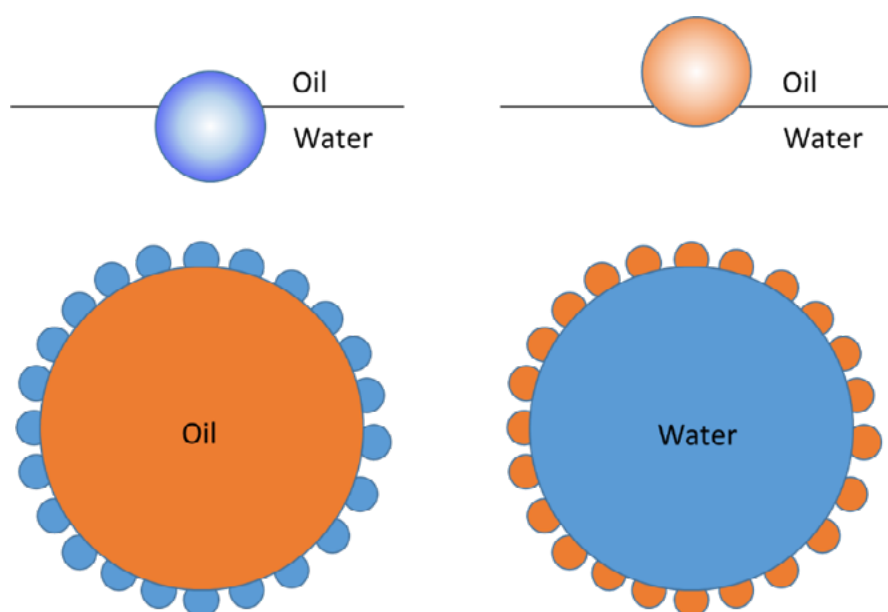


**Figure 4-12 If the particle favours water then the emulsion is O/W, if the particle favours oil then the emulsion is W/O. These particles probably partition too much into their preferred phase so would be inefficient emulsifiers.**

Their design rules are straightforward in principle. If the particles' surface is easily wetted by oil or water they are of no use as they will spend most of their time in the phase that wets them. Intermediate hydrophilicity tends to give an O/W emulsion and intermediate oleophilicity tends to give W/O emulsions. But in general particles with a 90° water contact angle seem to be the best all-round emulsifiers. The ideal is a Janus particle with one half that is hydrophilic and the other half that is hydrophobic, though these are not so easy to produce.

The reason why 90° is optimal is that particles that are too easily wetted by water or oil will tend to be quite happy away from the interface, reducing its stability. The Pickering app attempts to illustrate the phenomenon. In the view shown here, the wetting angle with water is 41° and instead of a tight shell around the oil drop, many of the particles are happily in the water – with a colour coding to emphasise that point. When the wetting angle is 90° the particles are coloured white and sit in an orderly array around the drop.

**App 4-12** https://www.stevenabbott.co.uk/practical-surfactants/Pickering.php
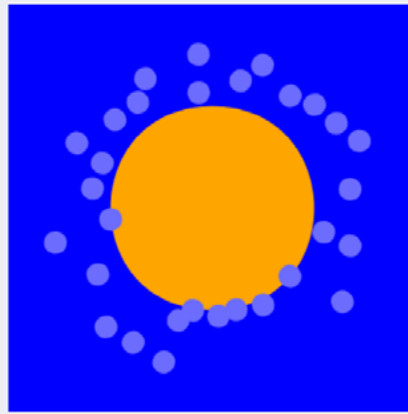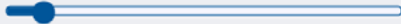
The other design rule is that the particles should be considerably smaller than the drops they protect, so it needs nanosilicas to create µm-sized drops.

The downside is that the big, lumbering particles (in comparison to a surfactant molecule) are slow to provide protection during the emulsification process. The upside is the same – once in place particles provide very stable emulsions.

For those who have to sell emulsions one attraction is that they are "surfactant free" which goes down well with consumers who have heard that surfactants are capable of doing bad things. On the other hand, the fact that they need nanoparticles is a problem for those who think that all things nano must be evil. When the particles are hydrophobised silicas the tendency might be to think of them as bad "chemical" nanoparticles. When the particles are waxes the tendency might be to think of them as being good, natural and "chemical free". One Pickering emulsion most of us are happy with is homogenised milk; the fat globules are stabilised by casein protein particles. The globules are sufficiently small (you get characteristic nanoparticle light scattering from the milk) that they do not cream because they are pushed around by thermal motion. The Pickering emulsion shell is tough and stops the drops from flocculating/coalescing so they do not form larger globules that would be able to cream.

# 5      Foams

We have already done a lot of the hard work and have most of the key ideas behind the science of foams. Indeed, to many surfactant scientists there is no significant difference between emulsions and foams. The science is the same, it is only the parameters that change. This chapter will not go to that extreme and so the language will tend to be foam-specific. But it is still profoundly true that the basic ideas are not so different.

We need to remind ourselves of the basic structure of a foam. The diagram (based, with kind permission, on an image from Prof Stubenrauch) shows the main features:
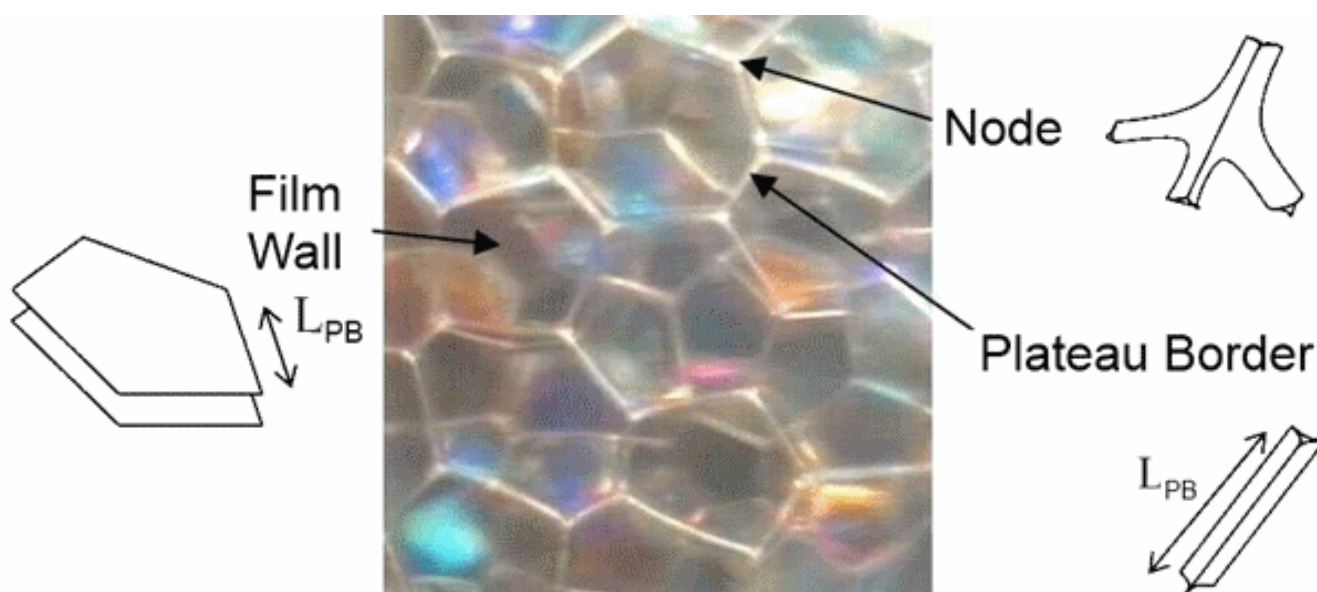


**Figure 5-1 The pretty foam bubbles are made from Film Walls, linked by Plateau Borders which meet at Nodes.**

As we all know, when a group of spherical bubbles meet so that most of the water has disappeared (a "dry" foam), the most energy-efficient form of packing involves polyhedral faces. One possible near-ideal arrangement is the tetrakaidecahedron Kelvin foam, a truncated octahedron with 6 quadrilateral faces and 8 hexagonal faces. Real foams tend, on average, to have 13 faces, one quadrahedral, 10 pentagonal and 2 hexagonal. Whatever the actual shape of the bubbles, the faces meet 3-at-a-time (therefore at 120°) along Plateau Borders and the borders meet 4-at-a-time at nodes (therefore at the tetrahedral angle of 109.4°).

Our concerns in this book are how/why those thin walls are stable (and how to deliberately increase or decrease that stability), how the foam behaves mechanically and how the liquid in the foam can eventually drain out leading eventually to the foam's collapse. Although I love bubbles, and there is lots that can be said about them, the focus here is on the surfactant aspects of them and the other aspects are outside the scope of the book.

For those interested more in the *applications* of foams, the excellent Foam Engineering[41] book edited by Paul Stevenson is highly recommended.

## 5.1    Foam basics

Because we all know that a foam is easily produced when a gas is bubbled through a liquid containing surfactants, the first question is why we need a surfactant. As we saw in the idealised model of creating an emulsion, the energy required to create the necessary surface area is trivial even without a surfactant. Halving that energy by going from 72 mN/m to 36 mN/m is not of great significance. The surfactant is required because the air/water/air interface is intrinsically unstable – there is no barrier to it collapsing because on collapsing it just creates more air and water interface. So air bubbling through water produces a kugelschaum (German for sphere-foam) that vanishes rapidly. Note that we are concerned with the air/water/air interfaces of foams and not the air/water interfaces of bubbles. You do not need a surfactant to create bubbles, you do need a surfactant to create foams.

So why are surfactants necessary for a stable foam? The answer to that question sets the scene for the entire chapter. There is not a single reason why surfactants are necessary. Everything needs to be right to obtain the correct foam that lasts the required amount of time. Here are the parameters.

- Elasticity. The first reason surfactants help create foams is that the surface becomes elastic. This means that the bubbles can withstand being bumped, squeezed and deformed. A pure water surface has no such elasticity and the bubbles break quickly. It also means that those systems which produce more elasticity (see the Elasticity section) will, other things being equal, produce more stable foams. In addition, as discussed in the Rheology section, a wall which is both stiff and elastic generally provides a foam with a greater ability to resist a pushing force and therefore a higher yield stress. Smaller bubbles also give a higher yield stress

- Disjoining pressure. The second reason that surfactants help create foam is that they give immunity to an intrinsic problem. Liquid in the film walls is naturally sucked out of the walls into the edges. This is nothing to do with drainage (as explained in Drainage, the walls contain an irrelevant fraction of the liquid), it is just simple capillarity caused by the curvature of the bubbles. The capillary pressure will keep driving liquid out unless a counter pressure ("disjoining pressure") acts against it. This can be produced by charges on the surfactant either side of the wall and/or by steric interactions between surfactant chains. These effects are discussed in DLVO, but because the charge effect operates over large distances (50nm) compared to the small distances (5nm) of steric effects, in general ionic surfactants are better at creating stable foams.

41   Paul Stevenson (Ed), Foam Engineering: Fundamentals and Applications, Wiley, 2012

- Resistance to ripening. The Ostwald ripening effect means that small bubbles shrink and large ones grow. As the Ostwald section shows, this is partly controlled by the gas (bubbles from $CO_2$ fall apart quickly, air/$N_2$ is slower and $C_2F_6$ much slower) but also by how good a barrier to gas diffusion the "wall" of surfactant at the surface provides.

- Resistance to drainage. The more water around the foam the less risk (in general) of it becoming damaged. So a foam that drains quickly is more likely to become damaged. As we will see, to resist drainage you need high viscosity and small bubbles, though the surfactant wall has some effect on the drainage process with stiffer walls giving (usually) slower drainage.

- Resistance to defects. If an oil drop or a hydrophobic particle can penetrate the foam wall it can cause the wall (and therefore the foam) to break. Although there are plausible and simple theories (discussed in AntiFoams) of Entry, Bridging and Spreading coefficients they turn out to be of limited predictive value. They are necessary but not sufficient. The key issue is the Entry Barrier. When this is high the foam is resistant to defects. Again it is the surfactant that provides the Entry Barrier and again ideas of elasticity are important in terms of formulating for a foam that is either desirably resistant (when you want a foam immune to hydrophobic defects) or desirably susceptible (when you want low foam).

Each of these effects is discussed in turn.

## 5.2    Elasticity

Although elasticity appeared many times in earlier chapters, there was no serious attempt to use an app to describe it. That is because elasticity is much more elusive and confusing in emulsions. As we will see, the equations require a thickness "h" and, for easy study, a suitably long timescale to allow meaningful measurements. Neither is readily obtainable for much of what is of interest in emulsions, while both exist for the types of foam and their timescales that are the basics of foam science.

There are a number of ways to talk about foam elasticity. I find the Wang & Yoon[42] approach particularly clear and helpful, though the prediction of a maximum (rather than a plateau) is, apparently, still controversial.

The elasticity of a material is stress/strain, i.e. the force required (stress) to change the dimensions by a certain amount (strain). For foams the elasticity, E[43], depends on how the surface tension γ changes with film surface area A. This can also be expanded to give terms with respect to surfactant concentration, c:

42   L Wang, RH Yoon, Effects of Surface Forces and Film Elasticity on Froth Stability, Int. J. Min. Proc., 85, 2008, 101–110

43   Although elasticity is often shown as ε, in this foam chapter ε is used for liquid fraction (and in DLVO for dielectric constant) so elasticity is shown as E to reduce confusion.

$$E = 2A\frac{d\gamma}{dA} = 2A\frac{d\gamma}{dc}\frac{dc}{dA}$$

5-1

For those who like to think in terms of "effects" the restoring force is the Gibbs or Marangoni force (experts will argue which is which) - the bigger the Marangoni force the more elastic a bubble is. In the expanded version, the restoring force is a combination of the change of surface tension with concentration δγ/δc and the change of concentration with area, δc/δA.

From the basic insight of what E is, it is possible (via a chain of reasoning described in the paper) to define E in terms of measurable quantities: surfactant concentration c; the bubble film thickness h; the maximum surface excess (adsorption density) $\Gamma_m$; the equilibrium adsorption constant K and RT, the universal gas constant times temperature:

$$E = \frac{4RT\Gamma_m^2 K^2 c}{h(1+Kc)^2 + 2\Gamma_m K}$$

5-2

As we have seen, the key values $\Gamma_m$ and K can be obtained from fitting the standard plot of γ versus concentration to the Langmuir-Szyszkowski equation as discussed in the first chapter:

$$\gamma = \gamma_0 - 2.3RT\Gamma_m Log(1+Kc)$$

5-3

Via differentiation of the equation for E, the maximum elasticity is obtained at a concentration c* given by:

$$c^* = \frac{1}{K}\sqrt{1 + \frac{2\Gamma_m.K}{h}}$$

5-4

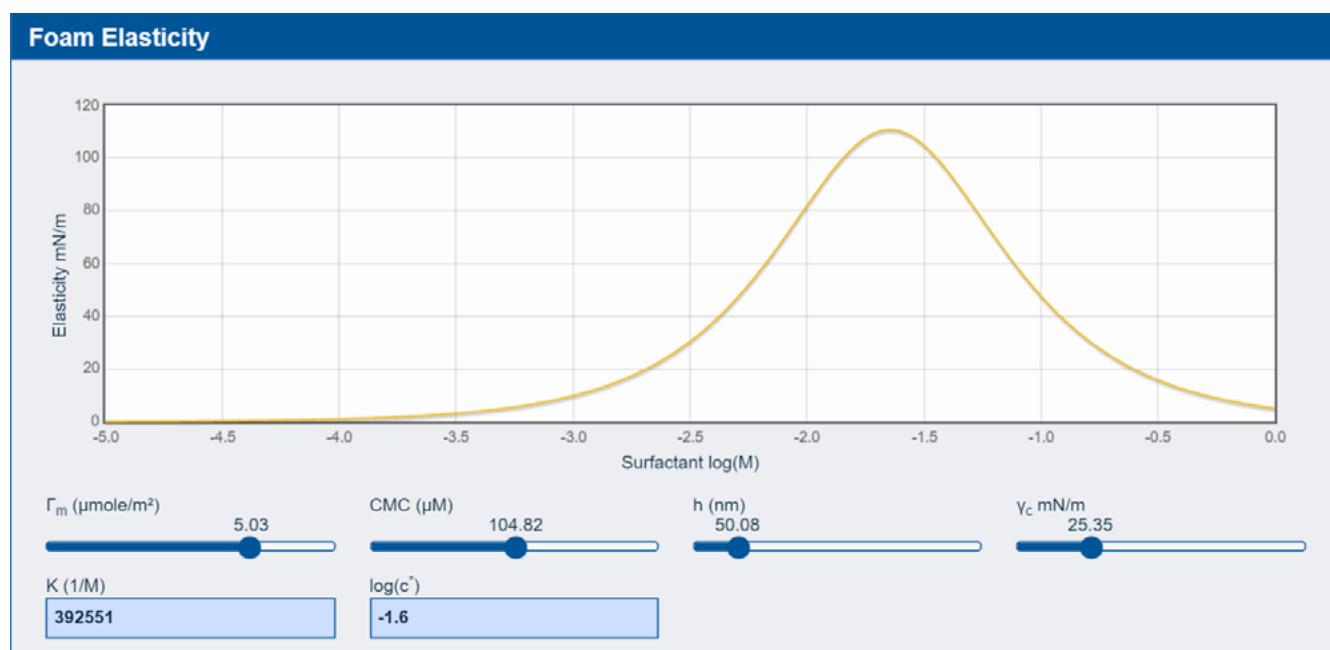If you prefer to think in terms of CMC then K can be calculated from the CMC and Δγ=γ₀-γc

$$K = \frac{e^{\left(\frac{\Delta\gamma}{RT}-1\right)}}{CMC}$$

5-5

There is nothing in this which is specific to foams. So why have we not had these equations before? The reason is "h", the thickness in the elasticity equation. For foams we know what this is. For most other situations it is very

hard to define. What, for example, is the thickness to be used for the layer between two drops? Give it a large value, say 1μm, representing a "typical" distance and the elasticity is small. Give it a value of a few nm, representing the moment of near contact, and the elasticity is very high. Add to this the dynamic behaviour when emulsion drops are colliding at potentially high speed and even the value of γ is unknown. For a foam studied over timescales of seconds or hours, the Langmuir-Szyszkowski isotherm gives us the necessary information.

Now we can look at the equations in the app. The biggest difficulty is defining the units. Here $\Gamma_m$ is defined as μmole/m², CMC as μM and h is in nm. There is little agreement on the perfect units for K, which varies over a wide range, so units of 1/M are used as the least-bad choice.



**Foam Elasticity**

$\Gamma_m$ (μmole/m²) 5.03

CMC (μM) 104.82

h (nm) 50.08

$\gamma_c$ mN/m 25.35

K (1/M) 392551

log(c*) -1.6

**App 5-1**  https://www.stevenabbott.co.uk/practical-surfactants/Foam-Elasticity.php

Given that h typically varies only from, say, 30 to 50nm (or, as we will see, from 5 to 10nm) and that $\Gamma_m$ varies only from, say, 1 to 10 and that RT varies little over a typical temperature range from 25-60°C, the big effect on elasticity is K. This can vary from 10 to 10,000,000. So to get good elasticity (only one of the key parameters for a good foam) concentrate on ensuring a high value of K via both a low CMC and a low $\gamma_c$.

How do we know that h varies only over that small range? At the instant a bubble is created, h might be 100s of nm. But the film wall rapidly thins (via capillary flow) to its equilibrium value which is controlled mostly by the DLVO effects discussed next. If the pressure inside the bubble changes this must surely change h. Yes it does and there are people who have patiently measured the change. In general the changes are in the 10% range, far too small to be significant in terms of all the other assumptions being made.

Lowering the CMC increases the elasticity via its effect on K. So it might seem that the best foam will be created with the lowest possible CMC. In practice, such low CMC surfactants are often useless because they do not readily get involved with the key processes for making the foam in the first place.
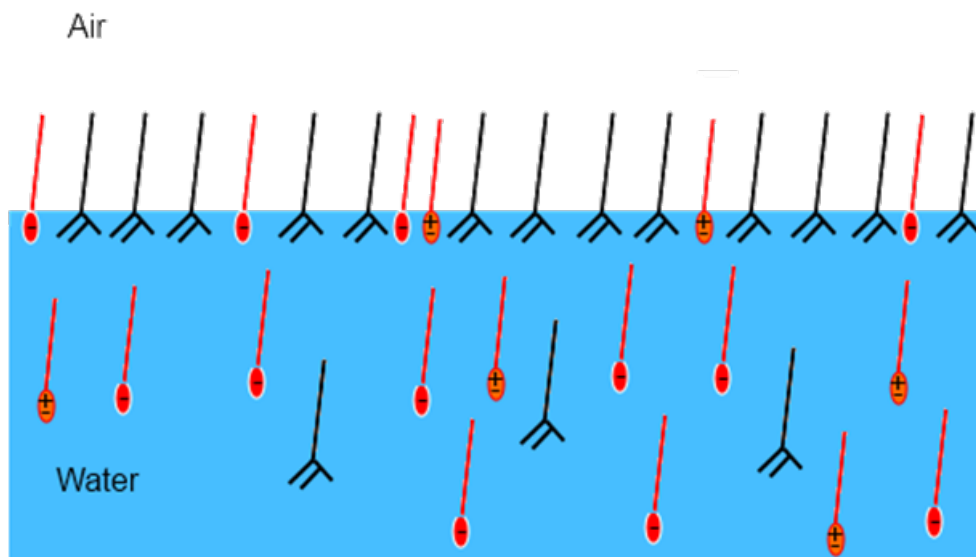


**Figure 5-2 The low level of (black) myristic acid in the bulk kicks out the higher levels of SLES (-) and CAPB (±) from the foam film, creating a tougher, more elastic foam.**

A well-known compromise, therefore, is to add a small percent of a low CMC, rather useless, surfactant such as myristic acid (black, neutral) to a typical foaming surfactant blend such as SLES (red, negative ionic) and CAPB (orange, zwitterionic). As the diagram shows[44], after some time, the slow, rigid, useless myristic acid kicks out a lot of the excellent SLES and CAPB from the foam film they generated, creating a much tougher foam. In surfactant science this is generally known as the transformation from a "Dawn" foam (named after the well-known US brand of dish washing detergent) to a "Gillette" foam (named after a shaving cream). Just as the flexible Dawn foam is unsuitable for shaving, so the rigid Gillette foam is unsuitable for washing dishes.

## 5.3    Foam DLVO

We have seen that foam films gain elasticity across a well-controlled thin film. Here we examine why those films are thin and surprisingly stable.

When a foam film is first created, its wall might be a few 100nm thick. The water in the film wants to do only one thing, to flow away driven by the modest capillary pressure across the bubble – e.g. for a 1mm radius bubble and 30mN/m surface tension, $2\gamma/r=60Pa$. The water would happily drain to 0 thickness and the bubble would pop. So there must be another force stopping

44   Based, with kind permission, on the image in N.D. Denkov, et al, The role of surfactant type and bubble surface mobility in foam rheology, Soft Matter, 2009, 5, 3389–3408

this from happening. In foam language this is called the disjoining pressure. The term is confusing because mathematically it can be positive ("dis"-joining) or negative ("con"-joining). But that is the terminology we have to use.

The curve of disjoining pressure, $\Pi$ versus thickness of the foam film, h, can be calculated readily from three independent terms, $\Pi_{vw}$ which is the van der Waals force (always attractive), $\Pi_{el}$ which is the repulsive electrostatic force and $\Pi_{st}$ which is the repulsive steric force:

$$\Pi = \Pi_{vw} + \Pi_{el} + \Pi_{st}$$

5-6

The terms themselves require a significant number of parameters which are not familiar to most of us but are shown here for the record and then taken care of within the app.

The van der Waals force depends on the Hamaker constant $A_H$ which is the generalised attractive force between any two surfaces. Most of us have no idea what $A_H$ should be, so 4E-20 J is a good-enough value to enter:

$$\Pi_{vw} = -\frac{A_H}{6\pi h^3}$$

5-7

The electrostatic force is rather more complicated. First we have to calculate an ionic strength, I, for a molar concentration of salt, c, with ionic charges of $Z_1$ and $Z_2$ given by:

$$I = 0.5c\left(Z_1^2 + Z_2^2\right)$$   5-8

Then we need a Debye $k^{-1}$ value (as $k^{-1}$ is confusing in a formula, we use D for Debye) which depends on I, on $e_0$, the permittivity of free space, $\varepsilon$, the dielectric constant, e the charge on the electron, $k_B$ the Boltzmann constant, T the temperature in °K:

$$D = \sqrt{\frac{e_0 \varepsilon k_B T}{2 N_A e^2 I}}$$   5-9

We need one more intermediate value, s from the surface potential, $\Psi$, where tanh is the hyperbolic tangent:

$$s = tanh\left(\frac{e\Psi}{4 k_B T}\right)$$   5-10

Finally we can calculate $\Pi_{el}$

$$\Pi_{el} = 64cRTs^2 e^{-\frac{h}{D}} \quad 5\text{-}11$$

The steric term is a little more uncertain. It assumes a steric length, b, below which the disjoining pressure is "infinite" (in the app it is set to $10^6$) and above which the following equation is an adequate approximation and q is a surface charge somewhere between 0 (no steric interaction) and 1 (maximum reasonable steric interaction):

$$\Pi_{st} = \frac{0.035qb}{h(h-b)} \quad 5\text{-}12$$

Before discussing the terms and their implications in more detail, it is worth playing with the app to get a feel for what is going on.



App 5-2 https://www.stevenabbott.co.uk/practical-surfactants/Foam-DLVO.php
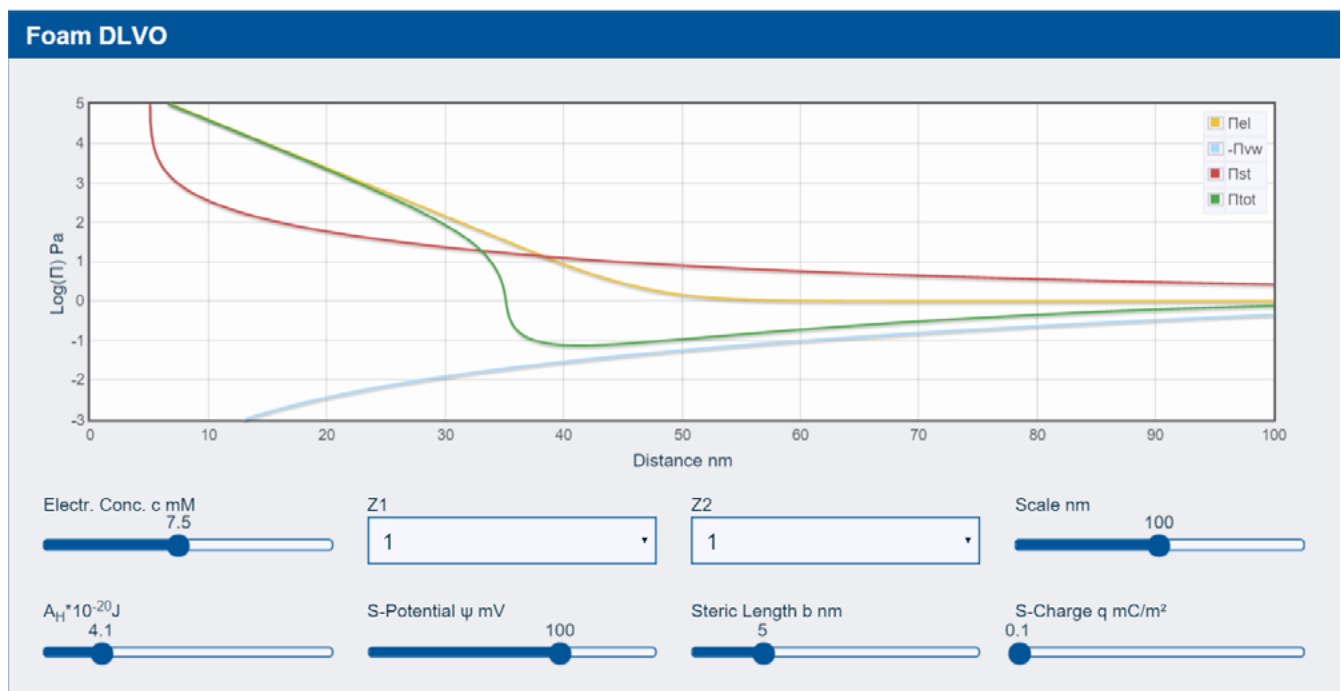
The first thing to notice is the 4 lines in the plot – the three individual terms and their sum. The disjoining pressures cover a large dynamic range so a log plot has to be used. This means that the negative values (which are impossible to encode in log format) have to be shown as the negative of a positive number, so -2 in the log plot means -100 not 0.01.

The aim, of course, is to have a large $\Pi_{Tot}$ at as large a distance, h, as possible to give the maximum chance of stability. A few hundred Pa is enough to give stability against the typical capillary pressures from a bubble – the few 10s of Pa discussed above.

For ionic surfactants enter a typical surface potential of 100mV. This is a *magnitude* as the sign of the surfactant charge is not important – negative charges across the film repel by the same amount as positive charges.

For non-ionics enter a small surface potential of, say, 5mV which comes from the water itself. While it is agreed that this potential comes from $OH^-$ ions at the interface it is not clear why these ions form at the surface at concentrations far higher than in bulk water.

The presence of added salts dilutes the mutual repulsion across the film so the stability of the bubbles decreases. More salt increases the ionic strength, I, and decreases the Debye distance D which controls the exponential fall-off of the charge effect with distance. The other term which affects D is the dielectric constant, ε, which needs to be as high as possible which, of course, water provides. The modest surface charges from non-ionics are especially vulnerable to added salts which may diminish the mutual repulsion and interfere with the mechanism for generating the $OH^-$ ions at the interface.

For non-ionic surfactants, a long tail will obviously help with steric repulsion and some charged "impurities" that gave a few 10's of mV surface potential will also help. It is not at all clear what the "Surface charge" means, nor what to do with a surfactant to maximise it.

The most obvious conclusion is that ionic surfactants will tend to give "common black films" which are stable in the 30nm range and that non-ionics will tend to give "Newton black films" which are thermodynamically stable in the 5-10nm range but are clearly more delicate in terms of shocks from the real world such as impurities or antifoams.

Assuming we have a reasonable disjoining pressure at a reasonable thickness and that the inherent elasticity is adequate, we can at least create a foam. The question is how stable is that foam over time. There are two (generally) undesirable destructive forces, Ostwald ripening and drainage. There is one desirable destructive force, antifoams which, frustratingly, don't work as well as simple theories say they should. Each is discussed in turn.

## 5.4    Ostwald ripening

Ostwald ripening in foams follows the same logic as that in emulsions. The pressure inside a bubble goes as 2γ/r so the pressure inside a smaller bubble is larger than that in a bigger bubble. If gas molecules can diffuse through the foam, the smaller bubbles will get smaller and the big ones bigger.

There are two ways to reduce ripening:

- Decrease the solubility of the gas so it diffuses more slowly

- Increase the resistance to diffusion across the film wall

The first technique is used by Guinness breweries to give their beers a long-lasting fine foam. The natural gas in beers is $CO_2$ which has a rather high solubility. Substituting 70% of the $CO_2$ for $N_2$ greatly reduces ripening. Those who want to study foams with essentially no risk of ripening create the foam with perfluoroethane which has negligible water solubility.

The second technique, not surprisingly, fits in with the ideas of elasticity and, especially, a close-packed surfactant layer such as the one involving myristic acid which transformed a rather flexible foam into a more rigid one, Dawn changing to Gillette.

So the principles are clear. The practice is tricky because the key parameters vary at any given time because the foam bubbles have a size distribution and they vary over time because that size distribution changes. Prof Tcholakova at U Sofia was able to create an algorithm[45] to handle the complexity and kindly helped in the creation of the app. Most apps give instant feedback as variables are changed. This one is so complex that you have to click the Calculate button.



**App 5-3** https://www.stevenabbott.co.uk/practical-surfactants/Foam-Ostwald.php

The core idea is that, looking at the graph on the left, you start with a known distribution of bubble sizes, in this case one with a Peak at 200µm and a

45  Slavka Tcholakova et al, Control of Ostwald Ripening by Using Surfactants with High Surface Modulus, Langmuir 2011, 27, 14807–14819

Gaussian Width of 100µm, then calculate for a given tMax, in this case 3600s and find, at that point, a new bubble distribution with a peak near 1000µm and a width (by eye) of 500µm. The graph on the right shows how the average size increases from 200 to nearly 1000µm over that time.

Before describing other aspects of the calculation it is important to think about the size distributions. As discussed in the emulsions chapter and as shown in https://www.stevenabbott.co.uk/practical-surfactants/Distribution.php there are different definitions of size, each of which has its own validity. In that discussion the Number distribution and the D[3,2] were a focus of attention. In this app you can choose which size metric to use, depending on your specific interest. The default is D[3,2].

Other inputs are: k, the permeability of the film wall, h, the foam film thickness, Ψ the air volume fraction (which cannot meaningfully go below 70%), D the diffusion coefficient of the gas through the aqueous phase and γ the surface tension. The width of the bubble size plot is controlled by dMax.

Because we tend not to know the permeability of a film wall, two other inputs are provided: $k_{ML}$ is the permeability of a surfactant monolayer (which can, in principle, be measured) and H is the Henry constant for the gas in water. From these, k can be calculated as:

$$k_{calc} = \frac{HD}{h + \dfrac{2D}{k_{ML}}}$$    5-13

You can read out $k_{calc}$ and manually enter it as your permeability k. This can be used, for example, to compare CO2, N2 and C2F6. Take a reasonable $K_{ML}$ of 0.05 and changing H respectively from 0.83 to 0.015 to 0.0014 changes $k_{calc}$ from 1500E-5 to 27E-5 to 3E-5. With the default time of 3600s the app produces nonsense with the CO2 because the calculations are restricted up to the maximum bubble diameter of 2000µm. Either increase this to 10000µm and wait a long time (the calculation has to go slower at high k values) to find that the bubbles reach 4600µm, or keep it at 2000µm and change tMax to 120s and find that the bubbles reach 1000µm in that time. For N2 the bubbles reach 700µm after 1hr and for C2F6 they reach 320µm.

Returning to N2's value of H, $k_{calc}$ will also change if $k_{ML}$ changes. Adding some lauric or myristic acid can reduce this by a factor of 10, giving 320µm bubbles after 1hr.

Interestingly, with the standard values reducing the thickness of the bubble film from 30nm to 10nm does not change $k_{calc}$ much because h is combined with 2D/$k_{ML}$ which is typically equivalent to 40nm. That is why we have apps – formulae

are fine but their meaning only comes to life when we see how the numbers change (or don't change) when we alter the input parameters. Some non-ionic foams such as those from the APGs are remarkably stable even though DLVO tells us that they should be rather thin. If the insight about wall thickness applies, then at least one contribution to stability might be that the APGs create a rather impermeable wall, giving a low value of $k_{ML}$.

Those who make bubbles to entertain children (I speak from personal experience) will know that added glycerol helps. The Tcholakova paper clarifies this. The first effect is that H decreases – air is less soluble in glycerol/water than in water. The second effect is that the viscosity of the water increases so the diffusion coefficient, D, decreases. Decreasing both H and D by a factor of 2 (the effects found in the paper) reduces $K_{calc}$ from 27 to 12E-5 so the bubble size after 1hr is 500μm instead of 700μm.

Because the film permeability is such an important value, measuring it independently is highly desirable. It is conceptually easy to do though it requires a careful experimental setup. Create a bubble at the end of a syringe in a dish of the surfactant solution and monitor how its radius, $R_t$, at time t changes from the starting radius $R_0$. For a solution of density ρ, with atmospheric pressure P and gravity g, k is given by:

$$k = \frac{3P}{2\rho g t} ln\left(\frac{R_t}{R_0}\right) \quad \text{5-14}$$

How does the app calculate the Ostwald ripening? The algorithm is necessarily complicated because everything depends on everything else. Every bubble is assumed to change size according to a master equation that depends on the known condition of the whole foam at a given instant. This changes the whole foam, giving different pressure distributions which in turn change how the foam will change in the next time step. If the change in a time step is too big then the assumptions behind the change are false, so the time step has to be reduced. Once the small-enough changes have been made then the conditions of the whole foam are recalculated and the next time step can take place.

What are these "whole foam" conditions? Clearly each bubble has an idealised radius but is squashed into a deformed shape. This means that both the surface area and the bubble pressure depend on the amount of deformation which, in turn, depends on the air volume fraction and also on the current size distribution.

How is the size distribution measured? The foam was carefully prepared in a small dish with a low foam height to avoid drainage. The system was carefully covered to avoid water evaporation. Then a video camera system with image analysis measured the projected area A of each in-focus bubble and the radius calculated as $(A/\pi)^{\frac{1}{2}}$.

Clearly the Tcholakova paper is an impressive achievement and is even more admirable because it does the careful work to match theory with experiment and challenges its own assumptions by doing the necessary tests with myristic acid and glycerol.

But those who create foams don't just have to cope with Ostwald ripening. Drainage is also changing key properties such as the air volume fraction. We need, therefore, to understand drainage too. As we will see, drainage is a complex process. In an ideal world the drainage and Ostwald calculations would be linked. In practice studying one (e.g. drainage using C2F6 to avoid Ostwald) or the other (e.g. Ostwald using small foam heights so that drainage is less important) is more than a challenge for most of us.
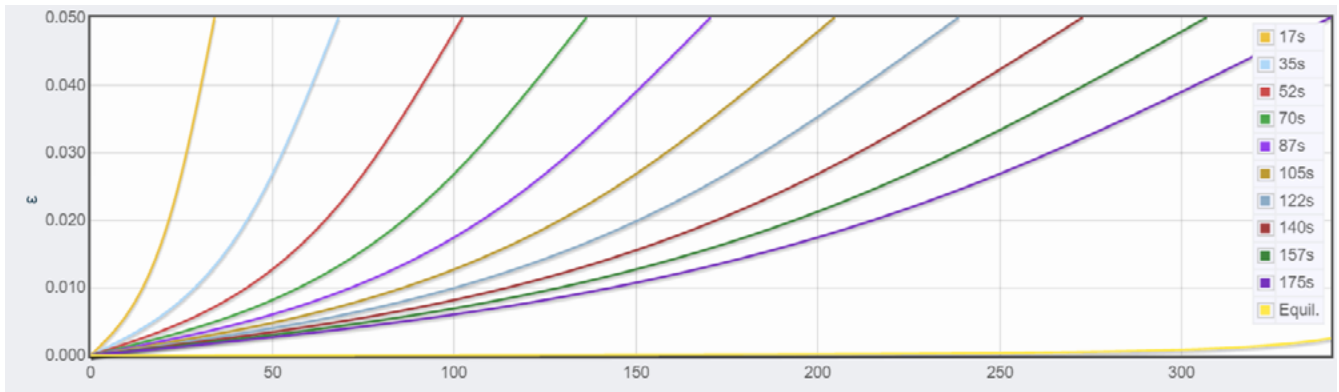
## 5.5     Foam Drainage

Take a glass cylinder containing a surfactant solution with a porous frit at the bottom. Pump air through the frit to create a foam. When there is enough foam, turn off the air. This is t=0. There will be a clear line of liquid near the top of the foam column. Over time the liquid drains through the foam and the liquid interface can be monitored as it falls, leaving a "dry" foam behind which may or may not break up (whether it does is of no concern to drainage theory). A key distinction is that in the very wet foam the bubbles are pure spheres. Once the % air exceeds the sphere close packing limit around 70% then the bubbles become polyhedral.

Such an experiment is called "free drainage" and represents our common understanding of drainage. Experimentally it is more challenging than "forced drainage" where liquid is pumped in at a controlled rate to the top of an existing foam, creating a steady-state condition more amendable to precise analysis. Here we will model free drainage, though taking advantage of the many experimental studies using forced drainage that helped develop and validate the theory.

The amount of water in the foam can be monitored with capacitance probes, with fluorescent additives, with light scattering or any other convenient technique. The foam world is divided between those who focus on fraction of air, $\Phi$ and fraction of water, $\varepsilon$ where, obviously, $\Phi = 1-\varepsilon$. Ostwald and foam rheology are more naturally thought of in terms of $\Phi$ while drainage is more naturally thought of in terms of $\varepsilon$.

The first challenge is to make sense of the experimental data. So we look at an image from the app before discussing the app and the science behind it.

**Figure 5-3 A typical output from the foam drainage app, showing the water fraction ε at depth z from the surface at different colour-coded times.**

The x axis shows z, the distance from the top of the column. At t=0 (essentially the y axis, so is not shown as a labelled line) the foam uniformly contained a liquid fraction, ε, of 0.05, i.e. 5%. After 85s there is exactly 0 at the exact top of the foam column and by ~25mm the foam contains the original 0.05. After 426s the top of the drainage zone has reached 150mm from the top and at 100mm from the top it contains 2.2% water.

Now we know what it is we wish to calculate we can see how to go about calculating it. The foam drainage literature is vast and I had to focus on app-able approaches so used a key paper from Koehler[46] and a group of papers from Saint-Jalmes[47]. The kind support by Prof Saint-Jalmes is gratefully acknowledged.

The first fact is surprising to most of us. Drainage has nothing to do with the water in the foam film walls. Even when those walls are new and "thick" at, say, 200nm, the volume contained in them is negligible. Anyone familiar with Poiseuille flow will know that pressure required to drive the flux in a channel of dimension r depends on $1/r^3$, so at 200nm the flux is very low. Yes, it is large in terms of getting the bubble walls down to their DLVO limit, but insignificant in terms of 5% water in the foam.

Instead the flux is along the Plateau borders, the triangular zones between faces. We will see that the nodes connecting the borders may or may not also be significant, even though they contain relatively little of the water.

The Plateau borders can be considered as pipes of length L and radius r. L can be derived from the bubble diameter D, it is simply L=D/2.7. The radius depends on the liquid fraction ε which is spread along a length L and we find that at time t the radius $r_t$ is linked to the liquid fraction $ε_t$:

---

46   Stephan A. Koehler et al, A Generalized View of Foam Drainage: Experiment and Theory, Langmuir 2000, 16, 6327-6341

47   Arnaud Saint-Jalmes, Physical chemistry in foam drainage and coarsening, Soft Matter, 2006, 2, 836–849

$$r_\tau = L\sqrt{\frac{\varepsilon_\tau}{0.17}} \qquad \text{5-15}$$

This means that the radius of the pipe decreases steadily as the liquid drains, so the rate of drainage decreases. Because drainage rate goes as $r^3$ (Poiseuille), it goes as $\varepsilon_t^{1.5}$.

That is the simple story. There is a complication caused by the fact that the pipe walls are lined with surfactant. If the surfactant is mobile then the wall itself can move and instead of Poiseuille flow which is zero at the walls (the "no slip" boundary condition) and fast in the centre, we have plug flow which is the same throughout the pipe and is overall much faster. It turns out that in the case of plug flow the bottleneck to flow is the nodes rather than the Plateau borders. So with relatively stiff surfactants that cannot flow, drainage is "channel dominated" and with relatively flexible surfactants that give plug flow, drainage is "node dominated".

How might you tell if your specific surfactant system is stiff or flexible in this context? The answer is to calculate an M value which depends on the bulk viscosity of the surfactant solution, $\mu$, and the surface shear viscosity of the pipe wall, $\mu_s$:

$$M = \frac{0.9\mu\sqrt{\varepsilon}}{\mu_s} \qquad \text{5-16}$$

When M<<1 (i.e. stiff walls because $\mu_s$ is large) flow is channel dominated, when M>1 flow is node dominated. The problem with M is that it is very hard to measure $\mu_s$ and few of us know where it lies in the general surfactant range of $10^{-4}$ to $10^{-2}$ g/s

It would, however, be reasonable to guess that the few % myristic acid that enhances elasticity and delays Ostwald ripening might change a node flow into a channel flow.

We now have enough information to describe the flow equations used to model drainage. There are three parameters needed to describe the flow.

- This is the motion of the drainage front itself – the borderline between the zone containing the original $\varepsilon$ and the foam that has started to lose liquid. The velocity is $v_f$.

- This is the velocity at which a front containing $0.5\varepsilon$ moves down the column. To imagine this, look back at the set of graphs and see that after 85s the 2.5% value is at 22mm from the top, at 170s it is 43mm and so

forth. These times and distances clearly depend on the velocity $v_m$ of the $0.5\varepsilon$ front.

- This is the equilibrium $\varepsilon(z)$ profile once the drainage front has reached the liquid phase.

The two velocities depend on whether the system is in channel or node dominated mode. Here are the relevant equations, where $\rho$ is the density of the liquid, g is gravity and $\mu$ is the viscosity. Another equation, from Stevenson, is included as it is used in the foam fractionation app later. Its overall behaviour is similar to channel mode.

| Factor | Channel Dominated | Node Dominated | Stevenson |
|---|---|---|---|
| Front Velocity $v_f$ | $v_f = \dfrac{0.007\rho gL^2\varepsilon}{\mu}$ | $v_f = \dfrac{0.002\rho gL^2\sqrt{\varepsilon}}{\mu}$ | $v_f = \dfrac{0.016\rho gr^2\varepsilon^2}{\mu}$ |
| Mid Velocity $v_m$ | $v_m = 0.5v_f$ | $v_m = 0.707v_f$ | $v_m = 0.5v_f$ |
| Equilibrium profile $\varepsilon(z)$ | $\varepsilon(z) = \dfrac{0.171}{(z_0 - z)^2}\left(\dfrac{\gamma}{\rho gL}\right)^2$ | | |

**Table 5-1 The foam drainage equations.**

Other than the surface shear viscosity, how else does the surfactant influence drainage? The basic answer is "not at all". The surface tension governs the final (capillary) profile and no doubt affects the bubble size for a given flow of air, but that is all. So drainage is very different from Ostwald where surfactant properties are super-important.

Finally we can look at the app. As you slide the sliders you will see that basically nothing happens. The Y-axis changes when $\varepsilon$ changes. The X-axis changes when the foam height H changes and the equilibrium curve of $\varepsilon(z)$ changes with $\gamma$ and bubble diameter D. The only other things that change are the time labels on the curves. If you swap from Node to Channel mode the shape of the curves change but that is all. First, have a play then we can discuss why the app does what it does.

**App 5-4** https://www.stevenabbott.co.uk/practical-surfactants/Foam-Drainage.php

Let us think through what the key parameters are doing, ignoring the equilibrium curve. D affects L which affects r – larger D means larger r which means faster drainage. H affects only the timescales because it takes $t_H=H/v_f$ for the drainage front to reach the liquid. Don't get too hung up on the exact measure of H, something that is hard to determine in practice. An error of 10% in H is simply an error of 10% in $t_H$. Density doesn't vary much, Viscosity is simply a linear effect – double μ and you double the time shown on any curve. Liquid fraction ε has a square root effect in Node mode, a linear effect on Channel mode and a square effect on Stevenson.

So the app tells us that a foam with large bubbles, lots of liquid and a low viscosity drains faster than one with small bubbles, little liquid and high viscosity. Who needs an app to state the obvious? First, it was not at all obvious that drainage should be so straightforward. Second, the Node versus Channel distinction is significant. If we started with the same 400μm Dawn foam and Gillette foam, the drainage front in the stiff Gillette foam will take 25% longer to reach the liquid than the Dawn foam – a significant difference.

But why does nothing much happen in the app? Think, first, of the drainage front. Its velocity is a constant and all the complicated things that happen in the drying foam behind it are irrelevant. We also have an end goal, reaching the liquid surface, which defines a master time $t_H=H/v_f$. So we can instantly calculate a set of equally spaced points down the foam column marking where the front is at intermediate times. We also know that at z=0 ε=0. Well, some experiments show this, others don't and it is very hard to calculate at that exact point. Even if the assumption is wrong near the start, it makes little difference to the overall

pictures, so we keep things simple. And we equally know $v_m$ the velocity defining the 0.5ε point. The app simply joins up the three points in a straight line for Channel mode and a "reasonable" graphics curve for Node mode. Given the practical difficulties of getting reliable experimental data, this simple approach seems to be more than adequate.

The equilibrium curve is again more illustrative than exact. It is very hard to measure, most of us don't have time to wait for it to be attained and other effects like Ostwald or foam breakage through evaporation are likely to intervene.

Writing a book and its apps is a compromise. I could fill this section with lots more fascinating equations, but in my view they would not help much. I could have implemented a more sophisticated algorithm for the app, but in my view the outcome would not have been much different. I must, however, put any blame for a light touch on myself. It has taken a vast amount of very hard experimental and intellectual endeavour from the experts in the field to reach what are fairly simple conclusions. Simplicity, in this case, is a great intellectual achievement on their part and any over-simplification is entirely my responsibility.

## 5.6    Foam Rheology

Scientific instruments are so powerful that they can measure just about anything we might want to know about. The question, then, is what *should* be measured? A delicate rheometer can be unleashed on a foam and can measure things like elastic modulus G', loss modulus G'', stress and viscosity with shear and yield stress. But why would you bother?

One way to answer that is to bring in a consumer test panel to assess some sort of foaming product. Consumers can detect subtle differences between different foams and will prefer the "feel" of some over others. This feel will, of course, depend on bubble radius R, on the air fraction Φ. The stability over time (Ostwald and/or drainage) might also be a factor. In addition, the human finger can detect very subtle differences in the way the foam reacts to touch and the scientific way to understand such differences is through rheology. Other examples are discussed in the Denkov chapter in Foam Engineering[48] from which the formulae below are taken: extrusion through nozzles and slits for food; transportation through pipes e.g. for oil recovery; flow through porous media, also for oil recovery. Prof Denkov's generous assistance in creating this section is gratefully acknowledged.

There are three parameters that most interest us:

- Modulus (how elastic the foam is);

---

[48]   Nikolai Denkov et al, Foam Rheology, Ch 6 in Paul Stevenson (Ed) Foam Engineering: Fundamentals and Applications, Wiley, 2012

- Yield Stress (how resistant it is to movement);
- Viscosity under shear (viscosity changing with velocity and gap of flow).

The shear modulus, G, depends on air fraction Φ (note that we have switched back to Φ from ε), surface tension γ, radius R (actually R[3,2]) as:

$$G = \frac{1.4\Phi(\Phi - 0.64)\gamma}{R} \qquad \text{5-17}$$

The 0.64 is the value of random ("jammed") packed spheres, though the calculated values close to that limit are unreliable. A "good" surfactant with a low surface tension will tend to give large bubbles so γ/R will be small, giving a weak foam. For a dry foam, Φ~1 so G=0.51γ/R. For those who want the elastic modulus, E (not to be confused with our earlier Elasticity, E), the Poisson ratio, ν, of foams is very close to 0.5, so E=2G(1+ν)=3G. The high Poisson ratio means that foam bubbles are "incompressible", i.e. their volume remains constant under compression. This makes analysis of foam rheology easier or, at least, less hard.

The yield stress, Y, which is the stress at which the "solid" foam will start to flow like a liquid, shows a similar behaviour.

$$Y = \frac{0.5(\Phi - 0.64)\gamma}{R} \qquad \text{5-18}$$

A foam which feels "rich" will tend to resist flow under pressure, i.e. will have a large Y, which can best be achieved via a small bubble radius. Again, a "good" surfactant will tend to give a poor performance because the radius will be large. For a dry foam, Y=0.18γ/R.

These values of G and Y are stated without a time-scale. They are values under "normal" measurement conditions rather than those at extremely low or high frequencies. Those with a suitable rheometer can measure G' and G" as a function of frequency. Those with patience can do experiments below Y and over long times and watch the creep that takes place as the bubbles get larger owing to Ostwald ripening. Although these are interesting effects they are not included in the app as they require input values that are not readily computed from fundamental parameters.

Once the bubbles start to flow past each other (i.e. the applied stress is greater than Y) we are interested in how the viscosity changes with the applied shear rate. We all know that shear rate is usually shown as $\dot{\gamma}$, gamma with a dot above it. Because it is easy to confuse this with surface tension γ, the letter S will be used instead.

Viscosity is a measure of the dissipation of energy during flow. If we ignore the friction of bubbles sliding against the wall of a pipe, for bubbles there are two forms of dissipation:

- $\tau_{VF}$ Sliding friction (bubbles sliding past bubbles);
- $\tau_{VS}$ Surface dissipation (bubbles expand/contract during motion and dissipate energy).

Bubbles with nice freely-moving, "fast" surfactants show little surface dissipation. So Dawn foams show only sliding friction while Gillette foams also show surface dissipation.

The sliding friction is due, in principle, to two effects. The first is "standard" viscosity as it involves shear across a gap, h, between two bubble walls that are sliding past each other. The second is caused by the (Poiseuille) flow of liquid between two bubbles that are being pressed together. It turns out that the sliding term is more important than the thinning term.
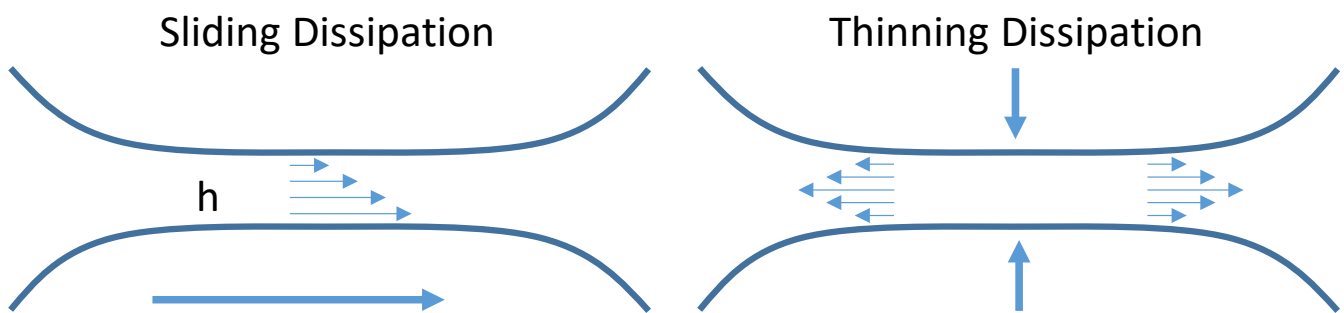


Figure 5-4 The two types of dissipation within the liquid between two bubbles. Both are included in $\tau_{VF}$ but the sliding mode is more important than the thinning mode.

The equation describing $\tau_{VF}$ can be described in a quasi-universal formula based on the non-dimensional capillary number (Velocity.Viscosity/Surface Tension) and scaled by γ/R. However, for those of us who have specific questions it is hard to go from normalised, non-dimensional formulae, so the specific, dimensional, equation is used here. Where μ is the viscosity of the liquid, S is the shear rate then:

$$\tau_{VF} = \frac{1.16\gamma}{R}\left(\mu S \frac{R}{\gamma}\right)^{0.47} \Phi^{\frac{5}{6}} \frac{(\Phi - 0.74)^{0.1}}{\sqrt{1-\Phi}} \qquad 5\text{-}19$$

The app brings the formula to life:

**Foam Rheology**
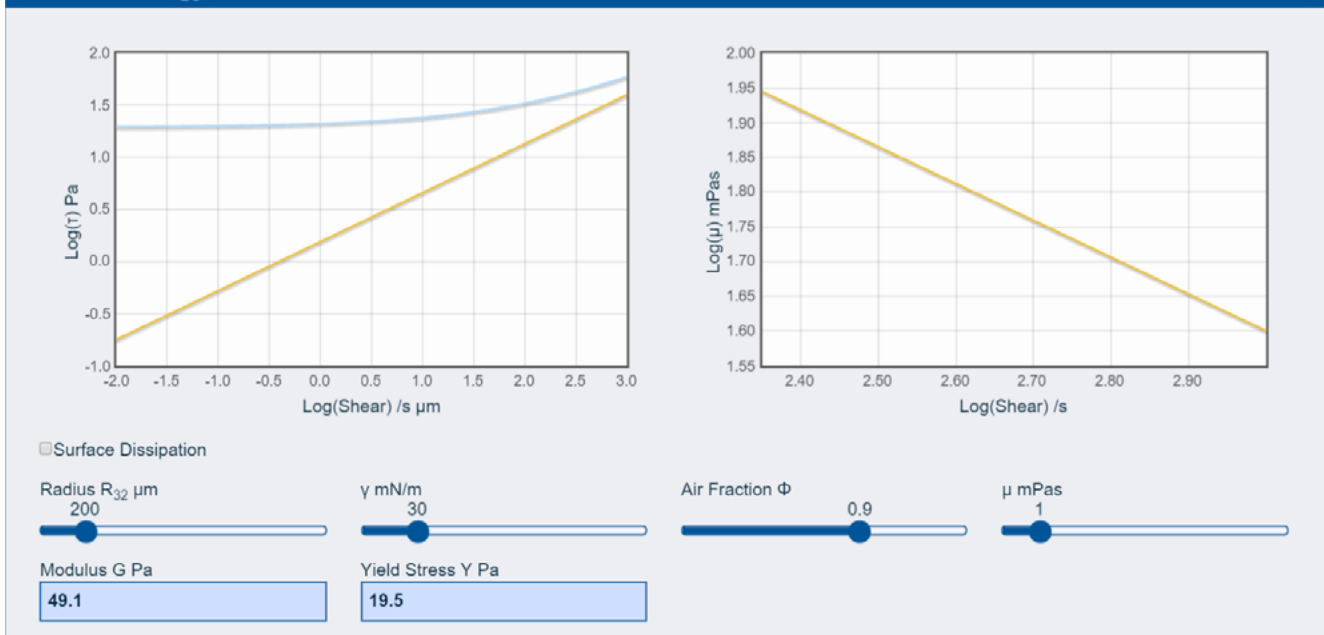
The two graphs represent two ways of looking at the same data. The graph on the left shows the strain dependence on shear, the sort of plot rheologists like to see. The graph on the right shows viscosity ($\tau$/S) dependence on shear, which is more familiar to most of us and which shows the familiar shear thinning effect.

Why does the shear plot have a second curve in blue? This is a more realistic plot because it takes into account the yield stress, Y. At the lowest shear rate the log value is 1.28, i.e. the 19.5 Pa calculated value of Y. In other words, you can't get data at that shear rate because the calculated shear stress is less than Y. It is rather hard to imagine what the viscosity plot should look like so it contains only the curve matching (calculated from) the ideal stress curve.

What are the calculations saying? Basically that a plot of $\log(\tau_{VF})$ or log(viscosity) versus log(S) should be a straight line, with larger surface tensions and smaller radii providing larger stresses. That is intuitively obvious. The effect of $\Phi$ is potentially very complex as it approaches 1, so the app limits it to 0.99. Clearly a drier foam is more viscous as the gap, h, between bubbles is necessarily smaller.

The shear modulus, G, plays no part in these discussions – it is calculated in the app because we have the formula.

So far the calculations have only been about the sliding friction. What about the surface dissipation? The formula is much simpler:

$$\tau_{VS} = \frac{9.8\gamma}{R}\Phi^2 B.S \qquad \text{5-20}$$

The key parameter is B. It is currently not calculable from standard parameters. For low modulus, flexible bubbles, B=0. For the most rigid bubbles so far measured it is 3E-3. In the app it is set to 2E-3 if the Surface Dissipation option is selected and 0 otherwise. For the default settings the effect is quite dramatic:
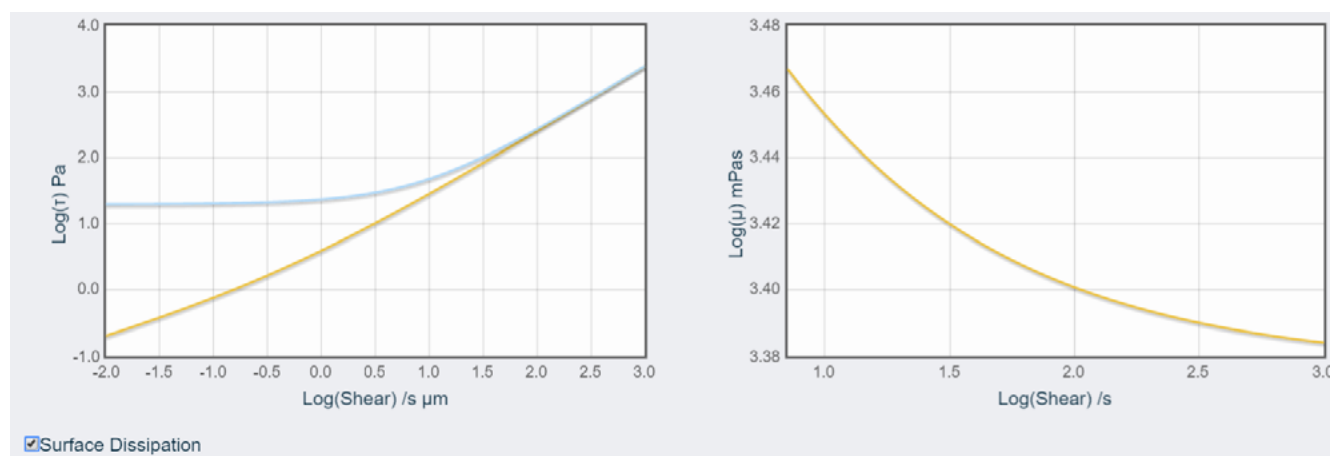


**Figure 5-5 The effect of turning on the Surface Dissipation option with all other parameters unchanged.**

The maximum stress is now 2000 Pa instead of 40 Pa. Adding a few % myristic acid (to use our standard example) gives a 50x increase in stress and viscosity.

As was the case with foam drainage, this exploration of foam rheology makes it look easy, with the conclusions of the models being rather obvious, that small, dry bubbles with stiff surfactants are more resistant to flow. That "simplicity" is only because I can choose to implement theories that have taken a long time to be produced and verified and can ignore the many other complexities that are less easily packaged into app format.

## 5.7    Antifoams

When you want a nice stable foam you often fail to get one. Conversely, when you absolutely want no foaming you often find a stubbornly stable foam that interferes with your process. It seems astonishing to most of us that one cannot simply add a standard antifoam to make all the problems disappear.

Even if we do find a reliable antifoam it might do nothing for some minutes and then the foam disappears. Why doesn't it simply work at the start?

These questions are all the more puzzling because for years there was a simple, compelling theory about how to create an antifoam and the science could point to no obvious difficulty other than the practical issues discussed in the next section.

Although there is an app for the simple science there is as yet no app for designing the perfect antifoam. Instead there is a recent understanding, again

from the Denkov group[49], about why antifoams have been so puzzling. Armed with this understanding the careful formulator might arrive more quickly at an adequate reduction, if not removal, of unwanted foam.

Antifoams can be pure oils, pure hydrophobic particles or blends. Some of the scientific details (e.g. the Bridging Coefficient does not apply to particles) vary somewhat, but the take-home message is the same for both so the different systems will mostly (there is an important exception) be treated as one and the language may stray between "drop" and "particle" to describe the antifoam.

Before embarking on the science of these sorts of antifoams it is worth recalling that there is a different class of antifoams – those that destroy the surfactant itself rather than the foam. A typical such antifoam is hard water when the surfactants are fatty acid salts such as sodium stearate. The calcium ions in the hard water cause the stearate to become insoluble and the foam disappears. Similarly, a small amount of a cationic surfactant can rapidly destroy an anionic surfactant foam. Finally, adding a lot of salt to a foam with a large electrostatic disjoining pressure can reduce the disjoining pressure, thin the foam and make it more susceptible to collapse.

And it is worth noting an intellectually interesting and scientifically proven alternative to adding antifoams is discussed below – just blow dry air over the foam to hasten evaporation of the bubbles so they burst. The idea comes from Prof Paul Stevenson and is covered in his foam fractionation book discussed below.

For those who wish to explore the science in much more depth, the definitive book by Prof Peter Garrett, *The Science of Defoaming: Theory, Experiment and Applications*, is highly recommended. The book is especially good because the author combines academic rigour with reality from his decades of antifoam work in industry. The "theory lite" discussion that follows has benefitted greatly from discussions with Prof Garrett for which I am deeply grateful. The "science bit" section captures some key basic ideas from the book which are simultaneously profound and deeply confusing to non-experts.

### 5.7.1  Practical issues

Although the issues in this section are neither profound nor especially interesting, they contribute a lot to the failure of many antifoams.

The first is that the antifoams should be strongly hydrophobic, which is no problem to the antifoam designer. The problem is that the antifoam should be present as lots of small drops/particles. How does one get a strongly

---

49   N.D. Denkov, K.G. Marinova, Antifoam effects of solid particles, oil drops and oil-solid compounds in aqueous foams. Chapter 10 in "Colloidal Particles at Liquid Interfaces" (B.P. Binks & T.S. Horozov, Eds.), Cambridge University Press 2006; pp. 383-444

hydrophobic oil to disperse throughout the foam rather than sit as a film along the surface? One answer is to use a surfactant – to create an emulsion form of the oil. If it is an especially good surfactant system then the oil is now dispersed but completely useless as it is effectively a hydrophilic drop thanks to the protective shell of surfactant.

This links to another issue. An antifoam might work wonderfully in the presence of 0.1% surfactant and be completely useless at 0.2% surfactant – for the obvious reason that at 0.2% the surfactant is sufficiently able to engulf/emulsify the antifoam droplets. This partly explains why silicone or fluoro antifoams tend to be more successful than other oils. The oils themselves would have no problem killing the foam but are more easily emulsified than the silicones, for reasons that were explained in the HLD chapter.

The other problem is that for slow-acting antifoams (the reasons for this are discussed below) in continuous processes such as bio-fermentation, the bubbles that pop near the surface will leave their antifoam at the surface with little chance of being mixed into the film walls of future bubbles. So the antifoam effectiveness within the bulk drops off over time. This means that a continuous feed of antifoam is required into the bulk, and because we need small particles, that feed might require agitation to disperse the antifoam and that agitation in turn might create more foam.

### 5.7.2  3 Coefficients

The plausible science that guided many formulators of antifoams looks at the three relevant surface tensions or interfacial tension. The antifoams are oils (or hydrophobic particles) and the oil has to be in the right place at the right time to destroy the foam:

- $\gamma_{AW}$ – the air/water surface tension
- $\gamma_{OW}$ - the oil/water interfacial tension
- $\gamma_{OA}$ – the oil/air surface tension

In turn these give three combinations that describe where the antifoam oil might be. Each shows a different balance between the competing interfacial energies:

- Entry coefficient: $E = \gamma_{AW} + \gamma_{OW} - \gamma_{OA}$. If E is negative the aqueous phase completely wets the oil drop, stopping it from getting to the foam interface.
- Bridging coefficient: $B = \gamma_{AW}^2 + \gamma_{OW}^2 - \gamma_{OA}^2$ . If B is positive and if E allows the drop to enter the wall, and if the oil drop radius is large enough to bridge the wall, then the bridge will be unstable and the film wall will break. Note that if B is positive then E must also be positive (but you can have a positive E and a negative B). There is a further restriction, B must also be less than $2\gamma_{OW}\gamma_{AW}$.

- Spreading coefficient: $S = \gamma_{AW} - \gamma_{OW} - \gamma_{OA}$. If S is positive then the oil can spread along the foam, contributing, perhaps, to its breakup. There is major disagreement about whether oil spreading is important and there are studies both confirming and refuting the idea that S is important

So the recipe for a good antifoam is simple: positive values for E, B and S. The app does the calculations for you:
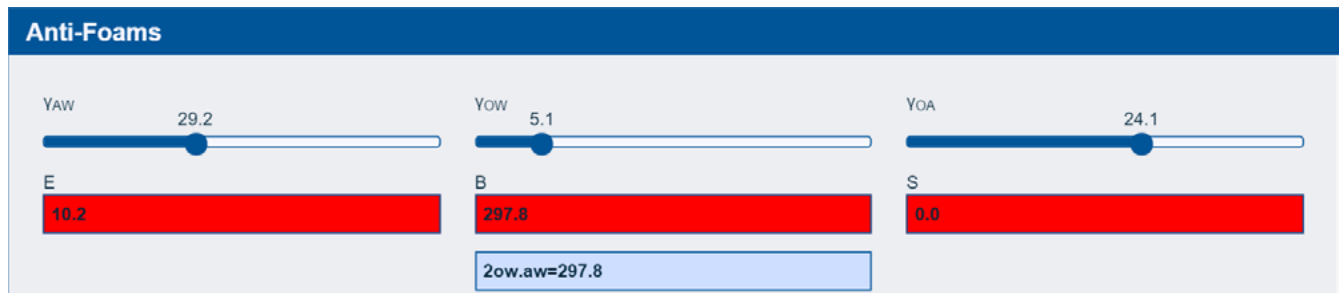
The red colour indicates "bad" in terms of the foam lifetime, i.e. a potential good antifoam. If either E or B is green that means that the antifoam cannot attack the foam. As mentioned, S is controversial. In reality, most antifoams, especially the silicones, have no problem providing positive values necessary for a good antifoam, though the extra requirement for a *maximum* value for B adds some complications – you will find it quite hard to get all three boxes to be red.

These ideas are easy to grasp, the three γ values are easily measured and the calculations are simple. There are, however, a number of big problems.

- These coefficients are most easily measured with pure oil and "pure" water+surfactant. However, often more relevant are the *equilibrium* values of water saturated with oil and oil saturated with water with the surfactant going where it wants to go. So it might be that the *initial* entry coefficient, measured on pure liquids, meets the criterion but the *equilibrium* value (which is generally much smaller) fails to meet the value.

- Because it takes time to reach the equilibrium value, and because of the issues of dynamic surface tensions, there is no reason why an antifoam that works well when foam is generated vigorously (short timescales), e.g. with harsh shaking will work well when foam is generated gently (long timescales) e.g. with gentle bubbling.

- The spreading coefficient is especially confusing. At *equilibrium* it is (almost by definition) equal to zero.

- The predictions are **highly unreliable in an asymmetric way**. If they say that a given oil will not be a good antifoam for a system the predictions are correct. But the opposite is often not the case: many predictions consistent

with good antifoams yield poor antifoam behaviour. Clearly something is missing.

- One extra factor must be that the oil drop (or particle radius) is large enough to break through the wall. Very small drops/particles will be ineffective. Yet even when the size is large enough, the antifoaming behaviour is unsatisfactory.

- Hydrophobic particles on their own tend to behave similarly to oil drops on their own (though the detailed mechanisms must be different as the particles cannot flow like oils). The combination of particles and oil is generally far more potent. The oil seems to set the scene but is hampered by the issues discussed below. The effect of the extra particles is to enable the required breakthrough.

- Particle shape can have a profound effect, with jagged particles likely (but not always) to be more effective.

### 5.7.3 The Entry Barrier problem

The problem with the Entry Coefficient and Bridging Coefficient arguments is that they are thermodynamic. The reason most antifoams fail is kinetic. There is an entry barrier to the antifoam particle so although it would destroy the foam wall if it could get into it, it sits outside doing nothing.

In fact, it is worse than that. Here is an image of an undoubtedly potent antifoam particle sitting happily inside a foam. The image is of some coffee oil in an espresso foam and is reproduced with kind permission of illycaffé. This is an example of where we are happy that antifoams fail, otherwise we could not enjoy the magic of a crema on our espresso.
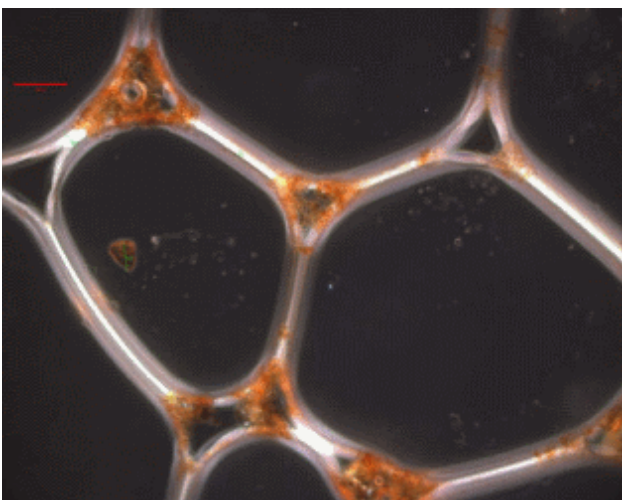


**Figure 5-6 Failed antifoam particles sitting in the Plateau borders and nodes. Image courtesy of Dr Luciano Navarini, illycaffé.**

The image shows that the antifoams are sitting happily in the large (by the standard of the foam film walls) Plateau borders and nodes. Only when the

borders/nodes drain to a sufficiently small size will the particles be forced by capillary pressure into the walls where they will burst the bubble.

Why are the particles sitting in the Plateau border? Statistically they were far more likely to have been found in the film wall at the moment the foam was created. The wall at that instant is relatively thick and there is rapid capillary drainage into the Plateau borders (not to be confused with foam drainage which is flow through the borders). Because of the entry barrier and the relatively low capillary pressure (large fluid thickness in the wall) very few of these particles burst the walls and were instead swept into the Plateau borders where they are relatively harmless. It is a sad fact that most of our antifoam is doing nothing except sitting comfortably in the Plateau borders and nodes.

How strong is the entry barrier? The answer is "frustratingly strong" and it can be measured using the conceptually beautiful Film Trapping Technique: FTT.
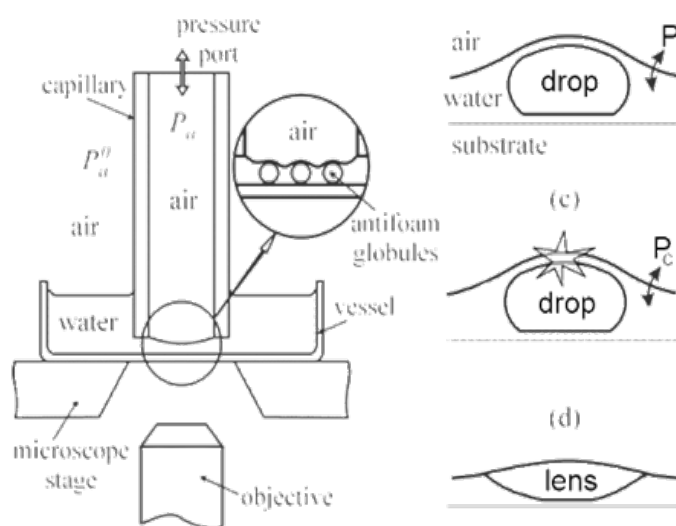


**Figure 5-7 The Film Trapping Technique. Image courtesy of the Denkov group.**

Insert a small glass tube into a surfactant solution that contains small drops (or particles) of your antifoam. With carefully controlled air pressure, force the liquid/air interface into contact with the antifoam drop. Thermodynamically the drop should burst the thin film of surfactant and water, but you can continue to increase the pressure and the interface wraps itself around the drop. At some critical pressure $P_c$ thermodynamics overcomes kinetics and the drop breaks through the interface.

So to design a good antifoam all you need is an FTT setup so you can tune your antifoam to your surfactant and find the smallest possible $P_c$ for bursting the bubble. Unfortunately it has so far proved impossible to make a general-purpose

FTT setup that can be used routinely by non-experts in an average lab and we have to rely on the U. Sofia experts for most of our information on $P_C$ values. It can only be hoped that someone with sufficient ingenuity and resources finds a way to make this important technique more generally available to development labs. It seems to me to be an excellent commercial opportunity for one of the companies that already provide tools for measuring surface phenomena.

It will come as no surprise to the reader that $P_C$ for a Dawn foam is much lower than that of a Gillette foam. In the absence of one's own FTT setup the working assumption has to be that a good antifoam depends less on the antifoam (assuming you have a reasonable oil, reasonably dispersed without being protected by whatever surfactant is in your system) and more on ensuring that the surfactant causing the foam is of low elasticity or, perhaps, has a disjoining pressure curve that ensures a stable zone of only a few nm rather than 30-50nm, so that it is more susceptible to small fluctuations that can pierce a 5nm film and not a 50nm film. Presumably an antifoam will work better in a system with an ionic surfactant if there is plenty of added salts to increase the ionic strength and reduce the electrostatic stabilisation.

When we discuss the making of a good latte we will touch on an issue that has appeared in other chapters. Many food surfactants are spectacularly bad surfactants when judged by many of the standard measures. And making good foams with such bad surfactants is not easy, as anyone who has tried to make a meringue can attest. Yet once made they can be surprisingly stable even though they contain large amounts of potent antifoams – fat globules. A measure of $P_C$ for a food foam will probably show a very high value. Those big, lumbering surfactants are poor in terms of creating a foam but are strongly resistant to rupture by a mere drop of oil.

### 5.7.4 Fast and slow

We can now see why there are (sometimes) fast antifoams and (often) slow antifoams. Note that the *physics* of fast and slow modes is the same - entry and bridging coefficients, entry barriers etc.; it is the *location* of the rupture of the foam which differs between the two modes. A fast antifoam works within the film wall during the short, vulnerable period when statistically most drops/particles are in the wall area. If the entry barrier is rather too large then the drops are swept into the Plateau borders and nodes where they can do no damage until the borders become small enough for the resulting pressure on the particle to exceed $P_C$. This then takes us back to foam drainage where times are measured in many seconds, so this form of antifoam action is "slow". Nothing happens for many seconds before there is enough drainage for the foam to burst. To repeat the point made previously, drainage has nothing to do with removal of water from the film walls, it is connected only to removal of water from the Plateau borders. Here we hit a problem. For fast drainage we want large bubble diameter D which gives a large value for r ($r \approx D.\varepsilon$), the effective radius of the

Plateau border. Yes it is precisely this large r which is stopping the antifoams from functioning. So, other things being equal, smaller bubbles will break faster than larger ones. Drier foams (smaller liquid fraction $\varepsilon$) will also burst faster. And because drainage times are proportional to viscosity, low viscosity liquids will burst faster. Surface tension plays no role in drainage though it obviously affects how much foam is created in the first place and the average diameter of that foam.

The trapping of the antifoam in the Plateau border gives us a dilemma. To pierce a film wall only requires antifoam drops/particles of, say, 100nm. If all the antifoam is in the form of 100nm particles then there is a high statistical chance of them bursting a bubble if $P_C$ is attained. But if all these particles are swept into the Plateau borders with radii of 10's of µm they have little chance of doing much because they are unlikely to be squeezed enough to break through the wall. So is it better to admit defeat and disperse the antifoam as large particles? The downside is that for a fixed volume of antifoam, a 10µm dispersion would have $(10/0.1)^3 = 10^6$ fewer particles. Aiming for 1µm dispersions gives $10^3$ fewer particles but might be the worst of both worlds. It is a difficult judgement call.

### 5.7.5  Disappearing antifoam

We have already hinted at the various ways in which the antifoam can disappear from active service. Here is a summary:

- Antifoam oils or particles are hydrophobic. This means that they can be difficult to disperse and added to the system. But of course, it is easy to disperse oils/particles simply by adding surfactants. So here is the first contradiction - if the surfactant in the system isn't good at dispersing the oil or particle then the antifoam may sit in useless lumps at the surface or at the bottom, but if the surfactant is too good at dispersing then the oil drops can become far too small to be effective and the particles might be totally covered by a surfactant shell, rendering them completely hydrophilic and useless.

- The oils/particles are sitting comfortably within large PB and the drainage is so slow (small diameter bubbles, rigid surfactant) that they never get the chance to break a wall or PB.

- It has often been noted that specifically for silicone oils containing hydrophobic particles the original excellent antifoam activity falls off over time. One explanation for this is that the particles "disproportionate". This means that if two oil drops each containing 4 particles happen to meet and fall apart again, one ends up with 6 particles and the other with 2 and after a while a few drops have all the particles and the rest have none. The causes, cures and generality of this specific failure mode are hotly debated.

- Particles are so slow at getting to where a bubble is being formed that most of them are sitting uselessly in the bulk solution with no chance to participate in either fast or slow antifoaming action.

- In many processes such as bio-fermenters, the antifoams were so good at getting to the bubble interface that after a while they have all been removed from the bulk and are sitting uselessly at the surface of the previously broken foam. Only a fresh injection of antifoam into the bulk will allow new bubbles to have a supply of antifoam.
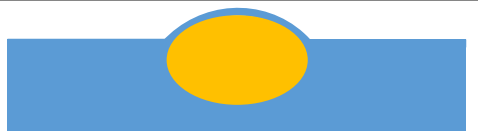
### 5.7.6  Now for the science bit

The formulae for entry, bridging and spreading coefficients and the idea of the critical pressure have all been introduced with no explanation of where they come from. Although the science behind them is not especially hard, it happens to be especially confusing. This is partly because it depends on ideas of curvature (capillary pressures require curvature) in various planes that are difficult to envisage. And is partly due to the fact that the nomenclature used is deeply confusing. Terms like "pseudoemulsion" and, even worse, "pseudo partial-wetting" were never designed to be user-friendly and are, to some extent, well-intentioned historical accidents. I am going to propose a more user-friendly nomenclature with little hope that it will be adopted. But one has to try.

Let us start with an experiment that was at one time quite popular but is generally banned for obvious reasons. Take some distilled water in some large open vessel so that edge effects can be ignored. Add a tiny drop of benzene. The surface tension goes down because you get a tiny water/benzene interaction at the surface. Notice the vague term "interaction". There is no doubt that the surface tension has gone down. There is also no doubt that this is not a classic monolayer that you would expect from a similar experiment with a surfactant. The precise nature is not important, because the interesting behaviour follows when more benzene is added. The extra benzene spontaneously forms into little lens-shaped drops. As you add more benzene you tend to create more lenses rather than bigger lenses. The curvature of the lens is governed by the usual balance of forces, but in this case it is not a balance of oil/water but oil and "water with its surface tension lowered thanks to the tiny amount of benzene at the surface".

Many oils behave differently. Something like hexadecane simply forms lenses, with the water between the lenses having its original pure surface tension. Something like silicone oil (PDMS) simply creates a uniform film across the surface right from the start, and as you add more oil the film simply gets thicker. Actually, there is evidence for some lens formation which may or may not be relevant to the antifoam behaviour of silicones, but such niceties are not the chief concern of this section and in any case the lenses are very thin.

Although the above behaviours can be found for oils on water, our interest is in oils on water with surfactant at the surface. All three behaviours can be found, with the complication that the oil is not just providing some "interaction" at the water surface but is interacting with the surfactant tail and changing the surface tension indirectly. This is a serious extra complication.

So, we have three behaviours and these have been given three names, along with a fourth term that will be discussed below:

| Behaviour | Name | Image |
|---|---|---|
| Lenses only | Partial wetting |  |
| Lenses with the benzene-like "interaction" and lower surface tension | Pseudo-partial wetting |  |
| Uniform oil film | Complete wetting with "duplex film" |  |
| Oil drop at the foam wall water/air interface | Pseudo-emulsion |  |

**Table 5-2 The official descriptions of the three possible fates of an oil at a surface, plus the strange term "pseudo-emulsion".**

To me "partial wetting" should be called "non wetting". But why refer to "wetting" at all. Why not call it lens formation because that is the key characteristic? I think that "partial wetting" is an attempt to distinguish the behaviour from those cases where the oil does not want to come to the surface at all – i.e. fails the basic entry coefficient test. But because we are concerned with antifoams which we know must come to the interface, just calling them "lenses" seems to be sensible.

Pseudo-partial wetting is well-intentioned. It is trying to say that the lenses are different shapes from "pure" lenses because the forces are different because the surface tension of the water is lower due to the "interactions" at the surface with a tiny amount of the oil. It seems to me that they should be called modified surface lenses.

The complete wetting is fine as a term, but the mysterious "duplex film" is unhelpful and should go. So the revised table should be

| Behaviour | Name | Image |
|---|---|---|
| Lenses only | Lenses |  |

| Lenses with the benzene-like "interaction" and lower surface tension | Modified surface lenses |  |
|---|---|---|
| Uniform oil film | Complete wetting |  |
| Oil drop at the foam wall water/air interface | Wall drop |  |

**Table 5-3 The more helpful descriptions of the three possible fates of an oil at a surface and replacement for the unhelpful "pseudo-emulsion".**

In terms of the detailed mechanisms of how oils can cause foam films to fail there are lots of fascinating intellectual differences between the three types. Although it is important for these distinctions to be made in the literature and although clever experiments can help disentangle what is going on, the net result shines little light on what one should do practically to get a good antifoam. Issues such as dynamic behaviour and entry pressure are, in my view, far more important than debates around the precise mechanisms which, in any event, seem not to be fully resolved. There are, for example, still people who believe that foams break via Marangoni effects and others who believe that Marangoni effects cannot be a root cause.

It is worth noting that the modified surface lenses relate to an important distinction made in the discussion of entry and bridging coefficients. Clearly the air-"water" and oil-"water" (and possibly air-"oil") surface tensions are very different if there is some surface interaction with the oil on the water (or "water" interaction with the surface of the oil) – especially because for foams the water has surfactant at the surface which has the double dynamic behaviour of first coming to the surface and second being modified by the oil. That is why the coefficients measured in equilibrium (when the surface can become modified) can be very different from (and generally much lower than) the "initial" coefficients based on measurements of pure oil and water.

We now come to the other confusing term: *pseudo-emulsion*. I cannot find any circumstance where this term is helpful. It is simply an oil drop pushing against the water+surfactant/air interface at the foam wall. The most important fact of antifoams is that getting the drop to break through into the foam wall is hard (requiring the entry pressure). So why not, in the context of antifoams, just call this drop of oil at the wall a wall drop? This ensures that every time it is mentioned the emphasis is on how to get the wall drop through the wall. For example to say that "hydrophobic particles within a pseudo-emulsion can exert extra pressure and break through" is true but confusing. To say "hydrophobic particles in a wall drop can exert extra pressure and break through" is much clearer.

That ends our discussions on antifoams so we can come to an issue that has been carefully avoided so far: the science of *making* foams.

## 5.8    Making foams

It would be reasonable of the reader to expect an app that unites all the above threads and predicts the volume of foam, liquid fraction and bubble diameter from a modest set of inputs. Regrettably I can find no way to create such an app.

The key problem is that creating a foam is very easy. Just about any method that creates a dynamic air/water interface can create lots of bubbles. Bubble size and foam density depend on multiple competing issues of creation, destruction, coalescence, each of which depends on the dynamic behaviour of the surfactant at the interface. And because most good foams are created with surfactant blends and small quantities of other additives, the "dynamic behaviour of the surfactant at the interface" is itself a hugely complicated set of issues.

There are many different answers to the simplest of questions such as the typical foam diameter to expect from generating a foam or the speed at which foam bubbles might rise in a column. An excellent summary of such knowledge[50] provides an overwhelming number of formulae to fit a large number of different cases.

Here is just one example of the complexity. Bubbles can be formed by blowing gas into a surfactant solution through a porous frit which has an average pore diameter from which an average bubble diameter might be calculated. However that diameter depends on the surface energy of the frit material, and the size distribution of the holes in the frit gives different bubble sizes depending on the gas pressure. At low pressure, bubbles are formed only via the largest pores, giving a diameter larger than the average and as the pressure increases bubbles form at the smaller pores. This larger density of bubbles from more sites in turn can affect the bubble diameter through crowding effects. And the plot of diameter versus pressure measured during an *increase* in pressure is different from the one measured when the pressure is subsequently *decreased*. The shapes of those curves will also shift depending on the surface energy of the frit. If even this "simple" system is complex there is little hope for general rules to guide us in our own foam forming systems.

### 5.8.1  Foam making essentials

Still, I've found that some words I wrote in 2014/15 have stood the test of time, so let's see what we *can* say about making foams, with some additions from 2020.

---

50   D. Exerowa, P.M. Kruglyakov, Formation and Structure of Foams. Pressure in the Liquid and Gas Phases of Foams, in Foam and Foam Films: Theory, Experiment, Application, Elsevier, 1997

As we know, the energy required to create a foam is inversely proportional to the surface tension, γ. The low surface tension certainly helps, but if γ changes from 40 (a "bad" surfactant) to 20mN/m (a "very good" surfactant) it's only halved the energy needed, which isn't all that significant. Consider low γ as necessary (after all, pure water can't form a foam) but not sufficient. So what things are required?

1. **Elasticity**. The first reason surfactants help create foams is that the surface becomes elastic. This means that the bubbles can withstand being bumped, squeezed and deformed. A pure water surface has no such elasticity and the bubbles break quickly. It also means that those systems which produce more elasticity (see the Elasticity section) will, other things being equal, produce more stable foams. As discussed in the Rheology section, in general a wall which is both stiff and elastic provides a foam with a greater ability to resist a pushing force and therefore a higher yield stress. Smaller bubbles also give a higher yield stress

2. **Disjoining pressure**. The second reason that surfactants help create foam is that the liquid in the foam walls is naturally sucked out of the walls into the edges. This is nothing to do with drainage (as explained in Drainage, the walls contain an irrelevant fraction of the liquid), it is just simple capillarity. The capillary pressure will keep pulling liquid out unless a counter pressure ("disjoing pressure") acts against it. This can be produced by charges on the surfactant either side of the wall and/or by steric interactions between surfactant chains. These effects are discussed in DLVO, but because the charge effect operates over large distances (50nm) compared to the small distances (5nm) of steric effects, in general ionic surfactants are much better at creating stable foams.

3. **Resistance to ripening**. The Ostwald ripening effect means that small bubbles shrink and large ones grow. As the Ostwald section shows, this is partly controlled by the gas ($CO_2$ falls apart quickly, air/$N_2$ is slower and $C_2F_6$ much slower) but also by how good a barrier to gas diffusion the "wall" of surfactant at the surface provides.

4. **Resistance to drainage**. The more water around the foam the less risk (in general) of it becoming damaged. So a foam that drains quickly is more likely to become damaged. As we will see, to resist drainage you need high viscosity and small bubbles, though the surfactant wall has some effect on the drainage process with stiffer walls giving (usually) slower drainage.

5. **Resistance to defects**. If oil or a hydrophobic particle can penetrate the foam wall it can cause the wall (and therefore the foam) to break. Although there are plausible and simple theories (discussed in AntiFoams) of Entry, Bridging and Spreading coefficients they turn out to be of limited predictive value. Once again they are necessary but not sufficient. The key issue is the Entry Barrier. When this is high the foam is resistant to defects.

These principles are so easy, yet creating foams efficiently is surprisingly hard. Why? The key issue is timescales. If a surfactant is marvellously elastic and has a strong disjoining pressure and is a good gas barrier and has a high entry

barrier it might (and usually does) fail to form a foam because it takes too long to reach the liquid/air interface and form its strong resistant domain so the foam has already collapsed. On the other hand, a surfactant that quickly reaches the surface to create an adequate elasticity and disjoing pressure will produce large volumes of foam - though the foam will collapse quickly, especially in the presence of oily impurities such as grease being washed from one's hands.

This leads us to the issue of Dynamic Surface Tensions. It would be wonderful to provide an app that fully described the complexities of DST and which therefore allowed you to produce a mixture with very rapid decrease of ST to give the fastest possible foaming behaviour. But my reading of the literature is that it is quicker to measure the DST behaviour using (most usually) a Maximum Bubble Pressure device (which creates bubbles over different timescales and therefore gives the surface tension at each of those timescales) than it is to attempt to describe the behaviour via theories. In particular, there are great debates about whether DST is limited by diffusion, by barrier entry and/or via the need to come out of a micelle before entering the interface. My reading of the excellent review by Eastoe[51] is that simple diffusion dominates and that the existence of micelles largely makes no difference because the timescale for a surfactant molecule to partition from the micelle is very fast even though the timescale for micelle formation/collapse is very slow. Of course one can find real cases of entry barriers and real cases of micelle-limited diffusion. But it is even more complicated. An extensive analysis from U. Sofia shows that there are 4 possible outcomes in systems containing micelles, two of which are indistinguishable (to the casual observer) from simple diffusion kinetics and two of which might be confused with barrier kinetics. Finally, distinguishing entry-barrier and micellar effects from the effects of small amounts of impurities in the surfactants is surprisingly difficult and for the practical formulator using commercial, unpurified surfactants there is little hope of understanding the subtleties of DST curves. The take-home message is "Don't formulate foams without measuring DST, but don't spend too much time theorising about why you get great results for some specific surfactant combination." I don't like writing such advice as I usually find that good models are the best way to avoid lots of lab experiments. However, the 2020 review paper, discussed below, contains a master-class on the relevant theory and concludes "The theory doesn't really help - just measure the DSTs" .

The harsh reality is that successful foaming agents tend to be mixtures, with all the complexities they induce. The ubiquitous SLES/CAPB (Sodium Laureth Sulfate/CocoAmidoPropyl Betaine) mixture happens to be made from two excellent fast foamers. The CAPB on its own produces a lot of stable foam, but is rather expensive. CAPB is especially good at creating a high entry barrier so is resistant to oils during the creation of foam. SLES on its own produces a lot

51      J. Eastoe, J.S. Dalton, *Dynamic surface tension and adsorption mechanisms of surfactants at the air/water interface*, Advances in Colloid and Interface Science, 85, 2000, 103-144

of relatively unstable foam. A mix of the two provides a good balance of cost, foam and stability. However, adding a small % of lauric or myristic acid has a dramatic effect on foam stability. It increases elasticity but also slows down bubble growth (Ostwald ripening) dramatically, so the foam remains small. This has a big impact on the ability of water to drain from the foam - drainage speed goes as Diameter² - and the drier the foam the more easy (other things being equal) it is to break it apart. The long-chain acids on their own are useless as foaming agents (and as sodium salts are of modest foaming ability as common soap, easily wrecked by hard water). The combination of SLES/CAPB/Long-chainAcid is a potent mix for creating a foam with small bubbles and a long life-time. Indeed, a simple way to transform a hand-soap to a shaving foam is to add a few % of the long-chain acid.

### 5.8.2  But what about my surfactant system?

The rules for creating a good, stable foam (or, indeed, the rules for making sure that such a foam is not created) are simple and clear. So why is it so hard to create new foam formulations? The answer is that if you have the right set-up to measure all the basics: CMC, $\Gamma m$, disjoining pressure v film thickness, interfacial elasticity and entry barrier then it's rather straightforward to make the best out of any set of surfactants and foam boosters you happen to want to use. The measurements can largely be automated so lots of formulation mixes can be screened quickly. One problem, as mentioned above, is timescales. Most measurements are made after comparatively long times so it needs extra time-dependent experiments to see if the appropriate parts of a surfactant blend will get to the surface fast enough to create a foam which then becomes stabilised as the slower components arrive to form a tougher surfactant layer. The other problem is that small additions of co-surfactants, foam boosters etc. can make a large difference, so it is necessary to carry out measurements on large numbers of samples. A robotic lab set up to do lots of high-throughput screening can do a lot of the hard work, but most of use don't have access to such a lab.

In the longer term, a theory that could predict the interfacial behaviour of mixtures of ingredients would make development of foam much more rational. But such a theory seems to be a long way off.

### 5.8.3  The view from 2020

I wrote the earlier text in 2014-15 and had no reason to update it till 2020. To my surprise, what I wrote has stood the test of time. I've not changed much. But a masterly review[52], backed by a serious amount of experiment and theory, allows

52    B. Petkova, S. Tcholakova, M. Chenkova, K. Golemanov, N. Denkov, D. Thorley, S. Stoyanov, *Foamability of aqueous solutions: Role of surfactant type and concentration*, Advances in Colloid and Interface Science 276 (2020) 102084

us to be a bit more specific. Again it is the team at Sofia, led by Prof Tcholakova, who have clarified the situation with five key points.

Although both non-ionics and ionics can produce excellent foaming, the non-ionics need to be above 95% of the full surface coverage of the interface (with a Gibbs Elasticity over 150 mN/m) before they will foam well - it's a sort of all or nothing. Ionics can start producing credible foam at 30% of their surface coverage (even with Gibbs Elasticity of just 50 mN/m), with a with a steady increase in production as you head to 100%. The reason is clear: steric stabilization of the foam interface works well, but only when there is near-full coverage; the interface can break easily if there is even a 5% gap in coverage. Charge stabilized ionics are much more forgiving.

The speed at which the surfactants generate the surface coverage is critical. Basically, if they get to the interface in a few 10s of ms, you'll easily get lots of good foam. This speed depends on concentration, CMC, surface mobility, salt concentratioin in no way that is readily extractable with 2020's theory/experiment (for some hints of the complexity, see DST-Choice, and read the master class on the theory within the paper, which concludes that it's not much help). This is sad in one way, but liberating in another. Just measure the dynamic surface tension at a 10ms timescale and tweak the formulation till you find a large reduction in surface tension. On a typical Maximum Bubble Pressure Tensiometer this 10ms timescale is measured at ~300ms (there's a fixed factor for any given MBPT device) because the real age of a 300ms bubble (it's expanding all the time) is only 10ms. The tradition from the Sofia school is to call the measured time (e.g. 300ms) $\tau_{age}$ and the scientific time (e.g. 10ms) $\tau_u$ for universal.

The foam at shorter timescales (in this paper, 10 shakes of their measuring cylinder) is not necessarily a reliable guide to the foaming after longer timescales (100 shakes). The faster-acting surfactants, not surprisingly, give more foam at short timescales, but the slower ones can catch up. As discussed in the next point, the foams tend to be self-limiting, so an initial advantage doesn't necessarily lead to a long-term advantage. Of course, for applications such as personal care, fast foaming is a requirement so this difference in performance is important. The point is that one has to be careful to distinguish different types of limiting factors.

This is only hinted at in the paper, but is linked to other Sofia work, with more published results promised. The amount and stability of foam gets limited by its own production method. To make more foam you generally need lots of smaller bubbles. These are created by whatever forces are able to trap air and squash, or shear bubbles so they get smaller. As the foam gets richer with smaller bubbles, it gets more viscous (depending on 1/Radius, see Foam Rheology), so at some point the forces are not large enough to deform the bubbles to something smaller. The effect depends somewhat on the rigidity of the interface and, therefore on the surfactant, but it's mostly dominated by the ability to create

the fine foam in the first place, i.e. the interfacial stability and speed of reaching it. This is why plenty of surfactants can produce similar amounts of foam as long as they are present at sufficient concentration to meet the previous two requirements. Looking back at plenty of other foam papers I see that there is lots of confusion of cause and effect because like wasn't being compared to like. And because there is (rightly) a separate focus on the foam stability, for which we have the other apps on this site.

The team deliberately used "as is" surfactants because their impurities show up rather interestingly in the data. Measurements of % surface coverage come, of course, from adsorption isotherms CMC and $\Gamma$ and these often show strange behaviour because of low levels of other components. This generally doesn't bother us, the surfactants are what they are, but they certainly complicate academic analyses when it's necessary to know, for example, whether you have 50% or 60% of surface coverage.

### 5.8.4  Foaming techniques

I had generally paid little attention to the different foaming techniques, but the remark in the previous section about foams being self-limiting made me realise that I've come across quite a few different methods.

1. **Shaking cylinder**. Put, say, 10ml of solution into a 130ml measuring cylinder and oscillate it, checking the volume of foam after a given number of shakes. If you get 90% trapped air then you are at 100ml, so finding whether you have 91, 92 ... gets tricky in a 130ml cylinder. My impression is that this sort of foam is relatively coarse, but I might be wrong
2. **Ross-Miles**. Put some test solution in the bottom of a tall cylinder. Now dropwise add more of the solution from the top. The drops smashing into the liquid below produce a foam. Measure the volume at the end of the addition, then, for stability, the volume after a few minutes. Amazingly, this is an industry standard test.
3. **Blender**. Just get a big blender and put in enough liquid to cover the blades. Whizz away and measure the volume by pouring the contents into a measuring cylinder. The fact that this can be done suggests that the foam is rather coarse, because a fine foam would be hard to pour..
4. **Planetary mixer**. Take you Kenwood Chef or equivalent with a wire whisk and watch what happens as the whisk turns on its axis while moving around on the other axis. A paper from the Sofia group shows a clear self-limiting effect once the foam gets thick enough to squash the surface waves which initially trapped the air, so this seems good for testing for the ability to create finer foams.
5. **Sparging column**. Blow air through a frit at the bottom of a column containing your foaming solution. You get some idea of the foamability and stability from the stable height of the foam, and/or you can measure the

weight of foam coming over the top in a given time. More details are available on the Foam Fractionation page.

6. **Micro-foam test**. I once had to measure foamability using mg of surfactant and µl of solution. This was remarkably easy to do with a steady stream of air blowing through a very fine syringe needle into the solutions in micro-titre plates. It's a very good high throughput technique (which is why we developed it) to distinguish low, medium and high foamers and short, medium and long-life foam. It's crude but amazingly effective.

7. **Compressed air foam**. Mix your surfactant solution with some high-pressure air, let it travel down a pipe, expanding as it goes, and burst out onto, say, an oil storage vessel in flames. I was once involved in a fire-fighting project that required the theory of such a foam and needed some measurements to parameterise the theory. Unfortunately the live experiments on a full-sized test rig failed because the rig burned down during one of the tests...

8. **Aerosol foams**. This is a variant of the previous one, on a smaller scale. The propellant in a can (typically a hydrocarbon gas blend) is beautifully mixed into the surfactant mixture so creates a mass of fine bubbles when it suddenly expands. A typical example is a shaving foam which has to be fine in order to have the high viscosity and yield stress to stay on the face.

9. **Hand rubbing**. I know that foaming has no significance in terms of washing - the craving for it is psychological, not physics. So I'd never bothered to see how much foam one could create with imaginative hand rubbing. It's quite a lot, but in my view not worth the effort.

10. **Shaving brush**. I had never understood shaving brushes. They didn't produce an interesting amount of foam and just seemed a complicated way of spreading soap over my face. But then I'd never bothered to learn how to do it. If you whisk away onto a blob of wet soap on one's hand, nothing much seems to be happening. That's because all the foam is in the brush. Just squeeze the brush in any way, and out comes a mass of very fine, stable foam, perfect for placing on the face. I was very impressed.

11. **Foaming net**. Take a few cm of a fine net and rub it hard between your hands with the wet soap. As with the shaving brush, nothing much happens if you don't know what you are trying to do - I had to go to YouTube to find out. If you pull the net between your fingers, a large amount of foam emerges. Repeat this a few times and you get an awesome amount of fine, stable foam. The fine net is clearly good at breaking up larger bubbles into smaller ones. Why anyone bothers to spend their time creating this mass of foam bubbles is not a question I am qualified to answer.

### 5.8.5 Measurements of key parameters.

I know of 3 common measurement techniques:

1. Obviously foam height, where appropriate, and the ratio of the total height to the amount of liquid in the bottom of the container, and how this changes over time.

2. A conductivity meter across a known gap, calibrated with the conductivity of the water used in the experiment, gives you a good idea of the volume fraction of air.
3. Put a large prism in contact with the foam and couple light into and out of it. A video shows a strong contrast between contact with water (white) and air (black) and it is then easy to use image analysis to measure the foam. Experiments have shown that the prism has a surprisingly small perturbation on the foam itself so the measurements are relevant. It's incredibly hard to get good image analysis from images of free foam because there's seldom reliable good contrast between walls and the rest.

## 5.9 Foam fractionation

This topic is important for those who enjoy a glass or two of champagne. Hopefully you know to use pre-nucleated, deep champagne glasses so that relatively few bubbles are produced at the lowest point of the champagne. These bubbles rise to the surface and explode. Foam fractionation is important because the aroma chemicals in the champagne are specifically attracted to the air/water interface and the bubble explosion sends them into your olfactory system where you can appreciate the subtlety of the champagne maker's art[53]. The reason for the pre-nucleation and deep glasses should be clear. It takes time to build up a concentration of aroma chemicals at the surface and the relatively few bubbles rising through the maximum distance provide the biggest chance of extracting the aroma chemicals.

For those who prefer a pint of Guinness beer then a health tip is that the protein content of the beer is 0.3% but in the foam it is concentrated to 75%, a factor of 240 in enrichment.

For those who appreciate the natural antimicrobial behaviour of nisin within their foods, the manufacturing process requires concentration of this polypeptide from within a dilute fermentation broth. The foam created within the broth attracts the nisin to the interface and once the foam is separated (and burst with an antifoam) the nisin has been concentrated by a factor of 10 or more, not bad for (on paper) such a simple process. This example is taken from the foam fractionation book of Stevenson and Li[54] and the theory here strikes a balance between the need for good theoretical models and the reality that these processes are far more complex than any model can hope to accommodate.

We already have the basic intellectual tools. The movement of the materials to the interface is governed by the sort of DST and isotherm equations discussed before. In particular the Ward-Tordai equation used in the DST Choice app is

53  Liger-Belair G,Cilindre C, et al., Unravelling different chemical fingerprints between a champagne wine and its aerosols. Proc Natl Acad Sci USA 106: 16545–16549 2009

54  X. Li and P.Stevenson, Foam Fractionation, CRC Press, 2014

the right tool – though it needs some modifications for the spherical geometry of bubbles. We also know about foam drainage which is vital for ensuring that we have the largest fraction of foam, with its walls covered by the desired substance, with as much as possible of the water removed from the process as it is now contains (relatively) little of the desired substance. However, even full application of these models is too complex, and unnecessary, for a good basic model.
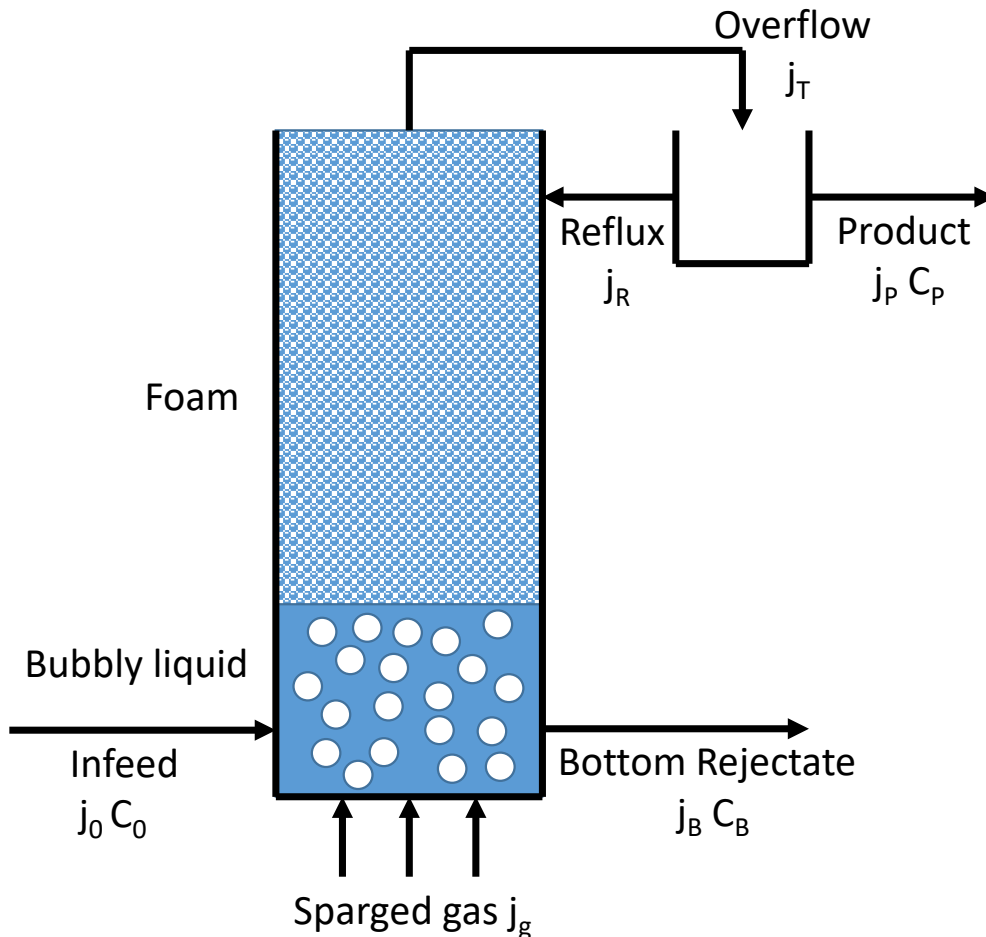


**Figure 5-8 The basics of foam fractionation (with optional reflux) after Stevenson & Li.**

As per the diagram (based on Fig 6.3 in the book) we assume that air is being bubbled into a column of liquid and the bubbles of radius r (Sauter mean for those who attend to such details) rise with a superficial velocity $j_g$. [The "real" velocity viewed standing outside the tower is $j_g/\Phi$ where $\Phi$ is the volumetric gas fraction in the bubbles. So when $\varphi$ is 0.1 the real velocity is $10j_g$.] Liquid with surface agent (e.g. nisin) at concentration $C_0$ is being pumped in at a rate $j_0$ and the depleted liquid at concentration $C_B$ is removed from the bottom at a rate $j_B$ and the product at concentration $C_P$ is removed from the top (i.e. in the foam) at a rate $j_P$. The *surface* flux $j_S$ is calculated from $j_P$ and r via $j_S=3j_g/r$.

The column of foam has an essentially constant liquid volume fraction $\varepsilon$ which can be calculated by numerically solving the Stevenson and Li drainage

equation shown earlier where $\rho$ and $\mu$ are the liquid density and viscosity, and g is gravity:

$$\mu \frac{\dot{j}_g}{0.032\rho gr^2} = \varepsilon(1-\varepsilon)^2 \qquad \text{5-21}$$

If $\varepsilon > 30\%$ the "foam" calculations start to be meaningless so a warning red appears in the app and the other calculated values can probably be ignored. The liquid flux in the foam $j_P$ (this is the product removal rate because it is this flux which leaves the column) can then be calculated via:

$$\dot{j}_P = \frac{\varepsilon \dot{j}_g}{1-\varepsilon} - \frac{\rho gr^2 m\varepsilon^n}{\mu} \qquad \text{5-22}$$

We can calculate the concentration in the bubble via the isotherm. If the concentration of the depleted liquid at the bottom of the column is $C_B$ then the concentration in the product is given by:

$$C_P = C_B + \Gamma_m \frac{KC_B}{1+KC_B} \frac{\dot{j}_S}{\dot{j}_P} \qquad \text{5-23}$$

Two more simple "balance" equations finish the job. Because $j_0 = j_P + j_B$ i.e. the input flow is the same as the sum of the product flow and the bottom flow, and because we know the first two, $j_B$ can be calculated. And $j_0 C_0 = j_P C_P + j_B C_B$ i.e. the total amount of the surface active agent going in is the same as that going out. By substituting $C_P$ in the equation above we can solve the equation for the one remaining unknown, $C_B$.

To complete the story we need two more values. In a real process we want the maximum enrichment, E, combined with the maximum recovery, R and the maximum throughput $j_0$. As with any other similar process such as distillation or solvent extraction there is always a compromise. It is generally easy to get reasonable E and R values at low throughput. At reasonable throughputs there is a trade-off between E and R and at high throughputs E is low. The new parameters are defined as:

$$E = \frac{C_P}{C_0} \qquad \text{5-24}$$

$$R = \frac{\dot{j}_P C_P}{\dot{j}_0 C_0} \qquad \text{5-25}$$

A common trick in all such processes is to "reflux", to send some fraction $\chi$ of the product stream back into the system. This returned material needs to equilibrate

with the stream of material rising through the column and it is common to define a number of "theoretical plates" or "theoretical units" in the system which depend on many factors. It is too complex to describe the theory here (please refer to the book), but you can choose a value of χ and an NTU (Number of Theoretical Units) between 1 and 4. Of course if χ=0, with no reflux, then you get the values calculated via the above procedure. Indeed, these values are the starting point for the iterative solver and are provided as useful reference outputs. When χ is non-zero then $j_P$ is replaced internally by $j_T$ for the flow from the top of the foam column and we need $j_R = \chi j_T$ for the reflux flow. Obviously our product flow is now given by $j_P = (1-\chi)j_T$. Because the effect of χ is described by the graph, neither $j_T$ nor $j_R$ is provided as an output, but can readily be derived manually for any specific χ.

At present it does not seem possible to know how many NTU one might have in a column so it is wise to assume at most 2, given that other assumptions such as perfect mixing and equilibration are unlikely to hold, making the idealised calculation over-optimistic. You can go up to 4 if you wish, but the numerics are more precarious. Although responsibility for numerical errors is mine, I warmly thank Dr Li for providing a fast and elegant numerical solver when it became clear that my own implementation was neither fast nor elegant.



**Foam Fractionation**

| Bubble r μm | 500 | C₀ mM | 0.05 | Viscosity μ cP | 1 | Density ρ g/cc | 1 |

| $j_0$ mm/s | 2 | $j_g$ mm/s | 5 | $\Gamma_m$ (μmole/m²) | 8 | K 1/M | 473000 |

| N units | 2 |

| Cₚ mM | 0.468 | $j_P$ mm/s | 0.185 | C_B mM | 0.0068 | $j_B$ mm/s | 1.815 |

| ε | 7.44% | $j_S$ /s | 30.0 | Enrichment | 9.4 | Recovery | 86.5% |

**App 5-7** https://www.stevenabbott.co.uk/practical-surfactants/Foam-Fractionation.php

The reflux idea is sound in theory but has the practical problem of having to break the foam and recycle part of it back into the foam column. Breaking the

foam can be harder than one might like, as is the provision of the pumps and pipework needed to redistribute the recycled material over the whole width of the column. An alternative is to provide "internal reflux", in other words to partially burst the foam near the top so that some of the liquid naturally trickles back.
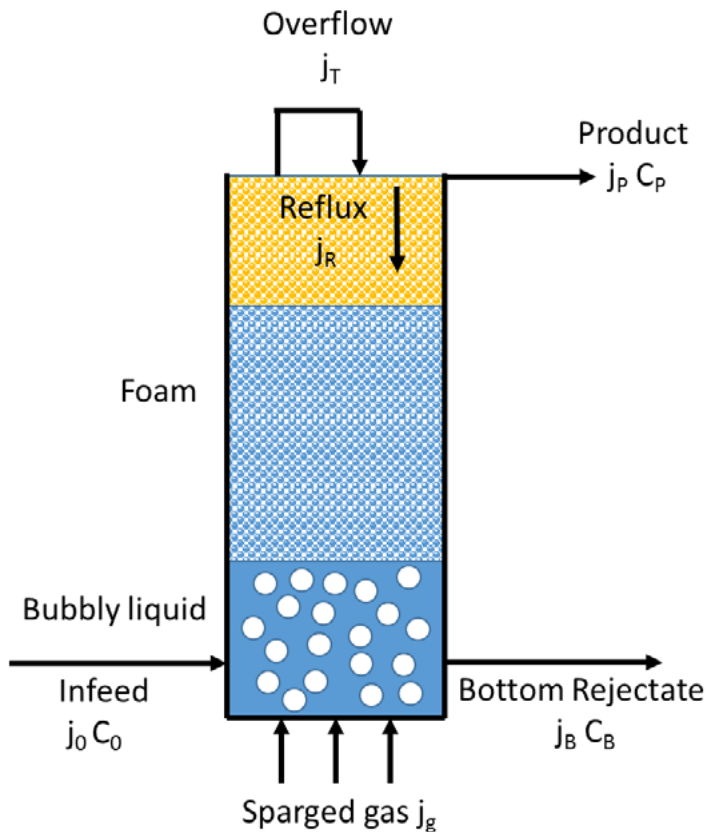


**Figure 5-9 Internal reflux shown schematically.**

Stevenson and Li have shown that the intellectually most satisfying way to burst some of the foam is to pass dry air over the surface. This causes rapid evaporation at the top of the bubble which sets up some Marangoni flows that upset the bubble structure and cause it to burst. But the required volumes of dry air might be too expensive for a process which is intrinsically low cost. A rotating blade can generally be set up to destroy some of the foam.

At the time of writing, foam fractionation is not wildly popular as a production technique. Part of the problem is that it is such a complex process, and intuition is such an unreliable guide that those who have tried it have become discouraged. There is also a tendency to over-complicate what is essentially a simple process, and the theoretical gains in enhancement and recovery can be offset by the practical needs for extra downtime to take apart, clean and re-assemble the more complex equipment.

In terms of this book, it is a fascinating bringing together of a number of the technical themes discussed earlier. And if it can help enhance your appreciation of champagne then that is a pleasant bonus.

# 6      Problems

We now have a powerful set of practical tools that can help us navigate through many parts of surfactant space. For example, the intelligent use of HLD-NAC or the equivalent PP or Helfrich torque approaches can provide much more insight than the classic, and deeply flawed HLB, Winsor R or Bancroft ideas. Although such tools are popularly supposed to be applicable only to microemulsions we have seen that they are key to understanding many aspects of macroemulsions via their influence on interfacial tension and through the updated Wedge theory. We can be proud about what the surfactant science community has achieved, even if we can be distressed that so much time has been wasted on HLB.

A formulation team armed with the modern tools can solve many issues much faster. But such teams everywhere are now hitting a key issue that can conveniently be called the "polar oil" problem.

## 6.1      The polar oil problem

Let us take some decane, some water and a convenient surfactant in the right region of HLD space so we know where we are. By tweaking salinity and temperature we can control whether we have Type I, Type II or Type III behaviour.

Now add some octanol. What happens? Embarrassingly, the answer is that no one really knows. Why is this?

Well, what is octanol? There are at least three answers.

- It is an "oil" because it feels oily and is hardly miscible with water.

- It is an alcohol and we know in general what the effects are of adding ethanol to a surfactant system – it makes the water slightly less watery so higher Cc surfactants are less repelled by the aqueous phase and the shift is to a lower HLD value.

- It is a surfactant because it has a polar head and a non-polar tail. It happens to be not a very good surfactant, but it is still a surfactant.

Some early versions of HLD had an extra F(A) term which was a function of the % alcohol. Some early experiments indicated that this was not a bad idea, and even pointed out that s-butanol was a "neutral" alcohol that could do useful things such as speed up the kinetics of emulsion equilibration (e.g. for measuring Cc or EACN values) without affecting HLD. But even s-butanol has been shown to have an "alcohol" effect in some surfactant systems. So there is no coherent way within HLD (or Helfrich torque) to deal with these polar oils and the F(A) term in HLD has largely been abandoned.

An excellent set of experiments[55] by Tchakalova's team at Firmenich has provided the data and a model that allows the complications of polar oils to be explored in detail.

The model is called CIT (Constant Interface Thickness) because a key assumption is made which says that any oils and polar oils that come into the interfacial layer act not to change the *thickness* but to change the *curvature*. This is undoubtedly a simplification, but it makes it possible to calculate the effects of each component in the complex balance of interactions.

The model first assumes that there are four key regions where an extra molecular species (additive) might reside. In the paper these are fragrance molecules such as eugenol but they can also be "linkers", "co-surfactants" or, indeed, "hydrotropes" in the meaning of molecules which alter the behaviour of the interfacial layer of a surfactant.

- O - Oil
- W - Water
- C - Micellar core
- I - Interface

Next the model assumes that for any given set of conditions one can calculate a packing parameter, PP based on three variables:

- V, the *effective* volume of the surfactant tail
- I, the *effective* length of that tail (typically 80% of the extended length of an alkyl chain)
- A, the *effective* area of the surfactant head.

TThe PP is defined, just as the earlier, discredited, CPP as:

$$PP = \frac{V}{A.I} \qquad \text{6-1}$$

Note that the word "effective" is included in the definition of each parameter. A stand-alone surfactant molecule has a V, I and A that can all be estimated from molecular models and from which the CPP can be calculated. But that is irrelevant to a real surfactant system where the V depends on how much (if any) oil and additive creep into the tail region, where I is averaged over a range of conformations and A depends on salinity (for ionics, where salts decrease head repulsion), temperature (for ethoxylates) and the extent to which oil and additive are inserted into the interfacial region.

---

55   Vera Tchakalova, et.al. Solubilization and interfacial curvature in microemulsions I. Interfacial expansion and co-extraction of oil, Colloids and Surfaces A: Physicochem. Eng. Aspects 331 (2008) 31–39

To implement the model, the two key assumptions are:

- All the surfactant is at the interface;

- The interfacial layer thickness is constant (hence CIT).

This model is distinct from the more common "wedge" model which assumes that there is no oil in the tail region and that any additive goes only into the tail, affecting V but leaving A unchanged. There are, of course, circumstances where this applies and in the app you can adjust sliders to emulate this mode (set τ and α to zero as discussed below). But CIT is surely more powerful because it is more general.

The master equations is:

$$PP = \frac{V_s + (\tau + \alpha\lambda)V_o + \lambda V_a}{l.\left(A_s + (\tau + \alpha\lambda)A_o + \lambda A_a\right)} \qquad \text{6-2}$$

What does this mean? The top line of the equation shows the intrinsic volumes of the surfactant tail, Vs, the oil Vo and additive Va and they are combined depending on:

- τ, the mole ratio, oil/surfactant in the interfacial region

- α, the mole ratio, oil/additive, of oil molecules attracted to the *interface* by the additive

- λ, the mole ratio, additive/surfactant.

So $\tau V_o$ makes sense as this is the amount of volume added by the oil to the tail and $\lambda V_a$ makes sense as this is the volume added to the tail by the additive. The $\alpha\lambda V_o$ is the extra oil (if any) brought into the tail by the additive.

The bottom line is similar except that we have Areas instead of Volumes and the whole thing is multiplied by the tail length, l.

The Radius of curvature, R, can be calculated from PP knowing that:

$$PP = 1 - \frac{l}{R} + \frac{l^2}{3R^2} \qquad \text{6-3}$$

So now all that is required is to enter the key values for each of the factors. These values can be measured experimentally from macroscopic measurements (these are difficult measurements – so here we are more concerned about the theory than the practicality), but for the app it is easier to enter them directly. The default values when you load the app for the first time are those for eugenol given in the paper.

**App 6-1** https://www.stevenabbott.co.uk/practical-surfactants/Curvature.php
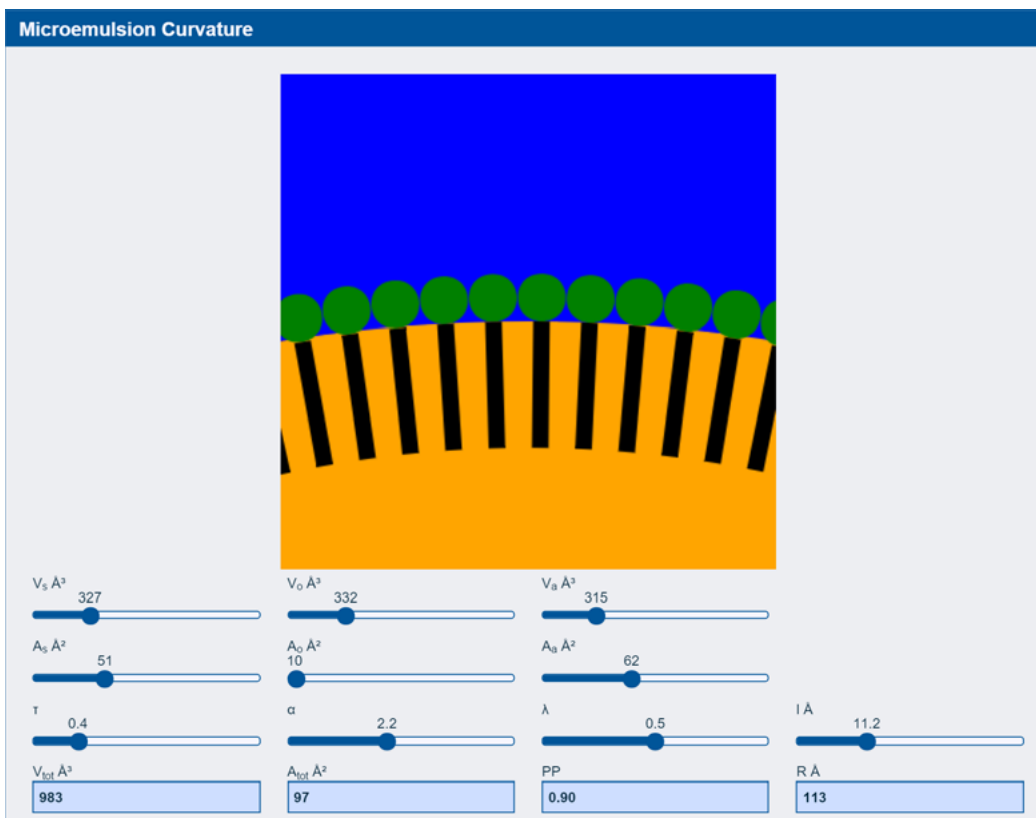
What is the app telling us? The outputs are the total tail volume, $V_{tot}$, the total head area, $A_{tot}$, the packing parameter, PP and radius of curvature, R, where a negative R means curvature in the W/O direction. When R is very large it is artificially limited to avoid the complex issues near zero curvature. The graphic shows the head and tail regions changing shape as the different sliders are adjusted and shows the overall curvature that results. The graphic is illustrative only - a 2D representation of a 3D phenomenon.

One aspect of the theory at first puzzled me - the effect of the additive is symmetrical on V and A, yet our intuition is that a hydrophobic additive should relatively increase $V_{tot}$ and a hydrophilic additive should relatively increase $A_{tot}$. But of course, the balance of $V_a$ and $A_a$ shifts between the two cases, in that $V_a$ will be larger than $A_a$ for hydrophobic additives and $A_a$ will be larger for hydrophilic additives. In addition, α will be small for the hydrophilic case, giving only a small amount of extra oil (which has a large V/A ratio) in the interface. Therefore not only are our intuitions confirmed but the theory gives a clear way to think about the multiple effects involved in a shift from a hydrophobic to hydrophilic additive.

## 6.2     Building on the CIT model

Now we have the CIT model, maybe we can all just get on and measure all those V and A values and start optimising our more complex formulations using this rational tool. Unfortunately it requires massive efforts to measure even scientifically pure systems of well-defined components. Most of us will never do

this. While we can hope that some method for calculating α, τ and λ will emerge, at present we can only make informed guesses.

So we have to be pragmatic and use simpler tools that are already available to us.

An example is again provided by Tchakalova[56] who set up surfactant and oil systems near HLD=0 which allowed the effect of (in their case) a fragrance oil to be studied in terms of its effect on curvature. Using an ethoxylate they could accurately determine the phase inversion temperature and then see whether the added molecule shifted it to a higher or lower temperature. Because a higher temperature in this case implies, from the HLD equation, a higher EACN or a lower Cc it means that at a given temperature the molecule is increasing curvature towards the Type II domain, favouring the tail region, though of course it is not known from this experiment how the various V and A factors are interplaying. A lower temperature implies a move in curvature to the Type I domain.



**App 6-2** https://www.stevenabbott.co.uk/practical-surfactants/Measure-EACN.php

The app puts these ideas into practice. The example shown starts with heptane as the oil ($EACN_{ref}$=7) and the PIT when 6.5% of the test oil is added is 38° which corresponds to an $EACN_{mix}$ of 6.4. Simple arithmetic then tells us that the $EACN_{oil}$ must be 1.8.

---

56   Vera Tchakalova and Wolfgang Fieberm, Classification of Fragrances and Fragrance Mixtures Based on Interfacial Solubilization, J Surfact. Deterg. (2012) 15:167–177

There is a temptation to say that the molecule which raises the EACN is more hydrophobic and one which lowers it is more hydrophilic. This might well be true, but it can be misleading. In EACN terms a smaller molecule tends to have a lower EACN than a larger one, so the trends we see might be more to do with size than any true notion of functional groups that create hydrophilicity. Remember that the single most important parameter for estimating an octanol/water partition coefficient is molar volume – bigger molecules (in general) are simply less soluble in water.

Armed with that caution it then becomes possible to make an informed judgement about how the molecule of interest is behaving in your system and how it might be interacting with the oil, the water, the tail and the head area. This is nowhere near as good as having all the CIT parameters, but given that we only have finite resources for any given formulation problem, it seems to be as good as it currently gets.

Doing the experiments with ethoxylates is especially easy because the phase inversion temperature is easy to measure using a simple temperature scanning apparatus. If your interest were with APGs then it would require a series of test tubes with varying EACNs (e.g. a toluene/hexadecane scan) in order to measure the shift in the HLD=0 point. With ionics it might require some salt scans.

Even these are rather too much hard work and an alternative is to get in the habit of systematically exploring the HLD=0 region with some sort of conductivity or electrical impedance setup as described earlier.

Here, in general, is how to do it assuming, as before, that the water contains at least a small amount of salt so that a continuous aqueous phase is easily identified by a significant conductivity.

We know that in Type I, oil in water, we have a conductivity of the bulk (saline) water reduced somewhat (we might not be able to measure the small difference) by the non-conducting oil drops. We know that in Type II, water in oil, we have no significant conductivity. And at Type III we have some weird phase that should show conductivity significantly different from Type I.

Now get a 50:50 mix of oil and (saline) water with your surfactant at a temperature, EACN, salinity that gets you close to HLD=0. Have the mix stirring in a tube with your conductivity meter and see where you are in HLD space. If there is no conductivity then HLD>0 and you need to reduce it by lowering temperature (if you are using an ethoxylate) or by adding a slight amount of a lower Cc surfactant. If the conductivity is close to that of your original water, then increase the temperature for an ethoxylate or add a little more salt for an ionic or a slight amount of a higher Cc surfactant for an APG.

Whatever your starting point was, if you have managed to flip the phase, see if you can un-flip it. Now we see why academics love ethoxylates. All they have to do is to change the temperature. With salinity we can easily go up but going down again requires more water and more oil to balance it. For those who used a Cc adjustment, un-flipping requires the addition of a high Cc surfactant.

Once you know how to routinely navigate around HLD=0 you can then start adding polar oils to see how they change things. Take two tubes, one that is conductive but very close to flipping to non-conductive and the other which is just in the non-conductive regime. Now add a drop of the polar oil to each tube. If it flips the conductive tube to non-conductive then you know that the polar oil is influencing the tail region; if it flips the non-conductive to conductive then you know that the polar oil is influencing the head region.

The point is that we are not writing an academic paper, we are trying to get a "good enough" way to see the effect. The scenario described here is one example of a "high throughput" mentality for getting lots of useful information for relatively little work. More examples are discussed in the final chapter.

As a community we are not yet able to solve the polar oil problem and although this seriously limits the applicability of HLD, it doesn't mean that we are helpless. With high-powered tools such as CIT experiments we can really get to grips with these difficult issues. And if we have neither the time nor the resources to do full CIT, by using HLD-style scans we can gain lots of insights remarkably easily – once we have invested in "good enough" systems to measure the impact of a polar oil on whichever surfactant system is of interest to us.

## 6.3    Polar oils as surfactants *and* oils

In 2019, a paper from the Acosta group[57] provided a new rational approach to the polar oil problem. While fully acknowledging other approaches, especially the CIT theory, the paper faces up to the fact that these oils really do have surfactant-like and oil-like behaviour.

A polar oil does not fit neatly into HLD-NAC because, by hypothesis, at low concentrations it tends to go to the surfactant interface and influence curvature like a surfactant, and at higher concentrations it acts like a pure oil, albeit one with a very low EACN.

Taking this hypothesis seriously makes it possible to understand how a given polar oil interacts within a given system. The first thing to do is to measure the change of HLD as the % polar oil (with respect to the real oil) is increased. In the

---

57    Amir Ghayour, Edgar Acosta, *Characterizing the oil-like and surfactant-like behavior of polar oils*, Langmuir, xxx, 2019

paper this is measured via changes in S* but the idea is general, and the app shows this curve in the first graph on the left:
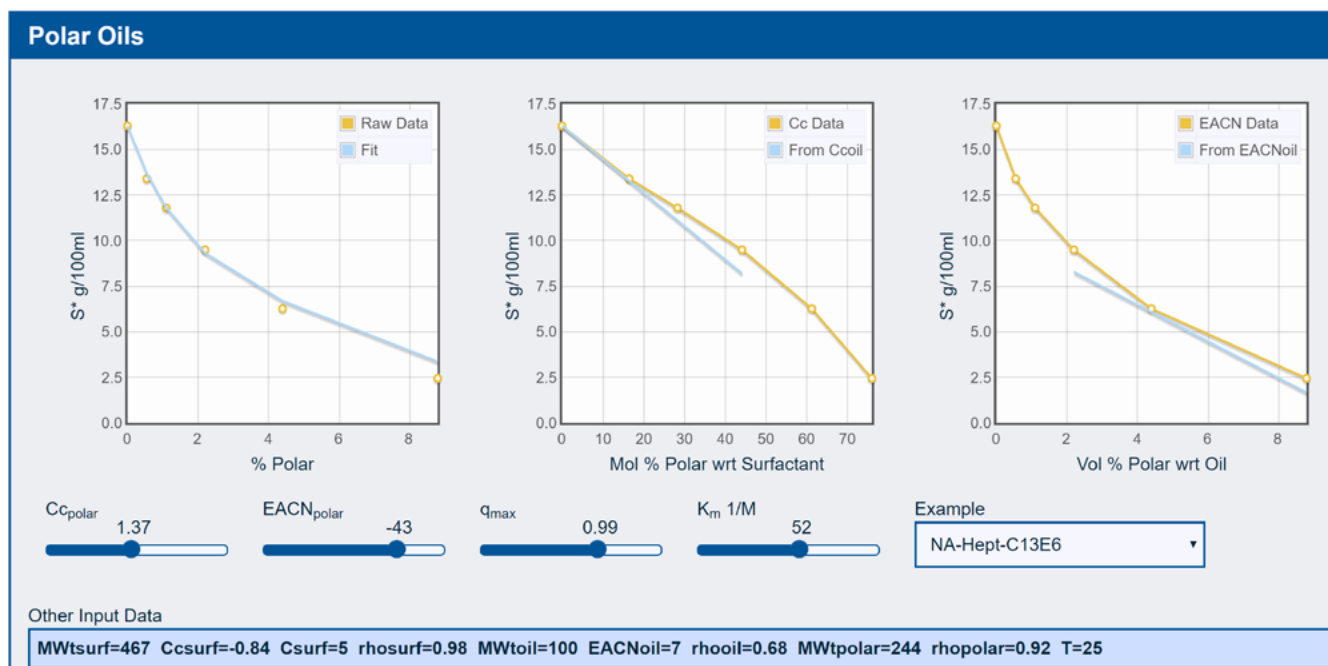
Now assume that the polar oil has migrated 100% to the surfactant at all concentrations and re-plot the S* data according to the mole fraction of the polar oil with respect to the real surfactant. By the Cc mixing rule, given an estimate of the Cc of the polar oil, we know the Cc of each mix so we can predict S* for each fraction of polar oil. If the results are a straight line then the assumption is correct; when it deviates (which it must do) we know that the polar oil isn't fully partitioning. From the slope of the linear part we can determine the Cc of the polar oil or, in the app, we change $Cc_{polar}$ till we get a good fit to the initial points.

Now assume the opposite - that the polar oil partitions completely into the oil. The plot is now of the polar oil as mole fraction of total oil. From the volume mixing rule and an estimate of the EACN of the polar oil, we can estimate the EACN of that mix. The assumption is only valid for high concentrations of polar oil, so the fit is to the final part of the curve.

We now have (via your informal slider fitting) estimates for the effective "pure" Cc and "pure" EACN for the polar oil. To fit the whole dataset we need to know the actual concentrations of polar oil in the surfactant and oil for each of the overall polar oil concentrations. For this we need two more parameters. The first is $q_{max}$ which tells us the saturated ratio of polar oil molecules to surfactant molecules at the interface. The second is $K_m$ where $1/K_m$ is the concentration of polar oil at which the transition takes place from surfactant-like to oil-like. These extra parameters are used to create a Langmuir isotherm, from which we know the concentration of the polar oil in the surfactant and in the oil at any

given concentration. Knowing these concentrations, and applying the respective mixing rules, we can adequately fit the whole curve.

What do these parameters mean? It seems that these oils like to partition to the interface (not surprisingly!) as shown by the fact that the $q_{max}$ values are in the 0.5 to 1 region. Any "polar oil" with a low $q_{max}$ would behave like a conventional, negative EACN oil. To put it another way, a definition of a polar oil (as opposed to an oil with a negative EACN) is one that partitions strongly to the interface. $K_m$ is similar, a very small $K_m$ means that you'd need a very large concentration before you'd see an effect, even if $q_{max}$ was large. Both these statements can be checked by sliding the respective sliders and looking at the predicted S* curve.

Given that they partition, what effects will they have? The app, as published, featured the 9 datasets in the original paper. It is hard to reach many conclusions about the general rules because only two polar oils (dodecanol and naphthenic acid) are being tested, against a small group of ionic and non-ionic surfactants. In addition, interfacial curvature is complex and subtle - small changes in geometry can have large changes on curvature. To see the whole picture we need to look at two different effects:

- The first effect (i.e. at small concentrations of polar oil) depends on where the molecule sits in the interface. For the naphthenic acid it is easy to imagine that the carboxylate head enters the sulfonic acid head area of the ionics, providing more negative curvature - so the Cc value is negative. For the ethoxylates, the partitioning to the interface is similar, but the curvature is in the opposite direction. As it is hard to imagine that the polar head prefers to stay more in the oil region, maybe the head is encouraging the ethoxylates to come together more tightly via some sort of hydrogen bonding around the carboxylic acid group. For the dodecanol the Ccs are positive for the one ionic and one ethoxylate. Here we can imagine that the small -OH head has no good reason to fill up space on the aqueous side, so the tail simply provides more positive curvature.
- The second effect (i.e. at large concentrations) requires us to think why a molecule has a larger or smaller negative EACN value because that is all that is happening. It is well-known that the single best predictor for EACN is the octanol/water partition coefficient logP so as a first approximation we can say that the EACN effect is a partition effect. However, the logP values of ~2.9 and 5.1 for naphthenic acid and dodecanol don't fit into any obvious pattern (we'd expect napthenic acid to be more negative than dodecanol) so this is early evidence for other, more interesting effects.

The implications of this approach are profound. We can now start to see how each polar oil interacts at the interface, working either as a surfactant or an oil. Because the app allows users to provide their own datasets, we can expect that the number of datasets will expand rapidly. With modest amounts of sharing of such datasets maybe we can start to see the bigger picture of which polar oils

interact in which ways with different types of surfactant. At the time of writing we are a long way from that sort of understanding - all I can do is offer the app as a service and hope that the community rises to the challenge.

## 6.4    Linkers, hydrotropes etc.

The surfactant literature creates a lot of confusion talking about things like "hydrophilic linkers" or "hydrotropes" to describe additives which increase the "efficacy" of a surfactant. As with so much of surfactancy, there are at least three confusions here.

- Many of such additives are simply changing the HLD (along the lines of the CIT model) and (more by luck than judgement) increasing the impact of a given amount of surfactant on the system.

- The additives may be changing the *kinetics* rather than the thermodynamics, so appear to give a successful formulation via some "linker" effect when they are simply changing the speed at which the desired phase is reached.

- These molecules can also change the ξ parameter, i.e. the ability of the surfactant to influence the oil and therefore the overall solubility of the system.

Given that these additives are typically polar oils such as octanol, our lack of a deep theory is unfortunate. This means that we have to rely on the CIT-style experiments to sort out the HLD part, adopt some patience to exclude kinetic effects and then find ways to understand the ξ part.

It is not surprising, then, that with three different effects and with most formulators unaware of HLD, the linker/hydrotrope literature is so confusing.

The starting point for discussions is ethanol. This is sometimes added to formulations to increase surfactancy, with explanations that make little sense. Because it is so easy to add, it seems quite a smart formulation option when it works. It is likely that ethanol is partitioning specifically into the head area of the surfactant, increasing its curvature towards a lower HLD. If the starting surfactant Cc was too high then this would fix the problem. However, it generates another problem. A theme discussed many times is that large head areas lead to inefficiency and instability because less surfactant gets packed into the same space. So although ethanol might improve surfactancy by fixing a bad Cc, it might make things worse by increasing the head area. It would have been far better to fix the original problem by tweaking the Cc to increase the proportion of low Cc surfactant in the blend.

2-butanol is a popular formulation additive in academic experiments on surfactants. It often seems to have a neutral effect on HLD and is added for

kinetic convenience. A good example is experiments with extended surfactants. These can be super-high efficiency with high ξ and high interfacial rigidity, ideal properties once the emulsion is formed. But equilibration can take weeks or months, whereas 1-2% 2-butanol gives equilibrium in hours. The undoubted practical benefit of the 2-butanol is offset by the increase in head area and therefore a decrease in surfactant efficiency.

The smaller alcohols were discussed because they are frequently used and cause confusion through lack of a language for thinking through cause and effect about modes of action. There is no doubt that higher alcohols such as octanol or decanol can be good general purpose improvers of surfactant efficiency, and here the word "linker" becomes appropriate. The idea was first explored systematically by the Salager team. The discussion here is based on the excellent review[58] by the Oklahoma team of Sabbatini, Acosta and Harwell which fully acknowledges the Salager contribution.

As we have seen, the simplest way to increase the influence of a surfactant over an oil (especially a large alkane such as hexadecane or a typical cosmetic oil) is to increase the tail length of the surfactant. This hits a limit of surfactant solubility and the appearance of horrid liquid crystal phases so cannot be implemented in practice. So why not provide a *virtual* increase in chain length by adding a molecule which partitions into the surfactant tail zone, opening it out and, perhaps, extending beyond it, therefore providing more room to influence the oil? This is a hand-waving way of speaking about the α parameter in the CIT model, the amount of oil brought into the tail by the polar oil. Because the additive is in some way linking the oil to the surfactant, it is called a linker and because it is partitioning into the tail it is called a hydrophobic linker.

For this trick to be effective, the polar oil should either have no net effect on the curvature (so its interaction with the head should be significant) or, better, should take a formulation that is designed to have an HLD value that is slightly too low so that the linker+oil provide an optimum curvature for the desired phase.

Sometimes it is not possible to pack enough linker in to increase the surfactancy to the desired level. Then a final trick can be used. An additive that partitions into the head region, though with a substantial fraction inside the early part of the tail, can help open up the surfactant to allow the hydrophobic linker to act further. Such an additive is called a hydrophilic linker.

58   David A. Sabatinia, Edgar Acosta, Jeffrey H. Harwell, Linker molecules in surfactant mixtures, Current Opinion in Colloid and Interface Science 8 (2003) 316–326
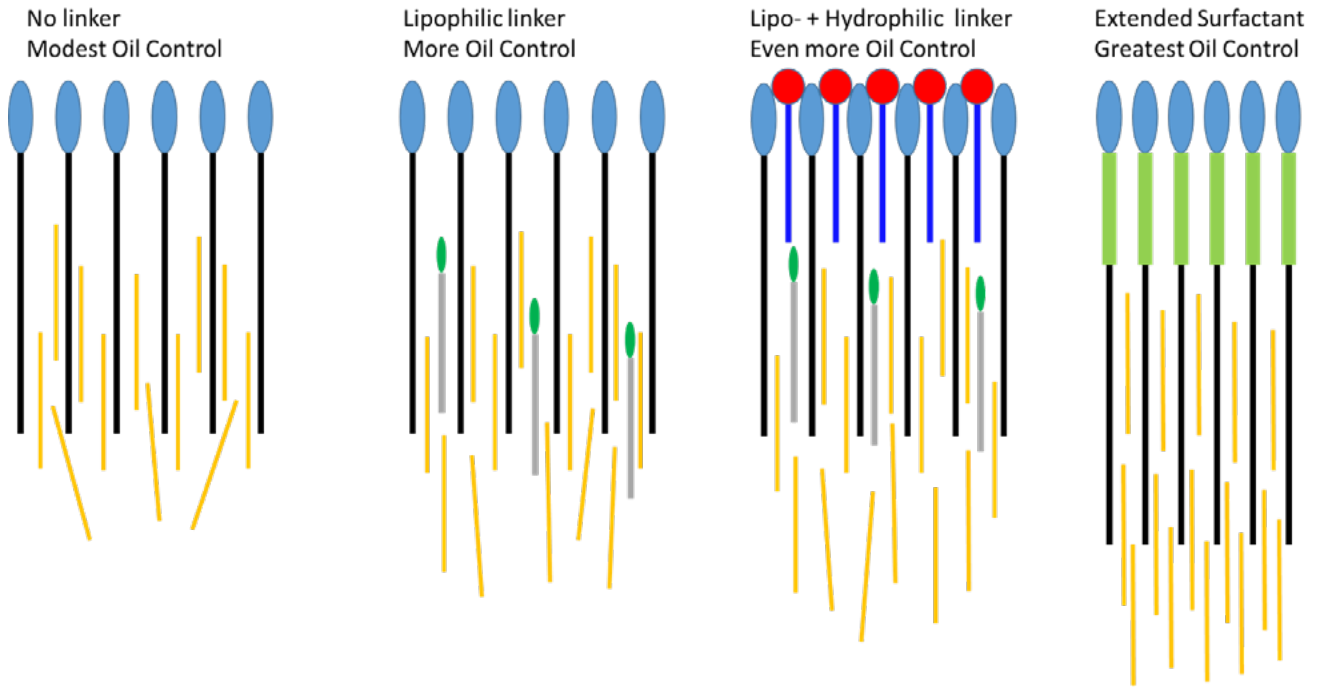
**Figure 6-1 From no linker to extended surfactant.**

The review gives examples of where the hydrophobic linker on its own provides most of the benefits, with other examples where the hydrophilic linker is required to maximise the overall benefits.

The choice of the hydrophobic linker is generally straightforward – a convenient long-chain alcohol that doesn't cause nasty phase problems with the surfactant of choice which should itself have as long a tail as possible. The choice of the hydrophilic linker is not so easy. The review shows that sodium naphthalene sulfonate (called a "hydrotrope" in the review, though that provides little enlightenment) has no beneficial effects on their specific system. The mono and dimethyl substituted sulfonate allows a doubling of the solubilization of the oil when combined with dodecanol so is called a hydrophilic linker. The HLD is changed so to get maximum solubilization the HLD has to be changed by other means, in this case salinity. The dibutyl substituted sulfonate increases solubilization even without dodecanol (though the effect is even stronger with the hydrophobic linker) so is called a co-surfactant.

The review also explores the effect on ξ, this time calculated in terms of the phase volumes of the system. $\Phi_o$ is the volume fraction of oil and $\Phi_w$ is the volume fraction of water in the middle phase, $V_m$ is the volume of the middle phase and $A_s$ is the interfacial area provided by surfactant + linker:

$$\xi = \frac{6\varphi_o \varphi_w V_m}{A_s} \qquad \text{6-4}$$

This definition of ξ makes the hitherto vague term "solubilization" clearer, a large ξ means a large middle phase for relatively little surfactant.

It is significant that the review was published in 2003. Although by most standards it is an eminently clear review, the work on NAC had only just been published and the implications of the model were not fully worked out. The CIT model from Tchakalova was still 9 years away. It seems to me that the whole area of linkers, hydrotropes (in this context) and alcohols could be greatly tidied up by a determined effort to disentangle the HLD effects, the kinetic effects and the "CIT" (i.e. the supposedly classic linker) effects.

At the time of the review, extended surfactants had been explored for a few years. Maybe the need for linkers has been and gone. Their HLD effects are better achieved by rational choice of the surfactant blend (e.g. the co-surfactant described in the review). The kinetic effects from smaller alcohols should be disentangled from their HLD effects and their downsides in terms of efficiency should be more clearly known. The true linker effects are presumably better served through the use of extended surfactants which reduce the complexity/ uncertainty of different linkers bobbing in and out of heads and tails.

Given that extended surfactants are still relatively rare outside EOR, the best advice is the same as that of dealing with polar oils in general. Set up your system in terms of classic HLD scans and make sure that you understand what your proposed linker is doing. If it is merely shifting HLD, then it is probably better to tweak the Cc of the surfactant blend. If it is shifting kinetics, it may well prove ultimately to be unwise as it will tend to reduce efficiency. If it is enhancing the phase volumes (so start to use the $\xi$ definition above) without changing HLD then rejoice in finding a very quick fix to your formulation problems.

## 6.5    From problem to opportunity

The four classic responses to problems are:

- Carry on doing what you know, even if you know that it doesn't work;
- Give up in despair and say that surfactant formulation is an art, not a science;
- Hope that someone else will solve the problems for you;
- See each problem as an opportunity.

There is certainly a lot of the first approach. The same tired explanations of how to use HLB have been going round for decades with consistent failure to deliver. The focus in training courses on micelles and CMC is usually of very little value to most practical formulations.

The idea that surfactant formulation is an art, not a science is popular with those in organisations who are the repositories of the art. This provides them with job security because managers are too frightened to risk the possibility of losing all that art. It is also the single biggest barrier to progress in formulation.

Those who rely on "art" have no methodology for bringing in fresh solutions. Given the complexities and problems discussed in this chapter along with a pervasive lack of understanding of surfactant fundamentals, it is not surprising that this approach is so popular, even though it is misguided and ultimately self-defeating.

The third approach has some merit. We can muddle along as best we can, then attend annual conferences in the hope that someone will get up and announce that they have found the right methodology to solve the key issues. If Practical Surfactants cannot identify a perfect tool then it provides a good-enough tool which is better than nothing.

But I favour the fourth approach. There seems to me to be a big opportunity for all of us to raise our game by focussing on what we already know (which is a lot) in two key areas, elasticity and curvature, all the while searching for ideas and data that will help us solve what we know we don't know.

### 6.5.1 $\Gamma m$, A, K and elasticity

All the effort (measuring $\gamma$ versus concentration) that has gone in to measuring the relatively unimportant CMC value can be re-worked to extract the three numbers, $\Gamma m$, A, K that really drive a lot of surfactant behaviour, albeit in an indirect way, because ultimately they control dynamic behaviour and elasticity.

Our formulation lives would be much more productive if we had a firm grasp of the dynamic and elastic behaviours within our formulations. Direct measurement of elasticity is notoriously difficult because it depends so strongly on frequency – with timescales varying from microseconds in an emulsion disperser to hours within a foam. Direct measurement of what is needed to break through an elastic wall (a Pickering emulsion drop, the Entry Barrier for antifoams, the act of flocculation and coalescence) is somewhere between hard and impossible.

Even worse, small levels of accidental impurities or deliberate additives can have large effects on the elastic behaviour – as shown by the transition from a Dawn to a Gillette foam with a few percent of myristic acid. Our linkers, hydrotropes and co-surfactants are all affecting the elastic behaviour in addition to their other effects on curvature and $\xi$. To these uncertainties we can add the problem of understanding why supplier A's version of a well-known surfactant gives different behaviour from supplier B's version.

As a community this gives us an opportunity. We can start to consistently measure what properties we can (starting with $\Gamma m$, A, K) and get into the habit of re-measuring them in the presence of suspected impurities or deliberate additives, and measuring them for different versions of the same surfactant. We can apply not-so-subtle pressure on surfactant suppliers to routinely provide this information. They have it already (if they don't there is something

seriously wrong with them), so why not give it to us? The more we build up a corpus of knowledge (either our internal company knowledge or, preferably, a wider community knowledge) the more we can identify trends that can be useful formulation guidelines (some of which are described in this book) while identifying black holes that need to be filled by those with the resources and motivation to do so.

Two communities would benefit from an industry-wide approach to grappling with the elasticity problem and they, in turn, would benefit the rest of the community.

The first community is those academics who have great ideas but insufficient funding. They have insufficient funding because surfactancy is generally perceived by the funding bodies as an old, tired topic long-since resolved via CMC and HLB. Funding bodies respond when they find that there are huge unmet needs that they can help to meet.

The second community is the manufacturers of test equipment. Let us take, as a specific example, the problem of measuring entry barriers for antifoams. Measuring entry barriers would be transformational to many different industries who face foam problems and for whom antifoams are a black art. If they could take their own foam and see which antifoams most easily penetrate the elastic barrier, or could tweak their own foaming system to reduce the barrier they could more quickly find a better solution. But until we all acknowledge how useful such a device would be, there is no incentive for the equipment suppliers to invest in creating such a device. It is a classic chicken-and-egg problem. We can short-circuit the dilemma by understanding the science of why it is so important and by asking the usual suppliers of surface equipment if they can come up with a solution.

### 6.5.2 Curvature and IFT

Whether it is HLD-NAC, PP or Helfrich torque, curvature is clearly at the heart of many formulation issues. As a community we can start to demand from ourselves and others that relevant issues should be discussed with whatever curvature language makes the most sense for us. The old excuse that "curvature on the nm scale has nothing to do with real emulsion on the µm scale" is discredited by three facts:

- Curvature, directly or indirectly, appears even in classical emulsion tools such as HLB, Winsor R, Bancroft, PIT.
- The revised "Wedge" theory from Kabalnov and others provides a coherent explanation of how nanometric curvature directly influences macroemulsion behaviour.

- Interfacial tension, in whatever emulsification model you choose, has a big practical effect (for good and bad) and the link between IFT and curvature is direct and profound.

The first benefit of a shift to a curvature-based approach to formulation would be a big reduction in confusion and a large boost to productivity. It is immensely frustrating to read an academic paper or listen to a conference talk where puzzle is piled upon puzzle simply because the author/speaker is lost in surfactant space without a guide to where they are in terms of curvature. Sometimes there is enough information to make it possible to recalculate things in terms of, say, HLD-NAC, and this then condenses much confusion into a few nuggets of useful clarity. Importantly, these nuggets can then become part of a wider corpus of knowledge. More often there is insufficient information to anchor the work in curvature space so, like countless other papers/talks, there is no ability to use the data for other purposes. If a paper has, say, Cc and EACN values then it is possible to take the data and get a good idea of what might happen in a situation closer to your own specific requirements. Without curvature data the paper speaks only to its specific situation and, unless you happen to need those exact conditions, is of no value to the rest of us. Good science builds on itself. Much work in surfactancy allows no such building so I would dare to call it, if not bad science, at least poor science.

### 6.5.3 Together

Although it was convenient to identify two separate themes, elasticity and curvature, it is clear that the two are intertwined. For example, NAC is linked to ξ which in turn is linked to interfacial rigidity and therefore to elasticity. Frustratingly, the desirable rigidity also means undesirably slow kinetics during formation of a system, coupled with either desirable long-term stability or (in terms of antifoams) undesirable indestructability.

So any work that focusses purely on elasticity or purely on curvature is likely to lead us astray. As a community we need to keep both in mind, though with different emphasis depending on context. A beautiful example of such an approach is the emulsion stability theory of Kieran and Acosta which blends fundamental HLD-NAC theory with coalescence theory. The first provides information on the generation of emulsion particles, via thermodynamics, the second uses estimates of elasticity to predict coalescence times and, therefore, emulsion stability.

There is always more that can be added to a theory. In the emulsion stability case, DLVO stability from charge or steric stabilisation can be added, along with Ostwald ripening theory which depends on oil solubility and, returning to the main theme, how rigid the barrier is around the emulsion.

The point is that armed with curvature and elasticity a lot can be predicted. DLVO and Ostwald can refine those predictions but are of little use on their own in this context because they can only operate within the constraints of the fundamental formation and destruction of the emulsion particles.

My feeling is that we are about to enter a golden age of surfactant science. There is enough information to prove the inadequacies of the old methodology. There is a store of theory, such as HLD-NAC, that is already helpful in practice and clearly working along the right lines. And, at last, there is enough realisation among the formulation community that there is a chance to do much better. Just a few years ago, mention of HLD was greeted either with blank looks or hostility (because it was "only a microemulsion theory" or "wasn't PIT"). Now it is becoming routine in many labs and the attitude is much more "how can we overcome its limitations?" That is a big step forward. There is also mounting pressure on surfactant suppliers to provide the necessary data. At the same time, other approaches such as PP and Helfrich torque are being worked up into methodologies that might either supplant or (more probably in my view) unite with HLD-NAC to give us the sort of formulation power we really need. I'm also encouraged by all the apps. Why? Because the most important thing for writing an app is a good appable theory. The broad range of apps is entirely due to the fact that some of the key academic groups have cracked many of the big problems and given us good-enough algorithms to work with. It is a common tragedy that much of the output from academic communities is effectively useless because formulators cannot take the complicate formulae and explore them. Now that we have the wonders of HTML5/JavaScript/CSS3, the barrier between academic and formulator is rather easy to overcome – especially when the academics (as I found with 100% success) are delighted to help bring their work to life if someone offers to write an app based on that work.

There is no good reason why it is me who is writing the apps. The younger generation of scientists are surely more "app literate". However, if you know of some useful theory and happen not to be able to write an app for it, let me know and I will be happy to write it and add it to Practical Surfactants.

Yes, we are entering a golden age of surfactant science.

# 7 Phase diagrams

Surfactant space is fiendishly complicated. What you might find in any formulation will depend, at the very least, on:

- % Surfactant and its Cc;

- % Oil and its EACN;

- % Water and its salinity;

- Temperature;

- % Polar oil and its properties.

So a description of surfactant space needs at least a 9-dimensional graph. Given that many of us struggle with even a 2-D plot and that 3-D plots are the limit of what we can cope with, it is not surprising that our views of surfactant space are limited to small slices of that space, with the hope that everything else outside that slice is irrelevant to our specific needs. Each of those slices is a phase diagram.

We have used HLD-NAC to great advantage to explore % Surfactant versus Temperature, Salinity, EACN and Cc in the set of fish diagrams. In this chapter we look at two more types of phase diagram:

- Binary diagrams (Fish diagrams are, of course, binary diagrams);

- Ternary diagrams.

In both cases they become much easier to grasp when explored with the various apps. Indeed, it was my own inability to understand ternary diagrams that led to the creation of the apps. At a chance meeting with Seth Lindberg of P&G I mentioned that I was thinking of writing a simple ternary diagram app so I could better understand them. It turned out that he was an expert on such diagrams but was frustrated at how hard it was to teach them to people like me. This led to a fruitful collaboration to which Seth added his awesome graphics skills. Just about everything in this chapter is a result of that collaboration. Seth was immensely patient with me as I continued to not "get" ternary diagrams and he also had to push me hard to implement the different versions described here and to sort out the bugs that arose every time a new feature was added. I am hugely grateful that he didn't give up till everything was right!
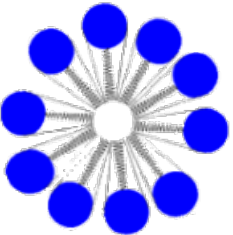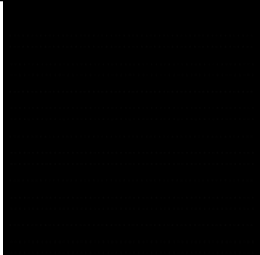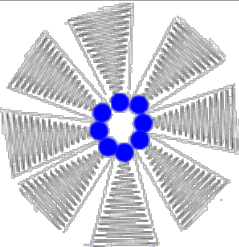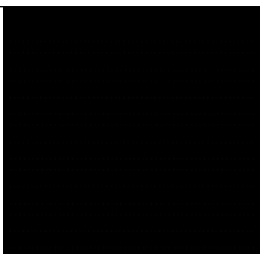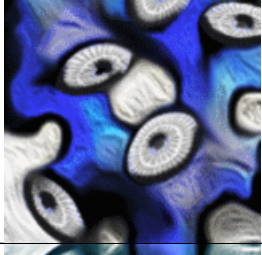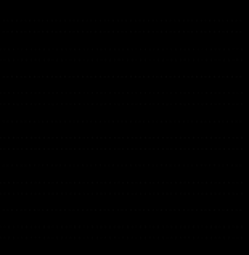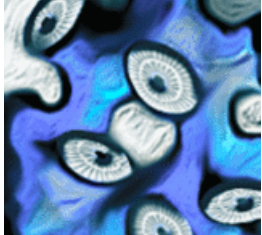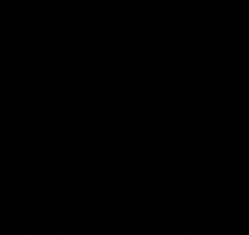
## 7.1 Phases
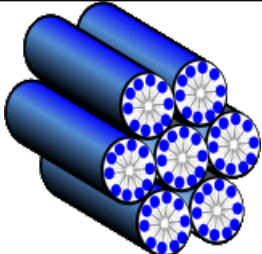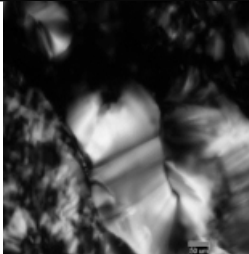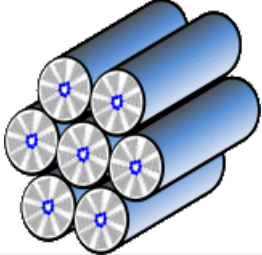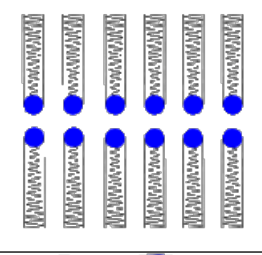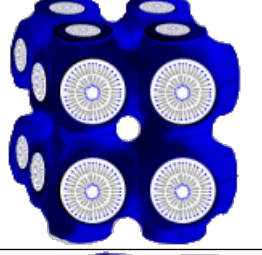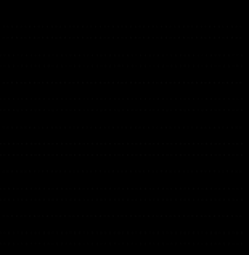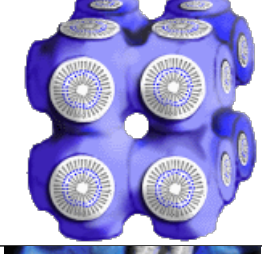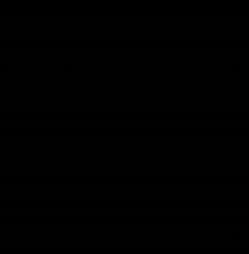
If we are going to have phase diagrams we had better agree on what those phases might be. Fortunately we have Seth's graphics which neatly describe the rather bewildering configurations that surfactants can take, along with the bewildering variety of their names.
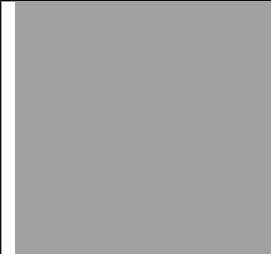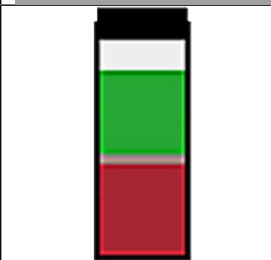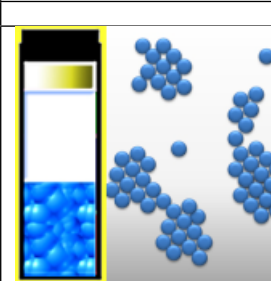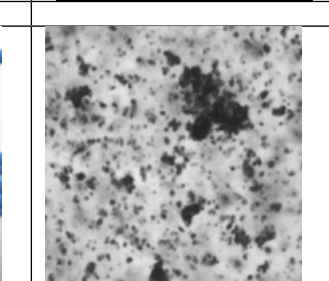
It would be nice to say something coherent about these phases and offer guides to predicting them and their behaviour when factors such as temperature and % Surfactant change. The simple, or perhaps simplistic, ideas behind Critical Packing Parameter, discussed near the start of the book, can be offered. But these only show an ideal trend throughout a surfactant class as the heads and tails change in relative sizes. As discussed in terms of the Tiddy paper in the CPP section, there are some broad trends that can be identified under ideal circumstances (e.g. comparing sets of ethoxylate surfactants) which depend on the extra packing force of micelles bumping into each other. But there are plenty of problems even in Tiddy's idealised set, so the ideas are not generally applicable to an arbitrary surfactant that may be of interest to you. Those who like DPD (Dissipative Particle Dynamics) might claim that this is a powerful approach towards prediction of phases. Those who dislike DPD claim that it tends to produce the answer you want. As I have no direct experience I cannot comment. Assuming that DPD are not the answer (if they were, we'd all be using the technique) I know of no general approach to predicting the phase behaviour. This means that the only option is the arduous task of experimentally scanning through phase diagram space and identifying whatever phase appears in the test tube.
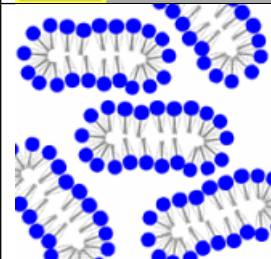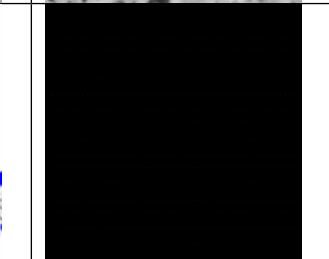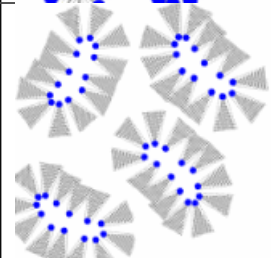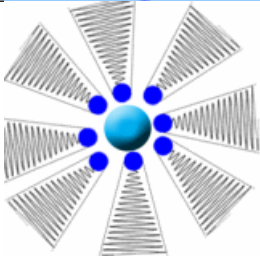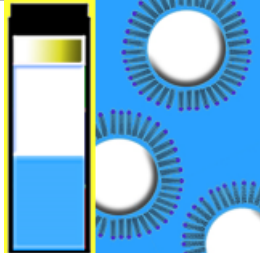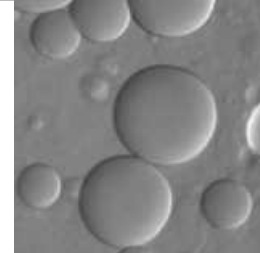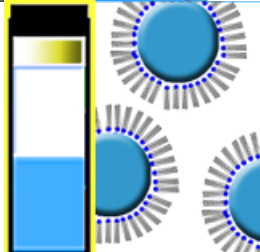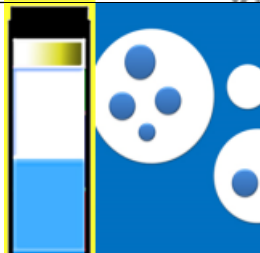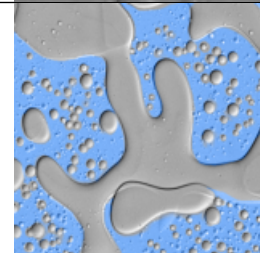
This immediately raises the question of how to tell what phase is present in the tube. If you are lucky, a glance under a polarizing microscope might reveal a defining character. If a phase is isotropic under the microscope yet of high viscosity (as judged by a poke on the cover slip) then it is likely to be cubic or bicontinuous. The ultimate decision comes from small angle x-ray scattering, SAXS (or neutron scattering, SANS), for those who have a machine to hand.

Here are the phases used throughout the apps. In each row, where applicable, there is an idealised graphic of the structure, an idealised view down a (polarizing) microscope and the name or names of that phase. If there is a SAXS/SANS pattern it is also recorded.

| Phase Graphic | Polarizer View | Phase Name(s) |
|---|---|---|
|  |  | L1/Micelles |
|  |  | L2/Inverse Micelles |

| | | |
|---|---|---|
|  |  | H1, M1, Hexagonal or Middle<br><br>SAX=1, √3, √4, √7, √9, √12 |
|  |  | H2, M2, Inverse Hexagonal<br><br>SAX=1, √3, √4, √7, √9, √12 |
|  |  | Neat, Lα, Lamellar<br><br>SAX=1, 2, 3, 4, … |
|  | | V1, I1, Q1,Cubic<br><br>SAXS=Multiple |
|  | | V2, I2, Q2, Inverse Cubic<br><br>SAXS=Multiple |
|  | | V1-Bicont, I1-Bicont, Q1-Bicont, Cubic-Bicontinuous<br><br>SAXS=Multiple |
|  | | V2-Bicont, I2-Bicont, Q2-Bicont, Inverse Cubic-Bicontinuous<br><br>SAXS=Multiple |

| | | |
|---|---|---|
|  | | Solid/XTLS/Crystal |
|  | | 2 Phase |
|  | | Three Phase |
| |  | Isotropic |
|  |  | Precipitate |
|  | | Worm-like micelles |
|  | | Inv.Worm-like micelles |

| | | | |
|---|---|---|---|
|  | | Microemulsion | |
|  | | Inverse Microemulsion | |
|  |  | O/W emulsion | |
|  |  | W/O emulsion | |
|  |  | Mixed emulsion WOW | |
|  |  | Mixed emulsion OWO | |

## 7.2    Binary diagrams

Although there are many potential binary diagrams, the most usual one shows the different surfactant phases that appear at different temperatures and % surfactant in water, with all other variables such as salinity constant or, in the case of oils, irrelevant.

Obtaining these diagrams is tedious. This is unfortunate because it would be useful for everyone if they were available for all common surfactants. But funding agencies won't fund such dull work and presumably those surfactant

suppliers who have done the work don't want to give away hard-won knowledge. Still, with modern high-throughput techniques how hard can it be to gather the binary diagrams of 10-20 surfactants?

The app allows you to explore 9 diagrams, 6 of which are systematic scans across the ethoxylate range. The data are from public domain reports (all sources are acknowledged) though the digitisation into appable format was carried out by Seth. There is an option to read in your own data files if you happen to have them. The format is explained in the app and need not detain us here other than to say that a minimalist approach is adopted – the boundaries are specified as crudely as possible because accurate digitisation of a complex phase boundary is (as we discovered) far too tedious.
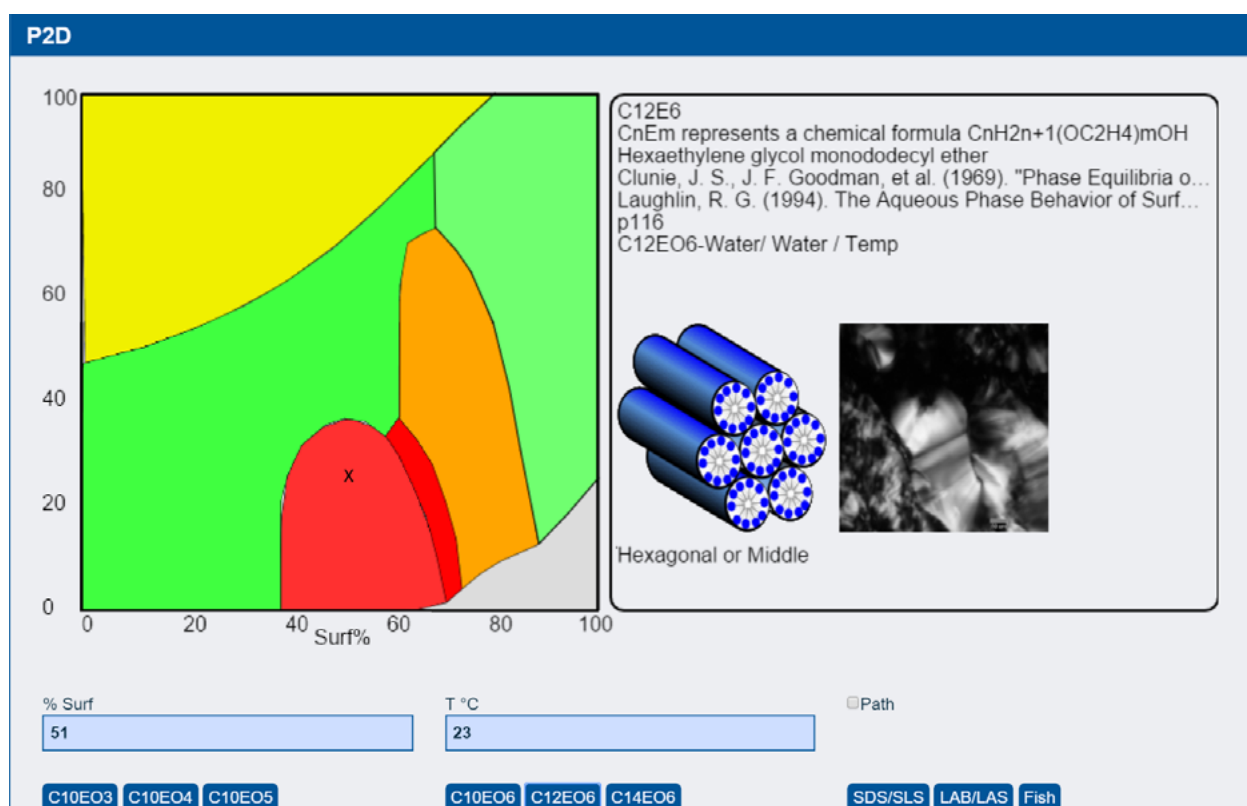
In the example shown the surfactant is C12EO6 and the mouse is at the point marked X where there is 51% surfactant and the temperature is 23°. This particular phase is hexagonal and the red colouration is a warning that this is a highly viscous phase, with orange being less viscous (in this case a lamellar phase) and green being low viscosity. The bottom right-hand corner is the solid phase. The slight blip along the left-hand edge of the hexagonal phase is an error from the simplified digitisation process.

What does the diagram show us overall? At modest temperatures and lower or higher concentrations the system forms the simple micelle phases we think of as being normal for surfactants. The limitations of CPP are apparent because this surfactant forms micelles at concentrations <67% and inverse micelles at higher

concentrations. At higher temperatures the system produces the yellow 2-phase region.

The practical issue for handling C12EO6 is the red and orange phases in the middle. If you start with a concentrated inverse micelle phase and naively dilute to get the usual micelle phase your plan will fail – you go into the cubic and/or hexagonal phases which are very high viscosity. Unless you have very powerful mixers the system gets stuck and you end up with a mess. This is not some academic nicety – these phases are a significant impediment to those who sell detergents or washing-up liquids. At some point (either in manufacture or within modern concentrated consumer products) the surfactant has to get diluted and if it passes through one of these bad phases the manufacturer or the consumer ends up with a big problem. Some surfactant formulations that would offer great performance at the dilutions used by consumers are impossible to produce or sell because the dilution goes through a hexagonal phase. Considerable research and development ingenuity goes in to avoiding such problems. These issues are explored further when we look at the Path option in the ternary diagrams.

## 7.3      Ternary diagrams

There are two issues with ternary diagrams. One is that they can be more complex than binary diagrams, involving things such as tie-lines. More fundamentally, however, the key issue is that most of us most of the time do not know how to read them.

### 7.3.1  Reading values from a ternary diagram

It should be very easy. Every point in the triangle has a unique composition, shown in the apps as W for Water, S for Surfactant and X for whatever else is being used (oil, co-surfactant, alcohol …). The triangle tends to have 3 sets of grid lines from each of the axes and you "simply" read off from the grid. The graph is redundant – once you know two of the values, the third is automatically known because all points add up to 100%. The problem is that everything is so ambiguous.

For example, look at this diagram in B&W. What is the value of the point marked O?
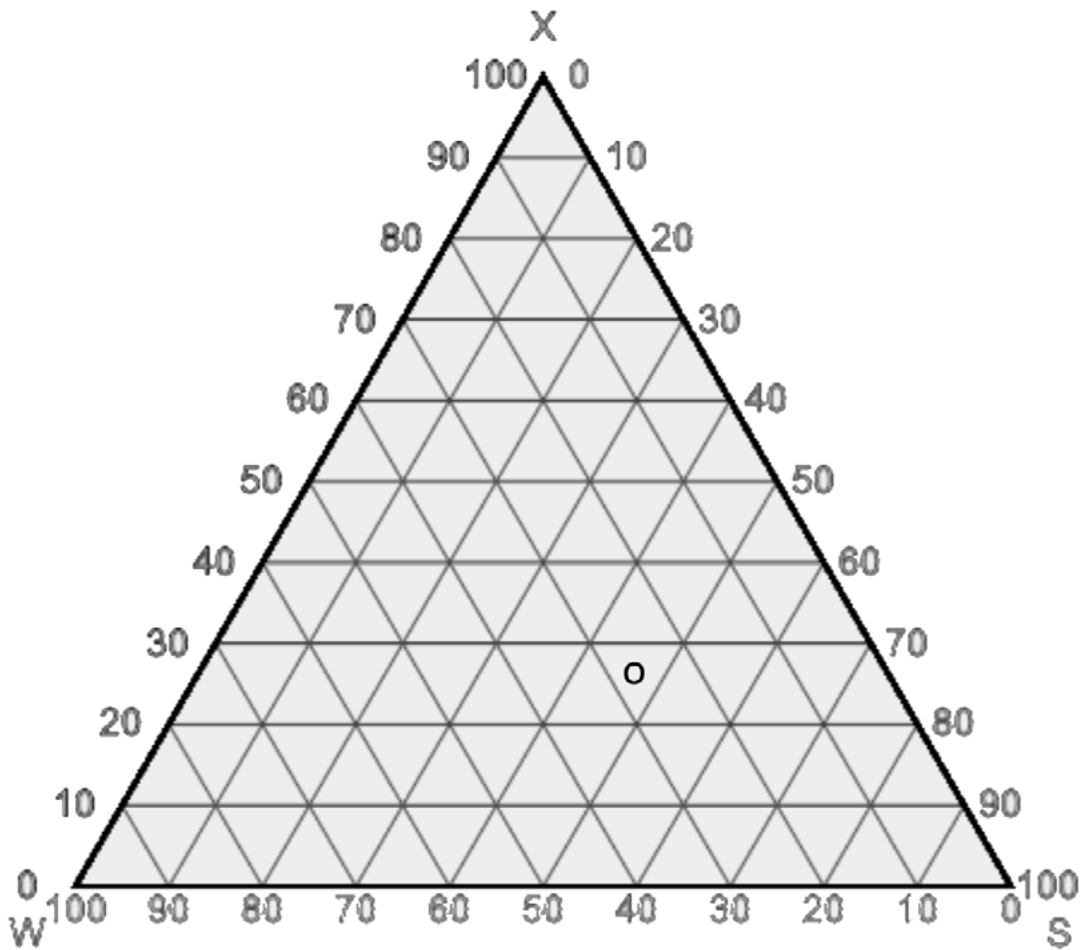
**Figure 7-1 O marks the spot. But what are the W, S and X values at that point?**

Even though I wrote the app, when I deliberately reproduced it in B&W and tested myself I made a classic readout error and got it wrong. The answer is W=27, S=46, X=27.

The problem is that along the bottom, W, axis the grid lines go in two directions. In this example, the O is either between the lines going right to left, so between 20 and 30, or between the lines going left to right, so between 50 and 60. How does one know which set of lines to choose?

Going to the app in colour it is easier to see why the answer is what it is:

**App 7-2** https://www.stevenabbott.co.uk/practical-surfactants/PDE.php
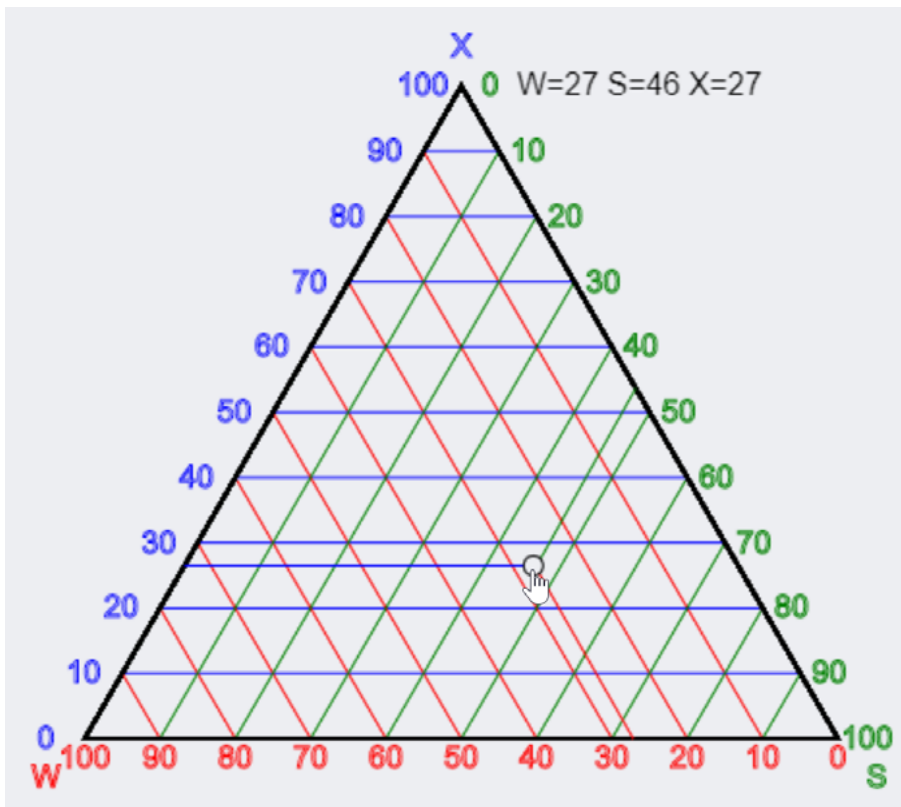
The TriPoint option is selected so the circle that marks the spot is connected back to the respective axes which are conveniently colour coded along with the connecting lines. When it is drawn like this there is no ambiguity, clearly the W is somewhere between 20 and 30, clearly S is between 40 and 50 and X is between 20 and 30.

In the absence of these colours, how does one know which set of grid lines applies to which axis? It seems even worse than that. Does the W apply to the bottom of the triangle or its left-hand edge? This is where we solve the riddle of these diagrams. In the app you can click the Flip option which will make the W axis go along the left-hand edge and the S axis run along the bottom. The secret is that it makes no difference to the readout. This means that the grid lines you choose for, say, W must be the ones that work whether the W axis is bottom or left. The only grid lines that meet that criterion are the ones going right to left from the bottom.

The confusion doesn't end there. Sometimes the ternary diagrams are labelled with axes along the edge rather than along the corner. This means we lose the ability to apply the trick above. So the other trick (which applies in both views) is to say that the grid lines to be chosen are always angled from low to high values. Going back to our original dilemma of the B&W diagram and looking at the lines starting on the bottom of the triangle (if we assume this is the W axis), the lines sloping towards the 100 value are those going from right to left, which means we read out the value as 27.

As a check, and to show the point that the graphs can be plotted in 3 equivalent ways, what is the value in the X spot?
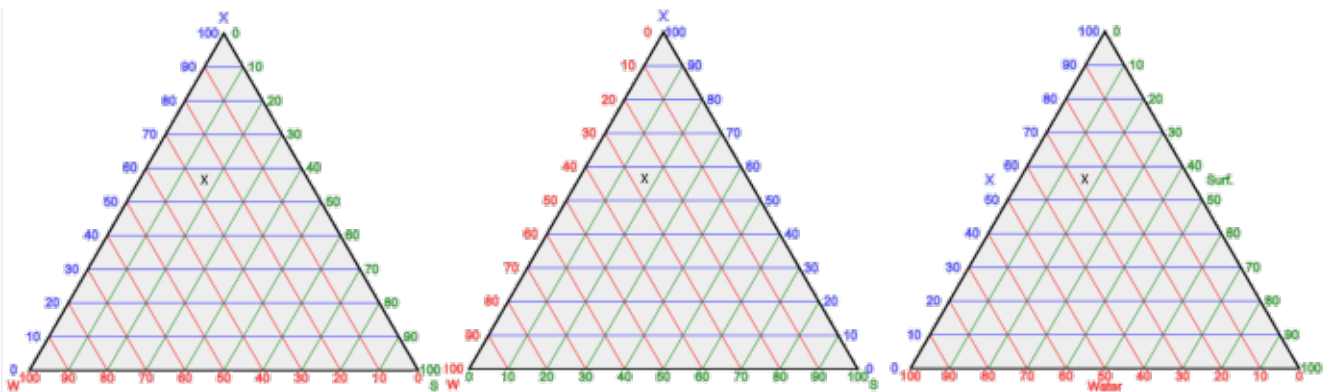


**Figure 7-2 X marks the spot. But what are the W, S and X values at that point in each of the three graphs?**

The answer in each case is W=27, S=17, X=56.

Unfortunately the confusion doesn't stop there. The same data can be plotted in 6 different ways depending on the choice of which axis applies to which component. Here are another three graphs. Can you read out the W, S and X values in each case?
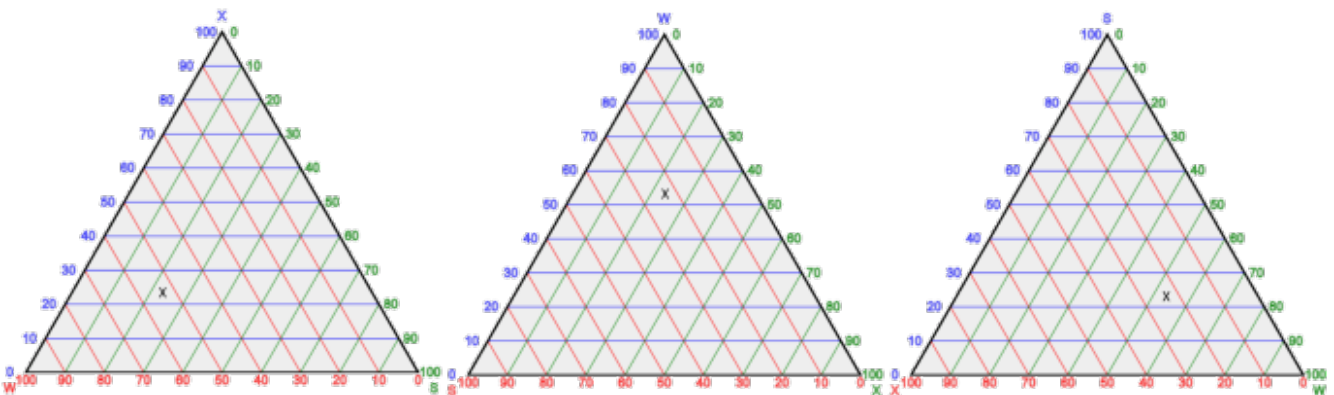


**Figure 7-3 X marks the spot. But what are the W, S and X values at that point in each of the three graphs?**

Again the answer is that they are all the same, W=54, S=23, X=23.

Hopefully, by playing with the apps, swapping around axes, turning various options on and off you will have built up an intuition so that reading points off any ternary diagram, no matter how it is presented, will become straightforward. It is especially useful to turn off the Lines option, leaving a bare graph with just the axis labels. Many academic papers publish diagrams without the grid lines so it is good to be able to read out values without their visual help.

### 7.3.2 Paths through phase diagrams

First we need to touch on an important issue raised in the discussion of binary diagrams. If one starts at some point in the diagram (e.g. a concentrated surfactant) and wants to go to another part (e.g. a dilute solution for use in a washing machine) the process passes through the phase diagram along a *path*. If some highly viscous hexagonal or cubic phase happens to be along that path then the result is disaster – the process will grind to a halt thanks to the viscosity of the phase. In the path explorer app you can explore what might happen as you can set the start and finishing points along with a blob in the diagram which represents the phase of interest.



**App 7-3** https://www.stevenabbott.co.uk/practical-surfactants/PDE1.php
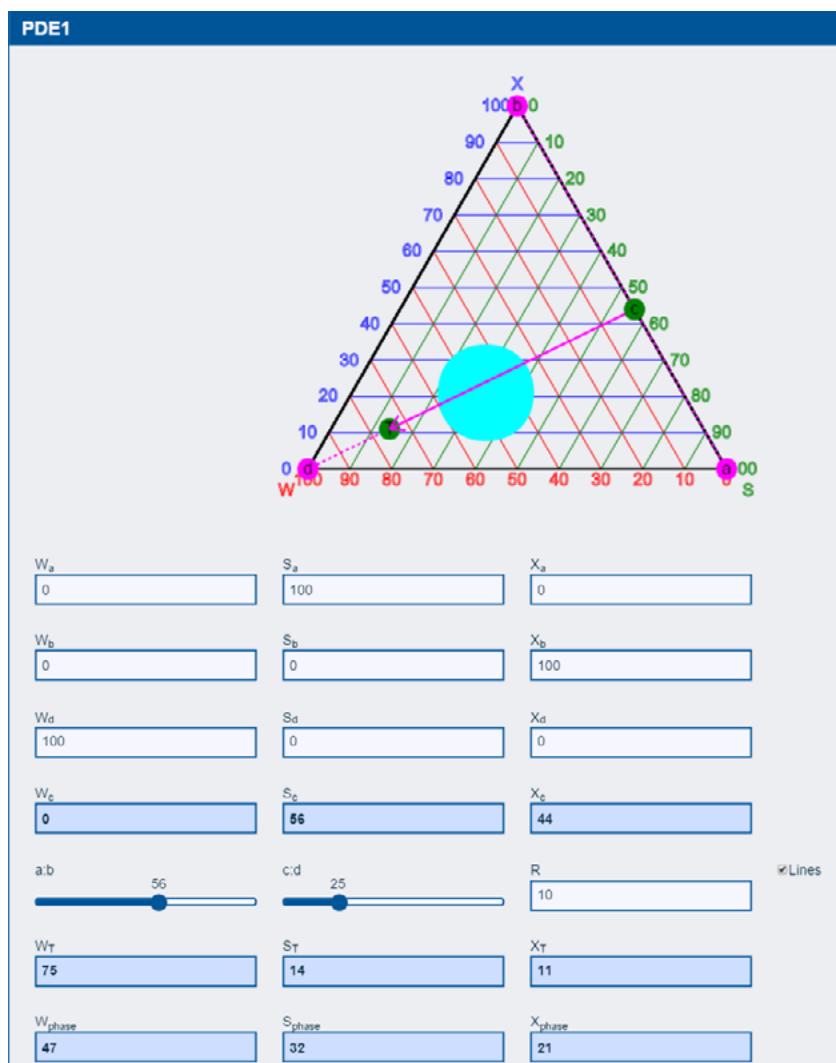
In the example shown, it would be a disaster starting at 56:44 S:X and heading to 74:15:11 W:S:X because a nasty phase at 47:32:21 lies along that path.

### 7.3.3 Diagrams with 2-phase regions

I could never understand why ternary phase diagrams often contained weird regions with lots of lines nor could I grasp why they were called *tie-lines*. The app helps to make sense of what is going on.



**App 7-4** https://www.stevenabbott.co.uk/practical-surfactants/PDE2.php
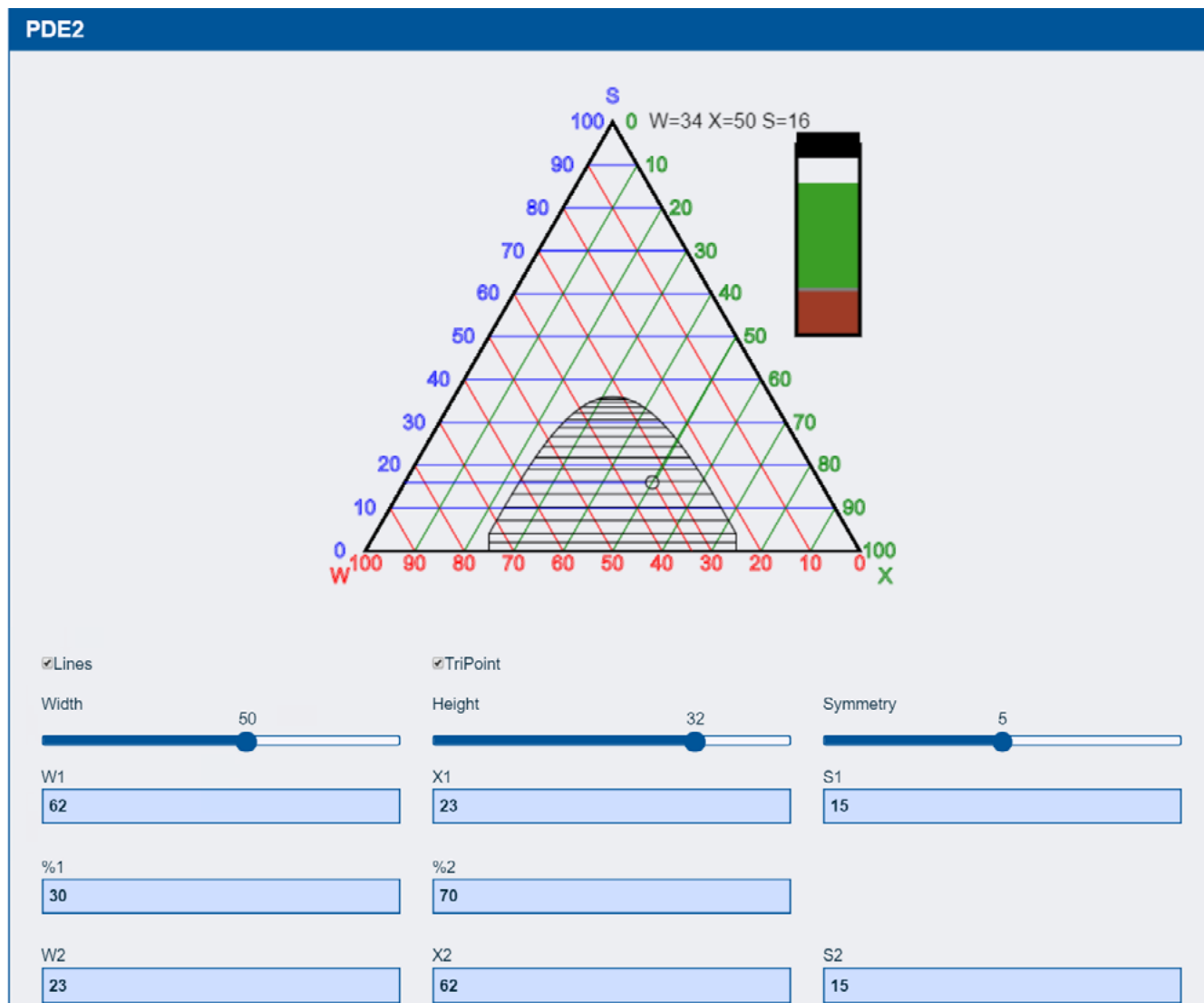
The diagram is typical of a water/oil (i.e. X=oil) mix. At low water in oil or low oil in water there is a single phase. At 50:50 O:W it needs a lot of surfactant (35% in this example) before achieving a single phase. Elsewhere the system splits into two phases.

These phase are *not* water and oil. In this example the composition (marked with the circle) contains 16% surfactant and 34% water and 50% oil. But it splits along the 16% surfactant line (the horizontal tie-line) to give, at the left end of the tie-line 63% water and 21% oil and at the right end of the tie-line 21% water and 63% oil. So the two phases in the test tube both contain oil and water, but

in different proportions. The reason we consider the tie-line is that if the starting composition had kept the same 16% surfactant (the definition of the tie-line in this example) but instead of W:O 34:50 it was changed to W:O 50:34, the two phases in the test tube would still be 21:63 and 63:21. How is it possible that two different formulations can split into exactly the same O/W and W/O compositions? The answer is that although the O/W and W/O *compositions* are the same, the *ratio* of O/W:W/O changes. In the example above, the W/O phase is 69% of the tube and the O/W phase is 31%. You can see this in the outputs and also visually in the test tube. If you slide along the tie-line in the app to 30% oil then the O/W:W/O ratio changes and now 79% is O/W and 21% is W/O. If you take it to the extremes then if you start with a 21:63 W:O ratio you end up with a single phase (i.e. 100%) of the 21:63 W:O and similarly for the O:W.

Naturally, if you start on another tie-line (in this simple example a different % surfactant) the O/W and W/O ratios at the two ends are different.

The app tries its hardest to make this as easy to understand as possible. The test tube always shows the ratio of the two phases and the outputs always show the phases at the ends of the tie-lines. The colours of the phases represent their purity – the closer you get to one corner of the triangle, the closer they approach pure red, green or blue. For simplicity of explanation, the tie-lines in the above example were horizontal. The app is happy dealing with asymmetric tie-lines (so the mixes at the end contain different % surfactant) and you can explore the implications both visually and with numbers.

So there is nothing difficult or mysterious about tie-lines. Most of us have been frightened of them because we were frightened by ternary diagrams and also because it was hard to grasp that the two phases that separated along a tie-line were not pure oil and pure water, that the ratios in the mixes at the ends of those lines are constant (though this should be obvious from the diagram) and that what changed as you slid along the tie-line was the proportion of the two phases – something that the phase diagram can't tell you.

### 7.3.4 Diagrams with 3-phase regions

Now we have worked out what is happening in a diagram containing 2-phase regions we're ready for diagrams containing 3-phase regions. Indeed, it is impossible to view a 3-phase diagram without understanding 2-phase regions. This is because there is a fundamental rule (called the Gibbs Phase Rule) which says that it is impossible to go straight from a 3-phase region to a 1-phase region – there must always be 2-phase regions in between. For those who are interested, there is an app (https://www.stevenabbott.co.uk/practical-surfactants/Gibbs.php) that describes the Gibbs Phase Rule but for the purposes of this book we just need to know this important principle of "no 3-phase without 2-phase".

We can now take a look at a nice symmetrical 3-phase plot. As with the 2-phase plot the app allows asymmetry but it is easier to sort out the basics with a symmetrical plot.



**App 7-5** https://www.stevenabbott.co.uk/practical-surfactants/PDE3.php
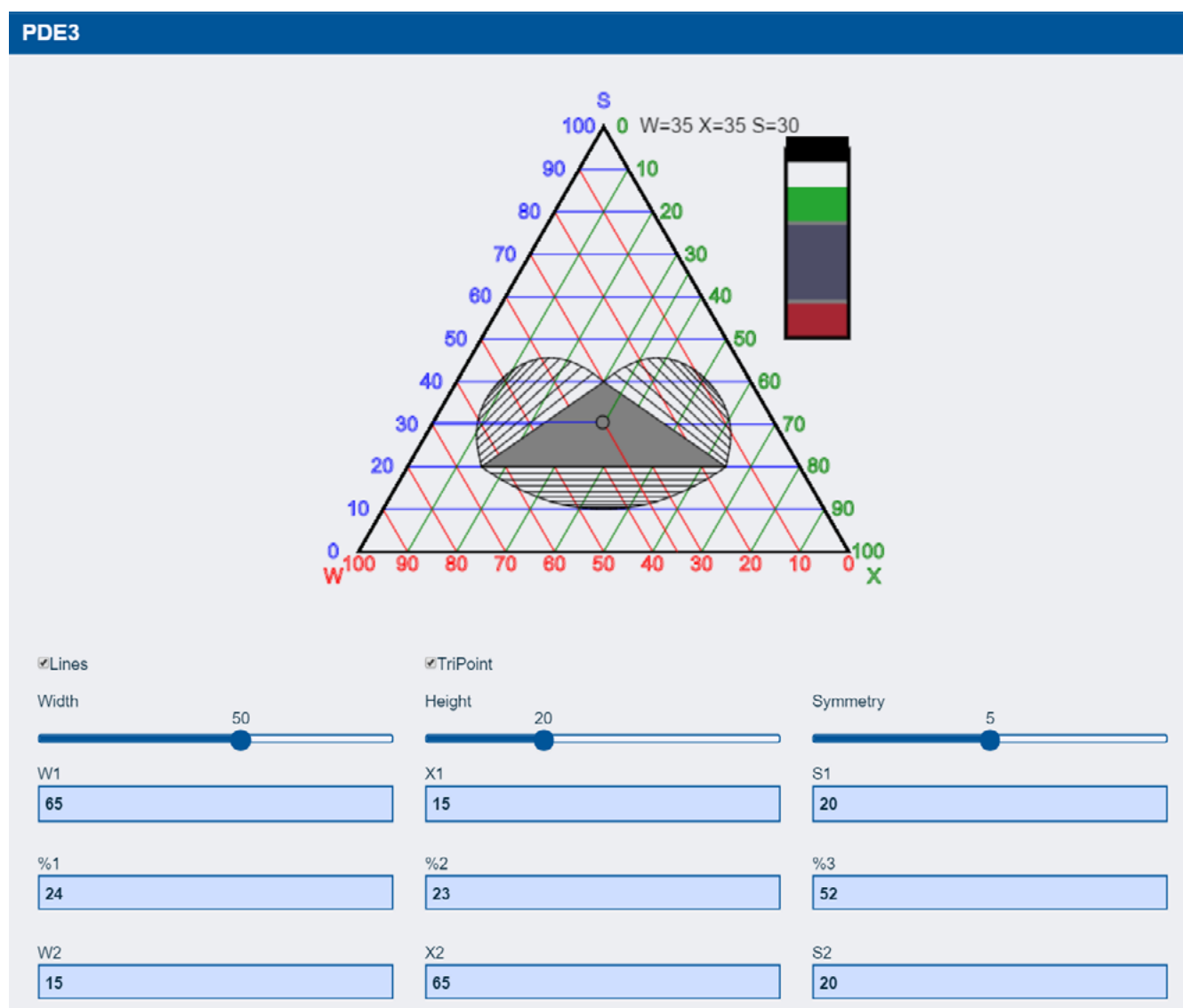
The circle in the middle of the grey triangle marks the point of current interest and the readouts provide all the necessary information.

The first clue is the test tube; it shows three phases, so we know the grey triangle is a 3-phase region. The other way to know this is to see that it is bounded by 3 2-phase regions, as required by the Gibbs Phase Rule. The output boxes, W1, X1, S1, W2, X2… tell us that the composition of the three phases, and with your ternary diagram skills you will be able to spot that they are the values of the three corners of the triangle. So if you start (as in this case) with W=35, X=35, S=30 you end up with those three phases in the ratio 26:26:48

If you move the mouse within the grey triangle, i.e. you start with a different original composition, the output boxes remain constant – i.e. the three points of the triangle are fixed. What changes (as with the tie-lines in the 2-phase plot) are the ratios of the three phases, shown in the test tube and also in the ratio output values.

If you move into a 2-phase region then you are in familiar territory – the rules are unchanged though now we have tie-lines at angles which means that the end phases are more complex mixtures than in the horizontal lines we explored previously.

As you get used to moving the mouse around and looking at the ratio of the phases as well as their colours (as a guide to their make-up) you will find that everything is self-consistent. As you glide from 3-phase into 2-phase or from 2-phase into single phase there are no sudden jumps or mysteries. It is all remarkably obvious and straightforward.

The real difficulty is, therefore, not in the diagrams themselves, but obtaining them in the first place. If binary diagrams are tedious to gather, ternary diagrams are doubly or triply so.

### 7.3.5  HLD-NAC and microemulsion diagrams

In a tour-de-force, Prof Acosta has shown that with a few modest assumptions, a ternary phase diagram can be constructed from HLD-NAC.

**App 7-6** https://www.stevenabbott.co.uk/practical-surfactants/MPDE.php
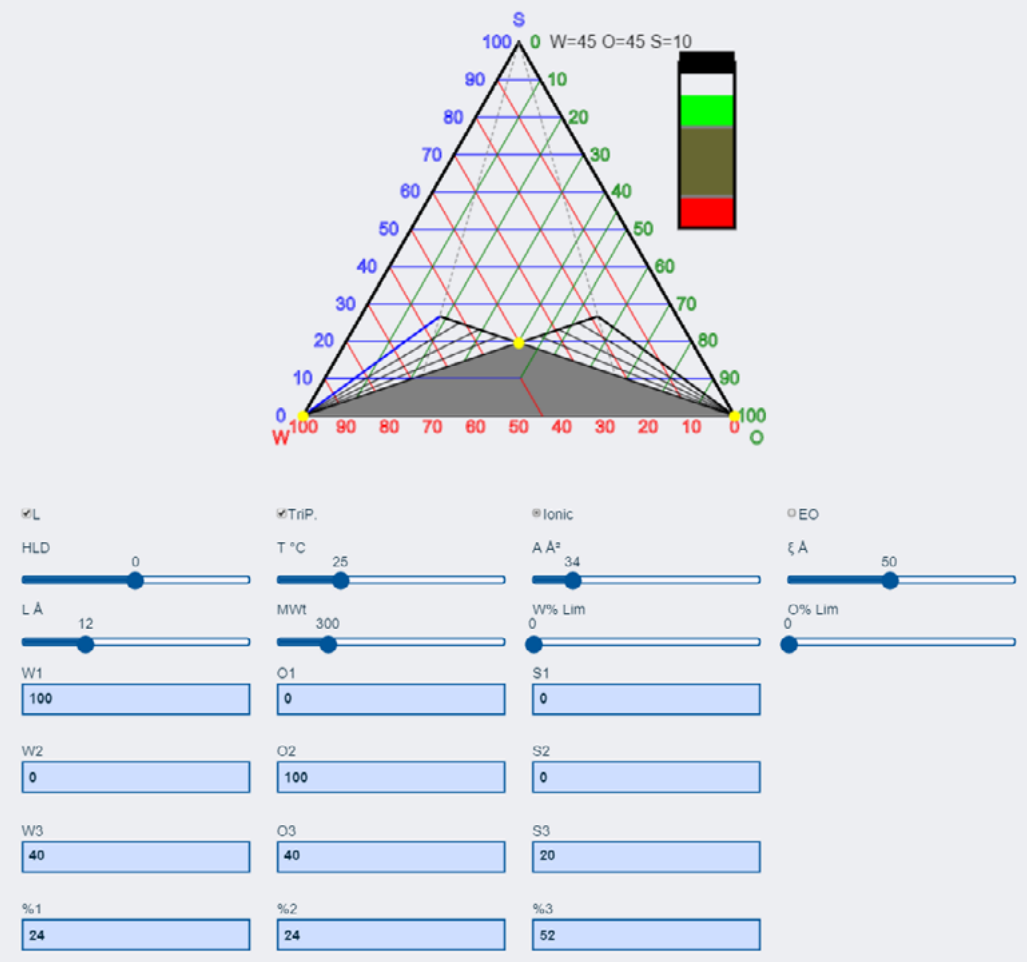
At the time of writing, the theory behind this app has not yet been published so it is discussed in a more phenomenological manner. First, it is a typical 3-phase diagram with the necessary 2-phase zones either side. Issues such as the composition of the three phases in the test tube along with the relative sizes of each phase are unchanged from the previous section. What is different is that the system can now be explored in terms of HLD, ξ, L and T, seeing how each of these factors changes the diagram. Changing HLD, for example, skews the diagram. Starting from HLD=0, changing T equally skews the diagram because it is changing HLD. Large ξ and L cause the system to shrink. At first this seems the wrong way round – until you realise that both changes make the surfactant more efficient so less is needed to reach the isotropic phase. Skewing the HLD to high or low values pushes the 3-phase region towards the edges till you end up with a single 2-phase region, exactly as one finds in reality.

### 7.3.6 General ternary diagrams

This step-by-step approach to ternary diagrams, focussing on specific regions, prepares us for the general ternary diagram full of complex phases.

**App 7-7** https://www.stevenabbott.co.uk/practical-surfactants/P3D.php
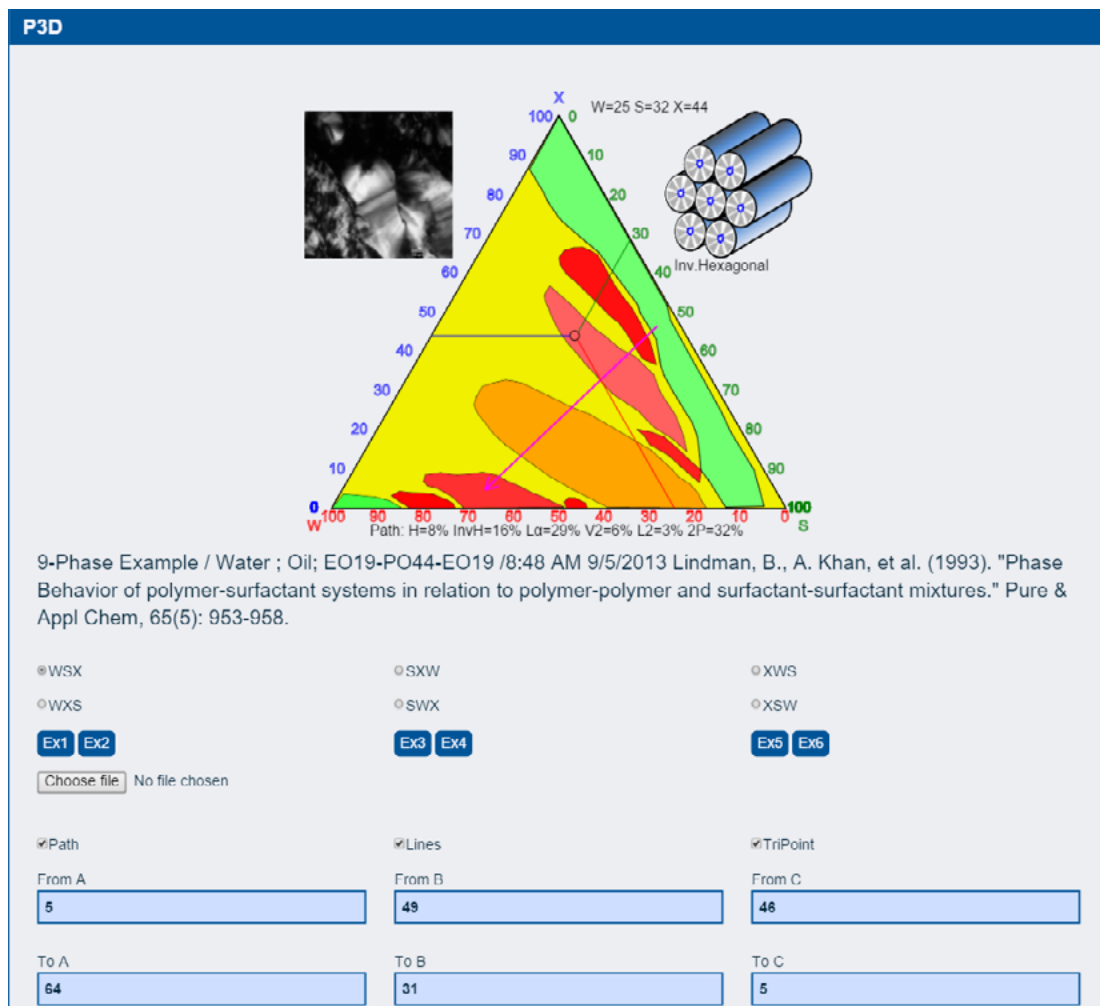
This awesome diagram (loaded from Ex1) features multiple phases and includes a path through the phases. The specific region being examined is a lamellar phase as is clear from the polarizing microscope image.

The amount of work needed to map these phases must have been huge so we are lucky to have such a dataset to explore. Once you can navigate around this diagram, most of those you meet in real life will seem almost trivial.

And that is the point of this chapter. Phase diagrams are a convenient way to pack in a lot of data. It is unfortunate that most of us don't naturally "get" what we're looking at so we derive far too little value from them. With the apps there is no longer an excuse for not understanding these diagrams.

I have to admit that with a year having passed between writing the apps and writing this chapter, I found it distressingly hard to navigate comfortably around them as I had forgotten much that I had learned. It was, therefore, a real pleasure to find that the apps helped me greatly in rebuilding my confidence in extracting information from binary and ternary phase diagrams.

# 8      Putting the principles into practice

Each of the sections below is relatively short. It turns out that with a sound set of principles, the practice is much easier to explain. Although I have not personally had to formulate in every area described here, I can personally vouch for the how much easier it is to formulate in many of these areas once the principles are in place.

## 8.1      Detergency

It is often said that detergents clean mostly via reduction of surface tension. This is clearly nonsense. Something like SLES does a great job at producing a low surface tension, is cheap and safe. So why doesn't everyone use it for all cleaning jobs?

To answer that question we first have to think through the meaning of "cleaning". For those soils that are general "dirt" the simplest way to clean is to make sure that water can get in and do what it does very well – reducing forces between particles. A monolayer of water will reduce van der Waals attractions by orders of magnitude because they have a $1/r^6$ dependency; the difference between a particle-surface separation of 1Å and 2Å is a factor of 64. So a low surface tension helps with quick wetting of the surfaces and with removal of the general dirt particles. Water's high dielectric constant of 80 also reduces charge-charge interactions strongly. And once the particles are in the water, DLVO charge-charge repulsion keeps the soil, which is generally anionic, suspended. There is a tendency for detergents to be basic to encourage particles that are close to neutral to flip into an anionic state for good general-purpose particle-particle repulsion. If the particle has some attraction for the tail of SLES then the sulfate groups help with separation. Non-ionics with modest EO chains cannot provide an equivalent steric stabilisation so are generally less effective, though things like Tweens with large EO assemblies can be good.

Of course there are difficult soils such as proteins and solid greases that are best tackled with chemistry – i.e. enzymes, and others that are best tackled via the oxidants in bleach. Those issues are outside our concern here which is about the rules for detergents that can remove grease.

Returning to the question, the reason we don't all use SLES is that removal of grease and oil is much more about super-low interfacial tensions and curvature than it is about reducing surface tension. Anyone can get a low surface tension with a cheap surfactant. Designing a good detergent system to give super-low IFT is far more difficult.

Another popular idea is that CMC is important for detergency. Again, CMC on its own, is not hugely significant, as a quick calculation will show. Suppose surfactant A has a CMC of 1mM and a $\gamma_c$ of 30 dyne/cm while surfactant B has

a CMC of 10mM and the same $\gamma_c$. If the MWts are comparable around 300 then a 1% solution would contain 30x CMC for A and 3x CMC for B. The surface tensions are the same so how would the user detect a difference if the main driver for cleaning is wetting? At 0.1%, A is still at 30 dyne/cm and B (assuming a typical Langmuir curve) is at 40 dyne/cm. Would the user notice a big difference? If the difference *is* significant then knowing just the two CMC values is of little help without knowing their MWt, $\Gamma$m, $\gamma_c$ and K. In other words, it is not the CMC values we need but the Langmuir curves plotted in % surfactant so we can make an informed decision.

Linked to the CMC explanation is the idea that the oil dissolves in the surfactant tails within the micelles. If you wash your oily hands with just about any soap or hand wash there is a good chance that this effect is in operation. We all use far too much surfactant in this mode of washing. The success of foam hand-wash dispensers which provide a much smaller amount of surfactant shows that most of the time we do not need large amounts of soap – the wetting/dispersing explanation is good enough for general hand washing. Informal experiments with a hand covered in olive oil shows that a foam cleaner is ineffective – it needs the large amount of soap to "dissolve" the oil. Although the "solubility" explanation works well in this case (and presumably in the case of pre-treatment of clothes before laundering) it simply cannot explain how the rather dilute detergent for laundry or dish washing can be effective.

The next most popular explanation of detergency is that it works via the "roll-up" effect. A patch of grease on a surface spontaneously rolls-up into a sphere that detaches itself from the surface. There are beautiful images of roll-up taking place during "cleaning" processes and it is a very compelling idea. Despite the fact that roll-up has been talked about for decades and is often taken for granted, there is a lot of confusion about its importance. To remove the confusion (and come to the not-too-surprising conclusion that it is IFT, curvature, and emulsification that are important) we need to look at some of the basics.

First, because they provide an especially cogent view of basic roll-up and because key features could be captured in an app, it is convenient to follow the approach of two roll-up papers[59] from the Kralchevsky group at U Sofia.

It is not a criticism of these papers that it looks as though any surfactant would be good enough. Their focus is on some very interesting drag behaviour at the interface which in turn raises fascinating questions about the interaction of water and surfactant molecules at the oil/glass interface. Here we just examine the basics.

---

59   V.L. Kolev, I.I. et al, Spontaneous detachment of oil drops from solid substrates: governing factors, Journal of Colloid and Interface Science 257 (2003) 357–363 and Peter A. Kralchevsky, et al, Detachment of Oil Drops from Solid Surfaces in Surfactant Solutions: Molecular Mechanisms at a Moving Contact Line, Ind. Eng. Chem. Res. 2005, 44, 1309-1321

There is also a warning that much of the excellent scientific work on roll-up relies on unrealistic timescales, minutes to hours, with drops of oil of mm radius that may not be realistic analogies to real-world soil. Again, this is not a criticism; it is vital that the fundamentals be investigated under controlled circumstances.



**App 8-1** https://www.stevenabbott.co.uk/practical-surfactants/Roll-up.php
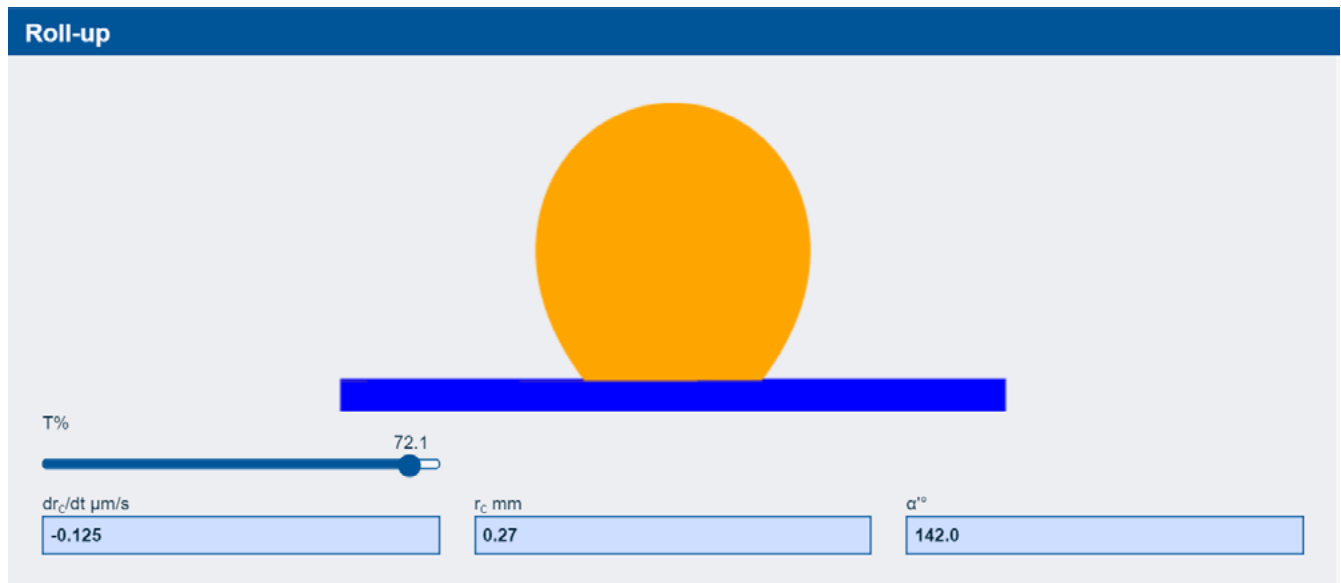
The approach starts with the conventional idea that surface phenomena involve three parameters:

- $\sigma_{OW}$, the interfacial tension between oil and water *in the presence of the surfactant;*

- $\sigma_{OS}$, the interfacial tension between oil and surface;

- $\sigma_{WS}$, the interfacial tension between water and surface.

The act of rolling-up involves a drag coefficient, $\beta$, which depends on the contact radius, $r_c$ of the drop with the surface, which, in turn, specifies a contact angle $\alpha$ which can be derived from $r_c$ and the known fixed volume of the drop, V. Note that the paper uses the "external angle" to define $\alpha$ while in general we are more comfortable with the "internal angle" 180-$\alpha$ , shown as $\alpha$' in the output. The speed of roll-up is described in terms of the rate of change of contact radius with time:

$$\beta \frac{\delta r_c}{\delta t} = \sigma_{ow} \cos\alpha + \sigma_{ws} - \sigma_{os} \qquad \text{8-1}$$

Because we want $\delta r_c/\delta t$ to be negative, $\sigma_{os}$ *must* be larger than $\sigma_{ws}$, i.e. the surface must be hydrophilic. Roll-up cannot take place on a hydrophobic surface.

Why is roll-up a slow process? At the start it can be relatively fast, but as the contact angle approaches the equilibrium contact angle $\alpha_{equil}$ (as it must) a reworking of the equation shows that:

$$\beta \frac{\delta r_c}{\delta t} = \sigma_{ow} \left( \cos \alpha - \cos \alpha_{equil} \right)$$

8-2

So the velocity decreases as the roll-up progresses, though with a sudden burst of speed at the instant of detachment because different drop dynamics are involved.

Their oil was hexadecane which has a viscosity of ~0.003Pa.s. The effective viscosity, β, is ~20Pa.s, though at higher temperatures it fell to ~1Pa.s. It turns out that β depends strongly on the details of the contact line. In their experiments which were on glass, there was a lot of resistance to water coming in to displace the oil; it required the surfactant molecules to move in and produce a glass/tail/head/water interface. However, if water had a chance to form a gel layer of silica-water then the drop could separate faster. This gel formed much faster at higher temperatures, so the effective viscosity was much lower.

The importance of the wetting behaviour of the glass in their experiments is further emphasised when it is revealed that a fresh, acid-washed surface gave such rapid roll-up that they could not study it via their technique.

Now we have explored the basics as generally understood, we can look at a more powerful approach that addresses all the key issues:

- Why do some surfaces show roll-up while others don't?

- Why is a general purpose surfactant much less good than an optimum one?

- Why do some drops give perfect roll-up and others give "snap-off" where a drop peels away from the oil, leaving a smaller drop of oil which in turn may (or may not) snap-off or roll-up?

The starting point is a single, dimensionless number that reflects the balance between gravitational effects (pulling the drop up and creating roll-up) and surface tension effects that can keep the oil on the surface. Some call it B for Bond number, others call it E for Eötvös number (pronounced urt -vursh where the u is as in "hurt" and the r isn't really pronounced). The theory used here[60] was described by Dr Jaideep Chatterjee of Unilever in a series of papers and the specific algorithm used is from the quoted paper.

60   Jaideep Chatterjee, Shape analysis based critical Eötvös numbers for buoyancy induced partial detachment of oil drops from hydrophilic surfaces, Advances in Colloid and Interface Science 99 (2002) 163–179

If we assume a characteristic radius R (chosen to be the radius at the apex of the drop if it were a simple spherical cap shape) and if the oil has a density difference Δρ, and interfacial tension γ then, given gravity g, E is defined as:

$$E = \frac{R^2 \Delta \rho g}{\gamma} \qquad \text{8-3}$$
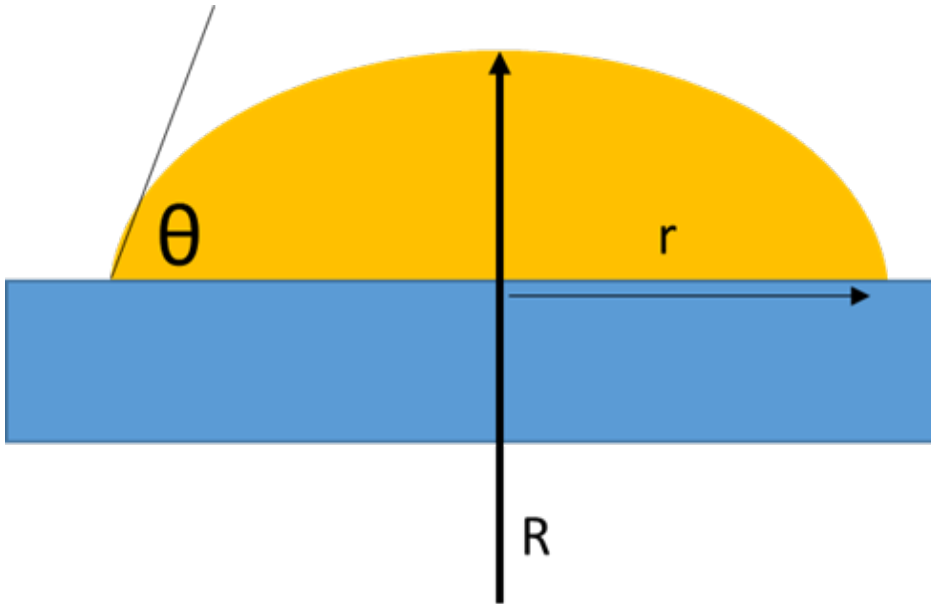


**Figure 8-1 The definition of the contact angle θ between oil and the surface and the two spherical cap radii R and r.**

R comes from the spherical cap formula for an oil drop of volume V with a contact angle θ, defined as the "internal angle" rather than the "external angle" of the Kralchevsky paper:

$$R = \left( \frac{3V}{\pi \left( 2 - 3\cos\theta + \cos^3\theta \right)} \right)^{0.333} \qquad \text{8-4}$$

The contact radius, r, of the drop again comes from spherical cap formulae:

$$r = R\sin\theta \qquad \text{8-5}$$

Using basic physics it can be shown that the drop will spontaneously separate from the surface at a critical value of E given by the following formula where A is somewhere between 3 and 6 depending on the starting assumptions:

$$E_{crit} = \frac{A\sin^2\theta}{2 - 3\cos\theta + \cos^3\theta} \qquad \text{8-6}$$

A plot of $E_{crit}$ immediately clarifies many of the mysteries of roll-up.
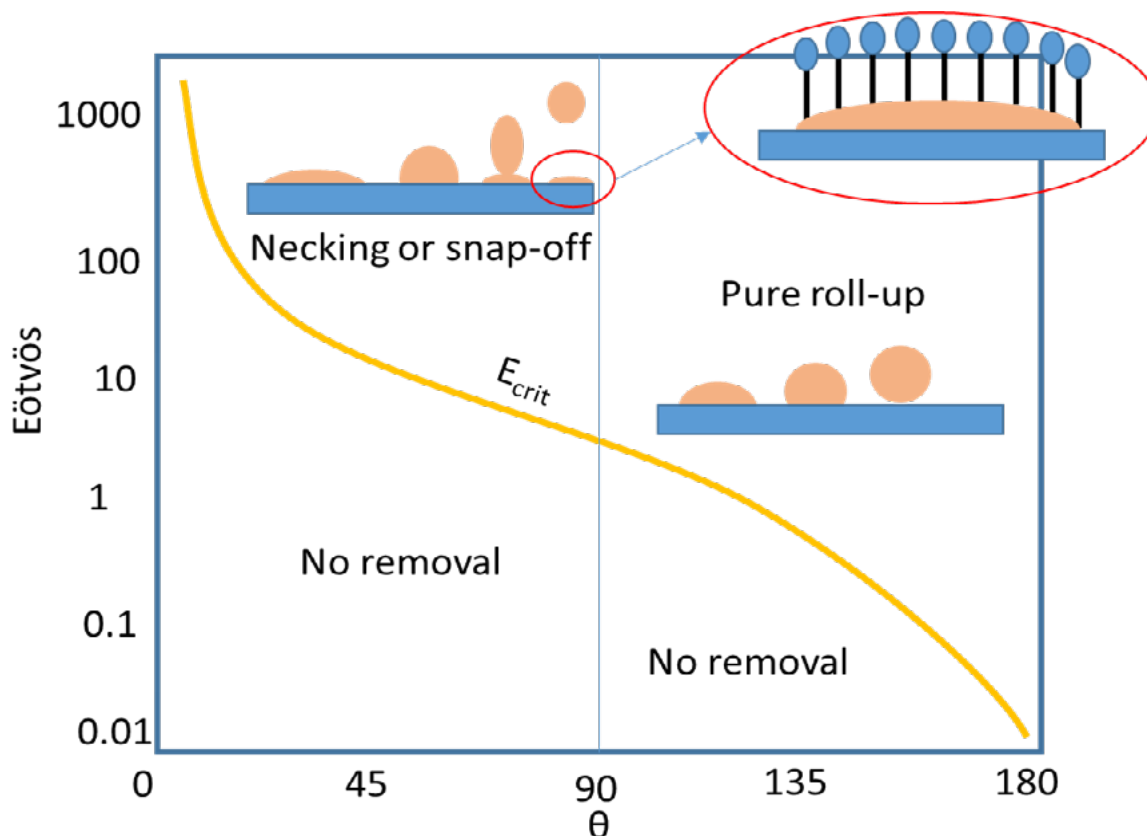
**Figure 8-2 How the Eötvös number influences roll-up and snap-off**

At low values of E there is, unsurprisingly, no removal. But the definition of "low" varies dramatically. For an oil drop on a very hydrophilic surface, i.e. with a high contact angle, $E_{crit}$ is low as it takes very little to remove the oil – which is why it is relatively easy to remove oil from cotton and why it was possible to study rather slow roll-up on ordinary glass but not on acid-treated (hydrophilic) glass where the process was too fast. Once the surface is relatively hydrophobic and contact angles with oil are relatively low, then it needs higher and higher values of E before roll-up can take place. As shown in the diagram, in these cases the chances are higher of getting snap-off rather than roll-off as pinning of the contact line is more likely.

One can get a high value of E with a large value of R, i.e. a large drop of oil. However, although there is an R² term in E, reducing the volume of the oil drop has only a modest effect on E. Because R depends on Volume$^{0.333}$ E has a Volume$^{0.666}$ dependency so halving the volume decreases E by only ~1.6. A large density difference Δρ will also help, but typical oils don't vary all that much.

Given that R and Δρ are not going to change by much, then the *only* way to get high values of E to get good detergency is via low or super-low IFT, which, in turn, can *only* be achieved when HLD~0.

There is a further lesson from the graph. For higher contact angles, classic roll-up is more likely and for the lower contact angles the chances are higher that the oil will separate from itself, giving snap-off, rather than giving pure roll-

up. Chatterjee readily admits that the theory is static, in that it calculates drop shapes and $E_{crit}$ based on the initial contact angle and does not take into account the varying contact angles shown in the initial roll-up app. Experimentally, on real-world surfaces with θ less than 90° there is a high chance of contact angle pinning so the approximations in the theory are more realistic. The magnified part of the image will be discussed shortly, having first explored the app.

Saying that the value of A in the $E_{crit}$ formula "is somewhere between 3 and 6 depending on the starting assumptions" needs further clarification. Chatterjee developed a more profound method for calculating $E_{crit}$ which confirms that A is near 6 for low values of θ and 3 for higher values, with a gentle transfer between them. The method is based on the fact that the drop shape can be calculated. Using a Runge-Kutta 4th order integration of the Young-Laplace equation that takes into account the buoyancy and IFT effects, the outline of the drop can be created and, obviously, the volume of that drop can also be calculated. A key insight is that when the volume of the *calculated* drop is different from the *known* volume then there has been a breakdown in the system. When this is done for a set of contact angles and values of E it becomes clear that this breakdown occurs at the same basic $E_{crit}$ as was calculated from the earlier formula, and that the variation between the factors of 3 and 6 is smoothly covered by this algorithm.



**App 8-2** https://www.stevenabbott.co.uk/practical-surfactants/Eotvos.php
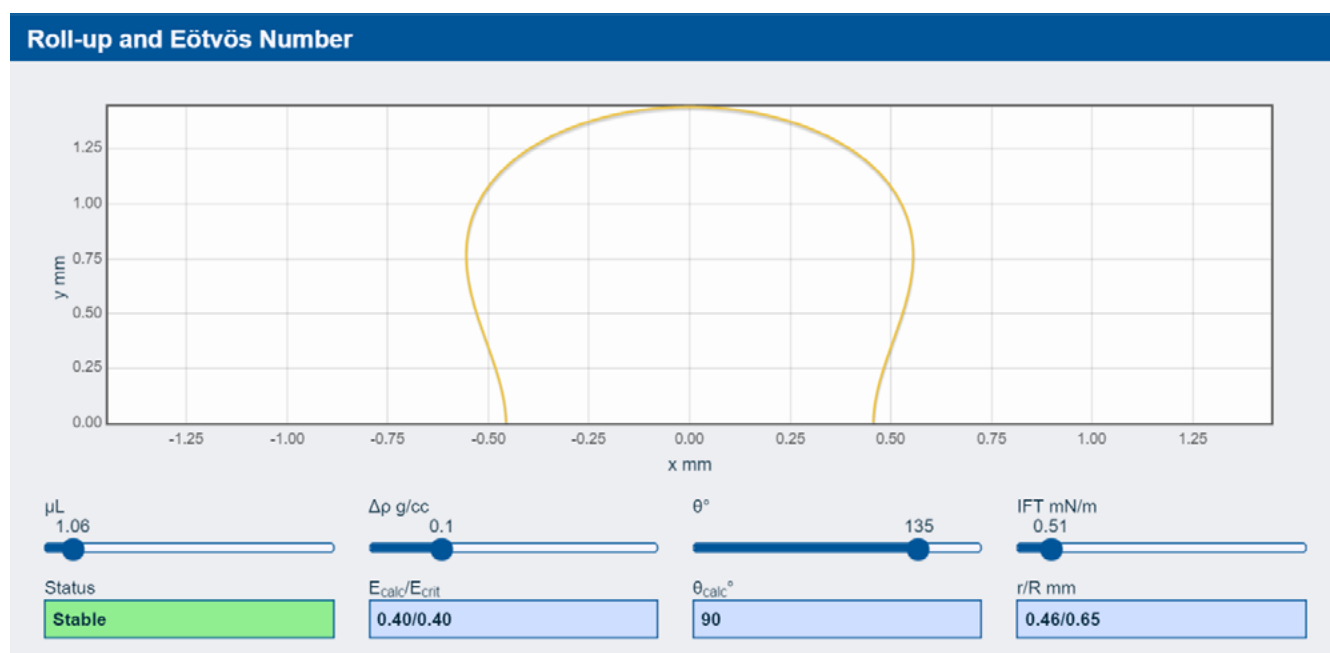
What is remarkable about the drop shape (as the papers point out) is that changing IFT at first makes almost no difference; i.e. normal changes between different surfactants are irrelevant. Only when the IFT falls to seriously small values does the drop shape change significantly. So most of the time the calculated drop volume is the same as the actual volume and the app shows the word "Stable". But when the IFT gets very low, $E_{calc}$ gets close to (the simplified)

$E_{crit}$ and the drop shape starts to go crazy and there is no numerical solution. At this point the app shows the word "Unstable" and the graph may show any number of bizarre shapes which are simply numerical artefacts. This "failure" of the calculation is exactly what we want – it shows where roll-up and/or snap-off are inevitable.

We can now address an issue that has been debated for some decades. Those who have understood that low IFT is important and know how to formulate reliably within that area have reported conflicting results. Some say the best detergency is at the minimum IFT (HLD=0) others say that the sweet spot is just a bit away from it at, say, HLD~-0.2. It turns out that the answer is a bit of both. If you examine the amount of residual oil after the wash process (without the rinse), in some cases the HLD=0 case gives disappointing results compared to HLD=-0.2. But following a rinse in clean, cold water, the results for the HLD=0 case are excellent.

It comes down to the fact that at super-low IFTs it is easy both for snap-off to take place (removing a lot of oil) but also for a nice even, surfactant-covered oil film to remain on the fabric. When this film enters a cold (if ethoxylates were used) saline free (if anionics were used) environment, it is no longer at super-low IFT and the surfactant-rich oil film bunches up and falls off.

The confusion arises because although the best roll-up must occur when IFT is super-low, if the oil is pinned to the fabric (e.g. by micro-roughness) then snap-off can occur rapidly, leaving the residual film. If the same experiments were carried out on a surface with the same initial contact angle but super-smooth there will be less contact angle pinning and therefore a better chance of complete oil removal at super-low IFTs.

Parts of this story have been known for a while, but the combination of the thoughts around $E_{crit}$, contact line pinning, starting contact angle and IFT around HLD=0 have not, to my knowledge, been brought together to create the whole story.
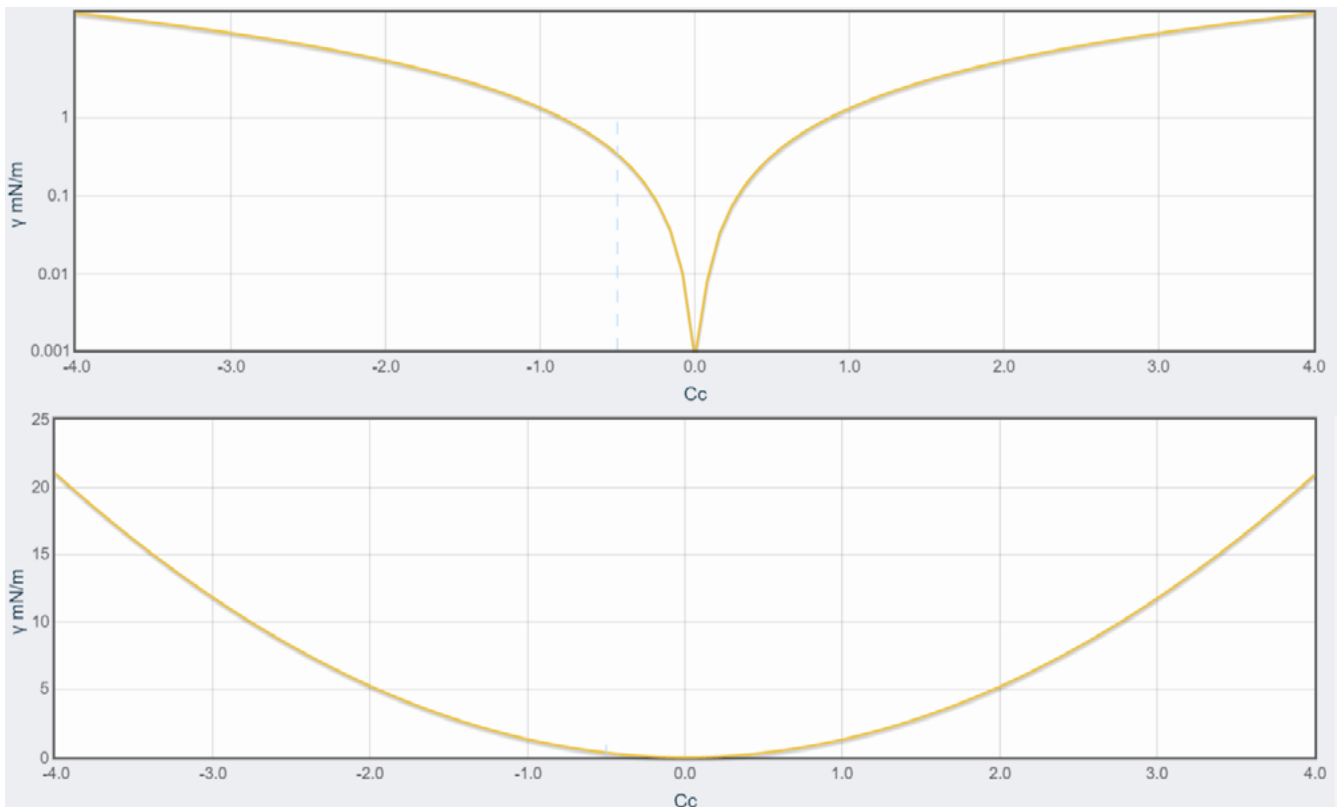
**Figure 8-3 The IFT app with the same data viewed in log and non-log format**

It is worth looking again at images of the app which shows how IFT changes with HLD. In the conventional log plot it looks as though small changes from HLD=0 (e.g. going to the ~-0.2 mentioned above) are catastrophic. But in a non-log plot it is clear that the curve is a gentle parabola, so things aren't quite so critical. Nevertheless, the fact noted above that the drop shape does not change much across a wide range of IFT and only starts to change significantly at low IFT values means that we cannot be too casual about attaining a low IFT.

Another problem with roll-up is the fate of that beautiful spherical drop. What is to stop it from splatting onto another surface, undoing all the cleaning work in one moment? Roll-up says nothing about covering the drop's surface with surfactant molecules to form a rigid barrier against re-deposition.

Given reasonably hydrophilic surfaces such as glass, crockery or a kitchen surface, something close to roll-up can be made to occur with just about any cheap detergent. A quick spray with a surfactant cleaner is enough to give some initial roll-up, and a wipe with a (microfiber) cloth is enough to remove the oil. The forces involved when a cloth comes into contact with surfactant + oil + surface are orders of magnitude larger than the genteel forces of pure roll-up. That is why a quick wipe with a cleaning cloth and surfactant spray is remarkably good. It requires microfiber cloths and/or sprays containing particles to get really good cleaning. This is because a normal cloth provides very little surface contact so very little physical cleaning. A microfiber cloth gives much more contact, as do the particles inside a modern household cleaner. I am old enough to

remember the first time the particle-containing cleaners came out. Their cleaning efficacy was astonishing (though at the time they also physically scratched the surfaces so formulators had to develop kinder particles). I also remember the first time I scientifically compared the cleaning of a conventional and microfiber cloth – the difference was astonishing.

We are almost there with a full explanation of how to get the best detergency from choice of surfactant. We know that any surfactant will provide adequate wetting and that any anionic surfactant is, in general, going to do a decent job with particulate soils and for hydrophilic surfaces you don't have to try too hard to remove oil. Once it comes to hydrophobic surfaces, low IFT can give great roll-up and snap-off, and after snap-off the low IFT system gives a good coverage of the remaining oil. This then raises the question of whether the low IFT system can provide direct emulsification of the oil without the need for roll-up? This is distinct from the case of direct emulsification by applying a large amount of not-very-good surfactant discussed earlier. The answer is that the distinction between emulsification and low-IFT roll-up is fuzzy. In a sense, we don't care because the results are hard to distinguish scientifically or practically.

So if you want good detergency you arrange that at the washing temperature you have super-low IFT. Because, for other reasons, we need anionics, the choice is between adding an ethoxylate system tuned for low IFT (Phase Inversion Temperature) at the chosen wash temperature, adding enough salt (via "builders") to allow the anionic to be at HLD=0 or, if you want to use APGs, just add APGs that give HLD=0 (or slightly less if you are concerned the super-low IFT doesn't remove all the oil in the wash step) under the wash conditions. In practice we see that many detergent labels inform us that the contents are "5-15% anionics and 5-15% non-ionics", so the formulations are a bit of both.

To get efficient detergency requires a system that creates high $\xi$ values. In academic labs this can readily be done via extended surfactants. In practice there are currently issues with the biodegradability of the propylene oxide chains in many of these molecules. If/when those problems are solved then there is a chance for significantly enhanced detergency for smart formulators.

If good detergency were just about solubilization then there would be little competition possible between suppliers of household detergents – they would all use essentially the same blend, depending on the required washing temperature. Like reduced surface tension, solubilization is necessary but not sufficient. And we already know the reason why, and have the principles to solve another part of the puzzle.

As discussed, although HLD=0 is the optimum for solubilization, it is also the minimum for emulsion stability. So any oil taken up by the surfactant has a good chance of being re-deposited on the dishes or clothes. That is one reason why

the best detergency is often found at HLD slightly less than 0. And that is why commercial detergents tend to contain polymeric surfactants.

Polymeric surfactants have awesome solubilization powers. In other words, they have massive ξ values. This isn't surprising – ξ is about coherence length and about stiffness, and polymeric surfactants provide both of those. So why don't we just use polymeric surfactants for everything? It is the same problem as with the extended surfactants. Their thermodynamics are wonderful, their kinetics are awful. They are far too slow at the main task of solubilization. Smart designers of detergents therefore aim to get the best of both worlds. They use the conventional small-molecule detergents to get the oil off the surface then rely on added polymeric surfactants to trap the oil to avoid re-depositing.

Other polymers are added not so much for their surfactancy but because they can provide DLVO-style charge or steric stabilisation. Anionic cellulosic polymers stick to cellulose fibres and repel anionic dirt particles. PET-PEO di-block polymers attach to PET fibres and provide a steric stabilisation via the PEO. And polyacrylics are the workhorse anionics for sequestration of unhelpful Ca ions, for charge stabilisation and for avoidance of redeposition.

If we ignore all the other important issues of cost, enzymes, bleaches, fragrances etc. there is just one final problem for the detergent designer. At some stage the detergent is going to be a concentrated liquid or a solid and in final use it is going to be dilute. As we saw in the phase diagram chapter, it is possible for a modestly viscous surfactant solution to be diluted with water to a hoped-for low viscosity solution but then get trapped in some hexagonal or cubic phase where it sits around in big globules that annoy either the production line or, even worse, the customer. These problems have become more severe as manufacturers have offered greener product with less water or filler. It is doubly sad that their sincere (and largely successful) efforts to go greener have been met with suspicion by the consumer (who think they are getting less for their money) and general ignorance amongst the scientific community about how tough a challenge the phase problem has been.

## 8.2    EOR

We have already discussed that the problem of pushing oil out of the ground with water pumped from the other side of the well is one of exceeding the "critical capillary number", Ca, below which the water simply passes through channels in the oil, achieving little.

$$Ca = \frac{U\eta}{\gamma}$$    8-7

Because capillary number depends on velocity U, viscosity η and interfacial tension γ, the oil company can choose to pump faster, to add polymers to increase the viscosity or to decrease IFT. If the oil can potentially rise under its own buoyancy then another dimensionless constant, the Bond number B appears. The physics of this is very similar to that of roll-up in detergency and the Eötvös number used there is simply B/2. For a density difference Δρ and drop radius r:

$$B = \frac{2g\Delta\rho r^2}{\gamma}$$  8-8



**App 8-3**  https://www.stevenabbott.co.uk/practical-surfactants/Removal.php

The ability to recover the oil, i.e. the fraction φ remaining in the ground, depends both on Ca and B via a number shown as CB in the app and often called the Trapping Number TN in the literature. CB=(Ca² + B²)$^{0.5}$. The dependency follows a typical sigmoidal curve, frequently fitted via a Van Genuchten equation based on $\varphi_{hi}$ which is the volume fraction at very high values of CB, $\varphi_0$ which is the original amount in the rock, a critical value of CB where things change rapidly and a power law number n which controls the sharpness of the transition:

$$\varphi = \varphi_{hi} + \frac{\left(\varphi_0 - \varphi_{hi}\right)}{\left[1+\left(\dfrac{CB}{CB_{crit}}\right)^n\right]^{1-\frac{1}{n}}}$$  8-9

Although this is simply a fitting equation, it has its roots in the physics of the complex flows of liquids through porous media. $CB_{crit}$ can be estimated (assuming full wetting of the rock) from the parameters of the pores in the rock such as the radius of the globules, $r_g$, and the radius of the passages connecting the holes, $r_p$, and the IFT $\gamma$. The starting point is that the pressure difference across a pore needed to force out an oil globule is:

$$\Delta P = \frac{\gamma \left(1 - \dfrac{r_p}{r_g}\right)}{r_p} \qquad \text{8-10}$$

Given a permeability term k (in fact made up of the intrinsic permeability of the rock and the relative permeability of the water) and an average length L of a globule, this translates into:

$$CB_{crit} = \frac{2k \left(1 - \dfrac{r_p}{r_g}\right)}{L.r_p} \qquad \text{8-11}$$

The power law dependency, n, will depend on the variability of the pores in the rock. However, in general the data are just fitted to a Van Genuchten curve because the key parameters are often unknown and procedures for extracting $CB_{crit}$ and n from the raw data are well-established.

For each specific well at each specific stage in the depletion of the oil, the tradeoffs are different. So one well might require expensive high-speed pumps, another expensive high pressure pumps (for the highly viscous polymer solution) plus the cost of the polymers, another might choose modest speeds and viscosities coupled with super-low interfacial tension and the cost of the surfactant.

For those who choose the surfactant-based route, the decision on the surfactant blend and the amount to be added is of $multi-million significance. It takes weeks or months for surfactant pumped in to reappear at the well-head, with or without lots of oil. So it is no surprise that the most enthusiastic users of HLD-NAC are the EOR industry.

The complexity of the system is enormous compared to the simple ideas of measuring EACN or Cc values. The oil is a mixture which itself contains surface-active agents such as asphaltenes and naphthenic acids. Salinities are often high and contain complex mixtures of cations, including calcium and magnesium which can cause problems with certain surfactants. Temperatures are high. And the surface of the rock can potentially attract the surfactant molecules, rendering

them useless, though in general the anionic surfactants (sulfonates) are not so badly affected.

Another interesting issue is pressure. A paper from Penn State[61] explores the effects of pressure on HLD. It is no surprise that the direct effects are modest, though some wells are at enormous pressure so the HLD shift cannot be neglected. More importantly, "live" oil contains methane gas. If the oil is measured at normal pressure, the methane is no longer present so the EACN is that of the oil. Down the well, the methane (which has EACN=1) reduces the EACN of the oil and this "pressure" effect is significant.

The EOR market is of huge significance for the surfactant industry and this has benefitted the rest of us in two ways. First, it has funded a large amount of the HLD work over the decades. Second, because it has driven the need for more effective surfactants, with the extended surfactants emerging as a logical solution to the need for long surfactants without the difficult liquid crystal phases that arise with long alkyl chains.

The key problem for most users of surfactants is that price is dominant over innovation. Any new surfactant would need to be produced on heroic scales for the price to fall to a level that would allow users to afford it – which means that the number of new surfactants is very small. We can only hope that the vast scale of surfactants used in EOR will mean that their price/performance will allow them to be used in other markets that can use their higher efficiency while accepting some problems of kinetics.

### 8.2.1  Soil remediation

There is an obvious parallel between EOR and the removal of water-insoluble pollutants from soil. Water+surfactant has to be pumped through the soil and by far the best way to encourage the pollutant to travel with it is to arrange for HLD~0 for the best prospect of low interfacial tension and solubilisation. The app shown in the EOR section is called "Removal" because the theory can be (and is) applied to soil remediation situations.

The problem is that the EACNs of many of the chloro pollutants are in the low-to-negative region so very different surfactants are required compared to EOR. Making the process cost effective means identifying surfactants available in bulk for other applications. The common anionics have Cc values rather too low, though if the soil is saline this might push them into the right region. It is rather sad to review the many soil remediation papers that have no idea of HLD. With little guidance as to how to steer the formulation a lot of expensive holes in the ground have been pumped full of surfactant to little effect. Those more familiar

61   S Gosh, RT Johns, A New HLD-NAC Based EOS Approach to Predict Surfactant-Oil-Brine Phase Behaviour for Live Oil at Reservoir Pressure and Temperature, Society of Petrochemical Engineers, SPE-170927-MS

with HLD at least have good formulation hypotheses and merely have to cope with the many other practical difficulties of soil remediation.

One popular set of general-purpose surfactants that work well for contaminants over a large range of EACNs are the sulfosuccinates. AOT, the ethylhexyl (or octyl) version has a high Cc for hydrophobic (large EACN) contiminants. Dihexylsulfosuccinate is close to Cc=0 so can be combined with AOT to cover the Cc=0-2.3 range. Diamylsulfosuccinate is more hydrophilic so can be used for the chloro pollutants.

### 8.2.2 Dealing with oil spills

If you have to deal with large oil spills there are many competing interests. Microbes are remarkably good at spotting a great source of food and left on their own much of the spill will disappear naturally. However these processes are too slow for those who have to cope with populist pressures to clear the spill quickly. The problem can be made to go away visually if the oil can be rapidly emulsified and a rough calculation based on the high salinity of sea water and the high EACN of typical oils shows that AOT is likely to be (and in fact is) an especially good starting point. Although the safety profile, by surfactant standards, is good (I have eaten some AOT for scientific research – it tastes awful) you can always find someone who will claim that it will kill something. In addition, an efficient emulsifier/solubilizer simply makes the visible problem disappear – the oil itself is still around. It then becomes a question of whether the rapid dilution of the oil simply spreads the harm further or whether it can actively minimise damage. Fine oil drops will be disposed of more quickly by microbes and might be broken down faster by the elements, thanks to the huge surface area, so the gross contamination of wildlife is avoided. But the fine drops might equally make it easier for the complex aliphatic and aromatic chemicals to enter smaller organisms causing potential direct harm to them and indirect harm further up the food chain.

There are no easy answers, though on average making the visual problem disappear is going to be more popular than the alternatives. Pelicans covered in oil have been icons of oil spill disasters. When the disappearance of the oil is 100% natural (there are cases where a vast storm has appeared and broken up the spill through massive energy input) there are few complaints, when it is a technical decision (e.g. AOT-based formulations on a famous major oil spill) there will be plenty of post-facto critics.

But if the decision is that emulsification is the least-bad option, it needs a good reference guide to the science[62] and a lab with High Throughput capabilities to be supplied with large amounts of the actual oil and sea water plus information

---

62   Edgar J. Acosta and Suniya Quraishi, Surfactant Technologies for Remediation of Oil Spills, Ch 15 in Ponisseril Somasundaran, Partha Patra et al, Oil Spill Remediation: Colloid Chemistry-Based Principles and Solutions, Wiley, 2014

on temperatures, so that some phase scans can quickly be conducted to find where HLD=0. The number of cost-effective, efficient (large ξ) and safe surfactants available is going to be small and it is probable that a blend of less satisfactory ones will be needed for optimal effect. Although this intelligent approach is not guaranteed to please everyone, it is guaranteed to reach the quickest set of technical options for those with the political, financial and environmental responsibilities to make an informed choice.

Readers may notice a certain evasiveness in my language in this section. The political, legal and popular internet emotions that swirl around oil spills do not encourage frank scientific discussion.

## 8.3    Cosmetics emulsions

Consumers are far from rational. At the time of writing, on one side of the Atlantic sulfonate surfactants are bad, on the other side ethoxylated surfactants are bad. In both cases the worries are non-issues boosted by credulous internet stories. Similarly, there is an irrational love of "natural" cosmetics and, by extension, "natural" surfactants.

The background angst about "chemicals" has also given the general impression that surfactants must be intrinsically evil. So many formulations are based on "emulsifiers"  or even "hydrotropes" which sound altogether more friendly.

At the same time, a key ingredient, mineral oil, is being increasingly demonised and various exotic-sounding oils are being added for no benefit other than that of announcing that the cosmetic is made from some obscure plant found in the wilds of the Amazon forest. Given that all these oils are mixes of triglyceride from C6 to C20, with and without some unsaturation, the chances that oil A is better than oil B are slim, though the "feel" of the product with oil A may indeed by superior to that from B, though why that might be so is usually unclear.

So the formulator is surrounded by irrational demands on the raw materials which in turn can be of highly variable quality. A "good" oil today might be an "evil" oil tomorrow if some famous personality has discovered that extraction of the oil is destroying the lifestyle of orang-utans.

Then there are intense cost pressures on formulators. Although cosmetic products project luxury and glamour, the formulators are always under intense pressure to use the lowest-cost raw materials which might mean less sophisticated surfactants and greater batch-to-batch variability.

Finally there is the issue of efficacy. Because I have been strongly involved in the science of predicting skin permeation of formulations I have had many opportunities to confirm what I had previously only suspected – that many cosmetic formulations are perfectly set up to ensure that essentially none of the

active ingredients penetrate the skin. One of the many reasons for this is that even the modest levels of surfactant in the overall formulation become a high proportion of the formulation on the skin once the water (mostly) evaporates. A highly ethoxylated surfactant such as a Tween provides a good solubility home for many actives so they prefer to stay on top of the skin with the Tween than to migrate into the skin.

All these factors lead to a common conclusion: formulators need to get the maximum emulsion from the minimum surfactant.

The starting point is to know what the absolute minimum could be. This can be calculated readily from the emulsion surface area app that we saw previously:



**App 8-4** https://www.stevenabbott.co.uk/practical-surfactants/EmuSA.php
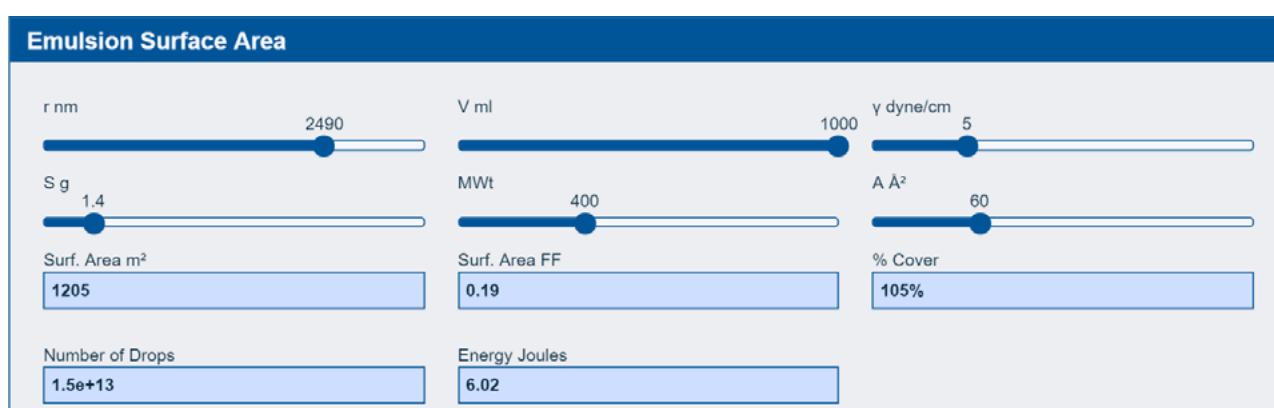
In this example, drops of 5μm diameter of a surfactant with a MWt of 400 and a head area of 60 Å² needs 1.4g/l to provide total coverage of the emulsion drops and, therefore, provide the greatest chance for stability.

This simple calculation is of great significance for the formulation team. What size drop is required? Smaller drops will tend to give a finer "feel" and are more stable against creaming (see the creaming app for the $r^2$ dependence on creaming velocity). But halving the radius doubles the amount of surfactant required – a big effect on the raw materials budget. Nanoemulsions are attractive in many ways, save for the large amount of surfactant required to cover the drops. For a given drop size doubling the MWt of the surfactant halves the effective number of surfactant molecules so doubles the weight of surfactant required. If surfactants were priced per mole this wouldn't matter, but they are priced per ton, so the smallest possible MWt is required. However, if halving the MWt also halved the head area it would give no benefit because large heads give greater coverage. Everything in formulation is a trade-off so it is good for the team to get an idea of these basic trade-offs with a simple app.

There is another trade-off. Although a large head area allows fuller coverage, as we have seen many times it leads to a less well-packed interface so the

elasticity and general toughness tends to be less and the interfacial tension tends to be higher, which means that it is somewhat more difficult to create a fine emulsion.

The calculation of efficiency assumes that all the surfactant wants to be at the interface. If a very hydrophilic surfactant is used then it will tend to be wasted within the aqueous phase. This leads directly to the issue of the optimum Cc value. A surfactant like SLES has a Cc <-2 so is completely unsuitable for providing any form of solubilization or low interfacial energy with respect to the oil and is likely to be found "wasted" in the aqueous phase. This is a problem with many of the common ionic surfactants. It needs something like AOT or its close analogue dihexylsulfosuccinate to be in the right sort of area for a closer HLD match to a typical oil. The simple CxEy ethoxylates are easier to obtain in the correct HLD area, but some of the sugar-based ethoxylates are at wild extremes of high and low Cc, making them unsuitable without intelligent Cc blending. The ethoxylates, of course, allow the standard PIT trick as long as you are happy to heat your formulation to 50 or 60° and then to cool it again. Those who use waxes might choose deliberately to use such heating cycles in order to melt the wax to incorporate it into the emulsion; others might want to choose different PIF techniques to avoid the energy cost.

The APGs are one of the few "natural" surfactants with Cc values that are already in the right range for efficient formulation. Because they are now used on a large scale they are relatively cost effective. Their big drawback to those who only know about PIT is that they have no significant change of solubility with temperature ($\alpha \sim 0$ in the HLD equation) and salinity effects are also too modest to allow saline PIFs. So those who want to formulate with APGs have to find other ways to create efficient emulsions via HLD=0 inversions (e.g. clever Cc mix tricks) or via the 25/75% inversion regime which is harder to master but potent in the hands of those who know how to exploit the effects.

The section on detergency discussed how useful polymeric surfactants can be (via their enormous ξ parameters) whilst also seeing how useless they are in kinetic terms. Detergent formulations cleverly combine the kinetics of small surfactants with the thermodynamics of the polymeric surfactants to stop redeposition. The smart formulator of cosmetic emulsions should be able to use an inversion regime with normal surfactants then "fix" the emulsion with a modest amount of a potent polymeric surfactant.

So far we have only considered the surfactant. The oil is another area requiring knowledge of the various trade-offs. One of the sad facts of life is that it is difficult to influence large oil molecules, i.e. the ξ value for a given surfactant is lower for high MWt oils. So the formulation team has a strong interest in aiming for the lightest possible oils consistent with the other demands of the end product. For a given (average) MWt the other parameter to be tuned is the EACN. If the surfactant blend cannot be moved for other reasons, then getting

the right HLD depends on optimizing the EACN. This would be rather easy to do if it were not for the polar oil problem. It is commonly observed that formulations need a dash of, say, dodecanol (lauryl alcohol), to fine-tune them. As we saw in the polar oils section, it is unclear whether the dodecanol is an alcohol, an oil or a (poor) surfactant. It might even be a "linker".

Although the EACN/polar oil problem has no easy solution, the practical implications are that for your particular system it is wise to use phase scans to see what effect the potential additives might be providing. It could simply be changing HLD or it might be leaving HLD unaltered while changing the solubilization (changing the phase volumes) for better or worse. Once you (or even better, your robotic systems) are set up to do routine phase scans much of the conflicting folklore of your favourite additives will be changed into science. An additive that changes just HLD or just solubilization might be easy to use; one that changes both might provide plenty of confusion.

The viscosity effects of the oil in terms of conventional emulsification can be understood using the emulsion drop size modeller [https://www.stevenabbott.co.uk/practical-surfactants/EDSM.php](https://www.stevenabbott.co.uk/practical-surfactants/EDSM.php). When the oil/water viscosity ratio exceeds ~4 then the effects described in the section on Critical Capillary number start to make emulsification via dispersion impossible. And the viscosity effects in the inversion regime which change the transition point going from W/O to O/W have been discussed in the inversion section.

Those who want to stabilise their emulsion using the high viscosity of lamellar phases (which in turn appear at a specific salt concentration) are following a long tradition. It seems to me to be an odd way to do things as it restricts the surfactants that can be used and requires a relatively high % of surfactant to do the job of a traditional thickener. Of course, for shampoos which have to contain relatively high concentrations of surfactant this trick is a supremely cost-effective way of giving the high viscosity that consumers associate (wrongly) with a superior product. For those who want to use such tricks, the various phase diagram apps are of use, though because of the need to map the strong effect of added salt it will need multiple diagrams to make sense of the complex space. Stacking individual ternary scans on top of each other creates a prism scan diagram for those who have the software to produce them. As no one has (yet) demanded a prism scan app and as the task seems difficult and as I would need a lot of raw data for illustrative test cases I have not written a relevant app. That might change if someone would like to help.

There is a further use for exploring phase diagrams. It is relatively easy in the lab to create an emulsion formulation using a small-scale disperser and adding the ingredients in any convenient manner. The needs of production may require the different ingredients to be added in a different manner and dispersed at different ratios and speeds. If you have a reasonable grasp of the phase diagram then there will be fewer surprises in the transition from lab

to production. In particular anything which provides viscosities and/or phase volumes near a phase inversion point (25/75%) will be prone to the creation of the wrong emulsion during that part of the production process. Sometimes the subsequent steps effortlessly undo these wrong phases. More often than not they create chaos and confusion, especially for those who don't have a grasp of the "Salager" inversion diagrams.

The lab can offer a rather powerful QC tool to production. Assuming that phase scans are routine, each batch of surfactant (blend) or each sample of a substitute blend eagerly offered by procurement because it is "the same stuff but 10% cheaper" can be tested for Cc (surfactant) or EACN (oil). Any significant changes provide an early warning of trouble ahead. The techniques are amazingly quick and simple for looking at the heart of raw materials issues. Identifying a minor shift in the chain-length distribution of a C10-14 ethoxylate is difficult, identifying that such a shift has changed the Cc by 0.3 is easy, and allows a rational reformulation via the other HLD parameters to compensate. If suppliers of surfactants or oils did such tests routinely they could fix the issues in their plant, saving everyone a lot of formulation trouble.

This description of the science of cosmetic emulsion formulation does not immediately lead to a magic formula that will guarantee success. The aim is to give the formulation team a background of solid scientific principles that will stop them from wandering through surfactant space without a map or compass. Observers of cosmetics formulations often note that they are all the same. There are a number of positive reasons for this. But a strong negative reason is that most formulations are based on previous formulations that worked, for reasons that no one is sure of. There is often great fear of changing anything within a formulation (things often get added but seldom removed from cosmetics formulations) because the likely result is a failed emulsion for reasons that no one understands.

With the principles described here, the formulator can start exploring new territories with some confidence, removing useless stuff from old formulations and trying out different surfactants and their blends with a better grasp of what the outcomes might be. The approach provides tools for rationally tweaking a modestly successful formulation so that it is excellent in the lab and capable of being put into production with, hopefully, less surprises and lower energy costs. All this can be achieved using less surfactant used more effectively.

And if your suppliers are not providing you with Cc and EACN data, start demanding it, if for no other reason than to make sure they stick to their Cc and EACN specifications thereby reducing the uncertainties in the formulator's life.

## 8.4     Emulsion Polymerisation

I confess to not properly understanding the old ways of making emulsion polymers via "macroemulsions". Although one might think that an emulsion polymerisation is simply the polymerisation within an oil drop, the classic method often involves water-soluble initiators and the polymer particles tend to grow from small seeds rather than via polymerisation within an oil drop. The stabilisation of the growing particle is therefore a difficult issue as it has to steal surfactants and dispersants from the emulsion drops which themselves are merely reservoirs of monomer for the growing polymer chains. The final particle size is not directly dependent on the emulsion drop (i.e. 10µm drops do not create 10µm polymer particles) and, indeed, macroemulsion polymerisation tends to produce particles over a wide range of sizes. Clearly the sorts of theories required are much more about nucleation, diffusion, partitioning and DLVO-style stabilisation than about "emulsions". One way or another these complex processes have been optimised for producing emulsion polymers on gigantic scales and the sensible conservatism of those large-scale producers is not going to do much to change processes that deliver great products at affordable prices.

A key plus point for the macroemulsion method compared to the naïve view of how emulsion polymers are formed is that initiation of a polymer chain is a relatively rare event so the MWt of the polymers tends to be satisfyingly high. This point will be revisited shortly.

Instead, this section is devoted to the more modern micro- or mini-emulsion techniques that are conceptually simple: an emulsion drop is formed, the initiator starts the reaction in the drop and the drop becomes polymerised. Miniemulsions are defined as being sub-micron size and microemulsions tend to be below 100nm, though there is no obvious boundary in size. Confusingly, those outside the emulsion polymerisation world would tend to call emulsions in the few 100nm size range "nanoemulsions". So we have "microemulsions" that are truly nano, "nanoemulsions" that are mid-nano, "miniemulsions" that span the range of nanoemulsions up to macroemulsions and "macroemulsions" that are micron sized so should be called microemulsions.

Although simple in principle, the practice is full of confusion for the usual reason that most practitioners have little idea of HLD-NAC theory. It is easy to find emulsion polymerisation papers that refer to HLB. It is hard to find any that refer to HLD.

The naïve view of emulsion polymerisation within miniemulsion drops would assume that the initiator is oil-soluble so initiates from within the drop. In practice this generally leads to low MWt polymers because of multiple initiations of separate chains, so water-soluble initiators which are much less efficient in

initiating polymerisation within the drops are the norm to ensure higher MWt products.

In macroemulsions the concentrations of initiator and (optional) inhibitors have a significant effect on the outcome of the polymerisation because there are so many competing equilibria. For miniemulsions the final properties have only a small dependence on starting conditions (assuming a good emulsion has been formed) so provide a much more reproducible outcome with far less black art.

So we can sweep away all the other considerations about miniemulsion polymers and focus on the one thing that has a huge effect on the outcomes – the emulsion itself.

Because the general literature on miniemulsification is so unsatisfactory due to its lack of understanding of HLD-NAC, it is good to start with one principle that is used extensively and is well-founded in theory.

If you prepare a nice miniemulsion then proceed to polymerise it, there is a good chance that the outcome will be polymer particles much larger than the original miniemulsion. The cause, of course, is Ostwald ripening and we already know a root-cause way to stop the smaller particles getting smaller while the bigger ones get bigger. By adding a few % of something that is highly insoluble in water while being adequately soluble in the monomer, Ostwald ripening comes to a halt due to the osmotic pressure difference between drops with smaller (large drops) and larger (small drops) concentrations of the additive. In the lab the standard additive is hexadecane. 1-2% is more than enough to stabilise the miniemulsion for the few hours needed to make the polymer. Because hexadecane has no desirable contribution to make to the end product, in commercial emulsions it is common to add one of two things. The first option is a water-insoluble monomer. In the case of acrylate polymerisations something like stearyl (octadecyl) acrylate can be used. The second options is a water-insoluble polymer. For emulsion polymerisation of methyl methacrylate, 1-2% of PMMA dissolved in the methyl methacrylate is good enough to provide the needed stability against Ostwald ripening.

After that, confusion is the norm. Surfactants that are great for one monomer seem to be useless for another. Magic additives (often they are long-chain alcohols) that solve a problem for one emulsion fail to solve a similar problem for another emulsion. The effects of temperature are even more baffling, especially because there can be large temperature changes between making the emulsion (say in a disperser) and the polymerisation in a vessel which has to provide heat to initiate the process then cooling to restrain the exotherm. Guaranteeing a uniform temperature throughout the batch is tricky without strong stirring, but the viscosity of the emulsions can be high and vary throughout the process so it is difficult to ensure an even temperature distribution.

Using HLD-NAC the situation becomes much simpler. Via HLD it is obvious that a surfactant system that works well with one monomer will not work well with a different monomer if the EACNs of the monomers are different. So the first thing that emulsion polymerisers need to do is measure the EACN of their monomer.

Next comes a crucial decision on which surfactant type to use. Ionics are attractive for those who wish to stabilise their emulsion via the charge mode in DLVO. Long-chain ethoxylates such as Tweens are attractive for those who wish to use steric stabilisation. In both cases the stabilisation is required for the emulsion and the polymerised emulsion (now a colloidal or nanoparticle dispersion) and the criteria will depend on issues such as end use. Those who wish to use "green" surfactants might want to use APGs, though it is not clear how these relatively small, neutral molecules can provide adequate colloidal stability.

Having made the choice of surfactant type, it is then necessary to tune the HLD to be not too far from zero. This will make it easy to create small drops without heavy duty (power-hungry) dispersers. From the EACN of the monomer and the required emulsification temperature it is easy to choose the Cc to give (for O/W polymerisations) HLD~-0.3. Finding a blend of affordable/compatible surfactants may not be easy, but not finding it makes life very hard.

If it is truly impossible (for various reasons) to get the right HLD via surfactants then the intelligent use of polar oils is required. Because we know that the rules of polar oils are not too clear, we need to use phase scans intelligently to see what effect each oil is having. It might simply be changing HLD, it might be acting as a lipophilic linker, it might be acting as a co-surfactant or might be doing a combination of all three. There are no rules to say which mode is best for a given application. The only rule is to understand (as the emulsion polymerisation world do not) what the polar oil is doing so the level can be optimised, or a different polar oil chosen because its balance of properties is superior.

When we come to look at microemulsion polymerisation we will see that polar oils add greatly to the confusion in an already complex situation, so formulation via a rational surfactant blend is highly recommended.

If the emulsification temperature is the same as the polymerisation temperature then the HLD balancing job is done. If the temperatures are very different then the choice of surfactant type has a huge impact on what happens at the emulsification temperature. If APGs are used, the great news is that the temperature makes no difference – the system is easily controlled. If an ethoxylate is used and the starting HLD is somewhat less than 0 then there is a danger that at the higher temperature the HLD will become positive with the risk of inverting the emulsion – a catastrophe. If an ionic is being used then at the higher temperature the HLD will be somewhat more negative so the surfactant

system might be a little too inefficient, though it will probably be OK. As a precaution it might be best to start the ionic HLD rather closer to HLD=0.

It is no surprise to find that some miniemulsions are formulated via catastrophic inversion, taking advantage of the fact that the interfacial energy is minimal so fine emulsions are created with far less dispersion energy. Once the miniemulsion is formed, the rest of the process is the same as before.

Viscosities throughout the process can readily be modelled using the Yaron, Gal-Or app discussed earlier: https://www.stevenabbott.co.uk/practical-surfactants/EmuVisc.php.

Guided by these principles, making basic emulsion polymers will be very much easier than via the normal methods that attempt to navigate without map or compass. Once basic polymerisation has been mastered it is time to address another key compromise.

In general, finer emulsion particles are preferable, but because halving the radius doubles the amount of surfactant to cover the surface, there is an obvious downside of having too much surfactant in the final formulation. As with detergency and cosmetic emulsions, the intelligent option is to harness the efficacy of polymeric surfactants (or any other surfactant with a large head area) while coping with the fact that their kinetics are generally very poor. Attempting to do this before mastering the basic techniques is likely to lead to disaster. Cautious addition of intelligently chosen polymeric surfactants at the right stage in the creation of a well-controlled (via HLD) formulation is likely to be far more successful. For classical polyvinyl acetate emulsion polymers the use of polyvinyl alcohol as surfactants/dispersants has been a notable success. For miniemulsification of other monomers the choice of surfactant is less clear, though it doesn't take a genius to guess that PE-EO or PE-PPO co-surfactants will be effective for polymerisation of alkanes. My attempts to find literature formulations rationalised along these lines failed, partly because most quoted formulations are prepared with little thought along rational HLD or solubility lines. The overwhelming impression is of trial and error and then sticking to something that seems to work, without understanding why.

HLD alone will help miniemulsion formulators. For those wishing to make truly nano-sized emulsion polymers via microemulsions then HLD-NAC is required to master the first part of the process which is to create the microemulsion.

The joy of microemulsions in this context is that they are thermodynamic rather than kinetic. In principle just lightly mixing the ingredients is all that is required to create the microemulsion. The drop sizes are remarkably uniform for no significant effort – when thermodynamics are on your side, life is easy.

After that, things get hard. The problem with microemulsions, in this context, is that the interfacial tension is very low and that there are lots of particles which collide frequently. The system, therefore is one of constant flux which means that the polymerisation process is not as straightforward as we should like.

Whereas miniemulsions start and finish approximately the same size because each drop is more-or-less independent, with microemulsions the drops interact and, in particular, those drops with growing polymer chains tend to attract monomer from other drops – though the extent to which this happens depends on the relative solubility of the monomer in the growing polymer, a factor that ranges widely for different polymers. It is under-appreciated how insoluble many polymers are in their own monomers, though something like polystyrene is very happy dissolved in styrene.

It is unfortunate that the relative solubilities of initiators, monomers and oligomer radicals in the aqueous and (growing) polymer phases produce different outcomes because an elegantly simple approach by Morgan and Kaler[63] had promised to clarify what had been, up to then, a most confusing situation. By choosing hexyl methacrylate as their monomer they were able to make a complete analysis of the micro-polymerisation process, including providing the data to show that the maximum rate of conversion exactly matched the predicted value of 39%, a value that seems curious till it is expressed in mathematical form, $1-e^{-0.5}$.

At the heart of their overall scheme are two facts which lead automatically to a third one. The first is that the monomer microemulsion drops are more-or-less uniform in size, with a radius ~2.5nm. The second is that the polymeric particles are of a constant radius of ~25nm throughout the process (well, from the point ~20% conversion where they are measurable). The third fact is that the volume of a 25nm sphere is 1000x larger than that of a 2.5nm sphere, so at the end of the process there are ~1000 empty micelles (and there were originally ~1000 full micelles) for every 1 polymeric particle.

What this third fact means is that each polymer particle grows surrounded by lots of monomer-containing micelles and can therefore be considered to be growing independently of any other particle. It therefore doesn't matter if a polymer particle starts growing near the start of the process or at the end of the process, it is basically going to see the same general environment in terms of starting and stopping growth, which is why the particles are so uniform. For those who doubt this, a follow-up paper showed that adding more monomer at the end of the process simply created more of the same-sized polymer particles because the monomer was swiftly absorbed into the 1000x excess micelles and behaved as if had been there at the start of the process.

---

63   John D. Morgan, Kate M. Lusvardi, and Eric W. Kaler, Kinetics and Mechanism of Microemulsion Polymerization of Hexyl Methacrylate, Macromolecules 1997, 30, 1897-1905

Another key fact is that many of the polymer particles consist of a single polymer chain – so the MWt is enormous. This means that once an initiator has set off the oligomerisation of some random bit of monomer, the growing polymer chain feeds off the other micelles that are constantly colliding with it. The working assumption of their successful model is that polymerisation continues till the system runs out of monomer, i.e. no explicit termination events are required nor, in this case, are they plausible. I had naively thought that this would mean that a 10x increase in initiator might lead to a halving of the particle radius (i.e. $10^{0.333}$) but instead it just leads to faster conversion with only a slight reduction in radius (33nm to 28nm). It seems that the huge polymer itself reduces the chance of more monomers arriving, so particle size is self-limiting, though I have failed to track down any clear statement on this.

After years of further analysis, Kaler[64] was able to write an updated review that balances the essential simplicity of the nicest cases with the complexities of systems where chain termination can take place back in the aqueous phase and/or where the micelles are not simply providing a linearly diminishing feed of monomers but where more complex partitioning effects take place. To the conclusion that more work is needed to understand the partitioning events that complicate the process, I would add that a focus on the HLD-NAC aspects of the microemulsion would also lead to improvements in the process. What, for example, would be the effect of systematically changing ξ with its dual effects of increasing solubilisation (thereby reducing the amount of surfactant required) and interfacial rigidity?

Looking at more recent papers that cite Kaler's work it is clear that the need for HLD-NAC remains high. With much ingenuity, surfactants are changed, oils are changed, co-surfactants are changed and, not surprisingly, reaction rates and particle sizes change. The authors seem to have no idea what effect they might be having on interfacial tensions, rigidities and intrinsic radius derived from NAC. So a key element of the story is missing.

However, there is progress in two areas. First, the great limitation of the technique, the large % surfactant required, can be overcome using the idea that adding more monomer stepwise through the process mostly replenishes the micelles so that the process continues unchanged. Yes, the particle sizes at the end are rather larger than they might be, but polymer to surfactant ratios of 40:1 can be obtained [65], an amazing result. Second, the ideal of obtaining particles of diameter similar to the original microemulsion can be attained under the right

64   Jennifer O'Donnell, Eric W. Kaler, Microstructure, Kinetics, and Transport in Oil-in-Water Microemulsion Polymerizations, Macromol. Rapid Commun. 2007, 28, 1445–1454

65   Raul P. Moraes, Robin A. Hutchinson, Timothy F. L. Mckenna, The Production of High Polymer to Surfactant Microlatexes, Journal of Polymer Science: Part A: Polymer Chemistry, Vol. 48, 48–54 (2010)

circumstances[66]. For water-insoluble monomers the use of CCTP – Catalytic Chain Transfer Polymerisation can produce the desired result. Put simply, the addition of a specific cobalt complex changes the dynamics of the process described by Kaler. The cobalt complex is able to stop the polymerisation within a particle then migrate to a fresh micelle full of monomers, restarting the polymerisation afresh. So instead of 1 polymer particle for every 1000 micelles, the result (in the best case) is 1 for every 4 micelles and instead of 30-40nm, the particle is ~ 4nm.

The discussions so far have been about O/W microemulsions (Type I) of monomers that were mostly not soluble in water. For water-soluble monomers such as acrylamide in W/O microemulsions (Type II) the general behaviour is remarkably similar and is not discussed further.

There is one final question about polymerisation of microemulsions: what happens when HLD=0? The short answer is that in the best case one gets clear, watery, surfactant filled blobs of polymer which don't offer amazing excitement. There are two obvious problems. First, the polymer can phase separate, creating processes that over-ride the delicate fractal structure of the bicontinuous phase. Second, if the polymer exactly replicates the original structure, it is likely to contain a percolation network of channels. It is not obvious that this offers amazing polymeric capabilities.

## 8.5    Making inorganic nanoparticles

The preceding discussion sets us up for making very small, highly stable (i.e. well-dispersed) nanoparticles. Before doing so it is worth discussing why it is necessary to go to the trouble. As a specific example, let us try to make some barium sulfate nanoparticles without a microemulsion. It seems easy. Take some barium chloride and some sodium sulfate and with strong mixing, plus an adequate dispersing agent, react them together. In terms of the basic physics of particle nucleation the situation is ideal – the solubility product of barium sulfate is extremely low, so nuclei are formed very quickly and lots of nuclei produced very quickly translates into lots of small particles. The problem is that even in the hands of an expert team[67] with access to sophisticated high-shear mixers and excellent dispersing agents, the best that can be done is to get particles down to 40nm, with conventional techniques giving particles in the 100nm range along with the risk of agglomerates if the dispersing agent isn't well-tuned.

66   Niels M.B. Smeets, Timothy F.L. McKenna, The synthesis of translucent polymer nanolatexes via microemulsion polymerization, Journal of Colloid and Interface Science 383 (2012) 28–35

67   Aneta Petrova, Werner Hintz, Jürgen Tomas. Investigation of the Precipitation of Barium *Sulfate Nanoparticles*. Chemical Engineering & Technology. 31, (2008), 604–608, Ying Ying, Guangwen Chen, et. al. *A high throughput methodology for continuous preparation of monodispersed nanocrystals in microfluidic reactors*. Chemical Engineering Journal, 135, (2008), 209–215

Note that the final size (40nm) is similar to that obtained with most of the microemulsion polymerisations – although the system started with lots of small seeds (in the case of polymers, lots of microemulsion micelles) other factors meant that growth of existing seeds always exceeded growth of multiple seeds unless something special (e.g. an equivalent of the cobalt chain transfer agent) could get in the way. Small particles are very difficult to obtain.

But a microemulsion method of creating barium sulfate particles[68] could effortlessly produce them as monodisperse, 4nm particles essentially independent of any mixing process.

The trick is to make a W/O microemulsion of barium chloride and a separate W/O microemulsion of sodium sulfate. Simply mix the two together and the rapid precipitation of barium sulfate takes place at every encounter of the two different drops.

This is a perfectly general principle. Create a microemulsion of an iron salt that is soluble in one oxidation state and add an oxidiser or reductant that changes it to an insoluble state. In this case the oxidiser or reductant can be added to the bulk solution – there is no need to create two separate microemulsions.

This attractive methodology has some obvious drawbacks.

- Microemulsions require relatively large amounts of surfactant and the nanoparticle ends up being stabilised by this surfactant – but in general this is far more than is necessary for general stability and the surfactant might not be ideal for when the nanoparticle is used in subsequent applications.

- Those who don't know what they are doing in terms of microemulsions, i.e. those who have never encountered HLD, will struggle to find the right oil, surfactant, salinity combination to produce a microemulsion with the required drop volume.

Although HLD includes salinity, and although there are methods for correcting for different types and valencies of salts (see https://www.stevenabbott.co.uk/practical-surfactants/HLDExpert.php) there can often be specific interactions between, say, transition metal salts and surfactants.

The story is a familiar one. The basic idea is excellent. Those who try to formulate in this domain are confused because they don't understand HLD-NAC. The extra problems of specific salt interactions add to the confusion. So the field gets the reputation of being "difficult" and "to be avoided". Once the HLD-NAC message filters through to this community, the background difficulty of

_____

68   Dendy Adityawarman. Precipitation of Barium Sulphate Nanoparticles in *Microemulsion: Experiments and Modelling*. PhD Thesis, (2007), Otto-von-Guericke-Universität Magdeburg

formulation will decrease and only the specific issues remain, and these are no harder than others in the nano world.

Here I can speak from direct experience. I once need to make a microemulsion-based nanoparticle for a cosmetics application. My first attempt was successful because I followed a literature recipe. But I had to change the oil from a "scientific" one (decane) to a "cosmetic" one (isopropyl myristate). I looked in the liteterature to see how to do this and was shocked to find that people were relying on instinct, HLB, hand-waving arguments or plain luck. That is when I first encountered HLD. The moment I found the theory I realised that I could formulate rationally and was rapidly able to find the right conditions for making the desired product. It was that need for a usable theory that started me on the adventure of producing Practical Surfactants and writing this book.

## 8.6    Emulsion Separation

A lot of the time we try to find surfactants that will give stable emulsions. Sometimes we need to separate oil that has been emulsified into water (or vice versa) via surfactants that have been deliberately added for other reasons (e.g. in a different stage of a process) or are present as natural contaminants.

If the bad surfactants can be removed by some rational process then the emulsion can be broken. For example, a fatty acid sodium salt can be precipitated as a calcium salt. Judicious addition of a cationic surfactant can catastrophically interfere with an anionic surfactant. But such opportunities are relatively rare. So an alternative possibility is to add just the right amount of the right surfactant.

The trick is the same as that of making a fine emulsion – go to HLD=0 so that the interfacial tension is reduced to a minimum. This makes it easy for the oil drops to coalesce and quickly cream to the surface.

How does one take a process which probably has a complicated oil and an uncertain surfactant system and determine how of much of which surfactant is needed to reach HLD=0? The answer is to use common sense and some low-cost, simple surfactants. If the oil is something like crude oil in a lot of water, it is a good bet that its EACN is high and that the emulsion is O/W – something readily checked with a microscope. The current surfactant is likely to be in relatively small quantities and the effective Cc is likely to be small-to-negative, so adding small quantities of a convenient high Cc surfactant will probably do the trick. In this case AOT would be a safe bet.

On the other hand, if there is a small amount of water inside an oil it is likely to be W/O with HLD>0 so simple addition of a low Cc surfactant such as SLS or a Tween might do the job.

Reality is not as simple as that. Here are some typical complications, all drawn from the crude oil demulsification world because (with large costs for failure) this is where most of the research takes place:

- One key feature of crude oils is that the asphaltenes that act as modest surfactants are capable of forming rather robust phases that can be hard to break down – so an optimal surfactant addition might not lead to rapid demulsification as it takes time for the surfactant to disrupt the asphaltenes. If the asphaltene concentration changes, either because the oil is changing or the % water in the oil is changing then the right amount of surfactant to reach HLD=0 changes. Similarly, if salinity changes the HLD=0 point changes.

- If there is plenty of agitation in the system then an otherwise good demulsifier such as SLS will become catastrophic if it generates foams. In general, ethoxylates are less foaming than simple ionics (see the foam DLVO section).

- The pH of the system can have a profound effect on the amount of effective surfactant if carboxylates are involved – so adding acid to an alkaline systems may help with demulsification.

- To inhibit an anionic surfactant it needs a cation that will form some strong neutralising interaction with the anion. While it is well-known that in general cationic surfactants cause anionic surfactants to crash out, life is never that simple. It is possible to create interesting surfactant systems via anionic-cationic hybrids. So the choice of the cationic surfactant is not obvious. As far as I know, there is no reliable guidelines for choosing them. Because demulsification chemicals are a huge business opportunity, there are plenty of patents around specific cationics.

- Many emulsions are stabilised by hydrophobic particles from the crude oil – i.e. they are Pickering emulsions. The pure surfactant tricks mentioned here therefore cannot work and other particle-specific ideas are required. These can come under the vague word "flocculators" though how the flocculators work is a whole other book. They are often polymeric molecules so they can span multiple particles ("bridging flocculation", see https://www.stevenabbott.co.uk/practical-solubility/depletion.php).

Because the systems requiring demulsificaton are complicated, the package of products required to ensure rapid demulsification (acids, polymers, cationics, Cc controls, salinity controls) will be complicated. Whatever happens, getting to HLD=0 via a rational process is a good starting point, and trying to demulsify with no idea of where the system is within surfactant space is unlikely to be an efficient process.

## 8.7 Bicontinuous Microemulsions

There are many potential uses for crystal clear single-phase, Type III, bicontinuous microemulsions. As specific examples to focus the discussions we will cover their use in "aqueous" cleaners and in cosmetics formulations.

### 8.7.1 Microemulsion solvent cleaners

Solvents can be great for industrial cleaning applications, but are generally perceived as "bad" and are subject to ever-more-restrictive rules on volatile organic content (VOC). Purely aqueous cleaning with modest amounts of the right surfactant (as discussed in the detergency section) can be fine for many soils, but can be useless when faced with soil that requires solvency.

A popular, relatively recent, compromise is to deliver the solvent within an aqueous environment. This is perceived as user-friendly and environmentally benign. If the solvent is fully soluble in water then there is no issue, but it is unlikely to be a good solvent for a water-insoluble oil. If the solvent is insoluble is water then it is hard to use as it requires constant mixing. An effective solution for the water-insoluble solvents is to deliver it as a microemulsion.

Why not deliver it as an emulsion? If the drops are relatively large then the amount of surfactant required to cover the emulsion and stabilise it is relatively modest, a distinct advantage compared to the higher amounts required for a microemulsion. The reason why such emulsions are generally not successful is a direct result of the need to ensure that the emulsion is stable for practical use – the surfactant shell around the solvent drop provides a barrier to effective cleaning. The more stable the emulsion so that it can be shipped and used conveniently, the less effective it is as a cleaner.

Given a solvent of choice, creating a Type III microemulsion is trivial if the EACN of the solvent "oil" is known from a few test scans. Assuming the level of solvent is going to be in the 10-20% range, and using the FishTail app https://www.stevenabbott.co.uk/practical-surfactants/FishTail.php to get an idea of the required Cc and % surfactant for complete solubilization, it is easy to make some test formulations. Reaching this stage without HLD is hard, with HLD it is easy.

The first problem likely to be encountered with early test formulations is their tendency to foam. Some otherwise great microemulsions have had to be abandoned because either in the cleaning phase or (more likely) in the rinsing phase they created vast quantities of foam. When this first happened to me at a customer's site it was highly embarrassing. It may be possible to fix this by adding defoamers, but then these are often silicones and how would you remove last traces of silicones from your cleaned surface? So it is likely that the preferred surfactants are going to be non-ionics which tend to be less foamable,

for reasons discussed under foam DLVO. An attractive alternative is to use kinetically slow surfactants, such as the extended surfactants. These are much more effective (high ξ) in terms of forming a microemulsion, but good foamers tend to have fast surface kinetics, which these molecules certainly don't have. Similarly, some of the more polymeric (or oligomeric) surfactants such as the Pluronics might have a good balance of efficiency and low foaming while not being themselves providing too much "soil" on the surface if they cannot easily be rinsed off at the end of the process.

So life is easy! A few days' work has produced a super-efficient, low foaming, environmentally friendly microemulsion cleaning product. Except there is a problem. If the formulation contains 10% solvent that means that shipping it to the customer means shipping 90% water – environmentally and financially undesirable. So why not ship a concentrated microemulsion which can be diluted in-situ? Everyone is happy. Everyone except the formulator. It is highly likely that the concentrate will be some awful, un-pourable liquid crystal phase. Even if it is OK on its own, there is a very good chance that during dilution to the final working product something nasty will happen (e.g. a hexagonal phase or a complete phase separation) and the user does not get the crystal clear microemulsion they were expecting. Industrial users might be forgiving if there is some modest inconvenience in the dilution process – but if a competitor's product dilutes without problem then you have lost a competitive edge.

There is no easy answer to this because, as mentioned before, there are no available tools for predicting these liquid crystal phases. Relying on hope is not a strategy. So the best approach is to set up a high throughput phase screening system where surfactants, solvents and water can be mixed robotically and the phase behaviours in the domain of interest rapidly assessed. Everyone who tries this for the first time finds it difficult. Adding solvents and water is easy, adding surfactants which can be solids, viscous liquids or concentrated aqueous solutions is never as easy as we might like. Interpreting what happens in the tubes, whether automatically or (semi-)manually is also difficult at first. Once some practical experience has been built up, doing phase scans becomes an efficient and effective way to reach optimal formulations.

For those who are brave, an extra element can be added to the formulation mix. We know that polar oils can have interesting effects on curvature and we can be sure that they will modify behaviour within the phase diagram. By comparing a relevant set of scans across the desired concentration range using solvents containing modest percents of plausible phase modifiers such as long-chain alcohols it might be possible to build up a picture of what does or does not help solve a specific formulation issue. Because the polar oil also changes HLD it is important to make sure that the surfactant blend is tweaked to ensure HLD=0 at the desired final dilution, otherwise the data from other concentrations will be useless as the diluted formulation would be useless. This is all extra work. But being guided by HLD and by the routine use of phase scans the extra work is

always purposeful. You build up a consistent, coherent body of knowledge that can always be traced back to a few key principles. Those who formulate by trial and error have a bunch of datapoints that generally provide no general insight. If trial and error produces a lucky formulation, then when something in the system changes there is no way to rationally tweak the formulation to compensate. If the formulation has been reached via rational principles, then any unexpected change (a different raw material, a surfactant removed from the supplier's catalogue, a change in working temperature) can be corrected by rational means. Add some modest DoE (Design of Experiments) tricks to the process and you have a truly formidable formulation setup.

### 8.7.2  Cosmetic microemulsions

Glittering clarity goes down well with cosmetics users. Microemulsions have the clarity and also have a slight scattering tendency that can look very attractive. But under which circumstances might a consumer ever want surfactant/oil/water blends with the benefits of microemulsions?

We can immediately dismiss one area that is sometimes mentioned. It is often claimed that delivery of actives into the skin can be assisted by surfactants and for those who believe this it seems obvious that a microemulsion might be rather good in this respect. As it happens, almost all the key references on how surfactants might enhance skin penetration are fundamentally flawed and I know of no good evidence (with one exception) showing that surfactants help delivery of skin actives. The one exception is SLS. It happens that for reasons still unclear, SLS is especially good at destroying the top few μm of skin, thereby allowing actives to penetrate more quickly. This is a trick that can be used in pharma where the upsides of delivering a potent medicine outweigh the irritation caused by the action of the SLS. The knock-out argument for why other surfactants do not enhance skin delivery is that a surfactant only provides surfactancy in the presence of water, and most of the water in a typical cosmetic formulation evaporates in the first 10-20min whereas skin delivery takes place over hours. The state in which the formulation starts off, O/W, W/O or a microemulsion makes no difference – at the end the water has gone so what remains is just oil+surfactant.

My own experience analysing confidential real-world skin delivery data is that most surfactants are a positive barrier to delivery because at best they dilute the formulation as it sits above the skin and at worst they provide a nice solubility haven for the active and as the surfactants are too large to penetrate the skin, the active stays in the surfactant on the surface till it is rubbed off by the normal action of humans in touching their skin and transferring the formulation to their hair, clothes, worktop etc. The dilution effect is readily explained. Take a 10% oil in water emulsion where the oil contains the active. Use 5% surfactant to create the emulsion. 5% surfactant isn't much. But when the water has evaporated that 5% surfactant is now 33% of the formulation remaining on the skin, rendering

the formulation totally different from whatever oil+active was envisaged as being good for delivery.

I know of one exception to the above negative take on skin delivery. For those trying to use the follicular route, a well-crafted microemulsion can be just what is needed to penetrate into that oily environment.

To answer our original question, the need for a microemulsion arises because the shampoos and shower gels we use are themselves the cause of a problem. They remove the natural oils our bodies produce and without them the health of the skin and hair deteriorates. As a scientist I once decided to test this hypothesis by ceasing to shampoo my hair. As many others have found, after a week or so of disequilibrium my hair has remained in good condition and needs only washing with water or, in extremis, the lightest touch of a simple shampoo. I haven't "properly" shampooed my hair for more than 2 years.

For the majority who do not want to follow this drastic route (and ruin the fortunes of the cosmetics providers) the loss of natural oil has to be made up by providing oils artificially, even if the replacement itself is labelled as "natural". Cosmetics is a strange world.

Adding a significant amount of oil to a normal shampoo/cleanser results in a cloudy emulsion. The alternative, then, is to create a microemulsion, preferably one with the characteristic glitter.

The scientific considerations are identical to those of the solvent cleaners: get into the HLD=0 zone with the most efficient surfactant (so use HLD and $\xi$ ideas) and solve the "dilution" problem because neither phase separation nor the creation of hexagonal gel phases when the user applies the product is acceptable.

In fact the problem is *much* harder than that. Users make a regrettable association between foaminess and cleanliness (although I know this is a delusion, I find myself making the same association) so the range of usable surfactants decreases rapidly because only high-foamers are acceptable (the opposite to microemulsion cleaners). Add to that problem the need to avoid scaring consumers with sulfates, ethoxylates or, indeed, any "chemical" and formulation space is very restricted. The consumer desire for "chemical free" formulations may be nonsensical, but even the power of cosmetics marketing organisations cannot provide the rudiments of scientific knowledge that show that everything is made of chemicals and that natural chemicals have no intrinsic toxicity advantage over man-made ones.

That is why the HLD and phase scan approach is even more vital. The task is difficult even for a formulator with the relevant skills and a diminished range of

acceptable surfactants and oils. The task is near impossible for those who try to formulate by intuition or by trial and error.

## 8.8    Levelling of paints and coatings

Everyone knows that in order for aqueous paints or coatings to level to a smooth finish it is a good idea to add a surfactant. Surprisingly, the most basic levelling formula, Orchard theory, shows that this is a bad idea.
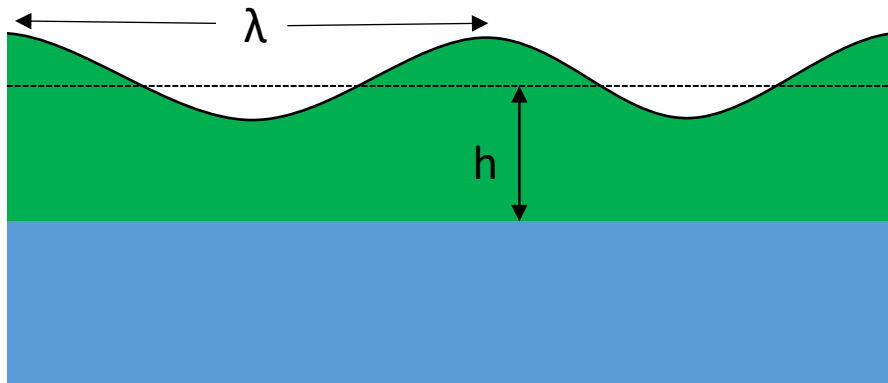


**Figure 8-4 The basic parameters of Orchard levelling theory, the coating thickness h and the wavelength λ.**

For a coating of thickness h, viscosity η, surface tension σ and with a sine-wave irregularity of wavelength λ, the time to level by 1/e is given by:

$$t = \frac{3\eta\lambda^4}{\sigma h^3} \quad \text{8-12}$$

Immediately we see that because the driving force for levelling is surface tension (gravity is several orders of magnitude too small), a higher surface tension levels faster. The app (which is in Practical Nanocoatings) lets you explore the implications of Orchard.



**App 8-5**  https://www.stevenabbott.co.uk/PracticalCoatings/Levelling.php
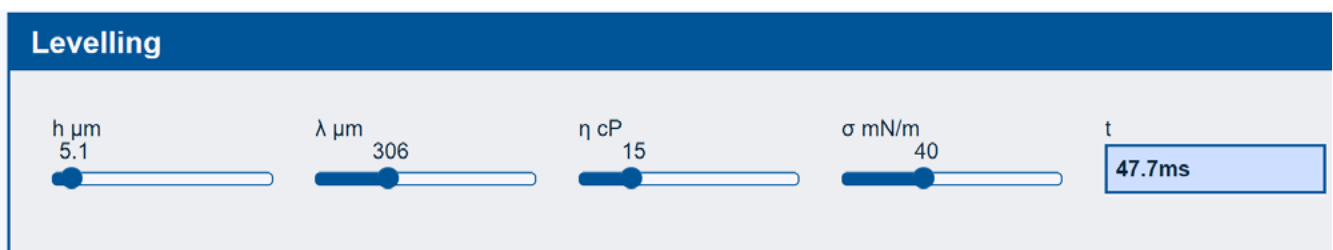
There are some key lessons from this simple formula.

- Because surface tensions of most coatings are going to vary only from 30-60 mN/m there is only a factor of 2 influence of surface tension. So playing

around with surfactants is unlikely to make much difference to this sort of levelling.

- Viscosity of coatings can vary by many orders of magnitude, so pay more attention to viscosity than to surface tension.

- The thickness of the coating makes a dramatic difference. If you halve the thickness then the time to level increases by a factor of 8, which is one of the many reasons thin coatings are hard to produce.

- More important than all the other factors is the wavelength of the defect – doubling this (i.e. making lines or pockmarks wider) increases levelling time by a factor of 16. This is why we should buy expensive paint brushes with fine bristles – the paint marks level out quickly. Cheaper brushes with bigger bristles give coatings that will not level unless we apply more paint (increasing h) – so what you save on the cost of brushes you lose on the cost of paint.

The dependency on λ is critical to practical coatings. Small, sharp defects visible at the moment of coating will quickly disappear. Any broad defects will tend to stay visible while the coating dries/sets. Surfactants can do nothing to help – halving the surface tension would simply double the levelling time, making things worse.

So why do we tend to throw in surfactants when we have bad coatings? There are two entirely separate reasons.

### 8.8.1 Surfactants to fix pinholes

It is common for a coating to acquire some sort of pinhole defect from dust or (more frequently) bubbles. If these pinholes self-heal then they aren't a problem. But sometimes the pinholes expand way beyond the original defect, turning a minor problem into a major one. Whether the defect self-heals or opens into a large pinhole depends on the balance of surface forces and depends on the coating thickness h, the diameter of the initial defect, d, and the contact angle of the liquid with the surface, θ:

$$Hole\ grows\ if\ \frac{h}{d} < 2(1 - \cos\theta) \qquad \text{8-13}$$

This can be explored in the app, again in Practical Nanocoatings, which in addition calculates the velocity, v, at which the hole will open or close, $v = \theta^3\sigma/\eta$:
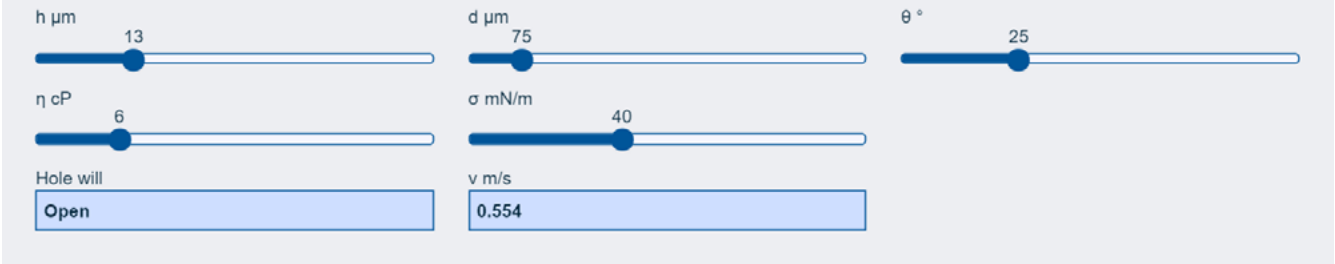
For pinhole free formulations you need thick coatings, small initial defects or a low contact angle, ideally 0°, i.e. fully wetting. There are two important points to raise:

A casually chosen surfactant might be good enough for a thick coating, but when Marketing ask you to halve the thickness it might well be that pinholes start to appear and the surfactant package must be changed.

The contact angle that counts is that at the moment the hole is formed, so your dynamic surface tension behaviour should deliver low contact angles as soon as possible. The DST Choice app https://www.stevenabbott.co.uk/practical-surfactants/DST-Choice.php explains why a "bad" surfactant such as a somewhat volatile acetylenic surfactant can be best for this sort of situation.

### 8.8.2  Surfactants to fix other levelling issues

The previous two sections are useful in their own right but don't touch on the main reason we add surfactants to coatings. This reason is to drown out Marangoni effects.

The word "Marangoni" tends to provoke two reactions. I remember the first time I heard it from a wise colleague who, as soon as I dashed in to tell him that my coating was full of hexagons, said "Ah, you have a Marangoni problem". I had no idea what he meant. The second reaction is to be told that the problem isn't Marangoni, it is Gibbs or Bénard. For those who like such debates please accept that I understand the differences but want to keep things relatively simple by calling them all Marangoni and using the term Marangoni number.

**Figure 8-5 Temperature-created and Composition-created Marangoni cells**

The figure shows what is going on. Over a coating thickness of h there is a temperature difference ΔT (e.g. the surface is cooled by evaporation) or a composition and surface tension difference Δσ (one solvent in a solvent blend is evaporating faster than the other). By chance, different regions have slightly different Δ values leading to small differences δT or δσ over a small horizontal distance. These differences are manifested directly or indirectly as surface tension gradients which then lead to fluid flow from zones of low to high surface tension. This flow has to be fed from somewhere and has to go somewhere and, rather quickly, local circulation zones are produced. These local zones, given a chance, will interact with each other and it turns out that the stable configuration is an hexagonal array of such cells. When you see them in a saucepan of nearly-boiling water you can call them Bénard cells.

The day I created 100m of hexagonal coatings (and much panic and consternation) was the day I discovered the power of tiny differences in surface tension that can get amplified into a highly visible defect pattern. For those who know of such things, such as my wise colleague, the hexagons immediately say "Marangoni". Many years later it was my turn to impress a customer. We were doing some custom coating for them and the product had lots of strange streaks in it. Because I had reason to believe that their solvent blend was unwise in terms of surface tension gradients I took the defects to the microscope. The streaks themselves looked like streaks, but at the start of each one was a neat hexagon. I didn't understand the streak formation, but the hexagons told us that the root cause was surface tension instabilities and the cure was to change the solvent blend. When we did so (using an alcohol more closely matched in boiling point to the other solvent) both streaks and hexagons disappeared.

For those who don't have the diagnostic of beautiful hexagons, the sight of a vaguely regular defect pattern (commonly called orange peel) is highly suggestive of Marangoni.

The tendency to form Marangoni cells based on temperature-driven surface tension changes is captured in the Marangoni number, Ma, that depends on the temperature difference ΔT across the coating, the change of surface tension with temperature, δσ/δT, the viscosity η and the thermal diffusivity α:

$$Ma = \frac{\delta\sigma}{\delta T}\frac{h\Delta T}{\eta\alpha} \qquad \text{8-14}$$

Its meaning is described and the wavelength of the resulting feature is calculated at https://www.stevenabbott.co.uk/Practical-Coatings/Marangoni.php. In essence the number captures the idea that a large temperature change, a large change of surface tension with temperature, a large thickness, a small viscosity and a small thermal diffusivity gives large numbers and therefore a high

risk of Marangoni defects. So thin, viscous coatings are more likely to resist the effect – both for thermal- and for concentration-dependent Marangoni.

Finally we can return to surfactants. The reason that surfactants often fix these defects is that the small temperature/composition differences that lead to small surface tension differences can easily be swamped by the surfactant. So surfactants don't help levelling (Orchard shows that they make matters worse), they stop the formation of uneven surfaces so that levelling is not required.

How does one choose the best surfactant for the task? Just as a "bad" surfactant can be excellent for rapid dynamic surface tension response, so a bad surfactant with a low elasticity is less likely to cause defects of its own. Surfactants can make matters worse if they start to amplify any local defects by being too slow to move around, thereby creating their own surface tension differences. Which is why many dispense with surfactants altogether and throw in a bit of silicone. This just swamps the surface with a rapidly-created, even, low surface tension and Marangoni does not stand a chance.

## 8.9    High throughput

A large amount of high throughput (HT) work turns out to be a faster way of wasting precious resources – chemical, equipment and human. HT without some guiding principle can provide a large amount of data that proves of little or no value. The fact that a robot filled the test tubes and, maybe, did the analyses doesn't make it any the less wasteful both in terms of direct costs and opportunity costs.

So a thread running throughout this book is that a few guiding principles such as using HLD or possessing phase diagrams can help us avoid large amounts of unproductive formulation time, human or robotic.

Assuming, therefore, that you will only unleash HT when you know why you really need it, here are some of the things I have learned in my encounters with the various techniques.

### 8.9.1  Phase scans for Cc, EACN etc.

The first day I spent creating phase scans was rather exciting. As the phases separated and I started to see my first transitions at HLD=0 I was delighted at what they could achieve. By the end of the second day I was making more and more mistakes as the mind-numbing tedium took hold. The first reason for using HT is to increase the reliability of the data. The robot needs to know in advance what it has to do, so the human does the up-front thinking of what is needed. The robot (unlike the human) then reliably does what it is supposed to do.

Not all scans look like the idealised versions in the Measurement apps because all sorts of weird phase behaviours (such as odd phases when SLS and AOT are blended) can occur. Sometimes, confusingly, the middle phase is found at the bottom of the tube, because the surfactant/oil/water phase happens to be denser than the pure water. And when you are exploring uncharted space your guesstimates of where to scan can be way off and you find 8 identical tubes with no hint of a transition.

But for suppliers of surfactants and oils it is not too hard to find a convenient scan system and then batch-to-batch variation can readily be found via the robot so that tweaks to the product can be made before shipping to customers who expect a given Cc or EACN.

When phase scans go well and there is plenty of raw material, tubes holding a few ml of liquid are convenient and the phases can be photographed automatically and phase volumes extracted by image analysis techniques.

What happens when only small quantities are available? It can be better to use the conductivity approach as discussed shortly.

### 8.9.2 Phase diagrams

The easy bit of creating phase diagrams is getting the robot to make the mixtures which span the phase space with few-enough steps to avoid overloading the system and enough steps to get the required accuracy throughout the space. The hard bit is making sure that what you see in the tubes is the thermodynamically stable phase (rather than some kinetic artefact) and then determining what that phase is.

All sorts of tricks are adopted to ensure the right phases are obtained. Tubes are repeatedly ultrasonicated/shaken/vortexed, heated and cooled, spun in a centrifuge and, eventually, left for days, weeks or months to ensure thermodynamic equilibrium. Then the contents of the tubes are checked with polarizing microscope, SAXS, SANS or $^2$H NMR. Given that all that hard work produces just one small slice through surfactant space, it is hardly surprising that few of us can be bothered to create them.

But with a slight change of mindset, the HT approach can produce lots of useful information in a short time. Although it is scientifically fascinating to get a detailed analysis of every phase in a given surfactant space, most of the time we only need to know whether we have the sorts of phases we want in the range of phase space of most interest to us. If we want to know whether there are some evil viscous phases along a phase path of interest, we don't care (much) what they are as long as we know where they are and whether we can get round them. If we want to avoid the problems of phase separation of a microemulsion

during dilution, we need to check if there are any phase changes in the region of interest.

With this mindset, it is possible to build up a lot of useful phase space information relevant to our needs. And if we have plenty of space for tube storage we can leave "interesting" tubes for weeks or months by which stage we will know if the exact phase is of great interest or, if we have totally abandoned this phase space for other reasons, we might just throw the tubes away unexamined. In other words, "good enough" is often good enough.

One other trick gives super-fast phase data, though without exact knowledge of the concentrations involved. It is called the penetration scan because the different concentrations are set up via penetration of a drop of water into a blob of surfactant.
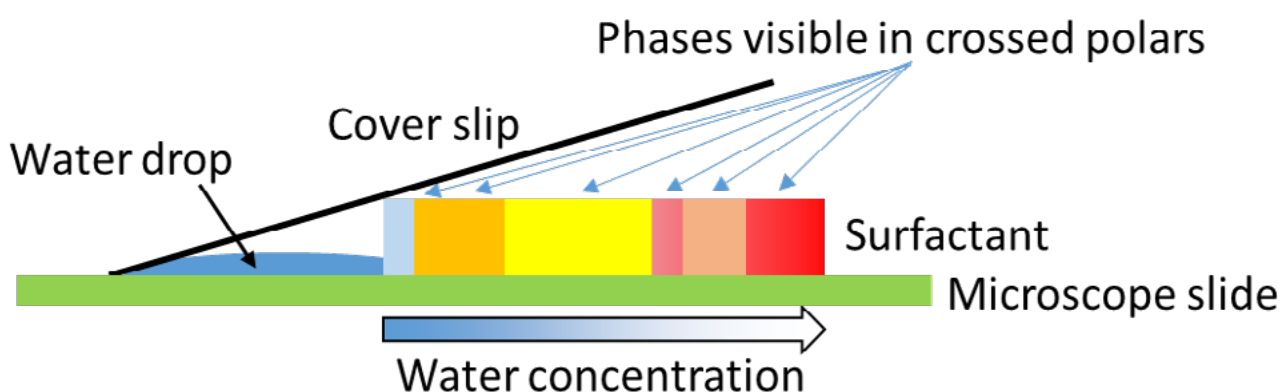


**Figure 8-6 The penetration scan provides phase data across the concentration gradient created by a drop of water placed at one end.**

This trick neatly gives an example of how a lot of information can be obtained quickly using small volumes of test materials. It just requires a bit of imagination and lateral thinking.

### 8.9.3 Small volume tricks

A foolish mistake on my part turned out to be most productive. I volunteered to help with HT identification of interesting surfactants generated by bacteria. The idea of screening hundreds or thousands of tubes containing a wide variety of bacterial strains did not worry me – until it was pointed out that the volume in each "tube" was 200µl. I should have realised the obvious, that these were 96-well plates.

Fortunately I was spared some embarrassment by being told of a brilliant technique for measuring the surface tension in 96-well plates. Dipping Wilhelmy plates into each well is not viable. Instead, those tubes containing a low surface

tension formulation could be spotted via a photograph[69]. The trick is to use polycarbonate tubes where pure water forms a flat surface via a 90° contact angle with the wall of the tube. The lower the surface tension, the more curved the surface of the liquid in the tube. Curvature of the surface distorts the image of a square grid below the tubes. So by measuring the distortion of the grid (using image analysis) the surface tension could be estimated with surprising ease.

As it happens, this ingenious technique has some drawbacks. I proposed instead to create lots of precise discs of some suitable material (a silicone rubber seemed ideal) which would float on high surface tension water and would sink once the surface tension was less than a critical value that depended on the circumference of the disk (force = $2\pi r\gamma$) and its thickness and density (force = $\pi r^2 g[\rho_p h_p - \rho_w h_w]$). I'm sure the technique could be made to work but we got distracted by a more pressing need, to identify high-foam surfactants. Before describing the technique we developed, we need to explore two background issues.

First, consumers claim they want natural surfactants and great foaming performance from shampoos and shower gels. In reality they overwhelmingly choose great foaming and low cost, unnatural surfactants. For those who insist on being more natural, the APG surfactants are about the only realistic choice. Unfortunately their foaming behaviour is not too attractive – they don't foam as well as the common SLES/Cocamidopropylbetaine combination and the foam is rather too "solid" and "sticky". The common experience with real biosurfactants (extracted from vats of microbes) is that they are generally useless foamers, and that the spectacularly good foamers such as surfactin, hydrophobin or latharin are impractical for other reasons. So finding a high-foaming biosurfactant would be exciting.

Second, almost no surfactant on its own provides consumer-acceptable foam performance. Generally the defects of one surfactant are fixed by combining with another surfactant that has a different set of defects. Or, as discussed in the foam chapter, small amounts of a very poor surfactant such as myristic acid can transform a light and frothy "Dawn" foam into a "Gillette" foam, i.e. a very solid, long-lasting foam. So searching for a biosurfactant that on its own gave outstanding foam performance would be a rather hopeless task. The search has to include a set of additives that might, plausibly, change a poor foamer into a great foamer.

Our challenge, therefore, was to do a lot of foam tests on a lot of potential biosurfactants themselves blended with lots of potential co-foamers. Given that industry-standard foam tests typically require a few 100ml of surfactant solution

69  Chien-Yen Chen, Simon C. Baker, Richard C. Darton, The application of a high throughput analysis method for the screening of potential biosurfactants from natural sources, Journal of Microbiological Methods 70 (2007) 503–510

and big flasks, shakers or frits, the task via the conventional route was hopeless as we only had milligrams of biosurfactants.

Instead we hooked up a tiny, controlled air supply to a very fine syringe needle and dipped the needle into a small volume of test surfactant (e.g. in a 96-well plate) and took a photo after, say, 30s of bubbling. The first time we tried this, all 96 samples gave visible bubbling and to our eyes there was not much to distinguish between them. After a few minutes discussing our failure we looked again at the 96 wells and all but one of them showed no remaining foam. Convinced that the one sample with bubbles was a fluke we left for lunch and re-did the experiment on our return. Sure enough, that one sample showed a long-lasting foam and the foam in the 95 others collapsed in a few minutes.

Later on we used this technique in a desperate search for a co-foamer for just about the worst-foaming biosurfactant we had ever seen. It was very easy to screen 10 likely co-foamers, all of which proved to have no effect. A different co-foamer was suggested. The test showed a slight but reproducible improvement – indicating that the crude technique could at least pick out some hope of foam. The next experiment was a rather far-fetched extension of the one that had had a small effect. When the tube bubbled over with a rich foam it was clear that I had accidentally used SLS rather than the biosurfactant, so the experiment was repeated properly – and produced the same result. Subsequent tests using industry-standard methods confirmed that the combination was an excellent foamer. Clearly we were lucky in finding that foamer (the details are being written up for publication), but the point is that adopting an HT mindset and turning a problem (small quantities of surfactants) into an advantage (super-fast screening) can be an interesting way forward.

There are many more such HT techniques to be worked out.

- As mentioned in the discussions about measuring Cc or EACN values via phase scans, instead of using large volumes of liquids and looking for the phase transitions in a set of test tubes, a tiny conductivity probe inserted into a small tube can readily distinguish between a conducting Type I (water as the continuous phase) and an insulating Type II (oil as the continuous phase) providing the system can be adequately mixed via a stirrer or shaker.

- Because we often want to know just the HLD=0 point and because we know that emulsions are their most unstable at this point, a quick scan varying EACN, Cc or salinity might yield, after a few minutes, a bunch of identical-looking tubes that convey no useful information plus one tube showing rapid separation into some ill-defined mess. There is a good chance that that tube has HLD=0. When you first try this technique you might not be convinced by the diagnosis. No problem, just leave the tubes

for hours or days. When you build up confidence, this trick can yield very high throughput identification of HLD=0.

- Those who want to optimize nanoemulsions can use particle sizers, but the difference between larger and smaller nanoemulsions drops can readily be assessed by the scattering colour, so a good lighting setup, a camera and a bit of image analysis can quickly distinguish the no-hoper formulations (creamy white or colourless) from the hopeful ones (some modest coloured scattering), and quickly identify trends across the formulation landscape (a shift to a red colour implies a shift to larger particles). The same trick can be used to monitor the stability against Ostwald ripening of nanoemulsions – simply watch how much or how little the colour changes over time.

- Video technology is available essentially free, so monitoring surfactant phase separations, foam life or nanoemulsions colour shifts can be done routinely – especially if stop-motion videos are taken. It would be easy to follow, say, 10 samples with one camera for one day at 1 frame per minute. Played back at 24fps, 24hrs of video takes just 1min to view. The hard part of getting good videos is not the camera, optics and software – those in a phone are generally good enough. The difficulty is always the lighting. Fortunately, LED lights now make it especially easy to provide strong lighting (at a desired colour) without overheating the sample.

- The stability of a cosmetic emulsion to high shear is an important aspect of its performance. It is surprisingly difficult to create significant high shear on a small sample using normal equipment. A remarkably effective HT method is to place a small drop on a microscope slide, put on a cover slip then slide it back and forth with a finger. If the thickness of the sample is 10μm and the sliding speed is 1m/s then the shear rate is $10^5$/s. Aside from its speed and simplicity, one advantage is that any damage to the emulsion is readily analysed by comparing the before and after images under the microscope. Another advantage is that this high shear is exactly what the formulation experiences when it is rubbed onto the skin.

### 8.10 Surfactants as Solubilizers

This section is tagged on a little uncomfortably as it has only tenuous links to the bulk of the book. However, it is a significant use for surfactants within pharma (the Tweens, for example, are regularly used) so is worth discussing. In this section the word "solubilizer" is used as a purely neutral word to describe the fact that the solubility of an insoluble molecule is increased. The name, fortunately, does not imply a mechanism. Indeed there are many different mechanisms that cause much confusion in the literature. Some of the ideas are explored in https://www.stevenabbott.co.uk/Practical-Solubility/HydrotropesIntro.php but here we consider the rather clear case of solubilization within surfactant micelles. Although the case is clear, the typical explanation of how the systems work is wrong.

The problem to be solved is the insolubility of many pharmaceuticals in water. The sloppy argument goes: "The drugs are hydrophobic, so they are insoluble in water, therefore we can dissolve them in the hydrophobic core of a micelle." The obvious flaw with this oft-stated argument is that there is no reason why many of these drug molecules should be soluble in, effectively, dodecane if the surfactant has a C12 chain. The app will show another reason why this naive idea is false. From the experimental increase in solubility it is able to calculate what percent of the core volume would be occupied by the solute. In many circumstances this is >100%, i.e. an impossible amount. Both arguments point to the truth: that the drug is often solubilized in some other region of the surfactant. Those familiar with Hansen Solubility Parameters will have no problem seeing that the large EO head of Tweens is a rather interesting solubility regime. A calculation similar to that of percent in the core allows you to see what percent of the surface area would be taken up with the drug. The experimental evidence is that this is satisfactorily below 100%.

There are two aims for those using solubilizers. The first is to increase the solubility of the drug. The second is to find some rational method for selecting the optimal solubilizer. My experience within the pharma industry is that it is largely a process of trial and error because the science is not well understood. The app, unfortunately, does not help directly in optimisation, but it helps to indicate where the solubilization is taking place which, in turn, might suggest other formulation options. The app is inspired by papers[70] from Bhat, Dar, Das and Rather at U. Kashmir



**App 8-7** https://www.stevenabbott.co.uk/practical-surfactants/Surfactant-Sol.php

---
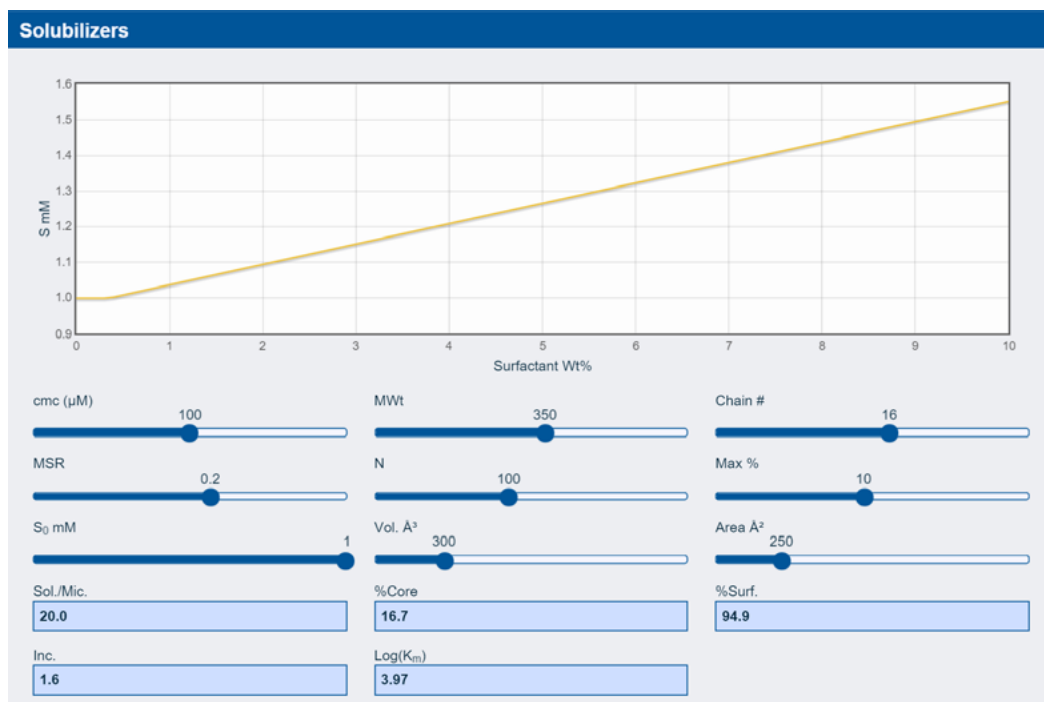
70  There are 2 other papers that can be found starting from this one: Parvaiz Ahmad Bhat, Ghulam Mohammad Rather, and Aijaz Ahmad Dar, Effect of Surfactant Mixing on Partitioning of Model Hydrophobic Drug, Naproxen, between Aqueous and Micellar Phases, J. Phys. Chem. B 2009, 113, 997–1006

Let us go systematically through the inputs. No solubilization takes place below the CMC because the process involves micelles. So the CMC is required. Because we tend to work in Wt % and CMC is in µM we need the MWt of the surfactant for the calculations. The size of the surfactant tail will affect the core volume so is specified here in terms of the number of carbon atoms in the tail. In any experiment the Molar Solubility Ratio (MSR) is determined experimentally. Here it is an input so you can simulate the experiments. The number, N, of surfactant molecules within a typical micelle clearly affects the solubilization capacity. For convenience, the app gives a list of surfactant properties for a number of common surfactants. Max % scales the X-axis of the graph and is the concentration used for calculating the increase in solubility. Because we are interested in the ratio of solute to micelle core volume and head area we need tbe reference solubility, $S_0$, of the solute in water. Finally, for the inputs, we need the volume and surface area of the solute which can be estimated respectively from molar volume and your favourite molecular drawing package.

The MSR is derived from the experimental solubility $S_S$ at surfactant concentration C as $S_S - S_0 = MSR(C-CMC)$, i.e. it is the solubility ratio corrected respectively for the solubility in water and for the CMC.

The calculations are straightforward:

- Solute/Micelle is simply N.MSR

- Core Volume and %. We know the volume of the core from the tail length (via the Tanford equation for length, l, from number of carbons), so $V=4/3\pi l^3$. The volume of the solute is an input (it is MVol/Avogardro) so the volume ratio can be calculated. Clearly this (and the area calculation) is a simplistic measure, but it provides a lot of insight for little effort.

- Surface Area and %. We know the surface area of the micelle from the tail length, $A=4\pi l^2$. The surface area of the solute is an input so the ratio of surface areas can be calculated.

- The Solubility Increase is the ratio of solubility at Max % to $S_0$.

- $K_m$ is the effective partition ratio of solute between water and micelle and is shown in log() format as it is generally a large number. It is calculated as $K_m = MSR/(S_0*0.01805*(1+MSR))$

The authors used this simple technique to describe the behaviour of three solutes. The first was naphthalene where 16 molecules were incorporated per micelle, a modest increase in solubility, so the % Core value is 8%, plausible for a solute which is most likely located within the core as per the classical view of surfactant solubilizers. For effective surfactants such as Brij nonionics, naproxen which is a substituted naphthalene, the % Core and % Surface are both high which conforms to our intuitions that naproxen will tend to be partly in the core and partly in the head. For erythromycin which is definitely *not* soluble

in dodecane the % Core value is laughably high, 300%, but the % Surface Area is reasonable.

Of course, the question of where a solute resides can be resolved by various sophisticated analytical techniques. Yet most of the time when we are trying a bunch of solubilizers we can easily get the data to give us MSR and all the other values, so the estimate of % Core and % Surface come almost free as a useful guide for discussions.

It is interesting to consider the impact of CMC and MWt for solubilizers. A low CMC gets the solubilization off to a good start. But as we tend to be interested in Wt% addition of surfactant, if a low CMC requires a large MWt, the gains in CMC are offset by the need to add more weight of surfactant. Similarly, a larger tail will give more Core volume and Surface area which should help, but again might give a higher MWt which offsets the gains. The app lets you play with these trade-offs.

The problem with solubilizers is that at a few % addition the increase of solubility is generally modest (factors of 2-5), yet adding high % of surfactants is generally seen as unattractive. This is because, for reasons as yet unclear, typical MSR values are in the range of 0.2.

The search for ways to predict the best solubilizer for any given system remains frustrating. This simple app at least provides some modest insights for modest effort.

## 8.11 The future of surfactant formulation practice

Everyone has biases about the best tools to use for rational formulation. It might be the case that many of my biases will turn out to be sub-optimal. If so, that is how science works and I hope that readers will help me to spot the error of my ways.

But I think that the some key principles will stand the test of time so as a formulation community we should follow them more attentively than we have in the past:

- Partition of surfactant from water to surface (K) and from water to oil is fundamental to most aspects of surfactancy yet is not routinely studied and is an under-utilised part of the formulation toolkit.

- More attention to Langmuir isotherms (with all three parameters, $\Gamma_m$, K and A) measured in our own labs or provided by surfactant suppliers will shift our focus from CMC to more fundamental properties.

- Elasticity should become a key part of thinking about many surfactant issues. There are many issues with understanding and measuring elasticity, but ignoring the issues does not help.

- HLB will die – not fast enough, but its demise is inevitable. Wise formulators should wean themselves off it sooner rather than later.

- HLB will be replaced by the best practical "curvature" tool. Today that is HLD-NAC but tomorrow it might be Helfrich torque, PP or some equivalent unknown to me. Pressure should be put on surfactant suppliers to supply the best-available parameters (which currently are Cc values) to enable us all to formulate more intelligently. Even if Cc is shown not to be the best number, because of the interconnectedness of curvature theories, the new numbers will readily be calculated from Cc values.

- Editors and conference organisers should strongly encourage papers to provide enough information that curvature-related results can be transferred to similar systems. The insights from a paper which has Cc, EACN, salinity and temperature data can be readily applied to other cases. Without those data the results are, sadly, useless for the rest of us as they cannot be applied to other situations. It is generally impossible to data-mine the surfactant literature because there is too little information from which extrapolate the results.

- The "polar oil" problem will be resolved sooner or later, presumably within the context of whichever is the best curvature tool. We should all do our bit to encourage academia to help solve the problem.

- Intelligent high throughput techniques using small volumes and various "good enough" tricks will enable us to better explore relevant parts of surfactant space. For example, the effects of small amounts of additives (which we know to be important) can be better quantified through HT techniques. Thus polar oils can be better understood in terms of effects on curvature, on $\xi$ and on kinetics only through intelligent exploration in the right parts of surfactant space. Similarly, HT foam scans can reveal important data on foamability, foam lifetime and rheology.

- Phase diagrams of relevant parts of surfactant space will become more normal once we have the HT tools to create them.

- Formulators will increasingly replace hand-waving arguments with numerical results. The apps here are one person's attempt to provide a set of useful tools. Modern technology makes it so easy to create them that it should become unacceptable for an academic paper *not* to offer an on-line implementation of their insights. In the past, complex numerical integrations required academic-level computational resources. Now they can be done on a smart phone. Why would an academic *not* want to see the theory brought to life? Large chunks of the literature lie useless because most of us don't know how to implement it. I've now created an App-Writer tool

that allows anyone who has just rudimentary Javascript programming skills to create their own powerful apps with all the sliders, graphs, responsive design capabilities of my own apps. Indeed, I now write all my new apps using my own App-Writer.

When these sorts of approaches become the norm, surfactant science will enter another golden age. It seems to me that the move towards this way of working is unstoppable. My hope is simply that the changes happen sooner rather than later and that this book can play a modest role in speeding up that change.